AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Clinical Microbiology®

MINIREVIEW

Towards a Universal Molecular Microbiological Test

Richard J. N. Allcock,[a,b] Amy V. Jennison,[c] David Warrilow[d]

School of Biomedical Sciences, University of Western Australia, Nedlands, Western Australia, Australia[a]; Translational Cancer Pathology Laboratory, Pathwest Laboratory Medicine WA, QEII Medical Centre, Nedlands, Western Australia, Australia[b]; Public Health Microbiology Laboratory, Queensland Health Forensic and Scientific Services, Archerfield, Queensland, Australia[c]; Public Health Virology Laboratory, Queensland Health Forensic and Scientific Services, Archerfield, Queensland, Australia[d]

ABSTRACT The standard paradigm for microbiological testing is dependent on the presentation of a patient to a clinician. Tests are then requested based on differential diagnoses using the patient's symptoms as a guide. The era of high-throughput genomic methods has the potential to replace this model for the first time with what could be referred to as "hypothesis-free testing." This approach utilizes one of a variety of methodologies to obtain a sequence from potentially any nucleic acid in a clinical sample, without prior knowledge of its content. We discuss the advantages of such an approach and the challenges in making this a reality.

KEYWORDS bacteriology, bioinformatics, diagnostics, genome analysis, high-throughput sequencing, hypothesis-free testing, next-generation sequencing, personalized medicine, virology

Accepted manuscript posted online 23 August 2017

Citation Allcock RJN, Jennison AV, Warrilow D. 2017. Towards a universal molecular microbiological test. J Clin Microbiol 55:3175–3182. https://doi.org/10.1128/JCM.01155-17.

Editor Colleen Suzanne Kraft, Emory University

Address correspondence to David Warrilow, David.Warrilow@health.qld.gov.au.

High-throughput sequencing (HTS; also known as massively parallel sequencing or next-generation sequencing), whereby massive amounts of data are generated in parallel, is a technology that has come to the fore over the last decade, and has been broadly applied in molecular biology and increasingly in medical settings with spectacular successes. There are three fields in which high-throughput genomics can make an impact: inherited diseases (mostly Mendelian), somatic disease (i.e., cancer), and general microbiology. While there has been substantial progress in implementing diagnostic approaches in inherited diseases and cancer, the application of whole-genome sequencing to microbiology diagnostics is still in its infancy. However, it should be possible to make rapid progress in this area as many of the lessons from the other fields are applicable to microbiology. In essence, the three fields have distinct but overlapping requirements. Diagnostic inherited genomics requires the detection of high frequency alternative alleles, usually in the form of single nucleotide substitutions or small insertions/deletions, at a given locus (i.e., hetero- or homozygous), most often from an easily obtainable starting material such as blood. Diagnostic cancer genomics requires the detection of intermediate and very low frequency alleles in a background of a highly variable amount of normal alleles, often from specific tissues that have undergone substantial chemical modifications through the process of formalin fixation, simultaneously complicating the process and making it much more expensive. Ultimately, diagnostic microbiology will require the sensitive detection, identification, and quantification of nonhuman sequences in a background of human DNA, as well as the detection of genotype-identifying alleles of those microbial sequences. These specific requirements will pose some unique challenges.

HTS-based diagnostic microbial detection offers at least five advantages over the classic microbiology paradigm. First, it is "hypothesis free" and makes no assumptions about the patient's condition; hence, it is less prone to error resulting from the lack of knowledge or experience of the requesting clinician. Ultimately, this will change the

role of clinicians from being "test selectors" to being "data interpreters." Second, detection is not dependent on prior knowledge of a microorganism. Hence, intrahost sequence variation of pathogens and previously undescribed microbes can be detected. Third, changes in strain sequence will not affect the diagnostic outcome as dramatically as they do with some current tests. For example, PCR primers and probes may be destabilized by a small number of nucleotide differences to their intended target. This would be particularly advantageous in the early stages of an emerging pathogen, when sequence information is minimal or nonexistent. Fourth, multiple agents across the full microbial spectrum contributing to disease can be simultaneously detected. Finally, HTS-based methods will allow the identification of nonculturable microbes. Culture has been the historical cornerstone of microbial pathology and is likely to remain important in specific circumstances. However, there is an increasing recognition that nonculturable organisms are also important pathogens. The culture-independent methods such as HTS offer the possibility not only to detect such microorganisms by obtaining microbial sequences from the clinical specimen but to also simultaneously substantially broaden our understanding of microbial pathology more generally.

