



# Advances in Clinical Sample Preparation for Identification and Characterization of Bacterial Pathogens Using Metagenomics

Nathan L. Bachmann<sup>1,2†</sup>, Rebecca J. Rockett<sup>1,3†</sup>, Verlaine Joy Timms<sup>3</sup> and Vitali Sintchenko<sup>1,3\*</sup>

<sup>1</sup> Sydney Medical School, University of Sydney, Camperdown, NSW, Australia, <sup>2</sup> Centenary Institute, University of Sydney, Camperdown, NSW, Australia, <sup>3</sup> Centre for Infectious Diseases and Microbiology–Public Health, Westmead Hospital, Sydney, NSW, Australia

## OPEN ACCESS

### Edited by:

Kai Zhou,  
Zhejiang University, China

### Reviewed by:

David William Cleary,  
University of Southampton,  
United Kingdom  
Olivier Vandenberg,  
Laboratoire Hospitalier Universitaire  
de Bruxelles, Belgium  
Elena Perrin,  
Università degli Studi di Firenze, Italy

### \*Correspondence:

Vitali Sintchenko  
vitali.sintchenko@sydney.edu.au

†These authors share first authorship

### Specialty section:

This article was submitted to  
Infectious Diseases - Surveillance,  
Prevention and Treatment,  
a section of the journal  
Frontiers in Public Health

**Received:** 18 August 2018

**Accepted:** 28 November 2018

**Published:** 12 December 2018

### Citation:

Bachmann NL, Rockett RJ, Timms VJ and Sintchenko V (2018) Advances in Clinical Sample Preparation for Identification and Characterization of Bacterial Pathogens Using Metagenomics. *Front. Public Health* 6:363. doi: 10.3389/fpubh.2018.00363

Whole genome sequencing (WGS) plays an increasing role in communicable disease control through high-resolution outbreak tracing, laboratory surveillance and diagnostics. However, WGS has traditionally relied on microbial culture in order to obtain pathogen specific DNA for sequencing. This has severely limited the application of whole genome sequencing on pathogens with fastidious culturing requirements. In addition, the widespread adoption of culture-independent diagnostic tests has reduced availability of cultured isolates for confirmatory testing and surveillance. These recent developments have created demand for the implementation of techniques enabling direct sequencing of microbial genomes in clinical samples without having to culture an isolate. However, sequencing of specific organisms from clinical samples can be affected by high levels of contaminating DNA from the host and other commensal microorganisms. Several methods have been introduced for selective lysis of host cells and/or separate specific organisms from a clinical sample. This review examines the different approaches for sample preparation that have been used in diagnostic and public health laboratories for metagenomic sequencing.

**Keywords:** metagenomic, culture-independent genome sequencing, bacterial pathogens, *Mycobacterium tuberculosis*, public health

## INTRODUCTION

The ability to perform high-throughput whole genome sequencing (WGS) on infectious agents has revolutionized research into microbial genomics (1–3). WGS has become the method of choice for subtyping of pathogens with epidemic potential and investigating nosocomial transmission, foodborne disease outbreaks, and antibiotic resistance (4). Further, WGS has been extensively applied to investigate etiology, evolution and transmission of respiratory illnesses (5) and sexually transmitted diseases (6) as well as the acquisition and dissemination of antibiotic resistance in hospitals (7, 8). The ultimate resolution provided by WGS has led to its widespread use by public health facilities, such as Public Health England, the USA Centers for Disease Control and Prevention and others to investigate outbreaks of foodborne-associated salmonellosis (9), listeriosis (10) and Shiga-toxin producing *Escherichia coli* (11). The current status of culture-based WGS in clinical settings and sequencing platforms are reviewed recently elsewhere (12–14). However, the

utility of WGS for communicable disease control is hampered by the reliance on the laboratory culture of the pathogen of interest.

Molecular-based culture independent diagnostics test (CIDT) methods offer significant advantages for diagnostic microbiology laboratories, the most important being a reduction in cost and an decrease in turn-around-time (15–17). The adoption of CIDT by diagnostic laboratories has led to significant increases in testing and diagnosis of enteric pathogens, but has reduced the number of isolates available for typing, resulting in notifications to public health registries without corresponding culture. As a direct result, our current laboratory surveillance systems are losing the capacity to detect clusters and identify the source of outbreaks and to recognize the emergence of high-risk or multi-drug resistant variants. The steady increase in CIDT-only notifications demands alternative methods for identification and characterization of bacterial pathogens with epidemic potential.

Metagenomics offers unbiased sequencing of all DNA in a clinical sample without culturing individual bacterial isolates [reviewed in (18)]. The applications of metagenomics have been largely discovery focused, probing the previously uncharacterized make up of environmental and human ecosystems and uncovering their unrecognized diversity (19). Furthermore, metagenomics has also been successful in identifying and characterizing emerging infectious agents (20). Metagenomic sequencing directly from clinical specimens eliminates the pathogen culturing step and could allow for comprehensive surveillance of antibiotic resistance and transmission dynamics. The application of metagenomics to diagnostics and outbreak tracing introduces many unique challenges regarding sample storage, DNA extraction, bioinformatics data analysis and reporting [reviewed in (21, 22)]. However, one of the largest barriers is to selectively target for microbial DNA in human samples. While bioinformatics tools can remove human reads, greater sequencing depth (which leads to an increase in cost) is required to obtain enough pathogen reads for identification of a causative agent and obtain information regarding resistance or strain type. This minireview will examine recent advances in techniques for overcoming this challenge, in particular how to deal with host and other contaminating DNA, with a specific focus on pathogens of public health importance.

