



# Metagenomics: The Next Culture-Independent Game Changer

Jessica D. Forbes<sup>1,2†</sup>, Natalie C. Knox<sup>1†</sup>, Jennifer Ronholm<sup>3,4</sup>, Franco Pagotto<sup>5,6</sup> and Aleisha Reimer<sup>1\*</sup>

<sup>1</sup> National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, MB, Canada, <sup>2</sup> Department of Medical Microbiology and Infectious Diseases, University of Manitoba, Winnipeg, MB, Canada, <sup>3</sup> Department of Food Science and Agricultural Chemistry, Faculty of Agricultural and Environmental Sciences, McGill University, Montreal, QC, Canada, <sup>4</sup> Department of Animal Science, Faculty of Agricultural and Environmental Sciences, McGill University, Montreal, QC, Canada, <sup>5</sup> Bureau of Microbial Hazards, Food Directorate, Health Canada, Ottawa, ON, Canada, <sup>6</sup> Listeriosis Reference Centre, Bureau of Microbial Hazards, Food Directorate, Health Canada, Ottawa, ON, Canada

A trend towards the abandonment of obtaining pure culture isolates in frontline laboratories is at a crossroads with the ability of public health agencies to perform their basic mandate of foodborne disease surveillance and response. The implementation of culture-independent diagnostic tests (CIDTs) including nucleic acid and antigen-based assays for acute gastroenteritis is leaving public health agencies without laboratory evidence to link clinical cases to each other and to food or environmental substances. This limits the efficacy of public health epidemiology and surveillance as well as outbreak detection and investigation. Foodborne outbreaks have the potential to remain undetected or have insufficient evidence to support source attribution and may inadvertently increase the incidence of foodborne diseases. Next-generation sequencing of pure culture isolates in clinical microbiology laboratories has the potential to revolutionize the fields of food safety and public health. Metagenomics and other 'omics' disciplines could provide the solution to a cultureless future in clinical microbiology, food safety and public health. Data mining of information obtained from metagenomics assays can be particularly useful for the identification of clinical causative agents or foodborne contamination, detection of AMR and/or virulence factors, in addition to providing high-resolution subtyping data. Thus, metagenomics assays may provide a universal test for clinical diagnostics, foodborne pathogen detection, subtyping and investigation. This information has the potential to reform the field of enteric disease diagnostics and surveillance and also infectious diseases as a whole. The aim of this review will be to present the current state of CIDTs in diagnostic and public health laboratories as they relate to foodborne illness and food safety. Moreover, we will also discuss the diagnostic and subtyping utility and concomitant bias limitations of metagenomics and comparable detection techniques in clinical microbiology, food and public health laboratories. Early advances in the discipline of metagenomics, however, have indicated noteworthy challenges. Through forthcoming improvements in sequencing technology and analytical pipelines among others, we anticipate that within the next decade, detection and characterization of pathogens via metagenomics-based workflows will be implemented in routine usage in diagnostic and public health laboratories.

**Keywords:** metagenomics, targeted-amplicon, food safety, public health, culture-independent diagnostic test, next-generation sequencing, antimicrobial resistance, molecular epidemiology

## OPEN ACCESS

### Edited by:

David Rodriguez-Lazaro,  
University of Burgos, Spain

### Reviewed by:

Dario De Medici,  
Istituto Superiore di Sanità, Italy  
Beatrix Stessl,  
Veterinärmedizinische Universität  
Wien, Austria

### \*Correspondence:

Aleisha Reimer  
aleisha.reimer@phac-aspc.gc.ca

† These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Food Microbiology,  
a section of the journal  
Frontiers in Microbiology

Received: 20 February 2017

Accepted: 29 May 2017

Published: 04 July 2017

### Citation:

Forbes JD, Knox NC, Ronholm J,  
Pagotto F and Reimer A (2017)  
Metagenomics: The Next  
Culture-Independent Game Changer.  
Front. Microbiol. 8:1069.  
doi: 10.3389/fmicb.2017.01069

## INTRODUCTION

The incidence and impact of foodborne illness constitutes a significant global issue to public health. Foodborne illness affects one in eight Canadians annually, resulting in an estimated 4 million infections, 11,600 hospitalizations and 238 deaths (Thomas et al., 2015); the leading causes of known foodborne infections include norovirus (65%), *Clostridium perfringens* (11%), *Campylobacter* spp. (8%), and non-typhoidal *Salmonella* spp. (5%). According to Health Canada, approximately 2.4 million cases or 60% of foodborne illness are attributed to unknown causes versus only 1.6 million cases or 40% causatively linked to 30 recognized foodborne microbes which include bacteria, viruses and parasites (Thomas et al., 2015). The detection of foodborne enteric pathogens and hence diagnosis of foodborne disease, historically (and still considered the gold standard) have been conducted via culture-dependent techniques, that is, the physical isolation of a bacterial pathogen. Though recognized for some time, the utilization of CIDTs has been increasing throughout the last decade, effectively transforming clinical and food microbiology laboratories (Janda and Abbott, 2014).

In clinical diagnostic laboratories, pathogen detection is increasingly contingent upon the analytic application of CIDTs, which include nucleic acid (e.g., PCR) and antigen-based tests (e.g., ELISA) among others. The extensive adaptation of CIDTs to clinical and food settings is largely reliant on inherent advantages over traditional culture-dependent diagnostic tests; use of CIDTs offers a considerably faster TAT which is crucial for (i) clinical decision-making and decreasing the unnecessary use of broad-spectrum or ineffective antimicrobials, (ii) early outbreak detection and control, and (iii) food industry release or recall of products. Moreover, most conventional CIDTs require less technical expertise and in the long-term may offer a more cost effective alternative.

The clinical use of CIDTs presents the potential to improve disease detection. First, CIDTs are reportedly more sensitive and specific than culture (Hanson and Couturier, 2016; Steyer et al., 2016). Second, reduced complexity of CIDTs allows for rapid testing thereby allowing for a higher throughput of biological specimens to be tested. Third, CIDTs can identify non-culturable or fastidious microbes such as *Campylobacter* spp. (Fitzgerald et al., 2016) or noroviruses (Jones et al., 2014). Lastly, CIDTs enhance the ability to identify polymicrobial or complex infections. Not only are CIDTs attractively useful in

the clinical microbiology laboratory but also in recent years their utilization has become widespread for routine and rapid detection of common pathogens in other sectors including public health and food safety laboratories in addition to *in situ* testing of food processing establishments and agricultural sites. Traditional diagnostics, however, rely upon tests that are tailored to the etiological agent associated with a particular syndrome.

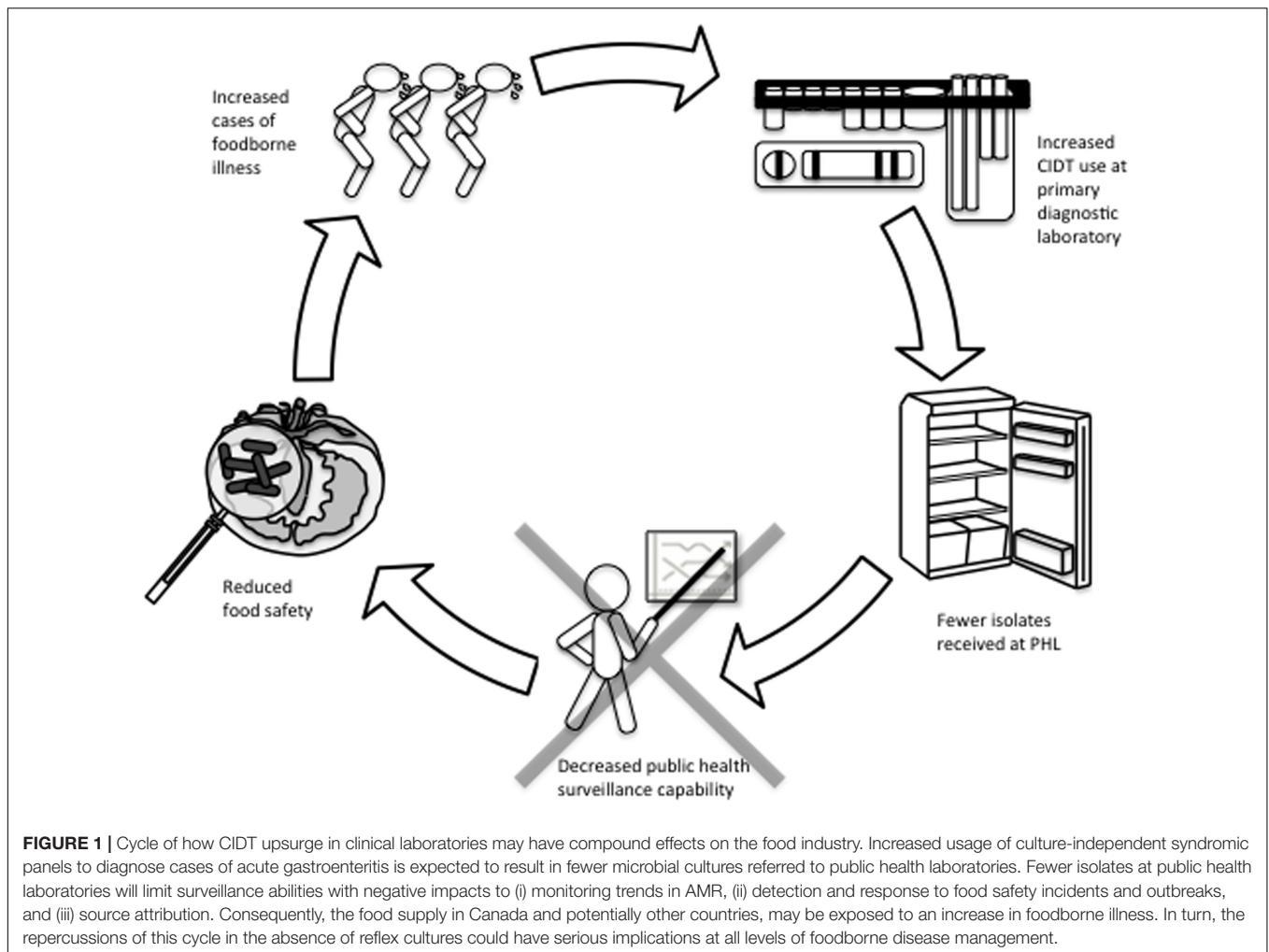
Next-generation sequencing technology has effectively transformed infectious disease research throughout the last decade. High-throughput laboratory techniques are bypassing onerous testing via complement or replacement of conventional microbiological, molecular and serological tests for identifying, typing and characterizing pathogens. WGS of cultured isolates has been extensively employed for pathogen characterization, outbreak detection, phylogenomics and microbial genome wide association studies and thus has progressed from the proof-of-principle phase to implementation in routine foodborne surveillance and outbreak response (Jackson et al., 2016).