In addition to these specific benefits, there are other consequences of such an approach. The additional data gained using HTS-based methods could provide information on virulence factors and antiviral resistance in the case of viruses (1) and on bacterial virulence-related genes and antibiotic resistance (2). From a laboratory management stance, a combined HTS-based test would be simpler, possibly cheaper, and in many cases, more time efficient than maintaining a battery of single-target tests. Human genetic testing has already set a precedent by combining multiple genetic tests into a single test using HTS technology (3). Further, as a clinical sample may include host nucleic acid (DNA and RNA), this new approach may ultimately be expanded to include simultaneous determination of the host's genetic background and immune response. Such applications have been broadly referred to as "precision medicine" or "personalized medicine," and while much of modern medicine has always been somewhat tailored to individuals, the increase in information content obtainable with modern genomic methods holds the promise of dramatically improving the level of personalization. As applied to diagnostic microbiology, this information will be useful when the prognostic outcome of an infection is dependent on the genetic background of the patient (4, 5) or when infection results in the activation of antimicrobial pathways such as interferon responses.

While there are clear advantages for an HTS-based genomic diagnostic approach, it is more complex and technically challenging than classic microbial molecular diagnostics. It also offers different advantages to laboratories depending on where they sit on the pathology spectrum, from routine medical microbiology to public health laboratory. In this review, we would like to consider the hurdles that this relatively new technology must overcome in order to become routine in the clinical microbiology laboratory. These will be discussed in 6 broad areas: (i) sample preparation and turnaround time, (ii) sensitivity and cost competitiveness, (iii) validation, quality control, and contamination concerns, (iv) bioinformatics, reporting, and interpretation of clinical significance, (v) ethical considerations, and (vi) future technological developments.

SAMPLE PREPARATION AND TURNAROUND TIME

HTS platforms require the construction of a library, usually to clonally amplify template nucleic acids and also incorporate ubiquitous primer sequences when using sequencing-by-synthesis. With current sequencing methodologies, this is a significantly expensive, time-consuming, often labor-intensive, and currently unavoidable part of the procedure. For example in our hands, manual cDNA synthesis (if required), library preparation, and sequencing of the most common platforms can be done in less than 48 to 72 h. Library preparation methods are increasingly automated, but in general, this either maintains or slightly increases the overall cost, and almost never reduces the overall time taken. By comparison, a standard PCR has a manual preparation time of a

matter of several minutes and total process times of substantially less than a single day. Given this, any HTS platform that takes longer than approximately 4 h is not competitive in terms of time with current routine amplification testing. Hence, sample preparation and sequencing are considerable time bottlenecks in preventing the wider adoption of routine HTS in microbiology laboratories. Technical developments will need to focus on streamlining, or preferably eliminating, library preparation and further reduce sequencing steps to reduce turnaround time. At present, there is substantial focus on the reduction of the scale of library preparation (i.e., being able to start from ever smaller amounts of starting material), and this miniaturization has the potential to substantially reduce library preparation costs but also time, if novel nanoscale approaches are utilized. In addition, there are efforts to reduce the clonal amplification steps for the current generation of equipment, but it is not yet clear whether they will reduce the overall process time enough to be comparable to standard techniques. As sequencing technologies continue to advance, there may even come a time when the clonal amplification step is eliminated altogether, as in some long-read/nanopore-based approaches (6, 7), and eventually, these might conceivably result in extremely rapid laboratory processes.

SENSITIVITY AND COST COMPETITIVENESS

One of the major hurdles that a universal microbiological test will have to overcome is its cost compared with routine PCR-based single-microbe testing. HTS has so far had the biggest impact in the diagnosis of Mendelian inherited disease and increasingly is making impacts in the cancer/oncology field. However, in those fields, the standard tests available were either relatively expensive to begin with or relatively uninformative by themselves, resulting in the requirement for many individual tests to be performed and a large per-patient cost. Hence, the introduction of HTS has seen rapid decreases in the cost of testing in these fields, combined with significant increases in diagnostic yield. Current microbiology testing differs in some important ways. For example, in our experience as a relatively low-throughput public health diagnostic laboratory, routine PCR-based testing is currently quite cheap (we estimate about US$30 per sample, including labor, automated extraction, amplification, data analysis, and quality controls), and we assume this to be a common experience of other similar-scale microbiological laboratories; for high-throughput microbiology laboratories, costs are considerably lower due to economies of scale. Hence, this presents a substantial hurdle to most current HTS approaches, which are likely more expensive than traditional approaches, albeit with large variation depending on the chosen platform, the desired number of reads, the level of sample throughput, and ancillary costs.