## METAGENOMIC SEQUENCING FOR PUBLIC HEALTH

The shotgun metagenomics has been applied for pathogen discovery or to uncover the etiology of an unrecognized infection (23, 24). Metagenomics has improved the diagnosis of central nervous system infections (24). Metagenomic will also be vital for understanding of other syndromes as unknown pathogens are estimated to be responsible for 20–40% of respiratory tract infections (25) and 40% of infections of the gastrointestinal tract (26). Currently, researchers have demonstrated that shotgun metagenomics has the equivalent sensitivity to diagnostic PCR-based methods, particularly when looking for pathogens in the

gastrointestinal tract, respiratory system and urogenital tract (Table 1) (33, 34).

The vast majority of sequencing reads produced during shotgun metagenomics are identified as human, which is on average 1000 times larger than the average bacterial genome. Direct shotgun metagenomics has limited utility in public health outbreak investigations due to the high cost of deep sequencing but it has been highly successful in diagnostics (32). As an example, metagenomics enabled the recognition of *Brucella* in cerebrospinal fluid (CSF) from a patient with partially treated meningitis where the causative agent could not be determined via traditional microbiological methods targeting common neurotropic pathogens (35). Although the CSF was positive for Epstein Barr virus and human herpesvirus 7, testing for bacterial pathogens including *Brucella* by serology and tuberculosis by culture were negative. When the patient's health failed to improve following antiviral treatment, the CSF was sent for metagenomic sequencing for a comprehensive identification of potential pathogens. Whole DNA extraction was performed on the sample, with half of extracted DNA been treated with turbo DNase followed by reverse transcription of the RNA to cDNA. The second half of the sample was treated with the NEBNext<sup>®</sup> Microbiome DNA Enrichment kit to enrich the microbial DNA. Both the enriched DNA and the cDNA were sequenced with Illumina HiSeq. From the raw data of the enriched DNA 0.0012% of reads corresponded to *Brucella melitensis*. It was noted that no *Brucella* reads were detected in the sequenced cDNA library. The patient was subsequently diagnosed with chronic neurobrucellosis and treated with doxycycline and rifampin with full resolution of symptoms in 2 weeks. This case demonstrates the utility of metagenomics sequencing as a supplementary test for notifiable conditions as *Brucella sp* was not originally cultured from CSF or blood cultures and the importance of microbial DNA enrichment as *Brucella* reads were also not found in the reserved transcribed RNA library.

## PROCESSING OF CLINICAL SAMPLES

The success of metagenomics is dependent on quality and quantity of DNA extracted from a given specimen. Different specimen types (e.g., feces, sputum, tissue, urine, etc.) present unique and specific challenges reflecting their matrix and concentrations of the target pathogen and resident microflora. Samples like urine and swabs will have low target pathogen cell counts making it difficult to extract a high concentration of genomic DNA (gDNA). In contrast, human stool samples are comprised of a complex matrix of fibers, enzymes, undigested parts, and other inhibitors which have to be removed during DNA extraction (36). Both stool and sputum samples will need to be liquefied using a tissuelyser or decongestion solution, respectively (37). Fortunately, in the age of PCR based diagnostics, there are a wide array of methods and molecular kits for extracting DNA from clinical samples (37). Commercial kits for DNA extraction and purification, such as Qiagen's DNA Stool Mini Kit and QIAmp Microbiome DNA Kit can be adapted to metagenomics sequencing (38). However, these kits

**TABLE 1** | Sample types and processing techniques for metagenomics of bacterial pathogens of public health importance.

Niche/sample	Pathogen	Sample type	Sample challenge <sup>a</sup>	Methods used	WGS technologies evaluated	Expected or demonstrated value	References
Food borne pathogens	<i>Salmonella enterica</i>	Fecal	Homogenization	Direct metagenomics	Illumina HiSeq	Strain characterization and outbreak tracing	(27)
	<i>Shiga-toxin producing Escherichia coli</i>	Fecal	Homogenization	Direct metagenomics	Illumina HiSeq	Strain characterization and outbreak tracing	(11)
	<i>Listeria monocytogenes</i>	Fecal	Homogenization	Direct metagenomics	Illumina HiSeq	Strain characterization and outbreak tracing	(27)
	<i>Campylobacter</i> species	Fecal	Homogenization	Direct metagenomics	Illumina HiSeq	Strain characterization and outbreak tracing	(27)
Respiratory pathogens	<i>Mycobacterium tuberculosis</i>	Sputum	Liquification	Probe based capture, host cell lysis	Illumina HiSeq, MiSeq an Mini Ion	Detection and resistance profile	(28, 29)
	<i>Bordetella pertussis</i>	Nasopharyngeal aspirate	Low bacterial load/DNA yield	None published	None used	Detection and strain characterization	
	<i>Legionella pneumophila</i>	Sputum	Liquification/Low bacterial load/DNA yield	None published	None used	Detection and strain characterization	
	<i>Streptococcus pneumoniae</i>	Sputum/swab	Liquification	16S metagenomics	Illumina HiSeq	Detection and strain characterization	(30)
Sexually transmitted pathogens	<i>Chlamydia trachomatis</i>	Urine, swab	Low bacterial load/DNA yield	Probe based capture	Illumina HiSeq	Strain characterization	(31)
	<i>Mycoplasma genitalium</i>	Urine, swab	Low bacterial load/DNA yield	None published	None used	Detection and Resistance profile	
	<i>Treponema pallidum</i> (Syphilis)	Swab	Low bacterial load/DNA yield	None published	None used	Strain identification, subtyping and drug resistance	
	<i>Neisseria gonorrhoeae</i>	Urine, swab	NEBNext microbiome kit	Direct metagenomics	Ion Torrent PGM	Resistance profile	(32)