Current public health infectious disease surveillance methodologies are generally reliant upon the frontline laboratory to refer pure culture isolates to their local or provincial public health laboratory. As we progressively enter a culture-independent era for infectious disease diagnostics, public health laboratories will inevitably receive fewer isolates. Hence, performing appropriate subtyping and AST assays to identify and track foodborne outbreaks will be difficult. Inadequate surveillance measures have the potential to negatively affect food safety via the inability to identify outbreaks and perform source attribution studies of contaminated foods. We expect this will result in an increase of contaminated foods lingering on the market thereby causing more cases of undetected foodborne outbreaks (Carleton and Gerner-Smidt, 2016). **Figure 1** shows a schematic representation of the effect of CIDTs on the food industry.

Employment of molecular methods that are highly informative will improve the all-encompassing issues associated between primary diagnostic and public health laboratories. Metagenomics for example offers the advantage of a less biased pathogen detection methodology through direct sequencing of the specimen's extracted DNA. This approach has the potential to capture a thorough representation of the microbial community (with some limitations; discussed below) thus eliminating the requirement for pure culture. Metagenomics and similar techniques traditionally have been applied to interrogate microbiomes of a particular ecological niche through sequencing of all nucleic acids recovered from a sample (Eckburg et al., 2005; Huttenhower et al., 2012; Mason et al., 2014). Moreover, a number of clinically relevant applications stand to benefit from such data – rapid identification of the etiological agent (known or novel) and gene content including virulence and AMR, or inferring functional pathways to elucidate multifaceted illnesses.

Laboratory methods that detect and identify pathogens serve two critical functions: clinical decision-making (individual level) and public health decision-making (population level). While traditional culture-based methods meet both of these needs simultaneously, CIDTs are exclusively designed to improve

**Abbreviations:** AMR, antimicrobial resistance; AST, antimicrobial susceptibility testing; CIDT, culture independent diagnostic test; CSF, cerebrospinal fluid; DNA, deoxyribonucleic acid; ELISA, enzyme linked immunosorbent assay; FDA, food and drug administration; FISH, fluorescent *in situ* hybridization; GMI, global microbial identifier; ITS, internal transcribed spacer; LAMP, loop mediated isothermal amplification; LFA, lateral flow assay; LLS, listeriolysin S; MALDI-TOF, matrix assisted laser desorption/ionization time of flight; MDR, multidrug-resistant; MLST, multilocus sequencing typing; MLVA, multilocus variable number tandem repeat analysis; NAAT, nucleic acid amplification test; NASBA, nucleic acid sequence based amplification; NGS, next-generation sequencing; PCR, polymerase chain reaction; PFGE, pulsed field gel electrophoresis; rRNA, ribosomal ribonucleic acid; SNV, single nucleotide variant; STEC, shiga-toxin producing *Escherichia coli*; SURPI, sequence-based ultra-rapid pathogen identification; TAT, turn-around-time; WGS, whole genome sequencing.



clinical decision making only. This leaves critical public health activities at great risk. A summary of techniques used in primary diagnostic labs is illustrated in **Figure 2**. Herein, we will review the current state of CIDTs as they relate to foodborne illness and food safety. We will furthermore discuss the diagnostic utility of metagenomics and comparable detection techniques in clinical microbiology, food and public health laboratories.

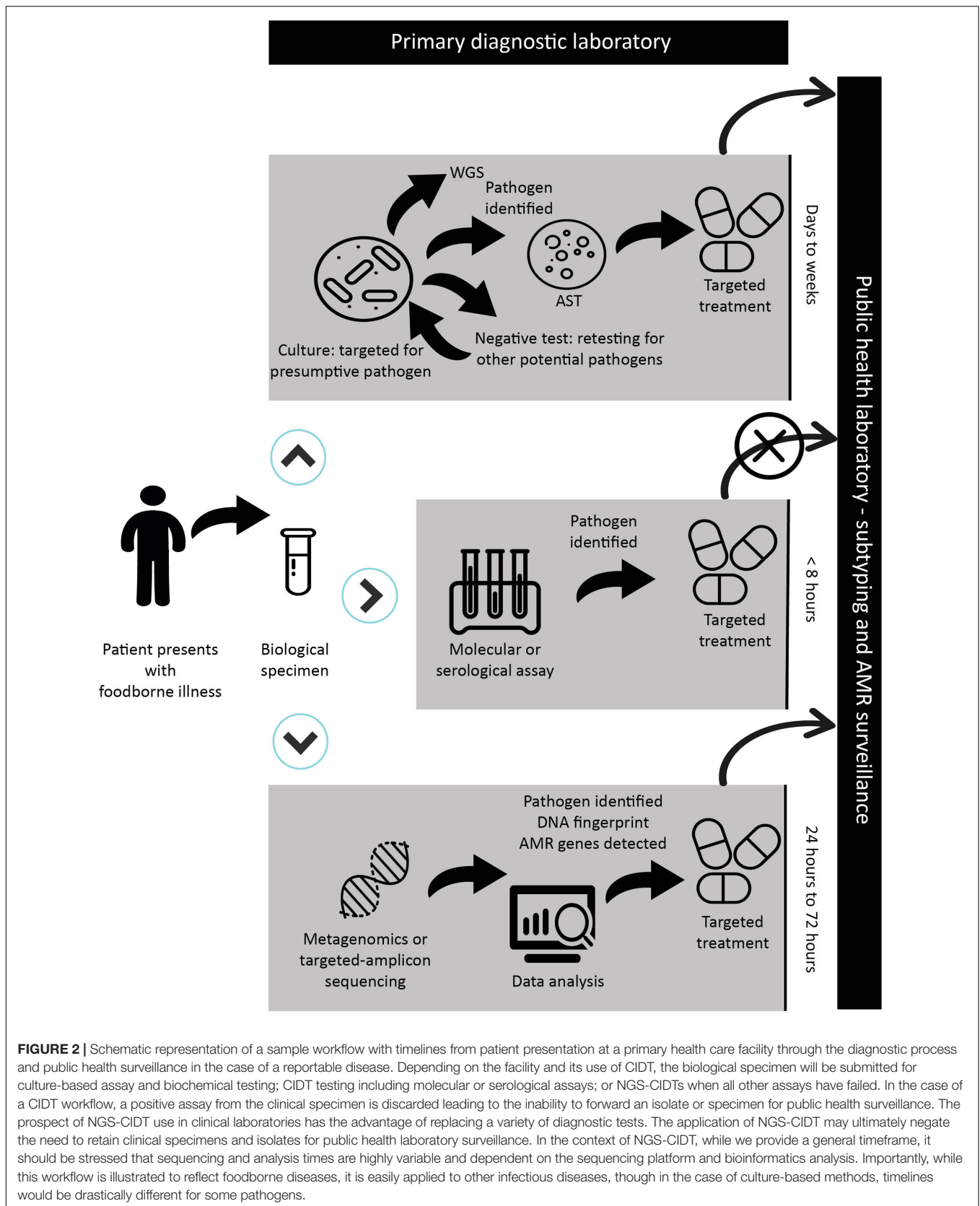
## CURRENT STATE OF CONVENTIONAL CULTURE-INDEPENDENT DIAGNOSTIC TESTS IN CLINICAL AND FOOD MICROBIOLOGY LABORATORIES

Traditional culture-based methods for detecting enteric pathogens in biological specimens, food processing environments or foods is dependent on the growth of viable and culturable microbes; biochemical tests and subtyping are further needed to confirm microbe identification and to establish genetic linkage, respectively. Though sensitive, relatively easy

and inexpensive, culture-based methods are laborious and ineffective for non-culturable or fastidious microbes. The diagnostic potential of CIDTs to circumvent the demanding task of culturing and subtyping pathogens is a widely attractive alternative.

## Conventional Culture-Independent Diagnostic Tests

Though globally, data is limited regarding the current use of CIDTs in clinical and food laboratories, many clinical laboratories in the US are reportedly in the process of converting to the use of CIDTs for the identification of enteric microbes (Shea et al., 2017). According to the CDC, significant increases in the amount of enteric infections diagnosed in the US entirely by CIDTs were reported in 2015 compared to year's prior (2012–2014). In particular, increases in positive CIDTs were reported for 4 enteric pathogens: *Campylobacter* (92%), *Shigella* (284%), *Salmonella* (247%), and STEC (120%; Huang et al., 2016). We expect the increased use of CIDTs is less dramatic in Canada, with suspected regional differences, due to health-care system differences, though there is a paucity



**FIGURE 2 |** Schematic representation of a sample workflow with timelines from patient presentation at a primary health care facility through the diagnostic process and public health surveillance in the case of a reportable disease. Depending on the facility and its use of CIDT, the biological specimen will be submitted for culture-based assay and biochemical testing; CIDT testing including molecular or serological assays; or NGS-CIDTs when all other assays have failed. In the case of a CIDT workflow, a positive assay from the clinical specimen is discarded leading to the inability to forward an isolate or specimen for public health surveillance. The prospect of NGS-CIDT use in clinical laboratories has the advantage of replacing a variety of diagnostic tests. The application of NGS-CIDT may ultimately negate the need to retain clinical specimens and isolates for public health laboratory surveillance. In the context of NGS-CIDT, while we provide a general timeframe, it should be stressed that sequencing and analysis times are highly variable and dependent on the sequencing platform and bioinformatics analysis. Importantly, while this workflow is illustrated to reflect foodborne illnesses, it is easily applied to other infectious diseases, though in the case of culture-based methods, timelines would be drastically different for some pathogens.



of published data to support this theory. For example, the Canadian health care system is not driven by private insurance, as is the case in the US, but rather, is publicly funded. To establish trends of CIDT uptake in Canada, surveys will be required to evaluate their usage in primary diagnostic laboratories. We anticipate that the percentage of infections diagnosed with CIDTs will continue to rise in clinical and food microbiology laboratories in developed countries in the coming years.

This past decade has seen a drastic expansion in the diagnostic application of available and validated CIDTs (Doggett et al., 2016; Shea et al., 2017). There is a wide degree of variation amongst implementation of CIDTs in diagnostic laboratories ranging from singleplex PCR assays to complete laboratory automation; thus, microbiologists are able to yield clinically actionable results of superior quality ultimately benefiting patient health and outcome. Conventional CIDTs can be assigned to one of two methodologies: nucleic acid-based methods and antigen-based tests. The use of nucleic acid-based assays to identify enteric pathogens presents with several advantages over culture-dependent tests in clinical and food settings. Their use allows for high levels of sensitivity and specificity and offers an added benefit in their ability to detect toxin-producing genes or other important biomarkers. Numerous nucleic acid assays are in routine use including PCR (singleplex, multiplex, quantitative, quantitative reverse-transcription, real-time), amplification such as LAMP and NASBA, DNA microarray, microfluidic chip and MALDI-TOF mass spectrometry. Alternatively, antigen-based tests for enteric pathogen detection include ELISA and LFA; reviewed in Zhao et al. (2014). A detailed description of conventional CIDTs is outside the scope of this review and we refer readers to an exhaustive review of CIDTs (Doggett et al., 2016).