Whereas the costs of routine PCR are low and fixed, a complicating issue for HTS is that cost is proportional to the amount of data generated, and the amount of data generated is a direct determinant of sensitivity (1). One factor not at play in conventional testing is that in HTS methods, the proportion of total reads that match the target relative to the host (i.e., human and commensal flora) will influence the choice of read numbers. Given that most microbiological tests have to be exquisitely sensitive, typically large volumes of the current short-read technology sequence reads will have to be generated as the number of microbial reads may only represent a very small proportion of the total sequence data. This is a major consideration when determining sensitivity and cost relative to more conventional amplification-based assays and will most likely be highly dependent on the specific anatomical site being sampled. For example, samples taken from throat swabs are likely to contain smaller amounts of contaminating human DNA than blood samples, and this will influence the required number of reads to achieve the required sensitivity. Hence, it may become necessary to develop methods to measure the amount of human DNA in a sample and use this information to determine the required number of reads in an individualized sample-specific manner. Another issue to consider is that this determination of minimum read number relates only to an ideal unbiased library preparation. "Random" preamplification methods are prone to stochastic and systematic biases (8–12). Laboratories should seek to

use methods as free of bias as possible. Only then can the read number be related to the probability of target detection.

One way to determine the minimum read number required is to do so empirically with clinical samples in which target material has been detected using HTS, in comparison with a test of benchmark sensitivity such as amplification-based methods. Where the sensitivity of HTS methodology approaches or is equivalent to that of standard amplification, it can be the basis of a "rule of thumb" for minimum read requirements. It is highly likely that this will vary with sample type (effectively a proxy for the proportion of "contaminating" human DNA), the size of the genome of the microorganisms present, and to a lesser extent, the choice of sequencing platform. For example, in one study, matches to viruses were as low as approximately one in a million reads (13). Similarly, studies have demonstrated that between 0.016% and 2% of the total reads are sufficient to reproduce a bacterial detection by nucleic acid amplification testing (14, 15).

How then does a universal sequence-based method compare with the routine testing in terms of cost? Studies showing amplification-equivalent sensitivity for virus detection have sequenced ranges from 1.5 to 3.5 million reads or 25 to 45 million reads (16), 5 million reads (17), 7 million reads (18), and up to 15 million reads (13). Using an example of 10 million reads, reagent costs for the most commonly used Illumina platform were estimated to be as high as $350 to $750 (MiSeq) to as low as $20 to $30 (HiSeq) (http://www.molecularecologist.com/next-gen-fieldguide-2016/). Hence, the sequencing-based diagnostic paradigm is approaching viability at the lower end of these costs; albeit, these costs do not include additional requirements such as labor, extraction, library preparation, amortization of the instrument, and controls. Of these items, the cost of library preparation is likely to be most significant and will add substantially to the overall cost of sequencing microbial genomes. The NIH National Human Genome Institute regularly updates sequencing costs, which include these additional items, and this currently stands at $21 per 10 million reads based on a 150-nucleotide (nt) read (https://www.genome.gov/sequencingcostsdata/). However, this estimate is based on sequencing a very large human genome, where the absolute cost of the sequencing is large compared to the library preparation costs. Microbial genomes are far smaller and hence will incur lower absolute sequencing costs, but the library preparation cost will likely remain identical. Hence, the overall cost per read in microbial sequencing could be 10-fold or more greater than that for sequencing a human genome. As is shown on the NIH National Human Genome Institute website, sequencing costs have annually decreased faster than the so-called "Moore's law" of the computing sector, which has become the yardstick by which the speed of technological developments are judged, and single-molecule-long read technologies are already demonstrating that library preparation can be simplified. While most pathology laboratories are not positioned to benefit from the economies of large-scale sequencing programs, progress to date indicates we might reasonably expect that further cost reduction is likely to make the new paradigm cost competitive in the next several years.