<sup>a</sup>Specific challenges during DNA extraction that needs to be considered. Fecal samples will need to be homogenized with a tissue lyser or Qiagen's DNA Stool Mini Kit, while sputum samples will need to be treated with a decongest solution for liquification. Urine and swab samples will have challenges associated with DNA yield due to low cell count.

are designed for PCR analysis and will need to be combined with host DNA depletion and/or bacterial enrichment. Thus, while metagenomics has the ability to provide universal detection of pathogens, in reality specific nucleic acid extraction, strategies for target enrichment or host nucleic acid depletion will be needed for each pathogen or particular specimen type (34).

There are three key challenges of sequencing microbial organisms directly from a sample: (i) contamination of host DNA and other microorganisms, (ii) low cell abundance of the target organism present in the sample, and (iii) presence of DNA amplification inhibitors and other confounding variables in the clinical sample matrix. Therefore, several complementary methods have been developed to reduce the presence of host DNA or separate out the microbial DNA. Afterwards usually an enrichment method is then used to address the low microbial DNA. Noteworthy, the DNA enrichment process can increase the risk of selection bias. The key pre-treatment methods for metagenomics are summarized in **Figure 1**.

## MICROBIAL SEPARATION

Microbial cells can be separated from the sample matrix using chemical, physical, physicochemical or biological mechanisms (39). One of the established protocol is immunomagnetic separation (IMS), which has been successfully used to concentrate targeted bacteria and remove inhibitors to improve the quality of the extracted DNA for amplification-based methods (40, 41). IMS has since been applied to successful sequencing of *Chlamydia trachomatis* directly from genital swabs (42, 43). This approach uses magnetic beads coated with antibodies specific to chlamydial LPS to bind and extract chlamydial cells in clinical samples from the site of infection. The DNA extraction is then performed on the isolated cells followed by DNA purification. As clinical samples often do not carry sufficient quantities of target bacteria to give enough DNA for genome sequencing, IMS has been combined with multiple displacement amplification (MDA) to amplify DNA using the  $\Phi$ 29 polymerase and random hexamer primers (44). The combined IMS-MDA protocol has provided high-quality genomic DNA for *C. trachomatis* directly from clinical samples (42, 43). However, it must be noted that the clinical specimens used in these studies were urine and cervical swabs. Both sample types contained minimal human and microbiome DNA reducing non-specific binding of cells to the antibodies. The application of IMS-MDA protocols to rectal swabs in order to support direct sequencing of sexually transmitted pathogens in high-risk populations, such as men who have sex with men (MSM) could be less effective.

In addition to microbial separation of a targeted organism with antibodies, approaches have also been used to separate host DNA from microbial DNA allowing for an unbiased study of the microbiome within the specimen. For example, the NEBNext<sup>®</sup> Microbiome DNA Enrichment kit (New England Biolabs) selectively targets methylation sites in eukaryotic DNA and binds it to magnetic beads. This approach was used to directly sequence 13 *Neisseria gonorrhoeae* genomes directly from