A growing number of US FDA – approved syndromic panels for multiple pathogen detection in addition to laboratory-developed tests have facilitated the upsurge of CIDT application for acute gastroenteritis in clinical laboratories. These panels employ PCR to detect unique DNA sequences to enable identification of enteric pathogens. Several commercially available gastrointestinal multiplex panels are in current use and include BDMax, FilmArray, Luminex, Prodesse, and Verigene. Gastrointestinal CIDT panels are advantageous over culture-based diagnostics for several reasons. First, the gastrointestinal panels are designed to detect multiple pathogens in a single assay, in some cases (e.g., FilmArray) up to 22 microbes. Commercially available assays differ in their microbial targets detecting distinct subsets of bacteria, viruses and/or parasites. Second, the various assays have different TAT – most ranging from 1 to 4 h – though even the most time-intensive assays (8 h) are considerably faster than stool culturing which can take 2–5 days dependent on the microbe being cultured. Third, multiplex panels retain the ability to detect polymicrobial infections, which is particularly relevant in clinical settings. Lastly, multiplex panels are highly sensitive (90–100%) and specific (98%) (Buchan et al., 2013; Navidad et al., 2013; Onori et al., 2014; Buss et al., 2015; Harrington et al., 2015).

## NEXT-GENERATION SEQUENCING AS A DIAGNOSTIC ASSAY

### Whole Genome Sequencing

Next-generation sequencing and other high-throughput laboratory techniques are circumventing laborious testing by active replacement or complement to traditional microbiological and molecular tests for identifying, typing and characterizing pathogens. An increase in read-length, output and quality of short-read NGS technologies has made it possible to apply WGS as a genomic surveillance system of foodborne diseases and outbreak management (Köser et al., 2012). In most cases, WGS provides higher discriminatory power than the combined effort of multiple conventional typing assays such as PFGE, MLST, MLVA, phage typing, virulence typing, and AST. Moreover, WGS is more reliable in a shorter time period and can be performed in a single comprehensive procedure thus allowing for rapid and sensitive pathogen identification and similarly for reporting to the food industry and government responsible for public health decision-making (Ashton et al., 2016; Jackson et al., 2016).

For some clonal microbes, molecular typing methods have proven unable to accurately discriminate genetically distinct isolates. As an example, in past outbreak investigations, highly clonal *Salmonella enterica* serovars such as *S. enteritidis* have been challenging to resolve using PFGE (the gold standard typing method). Taylor et al. (2015) reviewed seven epidemiologically supported clusters and revealed that SNV analysis was able to accurately discriminate cluster isolates from sporadic and suspect samples with high epidemiological concordance. PFGE results would have suggested that distinct outbreak isolates with no epidemiological link might have originated from the same source. The low resolution of PFGE was also exemplified in the 2008 listeriosis outbreak whereby suspected cases emerged with a mixture of two distinct but closely related PFGE patterns (Gilmour et al., 2010); WGS analysis revealed that isolates from both PFGE patterns were genetically similar apart from a large prophage responsible for the 40 kb band shift difference seen in the PFGE patterns of suspected isolates. WGS was only a research tool in those early days of public health genomics; however, its use in this large-scale outbreak was able to confirm that outbreak strains subtyped by either PFGE pattern were genetically similar apart from the insertion of a large prophage. The inclusion or exclusion of clinical strains associated with the outbreak and the resulting case definition can alter the case epidemiology and food safety investigation efforts.

Several NGS sequencing platforms are available; each with their own advantages and disadvantages differing in sequencing time, read length, cost and others, reviewed elsewhere (Mardis, 2017). At present, WGS, though still dependent on the presence of isolates, represents one of the predominant investigative tools to rapidly and accurately identify microbes, perform subtyping, cluster epidemiological relevant isolates in outbreak investigations and detect AMR genes and virulence profiles. The use of WGS has also proven exceedingly useful in retrospective outbreak investigations. In 2013, PulseNet USA incorporated

WGS as a surveillance tool for all *Listeria monocytogenes* isolates (Jackson et al., 2016) and is poised to begin WGS of *Campylobacter*, STEC and *Salmonella*; PulseNet Canada is similarly performing routine WGS of all *L. monocytogenes* and intends to begin sequencing all *Salmonella* isolates routinely in 2017. Moreover, PulseNet Canada also performs WGS on select cluster or outbreak investigations of STEC and other organisms.

## Diagnostic Metagenomics and Comparable Detection Techniques

Diagnostic approaches including culture and non-NGS CIDTs such as PCR or serology continue to represent the gold standard for infectious disease diagnostics. Each of these methods, however, is disadvantaged in that they represent a targeted detection methodology and hence *a priori* knowledge or hypotheses to identify the etiological agent in the sample are required. It has been suggested that the limited capacity of conventional diagnostics including culture and non-NGS CIDTs are partly responsible for failing to detect an etiological culprit in a considerable number of cases (Denno et al., 2012). High-throughput targeted-amplicon and shotgun metagenomics sequencing methods circumvent this limitation via broad-range detection of either a subset of microbes (targeted-amplicon) or all microbes (shotgun metagenomics). Further, diagnostic shotgun metagenomics offers an added advantage with the possibility to identify previously uncharacterized microbes or emergent and novel pathogens (Forster et al., 2016).

The dramatic increase of microbiome research in the last decade is effectively driven by the widespread usage of high-throughput DNA sequencing and the associated decrease in sequencing cost. In the context of the microbiome, NGS allows for the complete description of all genomic content of microbial communities (e.g., bacterial, viral, eukaryotic microbes) in a technique referred to as shotgun metagenomics (Sharpton, 2014). Though most microbiome studies are fundamentally designed to describe commensal populations (The NIH HMP Working Group, 2009) or alternatively, to investigate dysbiosis in distinct human body compartments affected by disease (Forbes et al., 2016), or to determine the efficacy of prebiotics (Alfa et al., 2017) and others, metagenomics can similarly be utilized for the identification of pathogens in clinical (Mongkolrattanothai et al., 2017) or food (Aw et al., 2016) samples. Like other molecular or serological CIDTs, the detection of microbes in metagenomics is independent of culture (Schloss and Handelsman, 2005) in contrast to WGS, which is also reliant on a pure culture (Hasman et al., 2014). Metagenomics and comparable detection techniques generate immense quantities of large-scale sequence data thus bioinformatics and/or computational approaches are required to assign sequences to particular microbes, microbial functions or other descriptors of relevance (Sharpton, 2014).

The majority of microbiome studies to date have focused specifically on the bacterial portion of microbiomes rather than characterizing the microbial communities across all domains of life in addition to viruses (shotgun metagenomics). As such, these analyses are often performed by high-throughput targeted-amplicon sequencing of a universal phylogenetically informative

genetic marker; the 16S rRNA gene is most commonly used though other markers are similarly able to discriminate between prokaryotes including *cpn60* (Schellenberg et al., 2016), *rpoB* (Case et al., 2007), 23S rRNA (Anthony et al., 2000), and others. Universal targets for eukaryotic organisms include the 18S rRNA gene (Kataoka et al., 2016) or alternatively, for specifically characterizing fungal populations the ITS is frequently used (Argenio et al., 2016).

While the science community has persistently been using the term metagenomics interchangeably to describe both high-throughput targeted-amplicon and shotgun metagenomics studies to profile microbial populations, it is important to make a clear distinction between both methodologies. Targeted-amplicon and shotgun metagenomics sequencing each present with advantages and disadvantages (Table 1). Features differentiating the two methods include targeted microbial community, associated costs and computational and technical expertise requirements among others. Thus, the decision to use one approach versus the other should be made with careful consideration and will be highly dependent on the research and/or diagnostic goals and hypotheses in question (Table 2). In this regard, targeted-amplicon assays are more apt for describing a specific group of microbes (e.g., bacteria) whereas a shotgun metagenomics approach is more suitable for characterizing the entirety of microbial DNA in a given sample limited by sequencing technology used and sample matrix with high host DNA-containing material.

## Metagenomics and Comparable Detection Techniques in Food Safety

Culture-based techniques are still considered the gold standard in the food industry, including sectors such as business operators, government regulatory agencies and national or international compliance testing. Countries such as Canada have policies in place that require quantitative (Pagotto et al., 2011a) or qualitative (Pagotto et al., 2011b) approaches to determine viable pathogens such as *L. monocytogenes*<sup>1</sup>. This is attributed to how foods are categorized based on the ability to support (or not), the priority assigned to high-risk foods and the ability of the pathogen to survive during the shelf life of the food<sup>1</sup>. Moreover, the target pathogen is often in such low numbers that in the presence of the background microbiota, enrichment is required – and culturing is the most effective strategy (Gill, 2017). While a molecular fingerprint *per se* is not currently required for compliance activity, it is invaluable to epidemiological investigations and for source attribution. As such, current approaches necessitate a physical isolate in order to generate molecular fingerprints. Further, it is important to note that more than one “molecular type” may be implicated in an outbreak and so, culture-dependent technologies continue to have an important role. This was shown in the *Listeria* cantaloupe outbreak where multiple serovars and five different molecular subtypes were found (McCollum et al., 2013). It is anticipated that bioinformatics will help play a role in being able to differentiate

<sup>1</sup>[http://www.hc-sc.gc.ca/fn-an/legislation/pol/policy\\_listeria\\_monocytogenes\\_2011-eng.php](http://www.hc-sc.gc.ca/fn-an/legislation/pol/policy_listeria_monocytogenes_2011-eng.php)

**TABLE 1** | Summary of the advantages and disadvantages to each high-throughput sequencing approach for unbiased detection.