VALIDATION, QUALITY CONTROL, AND CONTAMINATION CONCERNS

Any laboratory test should be properly validated, and microbial HTS is no exception. Standard tests have one target per test. How then are determinations of sensitivity and specificity to be made with a method that can detect potentially all manner of microbial genomes? Such questions are only now being properly considered. For the implementation of microbial HTS, the widespread and uniform adoption of validated and properly controlled tests will be needed. It is not in the scope of this review to cover this topic in any depth. Instead, readers are invited to read an excellent article on the subject by Schlaberg et al. (17). To briefly summarize the issue, first, methods should be established using spiked samples and the limit of sensitivity determined. Second, the sensitivity of the sequencing methods should be determined using various specimen types known to contain microbes with a range of different representative genome sizes and types: double-stranded DNA and single-stranded and double-stranded RNA. The

specificity of the bioinformatics component should first be determined initially *in silico* using artificially generated data sets and confirmed using samples similar to those used to determine the sensitivity.

Quality controls such as a microbe-free sample (negative control) and a spiked sample (positive control) should be included routinely. For spiked samples, a microbe outside the normal experience of the laboratory could be included to avoid potential contamination concerns. Additionally, internal controls may be added at an appropriate dilution to clinical material to further ensure the functioning of enzymatic steps. The outcomes of these controls will need to be recorded along with the performance of the run, including run yield, nucleotide quality scores, cluster density (if Illumina), and other relevant information (19). As results are obtained by sequence matches, the paradigm is entirely dependent on the integrity of the database. Hence, quality documentation should make a provision to record the source and exact build of the database used. All these requirements impose an added burden on the laboratory technician, and so the more these controls and documentation become incorporated into the protocols and software, the easier this technology will be to use and the wider its eventual adoption.

One issue that has been particularly problematic for the field of HTS in general is contamination (20, 21). The technology is notoriously contamination prone for a number of reasons. First, there is usually at least one amplification step in library preparation, and often more. Each amplification step increases the chances of contamination. Often, common primer sequences are used for amplification between samples and between separate library preparations. Due to the large numbers of reads sequenced per sample, on the order of millions, the detection of contaminating reads in the final library is exacerbated. Hence, much like the early days of PCR, the controls discussed above, and additional modifications to methods must be put in place to minimize or detect contamination.

Most laboratory technicians will be familiar with contamination minimization methods, such as having a linear workflow to prevent amplification product from entering reaction mix preparation areas, eliminating or minimizing samples with high template content, and using commercial library kits, disposable plastic items, filter tips, and molecular-biology grade water whenever possible. As many microbial pathology laboratories routinely perform PCR, they would be familiar with procedures to minimize contamination. However, one of the lessons of the implementation of HTS approaches in other fields has been that contaminants are extremely difficult to avoid absolutely. New approaches to reduce contamination will be required, such as the use of multiple sample barcodes/indices as seen in approaches involving ancient and degraded DNA (22), such as a regular rotation of barcodes/indices, meta-analyses of previously obtained results in the laboratory (for example determining whether a "positive" result for a rare pathogen could be explained by contamination from a sample recently analyzed in the same laboratory), and improved separation of preparative processes in the laboratory to ensure the minimization of contaminants.

BIOINFORMATICS, REPORTING, AND INTERPRETATION OF CLINICAL SIGNIFICANCE

HTS-based diagnostics are more complex, not only because of the vast amounts of data that must be dealt with in a practical period of time but also because the significance of the large number of matches obtained by the sequence analysis needs to be determined. Fortunately, a number of publically or commercially available software "pipelines" are now available to assist in the rapid analysis of large HTS data sets for diagnostic purposes. They all aim to simplify and expedite the processing and analysis of large amounts of sequence data. They typically have the following features: prior processing of raw reads to facilitate analysis for some pipelines, contig assembly, nucleotide or protein database sequence matching of contigs or raw reads, and presentation of the outcome in a suitable report format. The report should summarize the statistical significance of any match, the number and proportion of reads, and the taxonomic level this pertains to, for example, phylum, family, genus, or species. The