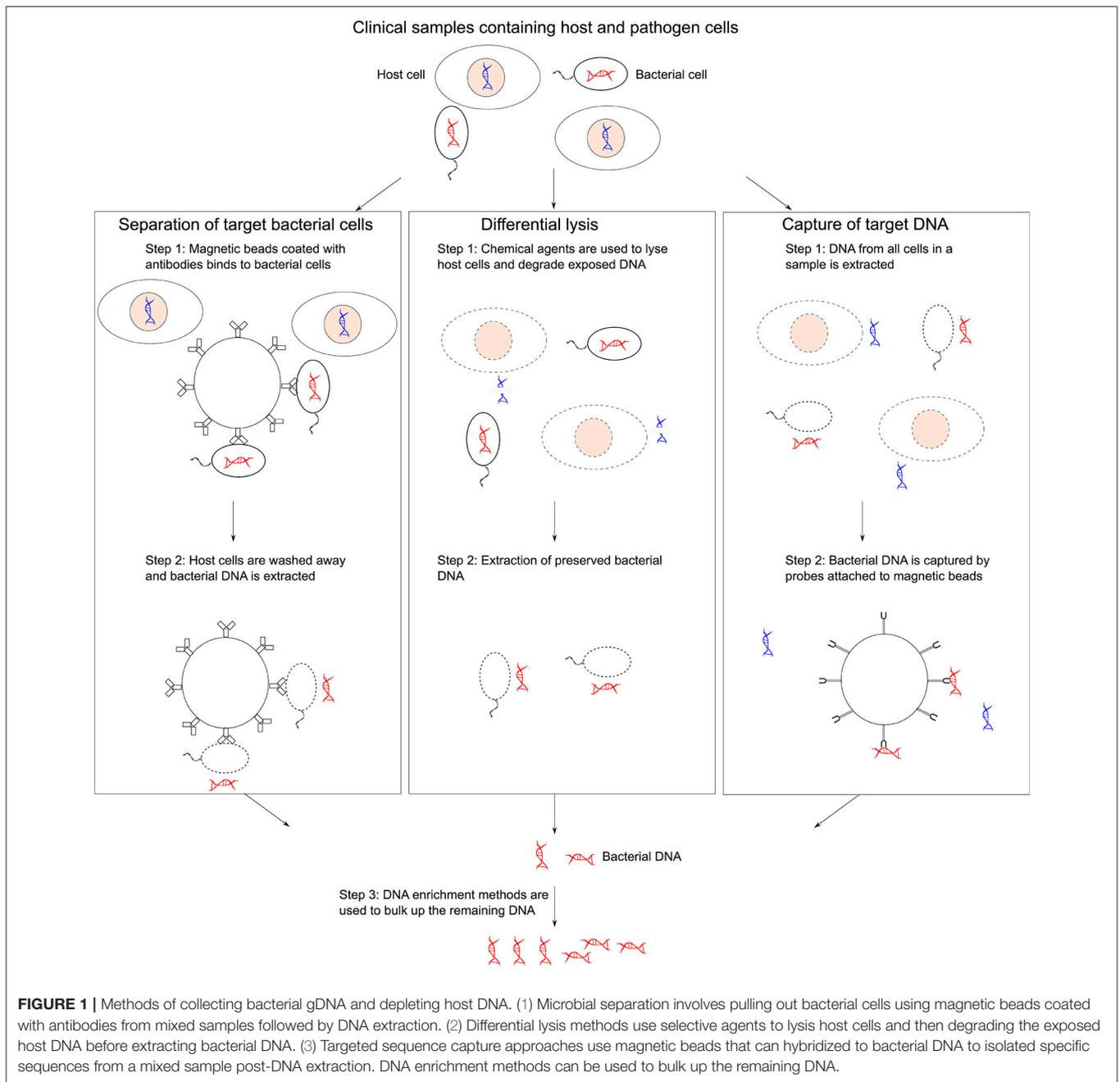
urine samples (32). Without enrichment, fewer than 1% of reads aligned to the *N. gonorrhoeae* reference genome, however, after using the NEBNext kit, the number of *N. gonorrhoeae* reads ranged from 2 to 43% of total number of sequenced reads. This provided sufficient coverage to obtain robust typing information for 11 of the 13 strains. While promising, this approach is still limited by the sample type (sterile vs. non-sterile site) and bacterial load as few reads were obtained for the pathogen of interest both in this example and from the CSF case mentioned previously (35). The success of this approach appear to be dependent on bacterial load and sample type.

## DEPLETION OF HOST NUCLEIC ACIDS

These methodologies focusing on depleting host cells from the original specimen by exploiting difference in cell surface structure between human cells and bacteria for selective lysis of host cells. Using a combination of selective cell lysis and DNase treatment, these techniques have had some success although high microbial load and a large original sample volume are often required (28, 45). One of the first applications of clinical metagenomics study to directly sequence *Mycobacterium tuberculosis* from respiratory specimens was based on osmotic lysis, where a large volume of sterile water is added to a sputum sample to increase the osmotic pressure and cause human cells to burst while leaving the more robust *M. tuberculosis* cells intact (45). A DNase enzyme was then used to degrade the liberated human DNA prior to *M. tuberculosis* DNA extraction. However, this approach will be severely limited for acquiring bacterial DNA from low load samples without enrichment (45). Another limitation is that some pathogens of interest, such as Gram-negative bacteria, are also susceptible to osmotic pressure.

A number of commercial kits are available for selective lysis of human or eukaryotic cells. Molzym's MolYsis Basic kit was used for direct sequencing of *M. tuberculosis* with greater efficacy than osmotic lysis only. The MolYsis kit uses a chaotropic buffer that selectively lyse human cells while keeping the bacterial cells intact (28). From this study sufficient DNA was obtained for antibiotic susceptibility prediction in 62% of samples. A critical advantage in this study was that genome sequencing was performed on a MinION platform (Oxford Nanopore), which has the ability to continue sequencing until sufficient coverage is obtained. Therefore, success rate can be lower when other, especially short-read sequencing, platforms are used because of static read coverage per sample. A comparison study on synovial fluid spiked with *Staphylococcus aureus* revealed that use of the MolYsis Kit produced a higher fold of microbial reads compared to the NEBNext<sup>®</sup> Microbiome DNA Enrichment kit (46).

Another study demonstrated an improved method of human cell depletion over the Molysis kit (47). The method sought to rely on mild detergents that are typically used to permeabilize mammalian cell lines for protein extraction or for the recovery of intracellular pathogens. Saponin, commonly used in hematology laboratories for hemolysis of human erythrocytes, was found to be the most effective in its differential effect on human vs. pathogen DNA in spiked cerebrospinal fluid (CSF) and



**FIGURE 1 |** Methods of collecting bacterial gDNA and depleting host DNA. (1) Microbial separation involves pulling out bacterial cells using magnetic beads coated with antibodies from mixed samples followed by DNA extraction. (2) Differential lysis methods use selective agents to lyse host cells and then degrading the exposed host DNA before extracting bacterial DNA. (3) Targeted sequence capture approaches use magnetic beads that can hybridized to bacterial DNA to isolated specific sequences from a mixed sample post-DNA extraction. DNA enrichment methods can be used to bulk up the remaining DNA.

nasopharyngeal aspirate specimens (40). In combination with Nanopore sequencing, saponin depletion can rapidly and accurately characterize the bacterial composition of the lower respiratory tract and provide antibiotic resistance data (48). This study showed that the saponin differential lysis method resulted in a  $10^3$  fold decrease in host DNA in bronchoalveolar lavage and endotracheal aspirates. This depletion efficiency depends on using both a high concentration of saponin (2.2–2.5%) and salt buffer (5.5 M NaCl). Nanopore sequencing also has the significant advantage of providing real time data acquisition and analysis compared to the Illumina platform, allowing for fast turnaround time of clinical results.