	Targeted-amplicon sequencing		Shotgun metagenomics sequencing	
	Advantages	Disadvantages	Advantages	Disadvantages
Microbial target(s) of interest	<ul style="list-style-type: none"> <li>Target is specific to a particular microbial group (e.g., 16S rRNA common for bacteria, archaea, 18S rRNA for eukaryotes, ITS for fungi, RdRP for RNA viruses).</li> </ul>	<ul style="list-style-type: none"> <li>Requires <i>a priori</i> knowledge for microbial group target.</li> </ul>	<ul style="list-style-type: none"> <li>Can sequence all DNA in a given sample (e.g., bacteria, archaea, eukaryotes, parasites, and viruses).</li> </ul>	<ul style="list-style-type: none"> <li>Virome assays require complex sample and nucleic acid work-ups.</li> </ul>
Abundance profiling	<ul style="list-style-type: none"> <li>Can use relative abundance changes to compare microbiomes across different samples or treatments.</li> <li>Can capture abundance of rare taxa provided that sequencing depth is sufficient.</li> </ul>	<ul style="list-style-type: none"> <li>Universal target chosen may be present in varying copy numbers across different taxa (e.g., 16S rRNA). PCR amplification bias, primer bias and errors.</li> <li>Absolute abundance difficult to impute.</li> </ul>	<ul style="list-style-type: none"> <li>Universal markers can be inferred from metagenomics datasets.</li> </ul>	<ul style="list-style-type: none"> <li>High abundance of host DNA can make it challenging to sequence low abundance microbial DNA.</li> <li>Low abundance taxa difficult to identify. Can be difficult to accurately bin each sequence to a genome.</li> </ul>
Taxonomic assignment	<ul style="list-style-type: none"> <li>Relatively easy to taxonomically classify sequences using a variety of validated tools and curated databases.</li> </ul>	<ul style="list-style-type: none"> <li>Databases can be self-limiting and have the potential to exclude novel microbes.</li> <li>Universal targets within microbial groups can give variable taxonomic classifications.</li> <li>Taxonomic resolution variable – species level identification should be interpreted with caution.</li> </ul>	<ul style="list-style-type: none"> <li>Plethora of software using phylogenetically informative gene markers.</li> </ul>	<ul style="list-style-type: none"> <li>High proportion of taxonomically uninformative sequences are discarded.</li> <li>Availability and access to comprehensive and curated databases across all microbial groups limited.</li> </ul>
Cost	<ul style="list-style-type: none"> <li>Low cost</li> <li>Can be carried out on most bench-top sequencers and sequencing platforms.</li> </ul>		<ul style="list-style-type: none"> <li>Can be carried out on most bench-top sequencers and sequencing platforms.</li> </ul>	<ul style="list-style-type: none"> <li>Can be cost prohibitive depending on the sequencing depth, sample type, and microbe(s) of interest.</li> <li>If high host DNA is expected or interest is in the low-abundance microbes or rare taxa, use of a higher throughput sequencer (Illumina HiSeq), may be required.</li> </ul>
Computational requirements	<ul style="list-style-type: none"> <li>Most analysis steps can be carried out on a modern desktop.</li> </ul>	<ul style="list-style-type: none"> <li>Large datasets (high sample number and/or sequencing coverage) may require access to a high performance computing cluster dependent on analytical pipeline chosen.</li> </ul>	<ul style="list-style-type: none"> <li>Cloud computing services are available for metagenomics data analysis for those without access to a high performance computing cluster.</li> </ul>	<ul style="list-style-type: none"> <li>High performance computing environment absolutely necessary.</li> <li>Cloud computing – potentially cost-prohibitive and might not have all available pipelines and/or software.</li> <li>Data privacy and sensitivity may prohibit the use of commercial cloud computing services.</li> </ul>
Technical expertise	<ul style="list-style-type: none"> <li>Moderate to high technical expertise is required depending on the analytical pipeline chosen.</li> </ul>			<ul style="list-style-type: none"> <li>High technical expertise required.</li> </ul>

**TABLE 2** | Overview of appropriate usage for each unbiased high-throughput sequencing approach.

Study goals/purpose	Suggested sequencing approach
Characterization of a particular microbial group (excluding viruses) in sample(s)	High-throughput targeted-amplicon sequencing; utilize shotgun metagenomics sequencing if interested in high taxonomic resolution above genus level.
Characterization of all microbial DNA in sample(s)	Metagenomics shotgun sequencing.
Pathogen detection	Dependent on the sample: <ul style="list-style-type: none"> <li>• If the etiological agent is suspected to be of viral origin a shotgun metagenomics approach is warranted.</li> <li>• If the sample type contains a high host DNA load (e.g., blood) should consider a targeted-amplicon or deep shotgun metagenomics sequencing approach. The latter may be cost prohibitive and require access to a high-throughput sequencer (e.g., Illumina HiSeq).</li> <li>• Low biomass samples (e.g., BAL/CSF), might require a targeted-amplicon sequencing approach initially. Shotgun metagenomics sequencing may not be able to sequence the infectious agent adequately (e.g., only a few sequences produced which may only yield a confounding signal).</li> </ul>
Functional profiling	Functional profiles can be inferred with a targeted-amplicon sequencing approach, however, results should be interpreted with caution due to the limitations of inferring gene function with universal targets. A shotgun metagenomics approach would yield more appropriate and reliable conclusions.
SNV or clonal isolate detection studies	Shotgun metagenomics sequencing.
Novel microbial identification and characterization	Targeted-amplicon sequencing relies on curated databases of known microbes and may not be able to adequately analyze novel microbes in an unbiased technique. Shotgun metagenomics would be recommended.

BAL, bronchoalveolar lavage; CSF, cerebrospinal fluid; SNV, single nucleotide variant.

the different molecular types present in a single food item (unlike a clinical specimen that tends to be contaminated with a single type). Molecular techniques that are currently used to compare isolates during an outbreak investigation, such as PFGE, ribotyping and gene-specific PCR, require an isolate as a starting point (Ronholm et al., 2016); these procedures would generally take approximately 7 days (Figure 3). If however, a high-sensitivity metagenomics methodology, capable of reliably detecting foodborne pathogens in samples with high-levels of background microbiota were developed, the time for outbreak recognition, causative agent sub-typing, secondary analyses for virulence and AMR genes and source attribution would be reduced (Figure 3). However, even after suitable techniques are developed, widespread use of CIDT as a replacement for culture-based techniques will still face significant challenges in food regulation and industry. International regulatory agencies with stakes in food safety such as Health Canada, US FDA, French agency for Food, Environmental and Occupational Health Safety (ANSES) are still in the initial phases of accepting WGS results as a replacement for more traditional techniques for outbreak investigations<sup>2</sup>. Some countries though, are much further along in this process than others (Allard et al., 2016). Using Canada and *Listeria* as an example, verification controls in processing environments is directly tied to consumer risk should the food be contaminated (Pagotto et al., 2011a). The ecology of *Listeria* in ready-to-eat products with respect to foods capable (or not) of supporting growth (or survival) may be elucidated through the use of WGS. Metagenomics has been used to help determine aspects related to enrichment (Ottesen et al., 2016); interference to even the ability of only detecting a single species when more than one may be present (Ottesen et al., 2016). A relatively

recent study investigated the efficacy of NGS techniques to detect food pathogens (Leonard et al., 2015). An FDA culture-based protocol was performed on spinach spiked with STEC in addition to shotgun metagenomics sequencing. In particular, the study aimed to address limits of detection, sensitivity and specificity levels. The authors reported an expected level of contamination (approximately 10 cfu/100 g) and were able to accurately detect strain-level and virulence information within 8 h of enrichment at a sequencing depth of 10 million reads.

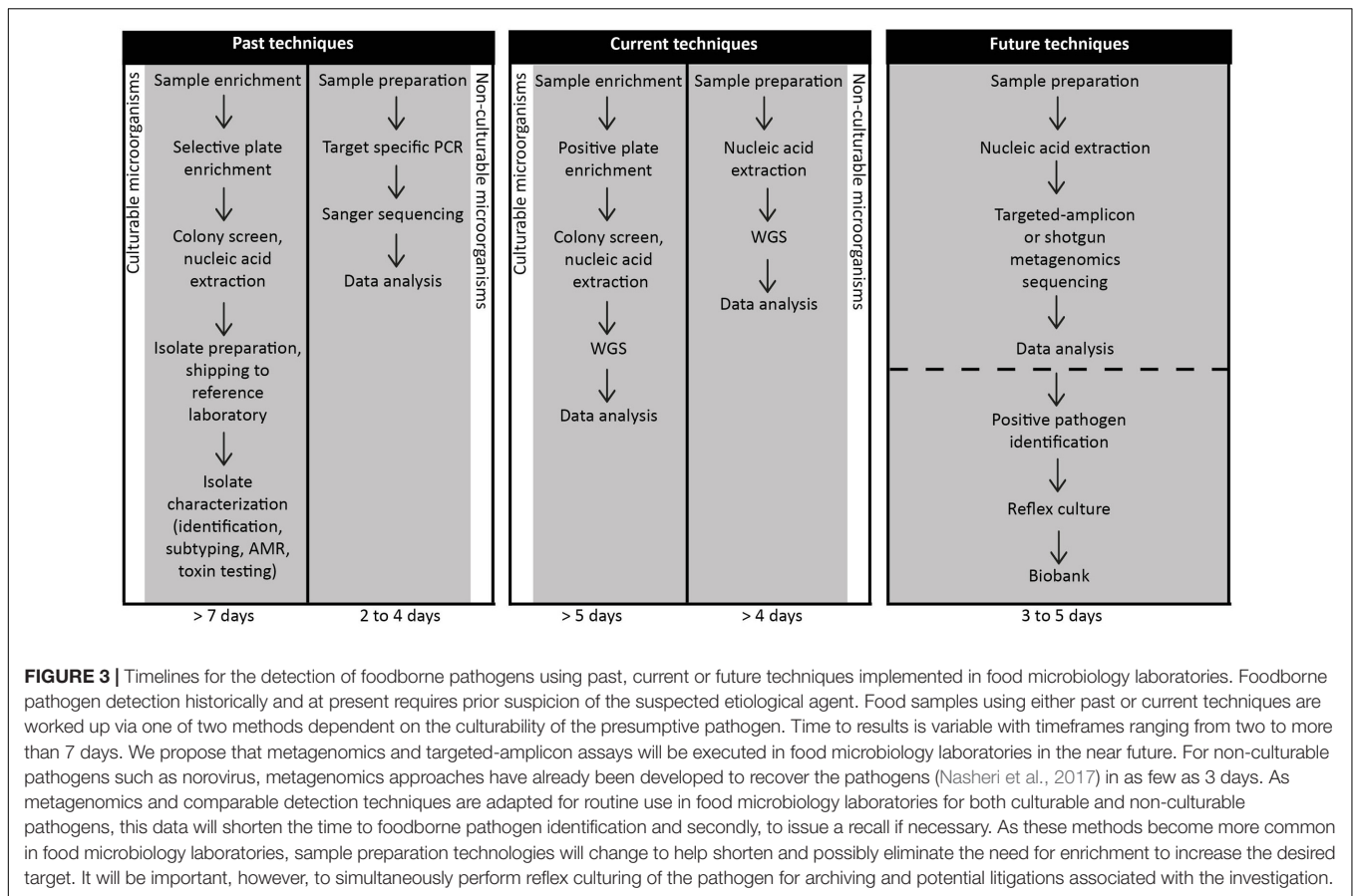
Next-generation sequencing-based approaches such as WGS rely on sequencing technologies, bioinformatics pipelines (Lambert et al., 2015) and high-quality reference databases (Allard et al., 2016). Therefore, widespread acceptance of WGS results, as a complete replacement for traditional molecular techniques must likely precede the introduction of shotgun metagenomics or targeted-amplicon approaches in food outbreak investigations and compliance testing. There are, however, many instances where metagenomics are being used in the food industry for rapid screening and research, but for regulatory testing purposes and for outbreak investigations, presumptive positive results must still be culturally confirmed<sup>3</sup>.