pipeline may also include genotyping or other analyses. These packages, used in combination with in-house or cloud-based computer clusters, have facilitated the ease and speed with which large sequence data sets are managed. In terms of making HTS more routine, the issues are that many pathology laboratories do not currently have access to the bioinformatics support to install and maintain such software pipelines and the hardware required to run them nor to the storage of the accompanying large sequence data sets. In this regard, improvements in computer hardware and software may circumvent some of the need for high-performance computing. For example, the recent development of the program DIAMOND has increased the speed of BLAST-like protein matching 20,000 times without a loss of sensitivity and with the capability of using large raw read files (23). While improvements in interfaces have also been made, much of the software features are accessed through command-line functions. Further improvements in user interface and in the simplification of reporting will be crucial in bringing the technology within the range of difficulty for those pathologists with more-limited computer capabilities. In addition, greater communication between scientists, pathologists, and clinicians will be required for the optimal interpretation of the complex output generated by these analyses.

As with the results of any pathology test, the report of the bioinformatics pipeline must be interpreted by a microbiologist for its clinical significance. Where the results conform to the identification of a single predominant microbial pathogen consistent with the clinical picture, this will be relatively straight-forward. Where matches are obtained to microbes where a causative relationship to the patient's symptoms has not been clearly established (24), this will be more complicated. This is problematic for pathology laboratories where there is constant pressure to provide accurate and easy-to-interpret results at a high turnover. The simplest way to deal with this issue is to enable the laboratory to set the type of information that they are comfortable extracting from the sample (i.e., routine diagnostics or exploratory), prior to the commencement of the analysis pipeline. Hence, designers must incorporate these opt-out and opt-in features into their software interfaces.

ETHICAL CONSIDERATIONS

Ethical guidelines come in to play when human specimens are used for research purposes. Ethical considerations for current routine microbiological testing are well established and, therefore, generally straight-forward. However, the specific hypothesis-based paradigm means that there are relatively few incidental findings. In contrast, HTS-based microbial diagnostics are inherently more complex and have greater ethical implications for the clinician, especially given they will generate a sequence for whatever DNA is present, including human DNA. In addition to the privacy considerations that this obviously raises, the potential to discover incidental findings unrelated to the original presenting condition is also possible (25–28). No consensus has yet emerged as to how to effectively deal with this issue; however, this must be adequately addressed if HTS diagnostic applications are to become routine. To date, there has been substantial discussion of ethics in the inherited disease and cancer fields, with a number of approaches having been trialed and then either discarded or modified (28, 29). Much of this experience will be relevant to the ethical issues around microbiology testing, and although beyond the scope of this article, we encourage readers to become familiar with this literature.

Where incidental findings relate to a genetic condition of the patient, one suggestion has been to filter out host sequences *in silico* to make any such findings very unlikely (27). One problem with this approach is that the pathologist needs to be confident that host sequences are properly filtered and that any filtration does not erroneously remove microbial reads. This can be challenging when the size of the human genome is considered and the chances of mismatch are high. Where incidental findings relate to a microbial diagnosis unrelated to the patient's presentation (e.g., HIV detection when another virus is suspected), this will present a major ethical dilemma for the clinician that is not easily resolved. Theoretically, a filter could be applied by the

software at the level of results reporting. However, this bypasses the question of what ethical and legal obligation the clinician has to disclose such an incidental finding to a patient (30). No doubt, consensus solutions to these difficult problems will eventually emerge. Until that time, ethical review committees will need to deal with such issues on a case-by-case basis.