## TARGETED ENRICHMENT OF PATHOGEN DNA AFTER EXTRACTION

The targeted enrichment of pathogen DNA commonly involves hybridization of complement probes to the target bacterial genome. Agilent Technologies developed one such kit (SureSelect<sup>XT</sup>) originally for the enrichment of specific regions of large eukaryotic genomes for deep sequencing of a selected subset of genes (49, 50). However, it has since been adapted for deep sequencing of viral genomes directly from clinical samples (51, 52). This sequence capture method involves designing customized biotin-labeled RNA probes that can hybridized to

a complete target genome sequence so that magnetic beads coated with the biotin-binding protein, streptavidin, can be used to extract the DNA (53). This approach can be suitable for culture-independent genome sequencing of *Chlamydia* species because strains within a species have highly conserved genome sequences with >90% nucleotide similarity. Therefore, probes can be designed using publicly available reference genomes to reliably capture target DNA for strains within the species. This has been successfully achieved for the identification of *C. trachomatis* in clinical swabs from patients with urethritis (31). Since then this approach has been applied to other *Chlamydia* species, including zoonotic pathogens *C. pecorum* and *C. psittaci* (54, 55). However, since the probe design depends on a reference sequence, this approach can be problematic when extended to pathogens with highly mosaic genomes, like *Salmonella* or *Escherichia coli*. Significant features in these pathogens, such as antibiotic resistance or virulence are often associated with mobile elements and genomic islands. Secondly, this approach may not be applicable for pathogen discovery as the emergent pathogen may not be captured and enriched by the probes.

Target enrichment has also been applied to the clinical situation for Tuberculosis diagnostics. In one case a clinical isolate from a patient with pulmonary tuberculosis was reported as resistant to rifampicin and low-level resistant to isoniazid based on initial results from Xpert MTB/RIF and Hain line probe assays (GenoType MDRTBplus v1.0), which target only *rpoB* and *inhA* genes in the *M. tuberculosis* genome (56). Therefore, the patient was placed on a combined drug regime including high-dose isoniazid. Metagenomic sequencing directly from the sputum sample identified a further *inhA* gene mutation that was consistent with high-level resistance to isoniazid and confirmed the absence of fluoroquinolone resistance. These findings informed treatment decisions with a shorten turnaround time by 12–14 days compared to conventional culture-based techniques. This study used the enrichment technology, SureSelect<sup>XT</sup> (Agilent Technologies) to capture and amplify mycobacterial DNA (57), which remains prohibitively expensive for use as part of routine diagnosis. In addition, this method requires equipment not commonly found in most diagnostic laboratories, such as focused ultrasonicators for DNA shearing. However, the cost of the enrichment can be reduced by only targeting regions that encode drug resistance (29).

## CHALLENGES OF IMPLEMENTING METAGENOMICS TO PUBLIC HEALTH

Risk of laboratory contamination and enrichment biases are of particular concern as the use of highly sensitive clinical

metagenomics becomes more widespread. Possible sources of DNA contamination include extraction reagents and columns, PCR reagents and library preparation reagents (58, 59). Another concern is background DNA contained within a WGS library preparation laboratory, particularly a laboratory routinely performing high-throughput sequencing on cultured isolates. Identification of pathogens that are routinely sequenced after cultivation in a clinical sample should be scrutinized as it may indicate sample contamination. However, this can be addressed by adequate quality control measures, such as inclusion of negative “no-template” controls for each preparation step, as part of every sequencing run, to monitor for contamination by commensal DNA (34).

Two significant objectives must be achieved to facilitate the implementation of metagenomics into public health laboratory surveillance and outbreak investigations. Firstly, the analytical sensitivity of metagenomic methods should be improved and validated (34). This evaluation should include testing of reproducibility and repeatability of testing which might require involvement of multiple laboratories. A key added value of metagenomics is likely to be its effectiveness on samples collected in earlier, most transmissible, stages of infectious disease. Secondly, the cost of consumables per sample should decrease significantly to support the uptake of this technology. Currently, DNA enrichment methods remain the most expensive aspect of metagenomic testing, with estimated cost up to \$350 USD per sample (57).

## CONCLUSION

Metagenomics has yet to be implemented into routine diagnosis and few long term studies have been conducted. However, future advancements in sequencing technology will most likely be the deciding factor as to when metagenomics will be integrated into public health surveillance. Sensitivity of shotgun metagenomics will improve as new sequencing platforms are released that can efficiently generate longer sequencing reads with higher depths. In addition due to the various sample types and data analysis methods available, laboratories will need to clearly define the intended clinical usage and range of pathogens that will be detected as this will greatly impact the choice of methodology for sample preparation and DNA extraction.

## AUTHOR CONTRIBUTIONS

NB and RR compiled and wrote the manuscript. VT and VS were involved in review and editing.