Metagenomics is currently being used in both the food industry and in research for many applications including: taxonomic profiling of microbiological food products and supplements, directing efforts to improve culture techniques, identification of non-culturable or fastidious pathogens and detection of co-contamination. Targeted-amplicon analyses for example, are well-suited to study, characterize and catalog changes in the bacterial populations that take place during fermentation reactions and long-term storage of traditional fermented foods such as soft cheese (Escobar-Zepeda et al., 2016),

<sup>2</sup><http://www.fao.org/documents/card/en/c/61e44b34-b328-4239-b59c-a9e926e327b4/>

<sup>3</sup><http://www.hc-sc.gc.ca/fn-an/res-rech/analy-meth/microbio/index-eng.php>





kimchi (Jung et al., 2011; Park et al., 2012), and kefir (Leite et al., 2012; Korsak et al., 2015). Targeted-amplicon analyses are also suited to evaluate the composition of probiotic supplements (Bergholz et al., 2014; Morovic et al., 2016). This technique, when used to evaluate probiotic supplements, could be of particular interest to regulatory bodies, since based on preliminary work, many of these products contain neither the number nor species of probiotic bacteria claimed on the label (Morovic et al., 2016).

Targeted-amplicon and shotgun metagenomics have also both been used to improve culture-based enrichment techniques by allowing for detailed characterization of population dynamics of the background microbiota during enrichment (Gorski, 2012; Ottesen et al., 2013, 2016; Jarvis et al., 2015; Zilelidou et al., 2016). When food items are added to a non-selective media, microbes from the background microbiota can co-enrich with pathogens interfering with detection and recovery of pathogens. For microbes from the microbial population closely related to the pathogen this can also be true of differential or selective media as well. As an example, tomatoes are well-known to be frequently implicated as the source of human *Salmonella* infections, however, isolating *Salmonella* from the tomato phyllosphere using a culture-dependent method has proven challenging (Ottesen et al., 2013). Shotgun metagenomics sequencing of enrichment media revealed that *Paenibacillus* spp.

readily outcompetes and may even kill *Salmonella* during the enrichment phase suggesting that an alternate enrichment media be used when investigating tomato contamination (Ottesen et al., 2013). A similar enrichment problem occurs with cilantro, where traditional pre-enrichment steps encourage the growth of Gram-positive Firmicutes instead of Proteobacteria (*Salmonella*), suggesting that an alternate pre-enrichment media should be developed for cilantro testing (Jarvis et al., 2015). The usefulness of targeted-amplicon and shotgun metagenomics sequencing in informing enrichment strategies has similarly been demonstrated in *L. monocytogenes* (Ottesen et al., 2016) and *E. coli* (Margot et al., 2016).

The use of CIDTs are much more readily acceptable for outbreak delineation and regulatory testing for non-culturable or fastidious foodborne pathogens such as viruses (Vinjé, 2014) and parasites (Buss et al., 2013). For example, noroviruses, which until very recently were non-culturable (Ettayebi et al., 2016), are currently detected by Sanger sequencing of partial regions of the polymerase and capsid sequences, directly from clinical samples (Vinjé, 2014). However, the same techniques are rarely successful on food samples since the viral titer in the food is generally too low (Gentry-Shields and Jaykus, 2015; **Figure 3**). Sanger sequencing of norovirus requires several PCR rounds to obtain usable information, and still only yields partial genome sequences (Verhoef et al., 2012). To successfully

delineate outbreaks with 100% specificity, a full capsid sequence is required (Nasheri et al., 2017). Currently, several WGS workflows exist for norovirus analysis, including the sequencing of several overlapping PCR fragments (Kundu et al., 2013; Won et al., 2013), target enrichment of the norovirus genome using custom designed RNA baits (Brown et al., 2016) and metatranscriptomics approaches (Bavelaar et al., 2015; Nasheri et al., 2017; **Figure 3**). The metatranscriptomics approach has the advantage of being rapid and further, a *de novo* assembly can be obtained. This approach is thorough and additionally, allows for the presence of other co-infections to be detected from the same sequence data (Nasheri et al., 2017). Metagenomics sequencing has also been used to investigate various food products for the presence of potential or emerging pathogens in the absence of a defined outbreak (Kawai et al., 2012; Aw et al., 2016). A shotgun metagenomics approach has recently identified several human and animal viruses in fresh produce (Aw et al., 2016), and it can be expected that this type of analysis will be repeated for other food products to provide greater insight into the scope of viral contamination of the food supply. Non-culturable parasites present similar challenges to non-culturable viruses; shotgun metagenomics has also provided a suitable solution for detecting parasitic contamination. Between 2008 and 2010, Japan experienced multiple outbreaks of gastroenteritis that included more than 1300 cases caused by an unknown etiological agent. Shotgun metagenomics was used to demonstrate that a parasite (*Kudoa septempunctata*) was the likely cause of the outbreak (Kawai et al., 2012).

The current regulatory dilemma is that it is impossible to tell directly from the presence of DNA if an organism is viable or not, and although DNA degrades over time, false positive rates of pathogen identification due to the detection of naked DNA have the potential to be quite high. This is where food differs from clinical samples. If pathogen DNA is detected in clinical samples, there is a very high-probability that the pathogen is alive and replicating in the patient. However, finding pathogen DNA in foods has quite a different interpretation. A well-known example is *Listeria* in smoked fish, where its DNA can be present due to dead cells but molecular methods alone would generate incorrect compliance activity based on a PCR-based method (Gambarin et al., 2012; Law et al., 2015). In addition, food-processing techniques (such as thermal, high-pressure, radiation exposure, and others) are known to kill bacteria, but leave detectable pathogen DNA in the food matrices. This problem with false-positive detection is a significant hurdle to overcome before metagenomics approaches will be useful in outbreak delineation or compliance testing in the food industry (Law et al., 2014). Moreover, when secondary analyses including virulence or AMR gene detection are applied to the metagenomics data, and particularly if the genes are known to be mobile, there is a current lack of bioinformatics pipelines available to accurately predict if the virulence or AMR genes belong to the pathogen, or rather to the background microbiota. It has been suggested that the alternative use of mRNA could contribute to solving this problem, although, its ability to persist in various samples has also been implied (Leimena et al., 2013).

## Diagnostic Case Studies Revealing the Usefulness of Metagenomics and Comparable Techniques

Numerous studies have investigated the capacity of diagnostic metagenomics for infectious diseases from clinical specimens; these studies have proven the methodology clinically useful from an analytical perspective and secondly, in an appropriate timeframe that can nevertheless yield positive patient outcomes (Nakamura et al., 2008; Wilson et al., 2014; Mongkolrattanothai et al., 2017). Both high-throughput targeted-amplicon and shotgun metagenomics sequencing have been performed on biological specimens for unknown pathogen detection either for research purposes or when all previous diagnostic tests have proven inconclusive. Importantly, the use of both high-throughput sequencing strategies to determine etiological causation is highly experimental due to inherent limitations (discussed in upcoming section) and typically only performed on select cases in diagnostic settings. Preceding widespread adoption and implementation of these methodologies, case studies are needed to evaluate concordance and compatibility with conventional diagnostic methods. Interpretation and reporting guidelines, in addition to criteria are likewise required prior to real-time use of this technology.

From a diagnostics perspective, metagenomics sequencing was first employed to detect human-associated viruses (rather than bacteria), which is unsurprising given the challenges associated with culturing most viruses. Throughout the last decade, viral metagenomics (viromics) has been used as a tool to diagnose cases of acute gastroenteritis and further, to determine etiological causation retrospectively in gastroenteritis outbreaks (Finkbeiner et al., 2009; Smits et al., 2014). This technique has proven particularly useful in determining causative agents whereby culture yielded negative results. It is hypothesized that novel viruses may account for a fraction of idiopathic gastroenteritis cases especially given that a large portion of global viral diversity has yet to be discovered (Anthony et al., 2013); hence methodologies incorporating “unbiased” and systematic high-throughput sequencing have the capacity for novel virus detection and discovery, which may have otherwise evaded routine diagnostic testing. In this regard, an early study aimed to characterize the viral populations present in pediatric diarrheal specimens (Finkbeiner et al., 2008). Interestingly, this study adapted a “micro-mass sequencing” approach. This included a minimal sample quantity (<100 mg stool), minimal sample purification and minimal sequencing (e.g., 384 reads per sample). Known enteric viruses including rotaviruses, caliciviruses, astroviruses, and adenoviruses were detected in addition to several sequences from at least nine putative novel viruses. Recently, a shotgun metagenomics approach was applied retrospectively to detect novel and known viruses associated with gastroenteritis outbreaks (Moore et al., 2015). Though the study failed to identify any novel pathogens, eight viruses, and one parasite were detected. The authors’ concluded that metagenomics could be useful to detect pathogens whereby routine testing has failed.

An early study tested the capability of shotgun metagenomics sequencing to detect bacterial pathogens during a case of acute gastroenteritis following culture based diagnostics where no candidate pathogens were identified (Nakamura et al., 2008). This case reported 156 *Campylobacter jejuni* sequences from a stool sample obtained when the patient was symptomatic versus no sequences identified in a sample obtained 3 months post-illness. This investigation represents one of the first proof-of-concept studies to be conducted in a clinical diagnostic setting.

A relatively recent case study emphasizes the clinical relevance of metagenomics as a diagnostic assay (Wilson et al., 2014). Briefly, a 14-year-old boy presented to a medical facility three times with complaints of fever and headache that ultimately advanced to hydrocephalus and status epilepticus. Though diagnostic workup was inconclusive on several occasions, an MRI eventually revealed encephalitis-like indications. CSF and serum specimens were subjected to shotgun metagenomics sequencing. Within 48 h of specimen receipt, bioinformatics analysis of CSF revealed a high abundance of sequences with homology to the *Leptospiraceae* family and mapped most closely to the pathogenic *Leptospira borgpetersenii* genome. The patient was therefore treated for neuroleptospirosis with intravenous penicillin G and markedly improved. In this context, treatment was initiated on the basis of metagenomics evidence and congruent clinical presentation prior to the completion of validated confirmatory testing. Five months post-treatment, targeted PCR and Sanger sequencing identified *L. santarosai* as the infectious agent. This case highlights various advantages and limitations associated with diagnostic metagenomics that will therefore be discussed in the following section.