TECHNOLOGICAL DEVELOPMENTS

As discussed above, continuing improvements in sequencing technology are expected to further reduce costs and eventually make this new diagnostic paradigm a reality. There are two possible means by which this might be achieved. First, increased numbers of reads from existing platforms, in combination with multiplexing of samples by barcoding, will reduce the per-sample sequencing cost. Second, the commercialization of fundamentally new chemistries enabling dramatic reductions in per-nucleotide cost on a smaller scale, and importantly, on more-affordable devices, may challenge established platforms. One technology that has been suggested will enable the latter is "nanopore"-based sequencing. This technology uses either modified biological or solid-state nanometer-scale channels to interrogate DNA or RNA molecules as they pass through the channel (31, 32). Biological nanopores are more advanced and have progressed to a quasi-commercial stage with the MinION from Oxford Nanopore Technology, which has already been used for the detection of microbes (6). Other advantages of this format are the small initial capital outlay, relatively simple sample processing, and also the long reads that can be generated from a single molecule. These advantages go some way to addressing the issue of cost competitiveness. However, the technology has been criticized for its high error rates and high per-nucleotide cost in comparison with more established HTS platforms (33). To circumvent the read error issue, researchers have utilized the much greater depth offered by amplicon sequencing (34, 35). However, chemistry improvements have reduced error rates significantly (36), leading to applications without prior generation of amplicons (6, 37). Solid-state nanopores are currently at the research stage and promise greater flexibility in pore design and stability and mass production potential (31). Further technical developments addressing these issues will determine whether this technology is widely accepted for laboratory-based microbial pathology.

REFERENCES

1. Govorkova EA, Baranovich T, Seiler P, Armstrong J, Burnham A, Guan Y, Peiris M, Webby RJ, Webster RG. 2013. Antiviral resistance among highly pathogenic influenza A (H5N1) viruses isolated worldwide in 2002–2012 shows need for continued monitoring. Antiviral Res 98:297–304. https://doi.org/10.1016/j.antiviral.2013.02.013.
2. Köser CU, Ellington MJ, Peacock SJ. 2014. Whole-genome sequencing to control antimicrobial resistance. Trends Genet 30:401–407. https://doi.org/10.1016/j.tig.2014.07.003.
3. Bradbury AR, Patrick-Miller L, Domchek S. 2015. Multiplex genetic testing: reconsidering utility and informed consent in the era of next-generation sequencing. Genet Med 17:97–98. https://doi.org/10.1038/gim.2014.85.
4. Kambhampati A, Payne DC, Costantini V, Lopman BA. 2016. Host genetic susceptibility to enteric viruses: a systematic review and metaanalysis. Clin Infect Dis 62:11–18. https://doi.org/10.1093/cid/civ873.
5. Patarčić I, Gelemanovic A, Kirin M, Kolcic I, Theodoratou E, Baillie KJ, de Jong MD, Rudan I, Campbell H, Polasek O. 2015. The role of host genetic factors in respiratory tract infectious diseases: systematic review, meta-analyses and field synopsis. Sci Rep 5:16119. https://doi.org/10.1038/srep16119.
6. Greninger AL, Naccache SN, Federman S, Yu G, Mbala P, Bres V, Stryke D, Bouquet J, Somasekar S, Linnen JM, Dodd R, Mulembakani P, Schneider BS, Muyembe-Tamfum JJ, Stramer SL, Chiu CY. 2015. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. Genome Med 7:99. https://doi.org/10.1186/s13073-015-0220-9.
7. Schmidt K, Mwaigwisya S, Crossman LC, Doumith M, Munroe D, Pires C, Khan AM, Woodford N, Saunders NJ, Wain J, O'Grady J, Livermore DM. 2017. Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. J Antimicrob Chemother 72:104–114. https://doi.org/10.1093/jac/dkw397.
8. Kim KH, Bae JW. 2011. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses. Appl Environ Microbiol 77:7663–7668. https://doi.org/10.1128/AEM.00289-11.
9. Li D, Li Z, Zhou Z, Li Z, Qu X, Xu P, Zhou P, Bo X, Ni M. 2016. Direct next-generation sequencing of virus-human mixed samples without pretreatment is favorable to recover virus genome. Biol Direct 11:3. https://doi.org/10.1186/s13062-016-0105-x.
10. Li L, Deng X, Mee ET, Collot-Teixeira S, Anderson R, Schepelmann S, Minor PD, Delwart E. 2015. Comparing viral metagenomics methods using a highly multiplexed human viral pathogens reagent. J Virol Methods 213:139–146. https://doi.org/10.1016/j.jviromet.2014.12.002.
11. Rosseel T, Van Borm S, Vandenbussche F, Hoffmann B, van den Berg T, Beer M, Hoper D. 2013. The origin of biased sequence depth in sequence-independent nucleic acid amplification and optimization for efficient massive parallel sequencing. PLoS One 8:e76144. https://doi.org/10.1371/journal.pone.0076144.
12. Yilmaz S, Allgaier M, Hugenholtz P. 2010. Multiple displacement amplification compromises quantitative analysis of metagenomes. Nat Methods 7:943–944. https://doi.org/10.1038/nmeth1210-943.
13. Graf EH, Simmon KE, Tardif KD, Hymas W, Flygare S, Eilbeck K, Yandell M, Schlaberg R. 2016. Unbiased detection of respiratory viruses by use of RNA sequencing-based metagenomics: a systematic comparison to a commercial PCR panel. J Clin Microbiol 54:1000–1007. https://doi.org/10.1128/JCM.03060-15.