## REFERENCES

1. Sintchenko V, Iredell JR, Gilbert GL. Pathogen profiling for disease management and surveillance. *Nat Rev Microbiol.* (2007) 5:464–70. doi: 10.1038/nrmicro1656
2. Kwong JC, McCallum N, Sintchenko V, Howden BP. Whole genome sequencing in clinical and public health microbiology. *Pathology* (2015) 47:199–210. doi: 10.1097/PAT.0000000000000235
3. Sintchenko V, Holmes EC. The role of pathogen genomics in assessing disease transmission. *BMJ* (2015) 350:h1314. doi: 10.1136/bmj.h1314
4. Gilmour MW, Graham M, Reimer A, Van Domselaar G. Public health genomics and the new molecular epidemiology of bacterial pathogens. *Public Health Genomics* (2013) 16:25–30. doi: 10.1159/000342709
5. Walker TM, Kohl TA, Omar SV, Hedge J, Del Ojo Elias C, Bradley P, et al. Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. *Lancet Infect Dis.* (2015) 15:1193–202. doi: 10.1016/S1473-3099(15)00062-6

6. De Silva D, Peters J, Cole K, Cole MJ, Cresswell F, Dean G, et al. Whole-genome sequencing to determine transmission of *Neisseria gonorrhoeae*: an observational study. *Lancet Infect Dis.* (2016) 16:1295–303. doi: 10.1016/S1473-3099(16)30157-8
7. Bradley P, Gordon NC, Walker TM, Dunn L, Heys S, Huang B, et al. Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. *Nat Commun.* (2015) 6:10063. doi: 10.1038/ncomms10063
8. Price JR, Cole K, Bexley A, Kostiuo V, Eyre DW, Golubchik T, et al. Transmission of *Staphylococcus aureus* between health-care workers, the environment, and patients in an intensive care unit: a longitudinal cohort study based on whole-genome sequencing. *Lancet Infect Dis.* (2017) 17:207–14. doi: 10.1016/S1473-3099(16)30413-3
9. Lienau EK, Strain E, Wang C, Zheng J, Ottesen AR, Keys CE, et al. Identification of a salmonellosis outbreak by means of molecular sequencing. *N Engl J Med.* (2011) 364:981–2. doi: 10.1056/NEJMc1100443
10. McCollum JT, Cronquist AB, Silk BJ, Jackson KA, O'Connor KA, Cosgrove S, et al. Multistate outbreak of listeriosis associated with cantaloupe. *N Engl J Med.* (2013) 369:944–53. doi: 10.1056/NEJMoa1215837
11. Loman NJ, Constantinidou C, Christner M, Rohde H, Chan JZ, Quick J, et al. A culture-independent sequence-based metagenomics approach to the investigation of an outbreak of Shiga-toxicogenic *Escherichia coli* O104:H4. *JAMA* (2013) 309:1502–10. doi: 10.1001/jama.2013.3231
12. Motro Y, Moran-Gilad J. Next-generation sequencing applications in clinical bacteriology. *Biomol Detect Quantif.* (2017) 14:1–6. doi: 10.1016/j.bdq.2017.10.002
13. Besser J, Carleton HA, Gerner-Smidt P, Lindsey RL, Trees E. Next-generation sequencing technologies and their application to the study and control of bacterial infections. *Clin Microbiol Infect.* (2018) 24:335–41. doi: 10.1016/j.cmi.2017.10.013
14. Gardy JL, Loman NJ. Towards a genomics-informed, real-time, global pathogen surveillance system. *Nat Rev Genet.* (2018) 19:9–20. doi: 10.1038/nrg.2017.88
15. Langley G, Besser J, Iwamoto M, Lessa FC, Cronquist A, Skoff TH, et al. Effect of culture-independent diagnostic tests on future emerging infections program surveillance. *Emerg Infect Dis.* (2015) 21:1582–8. doi: 10.3201/eid2109.150570
16. May FJ, Stafford RJ, Carroll H, Robson JM, Vohra R, Nimmo GR, et al. The effects of culture independent diagnostic testing on the diagnosis and reporting of enteric bacterial pathogens in Queensland, 2010 to 2014. *Commun Dis Intell Q Rep.* (2017) 41:E223–30.
17. Shea S, Kubota KA, Maguire H, Gladbach S, Woron A, Atkinson-Dunn R, et al. Clinical microbiology laboratories' adoption of culture-independent diagnostic tests is a threat to foodborne-disease surveillance in the United States. *J Clin Microbiol.* (2017) 55:10–9. doi: 10.1128/JCM.01624-16
18. Mullany P. Functional metagenomics for the investigation of antibiotic resistance. *Virulence* (2014) 5:443–7. doi: 10.4161/viru.28196
19. Bag S, Saha B, Mehta O, Anbumani D, Kumar N, Dayal M, et al. An improved method for high quality metagenomics DNA extraction from human and environmental samples. *Sci Rep.* (2016) 6:26775. doi: 10.1038/srep26775
20. Taylor-Brown A, Pillonel T, Bridle A, Qi W, Bachmann NL, Miller TL, et al. Culture-independent genomics of a novel chlamydial pathogen of fish provides new insight into host-specific adaptations utilized by these intracellular bacteria. *Environ Microbiol.* (2017) 19:1899–913. doi: 10.1111/1462-2920.13694
21. Gu W, Miller S, Chiu CY. Clinical metagenomic next-generation sequencing for pathogen detection. *Annu Rev Pathol.* (2018) 14:317–36. doi: 10.1146/annurev-pathmechdis-012418-012751
22. Martin TC, Visconti A, Spector TD, Falchi M. Conducting metagenomic studies in microbiology and clinical research. *Appl Microbiol Biotechnol.* (2018) 102:8629–46. doi: 10.1007/s00253-018-9209-9