Diagnosis of brucellosis historically has been a challenge as this illness presents with a continuum of clinical manifestations. In addition, diagnosis is further limited by current diagnostics via their inadequate sensitivity and specificity (Mongkolrattanothai et al., 2017). A recent report described a case whereby a shotgun metagenomics analysis of the patient's CSF was used to provide an accurate diagnosis (via *Brucella* spp. detection) and directed towards appropriate antibiotic therapy ultimately leading to a favorable patient outcome (Mongkolrattanothai et al., 2017). This study led to the development of a validated diagnostic assay in the CLIA-certified University of California, San Francisco clinical microbiology laboratory<sup>4</sup>.

A recent study aimed to assess the concordance between various diagnostic (PCR, qPCR) and NGS techniques (16S rRNA targeted-amplicon and shotgun metagenomics) for *C. difficile* infection (Zhou et al., 2016). Intriguingly, of PCR and qPCR positive *C. difficile* samples, this microbe was detected in 90.9% of samples via 16S rRNA analysis versus 86.3% with shotgun metagenomics. Moreover, *C. difficile* was co-detected with several known enteric pathogens such as norovirus and sapovirus which adds further credence to the efficacy of NGS CIDs to firstly, detect foodborne pathogens in an unbiased technique and secondly, to detect

polymicrobial infections. This study reflects current limitations of metagenomics techniques in that their sensitivity is still relatively low.

We mentioned previously that high-throughput targeted-amplicon sequencing is at the moment largely research based. A recent study described the usefulness of 16S rRNA targeted-amplicon sequencing in an exploratory setting (Almonacid et al., 2016). This study utilized 46 bacteria and archaea encompassing 15 genera and 31 species of microbes in the human gastrointestinal tract selected on the basis of clinical relevance including pathogenic, commensal, and probiotic prokaryotes. A bioinformatics annotation pipeline specifically tailored to have high prediction performance was applied. For example, taxonomies were reported based on 100% identity over the entire 16S rRNA V4 region and curated databases were procured for each taxon using various optimizing parameters such as sensitivity, specificity, precision and a negative predictive value. Applying a threshold of 90% for each parameter, 28 of 46 targets could be detected. Microbiome specimens from a cohort of 897 healthy persons were used to define a reference range that could be used to establish clinically applicable relative abundances for each target. The authors concluded that their assay accurately identified and quantitated all targets and known pathogens from each sample (clinical or synthetic). Thus, assays such as this may facilitate improvements to patient diagnosis, treatment, monitoring, and epidemiological study.

The use of diagnostic metagenomics is less established in the framework of parasitic infections; however, studies are beginning to investigate their role (or merely, presence) in the human gastrointestinal tract, in addition to elucidating genomic epidemiology. Moore et al. (2015) also detected a parasitic candidate – *Dientamoeba fragilis*. While the role of *D. fragilis* in causing gastroenteritis remains unclear it is suggested that in the absence of other enteric pathogens that this parasite could be considered a causative agent (Barratt et al., 2011). This study also revealed an interesting phenomenon: while *D. fragilis* was detected in 10.9% of undiagnosed outbreak samples, a higher frequency was reported (44%) in pediatric samples. A more recent study applied several diagnostic techniques (metagenomics, microscopy, and multiplex PCR) to four diarrheal samples as a means to detect multiple pathogens simultaneously (Schneeberger et al., 2016). Metagenomics detected 8–11 plausible enteric pathogens in all samples. Specifically with bacterial pathogens, diagnostic agreement between PCR and metagenomics was high though metagenomics did identify several bacteria not detected by PCR. Perhaps more interesting, however, was the finding that microscopy could detect some helminth and protozoan infections that metagenomics could not, again reflecting sensitivity issues inherent of metagenomics techniques. Parasitic metagenomics has also shown effective in the investigation of *Cyclospora cayentanensis* – a coccidian parasite responsible for several food and waterborne outbreaks worldwide (Cinar et al., 2015) and in a case study of Malayan filariasis caused by *Brugia malayi* (Gao et al., 2016).

<sup>4</sup><https://www.ciapm.org/project/precision-diagnosis-acute-infectious-diseases>



## Technical Challenges of Metagenomics Sequencing and Analysis

The diagnostic applicability of clinical metagenomics presents with its own shortcomings and hence is currently limiting the widespread use of this approach in clinical and food microbiology laboratories. In this regard, the aforementioned neuroleptospirosis case highlights many key points with respect to the utility of metagenomics in a diagnostic setting. First, Wilson et al. (2014) reported that a diagnosis would not have been possible via conventional assays. As is the case with many infections, differential diagnoses can be extensive; consequently, cultures, serology, and pathogen-specific PCR can be difficult in determining etiological causation in cases of this magnitude. These difficulties would particularly benefit from the use of unbiased diagnostic methods. Second, reference databases (e.g., NCBI, UniProt for protein and SILVA for ribosomal RNA) are a cornerstone of NGS-based diagnostics, however, the level of curation within these databases is variable. In the absence of adequately curated databases, accurate diagnoses will not be possible and could be detrimental to patient outcome should a misdiagnosis occur. In the neuroleptospirosis case, shotgun metagenomics and bioinformatics analyses utilizing the NCBI reference database identified *L. borgpetersenii*; confirmatory testing revealed the causative agent to be *L. santarosai*. In this context, query sequence homology to a reference sequence should not be misconstrued as causation. Third, the exceptionally short TAT led to appropriate therapeutic management and thus clinically actionable results, which corresponds to our next discussion point – perhaps most importantly is the developmental need for analysis tools and bioinformatics pipelines modernized for the requirements of diagnostic laboratories. For example, tools for pathogen detection must provide analytical results while concurrently minimizing analysis time to allow for reasonable diagnostic TAT. Though several analyses software or platforms for microbial identification are available (Segata et al., 2012; Naccache et al., 2014; Wood and Salzberg, 2014; Flygare et al., 2016), appropriate use typically requires bioinformatics or computational biology expertise often not available in diagnostic laboratories. One such computational analysis pipeline recently developed, SURPI (Naccache et al., 2014), is designed for unbiased pathogen detection from shotgun metagenomics sequencing data of clinical specimens applicable to infectious disease diagnostics, public health surveillance and outbreak analysis. Currently, practical utilization of this technology is hindered by several computational and technical challenges of analyzing data accurately and in a clinically actionable timeframe. Specific requirements include access to high performance computing, as well as laboratory technicians that are cross-trained in bioinformatics (ability to generate data) and in biology (provide interpretation of the data). There is an important need to develop clinically useful pipelines such that metagenomics can be implemented as a diagnostic tool.

Not only are challenges associated with the analytical performance, but sample preparation and sequencing methodologies must also be standardized. Due to the infancy of this field, standardizations are still in progress and various

ring trials and proficiency tests are underway to standardize various aspects pertaining to studies of this nature. For further information we direct readers to the following articles (Lesho et al., 2016; Mellmann et al., 2017). As outlined in **Tables 1, 2**, the choice of sequencing technique applied is largely dependent on the research or diagnostic goals while also considering advantages or disadvantages of each method. Other factors that can affect pathogen identification include as previously discussed, pre-enrichment steps, quality of DNA extraction, library preparation and chemistry. Furthermore, we are faced with questions pertaining to how to work with contaminants or host DNA, low biomass specimens, defining ideal read depths for particular biological specimens or food products and lastly, defining sequence number thresholds to confidently assign pathogenic etiology.

The presence of contaminants or host DNA poses serious challenges to metagenomics data analysis whereby in many instances, true (e.g., microbial) metagenomics data is overwhelmed by the abundance of host DNA. Considerations should be given to samples where eukaryotic cell content is expected to be high (e.g., biopsy). Similarly, variable host DNA content will be found in samples such as stool dependent on the consistency and quality (e.g., healthy, watery, or bloody). Studies have attempted to address the issue of contaminants (Salter et al., 2014) and host DNA (Hasan et al., 2016) with variable success. The primary concern, however, are methodologies surrounding removal of these products and their plausible effect on microbial DNA abundance or quality in a given sample.

While we frequently refer to metagenomics as an unbiased methodology, there are biases that need to be considered to ensure that all microbes of interest are captured adequately in the sequencing output. As an example, interrogating the DNA or RNA virome using a metagenomics approach (e.g., viromics) requires complex wet-laboratory procedures that are considerably distinct from that of a sample workup for a general metagenomics assay of 'all' microbes. Viruses have significantly smaller genomes as compared to prokaryotes or eukaryotes, thus their genomic content represents a microscopic fraction of total DNA or RNA in a given sample (Kleiner et al., 2015). Moreover, characteristics of viral particles are considerably diverse (e.g., size, density, structure, and others) and also include phages. In this context, an intricate series of sample and nucleic acid treatments must be conducted to ensure viral particles are amply interrogated with high sensitivity in a metagenomics assay (Thurber et al., 2009).

Within a given specimen, a broad range of taxon abundances are observed. Though shotgun metagenomics can produce near-complete whole-genome coverage of highly abundant microbes, etiological agents are not always among the most abundant. For example, while the infective dose (in colony forming units) is  $10^7$ – $10^9$  for enterotoxigenic *E. coli*, *L. monocytogenes* is  $10^3$  and *Salmonella* spp. is only 4–50 (Todd et al., 2008). Achieving sufficient sensitivity in shotgun metagenomics studies has proven to be a challenge, particularly with respect to low abundance microbes and gene level identification. For example, data mining for AMR, toxin or virulence genes has proven difficult. This was



well-established in a retrospective metagenomics study that aimed to investigate an outbreak of STEC O104:H4 (Loman et al., 2013). The authors' reported that the outbreak strain was recovered to near completeness (full genome breadth) from 10 samples (of 45) at >10-fold coverage and from 26 samples at >1-fold coverage. The shiga-toxin gene, however, which was identified in 100% of samples via culture-based methods was only detected in 27 of 40 (67%) STEC-positive samples. This study highlights some of the challenges of metagenomics related to gene level sensitivity. Further examination, particularly into the acceptable sequencing depth to ensure high sensitivity and specificity is warranted.

Metagenomics has already proven an effective complement to conventional diagnostics in complex cases and outbreaks (Loman et al., 2013; Ruppé et al., 2016; Langelier et al., 2017) though much remains to be elucidated prior to the widespread adoption of metagenomics in diagnostic and public health laboratories. Therefore, while high-throughput sequencing of biological specimens has a promising future, its utility is unlikely to become a standard and approved CIDT method in the imminent future.

## INFLUENCE OF CULTURE-INDEPENDENT DIAGNOSTIC TESTING ON HUMAN DISEASE SURVEILLANCE

### Pathogen Surveillance and Subtyping

Metagenomics sequencing of a clinical or environmental sample could offer a universal test for pathogen detection, clinical diagnosis, as well as a subtyping test for routine surveillance activities. Surveillance for foodborne infections historically has required the culturing of an isolate in order to perform the appropriate genotypic or phenotypic characterization, yielding a phenotype or genetic fingerprint capable of making detecting and resolving outbreaks possible. By tracking foodborne pathogens along the farm to fork to clinical specimen continuum, it is possible to monitor trends over time, track which foods are capable or implicated in causing illnesses and detecting outbreaks.