14. Graham RM, Doyle CJ, Jennison AV. 2017. Epidemiological typing of Neisseria gonorrhoeae and detection of markers associated with antimicrobial resistance directly from urine samples using next generation sequencing. Sex Transm Infect 93:65–67. https://doi.org/10.1136/sextrans-2015-052422.
15. 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. https://doi.org/10.1056/NEJMoa1401268.
16. Fischer N, Indenbirken D, Meyer T, Lutgehetmann M, Lellek H, Spohn M, Aepfelbacher M, Alawi M, Grundhoff A. 2015. Evaluation of unbiased next-generation sequencing of RNA (RNA-seq) as a diagnostic method in influenza virus-positive respiratory samples. J Clin Microbiol 53:2238–2250. https://doi.org/10.1128/JCM.02495-14.
17. Schlaberg R, Chiu CY, Miller S, Procop GW, Weinstock G, Professional Practice Committee and Committee on Laboratory Practices of the American Society for Microbiology, Microbiology Resource Committee of the College of American Pathologists. 2017. Validation of metagenomic next-generation sequencing tests for universal pathogen detection. Arch Pathol Lab Med 141:776–786. https://doi.org/10.5858/arpa.2016-0539-RA.
18. Cheval J, Sauvage V, Frangeul L, Dacheux L, Guigon G, Dumey N, Pariente K, Rousseaux C, Dorange F, Berthet N, Brisse S, Moszer I, Bourhy H, Manuguerra CJ, Lecuit M, Burguiere A, Caro V, Eloit M. 2011. Evaluation of high-throughput sequencing for identifying known and unknown viruses in biological samples. J Clin Microbiol 49:3268–3275. https://doi.org/10.1128/JCM.00850-11.
19. Endrullat C, Glokler J, Franke P, Frohme M. 2016. Standardization and quality management in next-generation sequencing. Appl Transl Genom 10:2–9. https://doi.org/10.1016/j.atg.2016.06.001.
20. Lusk RW. 2014. Diverse and widespread contamination evident in the unmapped depths of high-throughput sequencing data. PLoS One 9:e110808. https://doi.org/10.1371/journal.pone.0110808.
21. Strong MJ, Xu G, Morici L, Splinter Bon-Durant S, Baddoo M, Lin Z, Fewell C, Taylor CM, Flemington EK. 2014. Microbial contamination in next-generation sequencing: implications for sequence-based analysis of clinical samples. PLoS Pathog 10:e1004437. https://doi.org/10.1371/journal.ppat.1004437.
22. Murray DC, Coghlan ML, Bunce M. 2015. From benchtop to desktop: important considerations when designing amplicon sequencing workflows. PLoS One 10:e0124671. https://doi.org/10.1371/journal.pone.0124671.
23. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60. https://doi.org/10.1038/nmeth.3176.
24. Fredricks DN, Relman DA. 1996. Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates. Clin Microbiol Rev 9:18–33.
25. Clarke AJ. 2014. Managing the ethical challenges of next-generation sequencing in genomic medicine. Br Med Bull 111:17–30. https://doi.org/10.1093/bmb/ldu017.
26. Davey S. 2014. Next-generation sequencing: considering the ethics. Int J Immunogenet 41:457–462. https://doi.org/10.1111/iji.12155.
27. Hall RJ, Draper JL, Nielsen FG, Dutilh BE. 2015. Beyond research: a primer for considerations on using viral metagenomics in the field and clinic. Front Microbiol 6:224. https://doi.org/10.3389/fmicb.2015.00224.