23. Wilson MR, Shanbhag NM, Reid MJ, Singhal NS, Gelfand JM, Sample HA, et al. Diagnosing *Balamuthia mandrillaris* encephalitis with metagenomic deep sequencing. *Ann Neurol.* (2015) 78:722–30. doi: 10.1002/ana.24499
24. Brown JR, Bharucha T, Breuer J. Encephalitis diagnosis using metagenomics: application of next generation sequencing for undiagnosed cases. *J Infect.* (2018) 76:225–40. doi: 10.1016/j.jinf.2017.12.014
25. Gaynor AM, Nissen MD, Whiley DM, Mackay IM, Lambert SB, Wu G, et al. Identification of a novel polyomavirus from patients with acute respiratory tract infections. *PLoS Pathog.* (2007) 3:e64. doi: 10.1371/journal.ppat.0030064
26. Iturriza-Gomara M, Elliot AJ, Dockery C, Fleming DM, Gray JJ. Structured surveillance of infectious intestinal disease in pre-school children in the community: 'The Nappy Study'. *Epidemiol Infect.* (2009) 137:922–31. doi: 10.1017/S0950268808001556
27. Yang X, Noyes NR, Doster E, Martin JN, Linke LM, Magnuson RJ, et al. Use of metagenomic shotgun sequencing technology to detect foodborne pathogens within the microbiome of the beef production chain. *Appl Environ Microbiol.* (2016) 82:2433–43. doi: 10.1128/AEM.00078-16
28. Votintseva AA, Bradley P, Pankhurst L, Del Ojo Elias C, Loose M, Nilgiriwala K, et al. Same-day diagnostic and surveillance data for tuberculosis via whole-genome sequencing of direct respiratory samples. *J Clin Microbiol.* (2017) 55:1285–98. doi: 10.1128/JCM.02483-16
29. Doyle RM, Burgess C, Williams R, Gorton R, Booth H, Brown J, et al. Direct whole-genome sequencing of sputum accurately identifies drug-resistant *Mycobacterium tuberculosis* faster than MGIT culture sequencing. *J Clin Microbiol.* 56:e00666-18 doi: 10.1128/JCM.00666-18
30. Wright MS, McCarrison J, Gomez AM, Beck E, Harkins D, Shankar J, et al. Strain level *Streptococcus* colonization patterns during the first year of life. *Front Microbiol.* (2017) 8:1661. doi: 10.3389/fmicb.2017.01661
31. Christiansen MT, Brown AC, Kundu S, Tutill HJ, Williams R, Brown JR, et al. Whole-genome enrichment and sequencing of *Chlamydia trachomatis* directly from clinical samples. *BMC Infect Dis.* (2014) 14:591. doi: 10.1186/s12879-014-0591-3
32. Graham RM, Doyle CJ, Jennison AV. Epidemiological typing of *Neisseria gonorrhoeae* and detection of markers associated with antimicrobial resistance directly from urine samples using next generation sequencing. *Sex Transm Infect.* (2017) 93:65–7. doi: 10.1136/sextrans-2015-052422
33. Thorburn F, Bennett S, Modha S, Murdoch D, Gunson R, Murcia PR. The use of next generation sequencing in the diagnosis and typing of respiratory infections. *J Clin Virol.* (2015) 69:96–100. doi: 10.1016/j.jcv.2015.06.082
34. Schlager R, Chiu CY, Miller S, Procop GW, Weinstock G, Professional Practice C, et al. Validation of metagenomic next-generation sequencing tests for universal pathogen detection. *Arch Pathol Lab Med.* (2017) 141:776–86. doi: 10.5858/arpa.2016-0539-RA
35. Mongkolrattanothai K, Naccache SN, Bender JM, Samayoa E, Pham E, Yu G, et al. Neurobrucellosis: unexpected answer from metagenomic next-generation sequencing. *J Pediatric Infect Dis Soc.* (2017) 6:393–8. doi: 10.1093/jpids/piw066
36. Schrader C, Schielke A, Ellerbroek L, Johne R. PCR inhibitors—occurrence, properties and removal. *J Appl Microbiol.* (2012) 113:1014–26. doi: 10.1111/j.1365-2672.2012.05384.x
37. Paulos S, Mateo M, de Lucio A, Hernandez-de Mingo M, Bailo B, Saugar JM, et al. Evaluation of five commercial methods for the extraction and purification of DNA from human faecal samples for downstream molecular detection of the enteric protozoan parasites *Cryptosporidium* spp., *Giardia duodenalis*, and *Entamoeba* spp. *J Microbiol Methods* (2016) 127:68–73. doi: 10.1016/j.mimet.2016.05.020
38. Menu E, Mary C, Toga I, Raoult D, Ranque S, Bittar F. Evaluation of two DNA extraction methods for the PCR-based detection of eukaryotic enteric pathogens in fecal samples. *BMC Res Notes* (2018) 11:206. doi: 10.1186/s13104-018-3300-2
39. Stevens KA, Jaykus LA. Bacterial separation and concentration from complex sample matrices: a review. *Crit Rev Microbiol.* (2004) 30:7–24. doi: 10.1080/10408410490266410
40. Skjerve E, Rorvik LM, Olsvik O. Detection of *Listeria monocytogenes* in foods by immunomagnetic separation. *Appl Environ Microbiol.* (1990) 56:3478–81.
41. Seo KH, Brackett RE, Frank JE, Hilliard S. Immunomagnetic separation and flow cytometry for rapid detection of *Escherichia coli* O157:H7. *J Food Prot.* (1998) 61:812–6. doi: 10.4315/0362-028X-61.7.812
42. Putman TE, Suchland RJ, Ivanovitch JD, Rockey DD. Culture-independent sequence analysis of *Chlamydia trachomatis* in urogenital specimens identifies regions of recombination and in-patient sequence mutations. *Microbiology* (2013) 159(Pt 10):2109–17. doi: 10.1099/mic.0.070029-0
43. Seth-Smith HM, Harris SR, Skilton RJ, Radebe FM, Golparian D, Shipitsyna E, et al. Whole-genome sequences of *Chlamydia trachomatis* directly