Similar to trends occurring in infectious disease diagnostics, techniques for subtyping have enhanced the ability to distinguish between epidemiologically linked isolates from identical microbial species hence improving outbreak detection, surveillance and the overall understanding of microbial epidemiology. Molecular subtyping techniques in routine use include PFGE and MLVA; PulseNet Canada<sup>5</sup> (national real-time surveillance and outbreak response) has relied on these for nearly two decades but is presently transitioning to the use of WGS. The other national surveillance systems dedicated to foodborne disease in Canada, the National Enteric Surveillance Program<sup>6</sup> (weekly trend analysis at the species or serotype

level of bacterial, viral and parasitic enteric pathogens) and FoodNet Canada<sup>7</sup> (sentinel site-based surveillance system measuring burden of illness, attribution studies and food safety policy recommendations) are also transitioning from molecular subtyping to WGS.

## Genomic Epidemiology

Recent improvements of sequencing technologies and streamlined bioinformatics tools have not only advanced clinical diagnostics but are also transforming public health. WGS is increasingly implemented in epidemiological study, outbreak detection and surveillance of foodborne bacteria. Genomic epidemiology refers to the use of WGS to investigate epidemiological features. The listeriosis outbreak described above was the first application of WGS in during an active foodborne outbreak investigation (Gilmour et al., 2010). This study was seminal in bridging the gap between WGS and public health in real-time. Further, the initial study to utilize WGS for source attribution was performed during the 2009–2010 *S. enterica* serotype Montevideo outbreak (Lienau et al., 2011). The source was traced to red and black pepper that was used in the production of Italian-style spiced meats in a New England processing facility. WGS has since been executed in manifold outbreak and surveillance analyses.

Between 2010 and 2015, numerous severe illnesses associated with a complex multi-state listeriosis outbreak were reported and linked to two facilities of a large commercial ice cream producer as the source of *L. monocytogenes* (Jackson et al., 2016). This outbreak is highlighted here particularly due to the unusual length of the outbreak. Specifically, guidelines pertaining to listeriosis outbreak investigations have generally used a 120-day window (versus 16-days for other foodborne pathogens) for inclusion of suspected cases attributed to its psychrotrophic nature and viability in cured and processed food products with longer shelf lives. WGS routine implementation for all clinical, food and food processing environmental isolates, for current and retrospective cases led to strong evidence supporting a lengthy listeriosis outbreak. Traditional investigation guidelines and culture-based subtyping methods (e.g., PFGE) would not have been able to unequivocally link *Listeria* isolates to an outbreak cluster. The high discriminatory power of WGS combined with strong epidemiological evidence will inevitably lead to a higher proportion of detected and resolved outbreaks and a concomitant lower number of patients within each cluster, thus allowing for contaminated products to be removed from commerce more promptly.

## Metagenomics is Capable of Providing Informative Subtyping Data

Advancements have clearly been made with respect to pathogen detection via the application of molecular diagnostic techniques such as PCR. These advancements are driven by the capacity to bypass the need for pathogen culture and isolation. Current surveillance methodologies (e.g., PFGE, MLVA, MLST, and WGS) however, are reliant upon the presence of isolates.

<sup>5</sup><https://www.nml-lnm.gc.ca/Pulsenet/index-eng.htm>

<sup>6</sup><https://www.nml-lnm.gc.ca/NESP-PNSME/index-eng.htm>

<sup>7</sup><http://www.phac-aspc.gc.ca/foodnetcanada/index-eng.php>

Techniques used in public health surveillance for disease tracking and subtyping therefore necessitate adaptation to the culture-independent trend.

Rapid infectious disease diagnostics would particularly benefit from the ability to directly subtype pathogens from complex clinical specimens. While subtyping methods such as targeted-amplicon sequencing, FISH (Splettstoesser et al., 2010) and repetitive element sequence-based PCR (Hahm et al., 2003) may be able to interrogate and subtype microbes, their ability to do so at an adequate taxonomic resolution renders these assays less efficient in differentiating between subtypes of a specific species, compared to gold standard methods such as PFGE. Metagenomics sequencing of clinical specimens represents a plausible epidemiological and subtyping tool. Hundreds to thousands of sequence reads for a particular species are generated through metagenomics sequencing thus potentially providing sufficient informative data for subtyping. Genome coverage of a given microbe, however, is difficult to predict and often the pathogen may not be the most abundant microbe in a given specimen due to the pathogen's infective dose. Also, some serogroups are unable to be typed, for example "O rough" STEC (Chattaway et al., 2016). In recent years, the widespread usage of research-based metagenomics has coincided with the development and application of a plethora of novel analysis techniques, some of which are well-suited to type (e.g., interrogate the microbe below species level taxonomic resolution) microbial strains (Zagordi et al., 2011; Hong et al., 2014; Ahn et al., 2015; Cleary et al., 2015; Sahl et al., 2015; Joseph et al., 2016). Numerous analysis software have been developed and similarly been shown to achieve adequate sensitivity and specificity for pathogen identification. As of yet, metagenomics to our knowledge has not been utilized in a communicable disease tracking perspective.

## Changing Trends in Public Health Surveillance

Effective detection of outbreaks, particularly broadly disseminated outbreaks caused by the commercial distribution of contaminated foods is largely dependent on subtyping isolates from a sizeable proportion of cases. As diagnostic laboratories are in the process of shifting from diagnostic tests (culture-based or WGS) yielding isolates to CIDTs (molecular or serological), it is essential to maintain a system whereby subtyping can still be performed on a large percentage of positive specimens.

With changing diagnostic practices, several options to maintain the capability of foodborne surveillance, outbreak detection and source attribution are possible. First, clinical (and food) microbiology laboratories could perform reflex culturing of biological specimens (or food) that test positive via CIDTs such that positive isolates can still be submitted to public health laboratories for subtyping or other culture-based tests such as WGS (Cronquist et al., 2012). Second, clinical laboratories could alternatively submit biological specimens to public health laboratories that have tested positive using CIDTs. Lastly, culture-independent subtyping and streamlined bioinformatics analyses could be developed for both public health and clinical

microbiology laboratories; this would be particularly beneficial for specimens incompatible or not optimal with culture (e.g., fecal swabs; Kotton et al., 2006). Hence, while any of these methods would require substantial restructuring of national and international surveillance infrastructure it would in theory overcome the prospective predicament whereby public health laboratories would not have access to a considerable portion of isolates and thus capacity for precise outbreak detection and source attribution. If, however, culture were rendered obsolete prior to the implementation of any of the above-listed scenarios, it would inevitably be detrimental for public health surveillance. Additionally, the absence of isolates (and a biobank of historical isolates) will make retrospective studies of outbreaks challenging. In this regard, regulatory groups have been created in both Canada and the US to address the concern pertaining to a lack of enteric isolates available for further characterizations at public health laboratories. The US has created an interim recommendations document<sup>8</sup> to ensure that isolation is attempted or that positive CIDT specimens are retained – Canada is in the process of developing analogous guidelines.

## TRANSFORMATION OF ANTIMICROBIAL SUSCEPTIBILITY TESTING AND INFLUENCE ON ANTIMICROBIAL RESISTANCE SURVEILLANCE SYSTEMS

A limited number of conventional growth-based methods for AST have persisted in routine usage throughout the transformation of diagnostic microbiology. Included among these are disk diffusion (e.g., Kirby-Bauer) strategies and broth microdilutions. The latter of which has achieved gold standard status; thus, novel AST methods are compared to the efficacy of broth microdilutions from development through clinical trial. At present, AST is either accomplished via the above-listed conventional manual methods or growth-dependent automated AST systems including the Vitek System (bioMérieux, France), Avantage Test System (Abbott Laboratories, Irving, TX, United States), Phoenix (BD Biosciences, Cockeysville, MD, United States) and others, all of which are based on broth microdilution testing (Van Belkum and Dunne, 2013). More recent growth-based AST techniques have also been developed that generally employ innovative methods; these include MALDI-TOF mass spectrometry, microfluidics (NanoDrop BMD), isothermal microcalorimetry, real-time microscopy, and others.

## Whole Genome Sequencing Antimicrobial Resistance Gene Detection

With the rising usage of molecular CIDTs and NGS strategies in clinical diagnostics, speculation exists regarding the potential feasibility of NGS methods or other advanced technologies, as aforementioned, replacing growth-based AST. Assessments have

<sup>8</sup>[https://www.aphl.org/aboutAPHL/publications/Documents/FS-Enteric\\_Pathogens\\_Guidelines\\_0216.pdf](https://www.aphl.org/aboutAPHL/publications/Documents/FS-Enteric_Pathogens_Guidelines_0216.pdf)

been completed for some species [e.g., *Salmonella* (McDermott et al., 2016) and *Gonorrhoea* (Demczuk et al., 2015)] and it has been suggested that resistance genes can be accurately detected and further, the presence of genes and mutations associated with AMR has been shown to have a high correlation with phenotypic antimicrobial susceptibility profiles. As an example, WGS was recently applied to non-typhoidal *Salmonella* isolates (McDermott et al., 2016). The authors reported an overall 99% concordance rate between genotype and phenotype; for most classes of antibiotics, concordance was closer to 100% while lower for aminoglycosides and beta-lactams. Though inferring AST via WGS is becoming more common in diagnostic laboratories, to date, few studies have performed large-scale sequencing projects to investigate the utility of WGS to complement or replace conventional AST in routine laboratory workflows. Nonetheless, limitations concerning particular phenotypic-only traits that can't be determined exclusively from WGS or other NGS methodologies will need to be addressed.

## Metagenomics Antimicrobial Resistance Gene Detection

As discussed previously, conventional approaches used to determine AMR are generally reliant upon growth-based tests and further, are heavily targeted to testing human pathogens. The methodological spectrum for AMR gene detection must be expanded to overcome imperative limitations: relatively few bacterial microbes can be cultured (Eckburg et al., 2005) and commensals are thought to comprise a resistance gene pool (resistome) that can be transferred to pathogens (Davies and Davies, 2010). Hence, metagenomics has the capacity to overcome challenges associated with traditional AST. At present, both shotgun metagenomics sequencing and functional metagenomics have been utilized to interrogate the resistome.