28. Hehir-Kwa JY, Claustres M, Hastings RJ, van Ravenswaaij-Arts C, Christenhusz G, Genuardi M, Melegh B, Cambon-Thomsen A, Patsalis P, Vermeesch J, Cornel MC, Searle B, Palotie A, Capoluongo E, Peterlin B, Estivill X, Robinson PN. 2015. Towards a European consensus for reporting incidental findings during clinical NGS testing. Eur J Hum Genet 23:1601–1606. https://doi.org/10.1038/ejhg.2015.111.
29. Hegde M, Bale S, Bayrak-Toydemir P, Gibson J, Bone Jeng LJ, Joseph L, Laser J, Lubin IM, Miller CE, Ross LF, Rothberg PG, Tanner AK, Vitazka P, Mao R. 2015. Reporting incidental findings in genomic scale clinical sequencing–a clinical laboratory perspective: a report of the Association for Molecular Pathology. J Mol Diagn 17:107–117. https://doi.org/10.1016/j.jmoldx.2014.10.004.
30. Rahimzadeh V, Avard D, Senecal K, Knoppers BM, Sinnett D. 2015. To disclose, or not to disclose? Context matters. Eur J Hum Genet 23:279–284. https://doi.org/10.1038/ejhg.2014.108.
31. Feng Y, Zhang Y, Ying C, Wang D, Du C. 2015. Nanopore-based fourth-generation DNA sequencing technology. Genomics Proteomics Bioinformatics 13:4–16. https://doi.org/10.1016/j.gpb.2015.01.009.
32. Wang Y, Yang Q, Wang Z. 2014. The evolution of nanopore sequencing. Front Genet 5:449. https://doi.org/10.3389/fgene.2014.00449.
33. Mikheyev AS, Tin MM. 2014. A first look at the Oxford Nanopore MinION sequencer. Mol Ecol Resour 14:1097–1102. https://doi.org/10.1111/1755-0998.12324.
34. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, Oliveira G, Robles-Sikisaka R, Rogers TF, Beutler NA, Burton DR, Lewis-Ximenez LL, de Jesus JG, Giovanetti M, Hill SC, Black A, Bedford T, Carroll MW, Nunes M, Alcantara LC, Jr, Sabino EC, Baylis SA, Faria NR, Loose M, Simpson JT, Pybus OG, Andersen KG, Loman NJ. 2017. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nat Protoc 12:1261–1276. https://doi.org/10.1038/nprot.2017.066.
35. Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, Bore JA, Koundouno R, Dudas G, Mikhail A, Ouedraogo N, Afrough B, Bah A, Baum JH, Becker-Ziaja B, Boettcher JP, Cabeza-Cabrerizo M, Camino-Sanchez A, Carter LL, Doerrbecker J, Enkirch T, Dorival IGG, Hetzelt N, Hinzmann J, Holm T, Kafetzopoulou LE, Koropogui M, Kosgey A, Kuisma E, Logue CH, Mazzarelli A, Meisel S, Mertens M, Michel J, Ngabo D, Nitzsche K, Pallash E, Patrono LV, Portmann J, Repits JG, Rickett NY, Sachse A, Singethan K, Vitoriano I, Yemanaberhan RL, Zekeng EG, Trina R, Bello A, Sall AA, Faye O, et al. 2016. Real-time, portable genome sequencing for Ebola surveillance. Nature 530:228–232. https://doi.org/10.1038/nature16996.
36. Jain M, Tyson JR, Loose M, Ip CLC, Eccles DA, O'Grady J, Malla S, Leggett RM, Wallerman O, Jansen HJ, Zalunin V, Birney E, Brown BL, Snutch TP, Olsen HE, MinION Analysis and Reference Constortium. 2017. MinION Analysis and Reference Consortium: phase 2 data release and analysis of R9.0 chemistry. F1000Res 6:760. https://doi.org/10.12688/f1000research.11354.1.
37. Votintseva AA, Bradley P, Pankhurst L, Del Ojo Elias C, Loose M, Nilgiriwala K, Chatterjee A, Smith EG, Sanderson N, Walker TM, Morgan MR, Wyllie DH, Walker AS, Peto TEA, Crook DW, Iqbal Z. 2017. Same-day diagnostic and surveillance data for tuberculosis via whole-genome sequencing of direct respiratory samples. J Clin Microbiol 55:1285–1298. https://doi.org/10.1128/JCM.02483-16.