- from clinical samples without culture. *Genome Res.* (2013) 23:855–66. doi: 10.1101/gr.150037.112
44. Dean FB, Nelson JR, Giesler TL, Lasken RS. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.* (2001) 11:1095–9. doi: 10.1101/gr.180501
  45. Doughty EL, Sergeant MJ, Adetifa I, Antonio M, Pallen MJ. Culture-independent detection and characterisation of *Mycobacterium tuberculosis* and *M. africanum* in sputum samples using shotgun metagenomics on a benchtop sequencer. *PeerJ* (2014) 2:e585. doi: 10.7717/peerj.585
  46. Thoendel M, Jeraldo PR, Greenwood-Quaintance KE, Yao JZ, Chia N, Hanssen AD, et al. Comparison of microbial DNA enrichment tools for metagenomic whole genome sequencing. *J Microbiol Methods* (2016) 127:141–5. doi: 10.1016/j.mimet.2016.05.022
  47. Hasan MR, Rawat A, Tang P, Jithesh PV, Thomas E, Tan R, et al. Depletion of human DNA in spiked clinical specimens for improvement of sensitivity of pathogen detection by next-generation sequencing. *J Clin Microbiol.* (2016) 54:919–27. doi: 10.1128/JCM.03050-15
  48. Charalampous T, Richardson H, Kay GL, Baldan R, Jeanes C, Rae D, et al. Rapid diagnosis of lower respiratory infection using nanopore-based clinical metagenomics. *BioRxiv* (2018) 1–40. doi: 10.1101/387548
  49. Bundock PC, Casu RE, Henry RJ. Enrichment of genomic DNA for polymorphism detection in a non-model highly polyploid crop plant. *Plant Biotechnol J.* (2012) 10:657–67. doi: 10.1111/j.1467-7652.2012.00707.x
  50. Varela I, Tarpey P, Raine K, Huang D, Ong CK, Stephens P, et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature* (2011) 469:539–42. doi: 10.1038/nature09639
  51. Depledge DP, Palser AL, Watson SJ, Lai IY, Gray ER, Grant P, et al. Specific capture and whole-genome sequencing of viruses from clinical samples. *PLoS ONE* (2011) 6:e27805. doi: 10.1371/journal.pone.0027805
  52. Depledge DP, Kundu S, Jensen NJ, Gray ER, Jones M, Steinberg S, et al. Deep sequencing of viral genomes provides insight into the evolution and pathogenesis of varicella zoster virus and its vaccine in humans. *Mol Biol Evol.* (2014) 31:397–409. doi: 10.1093/molbev/mst210
  53. Gnirke A, Melnikov A, Maguire J, Rogov P, LeProust EM, Brockman W, et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol.* (2009) 27:182–9. doi: 10.1038/nbt.1523
  54. Bachmann NL, Sullivan MJ, Jelocnik M, Myers GSA, Timms P, Polkinghorne A. Culture-independent genome sequencing of clinical samples reveals an unexpected heterogeneity of infections by *Chlamydia pecorum*. *J Clin Microbiol.* (2015) 53:1–9. doi: 10.1128/JCM.03534-14
  55. Branley J, Bachmann NL, Jelocnik M, Myers GS, Polkinghorne A. Australian human and parrot *Chlamydia psittaci* strains cluster within the highly virulent 6BC clade of this important zoonotic pathogen. *Sci Rep.* (2016) 6:30019. doi: 10.1038/srep30019
  56. Nimmo C, Doyle R, Burgess C, Williams R, Gorton R, McHugh TD, et al. Rapid identification of a *Mycobacterium tuberculosis* full genetic drug resistance profile through whole genome sequencing directly from sputum. *Int J Infect Dis.* (2017) 62:44–6. doi: 10.1016/j.ijid.2017.07.007
  57. Brown AC, Bryant JM, Einer-Jensen K, Holdstock J, Houniet DT, Chan JZ, et al. Rapid whole-genome sequencing of *Mycobacterium tuberculosis* isolates directly from clinical samples. *J Clin Microbiol.* (2015) 53:2230–7. doi: 10.1128/JCM.00486-15
  58. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt ME, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* (2014) 12:87. doi: 10.1186/s12915-014-0087-z
  59. de Goffau MC, Lager S, Salter SJ, Wagner J, Kronbichler A, Charnock-Jones DS, et al. Recognizing the reagent microbiome. *Nat Microbiol.* (2018) 3:851–3. doi: 10.1038/s41564-018-0202-y
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Copyright © 2018 Bachmann, Rockett, Timms and Sintchenko. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.