The use of shotgun metagenomics sequencing to identify AMR genes is promising and studies are now beginning to elucidate their plausible role in surveillance. In this technique, metagenomics reads are mapped against a database containing a comprehensive catalog of known AMR genes; notable examples include CARD (Jia et al., 2017), ARDB (Liu and Pop, 2009), Resfams (Gibson et al., 2014), and MEGARes (Lakin et al., 2017). Alternatively, metagenomics reads can be assembled into contigs and next compared to a functional annotation database. A recent study surveyed the metagenomes of several ecological niches including the human gastrointestinal tract, water, animals and others for AMR genes (Fitzpatrick et al., 2016). A large abundance of AMR genes were detected in the human gastrointestinal tract and variably identified in other metagenomes. Nonetheless, the authors emphasized the importance of difficulties associated with shotgun metagenomics detection of AMR genes. First, similar to determining microbial abundance reference ranges (from a diagnostics perspective), limits of detection need to be established in order to have sufficient coverage with the capacity to detect rare AMR genes or in complex metagenomes. Second, methods to normalize data to overcome variable microbial diversity and genome sizes will also be required. Third, the biological features inherent of AMR genes

(e.g., often carried within the mobilome or other transmissible genetic elements) render sequencing and data analysis difficult due to their repetitive nature.

Numerous studies have applied shotgun metagenomics sequencing to detect AMR genes. The *C. difficile* study discussed above identified a total of 27 AMR genes and 55.6% of samples contained a minimum of 1 gene (Zhou et al., 2016). The most dominant AMR genes encoded cephalosporin (*Bl2e\_cfxa*; 25.9%) and tetracycline (*tetQ*; 25.9%) resistance whereas macrolide (*ermA*, *ermB*, *ermF*, *ermG*) resistance was variable ranging from 3.7 to 11.1% of samples. It was also shown that WGS of isolates predicted AMR phenotypes with high accuracy. A recent study explored the utility of shotgun metagenomics to detect MDR pediatric bacterial infections, specifically, methicillin-resistant *Staphylococcus aureus*, vancomycin-resistant *Enterococcus* and MDR Enterobacteriaceae (Andersen et al., 2016). The study included three cohorts – high-risk inpatients, low-risk outpatients and controls. Though the potential for MDR bacteria was increased in inpatients and outpatients compared to controls, no differences were detected between inpatients and outpatients. A noteworthy observation was that 53% of inpatients were colonized with an MDR bacterium that culture failed to identify.

Functional metagenomics on the other hand is an exceedingly powerful technique attributed to its capacity to discover novel and highly divergent AMR genes (Perry and Wright, 2014). This method involves cloning total community genomic DNA into an expression vector and transformation into a susceptible expression host (e.g., *E. coli*). The transformant library is then assayed for AMR by culturing on selective media – persistent AMR genes are sequenced and annotated. Therefore, this method is advantageous, ascribed to the possibility to determine genotypic and phenotypic traits.

As NGS-based methodologies like WGS are proving useful in clinical and public health laboratories, we expect that other advanced sequencing techniques will similarly be effective in amalgamating molecular AST and phenotypic susceptibilities thereby allowing for a complete bypass of microbial culture. In this regard, NGS supplemented with metatranscriptomics or proteomics can be more informative as it allows for description of both gene presence and expression (Cohen et al., 2015; Perez-Llarena and Bou, 2016). A combinatorial 'omics' approach incorporating any of the above-mentioned assays may be initially computationally laborious due to a paucity of streamlined analysis techniques and standardization. However, upon analytical validation, such an approach may still be less laborious than conventional culture-dependent testing.

## Current Condition of Antimicrobial Resistance Surveillance

Public health laboratories routinely track specific characteristics of bacterial pathogens that are implicated in infection, thus effectively allowing for increased understanding of bacterial pathogens and their epidemiology. Among these characteristics include monitoring AMR of common enteric pathogens such as *Salmonella* spp. and *Campylobacter* spp., as well as virulence profiles in for example STEC. At the Canadian public health



level, AMR surveillance is conducted through the Canadian Integrated Program for Antimicrobial Resistance Surveillance (CIPARS)<sup>9</sup> in combination with FoodNet. Importantly, though AMR in a clinical setting is largely associated with the utilization of antimicrobials for the treatment of infections, their use in agri-food production is also known to contribute to the resistant microbe pool. Resistant bacteria are associated with more severe disease (van Duin and Paterson, 2016) and poor patient outcome, hence monitoring for AMR is critical. Accordingly, AMR or MDR bacteria have caused several recent foodborne outbreaks (Cartwright et al., 2016; Gieraltowski et al., 2016; Kawakami et al., 2016).

Irrespective of how and when NGS technologies will be capable of rapidly and accurately detecting AMR and associated susceptibilities, difficulties in public health AMR surveillance will be apparent. With a growing number of clinical and food microbiology laboratories opting for utilization of CIDTs for diagnostic purposes, laboratories are bypassing the need to culture microbes; this may affect public health surveillance of AMR trends which continue to use isolates as the foundation for AMR surveillance programs and assessment of phenotypic resistance to antimicrobials. CIDT for enteric pathogens that yield no isolate and potentially no specimen for public health laboratories does pose an important challenge, however, actions can be taken to maintain adequate AMR surveillance. Gonorrhea, for example, represents a proof-of-principle for overcoming a lack of isolates in the context of surveillance. In particular, routine testing of gonorrhea has been performed by NAATs for some time due to the higher sensitivity and specificity. Guidelines have been documented regarding sentinel surveillance mechanisms to continue monitoring AMR trends<sup>10</sup>. Moreover, CIDT techniques could also be developed that have the capacity to test for known AMR though culture will still be systematically needed to detect novel and emerging resistance mechanisms.

## FUTURE CHALLENGES OF METAGENOMICS IN DIAGNOSTIC AND PUBLIC HEALTH SETTINGS

Whole genome sequencing as a complement to conventional culture-based, molecular or serological CIDTs in real-time serves as an empirical model for the use of future techniques (Jackson et al., 2016). From lessons learnt via WGS application, metagenomics and other NGS techniques should move into front-line laboratories in the near future. Initial proof-of-concept studies (Gilmour et al., 2010), retrospective analyses (Loman et al., 2013), validation processes and real-time implementation of WGS (Jackson et al., 2016) in public health surveillance and outbreak response will undoubtedly help enhance acceptance of the aforementioned metagenomics methodologies from the medical and public health community at large. Standardization is an essential element for implementation of any method set to become the next “gold standard.” Various WGS studies have

been undertaken of historical clusters and sporadic cases to determine concordance of WGS data with traditional subtyping and epidemiological data; in many instances, WGS provided higher discriminatory power than traditional subtyping methods and thus resulted in many occurrences whereby isolates were either included or excluded from an outbreak due to WGS and supportive epidemiological data, if present. To date, isolate cluster inclusion/exclusion criteria remains a moving target due to sizeable differences in biology amongst foodborne pathogens. Although WGS is part of the public health toolbox, it remains to be standardized in such a way that it can be applied to all foodborne pathogens. As a result, regular cluster audits to assess analysis pipelines for performance in addition to adequacy of quality control and assurance metrics for sequencing and all analysis outputs should be routinely conducted. Furthermore, the transition to add WGS as a public health epidemiology tool has been an internationally driven effort with organizations and partners from every field – primary health care, academia, industry and all government tiers. For wide adoption of new methods, involvement of the international community is key. Each of these aforementioned guidelines will be paramount to ensure cluster definitions remain robust in the context of genomic epidemiology. Another crucial aspect of implementing new technologies is knowledge translation. Currently, large efforts are being directed at educating and communicating the use of WGS in public health surveillance and outbreak response. Not only will laboratory technicians be required to advance their skillsets, but also policymakers and media will require training on how to interpret highly technical data and communicate it to the public (Jackson et al., 2016). As we enter this new era of “omics” in public health, the aforementioned criteria will be instrumental in guiding and implementing unbiased NGS-based CIDT methods in foodborne disease surveillance and outbreak response.

As mentioned previously, preceding widespread implementation of metagenomics sequencing and analysis in diagnostic and public health settings, bioinformatics encompassing algorithms, software and pipelines will need to be developed, validated and standardized through various ring trials and retrospective analyses. Perhaps the largest obstacle faced by the field of metagenomics is the lack of universal analyses or pipeline recognized as the status quo. In this regard, we previously discussed the significance of curated sequence databases to appropriate patient diagnosis. The GMI<sup>11</sup> consortium has begun to address this issue via the creation of a global database that house uniquely identifiable microbial genomic data in combination with high quality meta- and epidemiological data. This drive will assist in the global amalgamation of microbial WGS and metadata in an easily accessible portal ultimately enhancing global surveillance of infectious microbes and emergent pathogens. Moreover, such a global database will similarly be influential in data mining investigations as they relate to gene level comparisons such as AMR, virulence, and environmental fitness.

Procurement of regulatory approval for routine implementation of NGS-based technologies in diagnostic, public

<sup>9</sup><http://www.phac-aspc.gc.ca/cipars-picra/index-eng.php>

<sup>10</sup><http://www.phac-aspc.gc.ca/std-mts/sti-its/egsti-ldcits/>

<sup>11</sup><https://www.globalmicrobialidentifier.org>



health and food safety laboratories will be onerous owing to various difficulties inherent of such an assay. First, metagenomics dataset contain identifiable information. As discussed previously, many biological specimen types have high amounts of host DNA; thus, a large proportion of generated sequences will be of eukaryotic origin. In research- and diagnostic-based cases, host DNA is filtered from the dataset, however, issues are apparent with how best to protect patient privacy. There may be plausibility that genetically informative host sequence data could be used to screen against a panel of known disease-causing genetic variants. Providing patients with information pertaining to a potential genetic disease via such an assay is an ethical concern (Chrystoja and Diamandis, 2014). On the other hand, providing patients with their metagenomes may lead to precarious implications should patients seek to self-diagnose. Second, access to patent DNA from the human genome, food products and ingredients sequence data could pose unwarranted legal implications. The recovery of patented DNA sequences from food sources or ingredients such as genetically modified foods (crops or animals) may give rise to concerns of patent infringement (Chrystoja and Diamandis, 2014). Metagenomics may also be used for the screening detection of unauthorized GMO use and similarly aid regulation of the food industry, particularly with high import/export rate of global food products and ingredients. Additionally, metagenomics may allow for the detection of species fraud, food product mislabelling or incorrect claims (Corrado, 2016) that may be subject to increased scrutiny. Third, the detection of AMR or virulence genes may lead to undesirable outcomes. Specifically, rising concerns with AMR has led to the complete ban of antimicrobial sub-therapeutic use in food animals in the European Union. The potential to identify AMR in food products may inhibit producers and commercial food processors from agreeing to regulations or policies related to metagenomics testing of food products in fear of repercussions in food trade and export. Overall, should metagenomics become a validated assay in clinical, food and public health settings its impact may be far-reaching in the realm of ethical and legal implications.

## REFERENCES

- Ahn, T. H., Chai, J., and Pan, C. (2015). Sigma: strain-level inference of genomes from metagenomic analysis for biosurveillance. *Bioinformatics* 31, 170–177. doi: 10.1093/bioinformatics/btu641
- Alfa, M. J., Strang, D., Tappia, P. S., Graham, M., Van Domselaar, G., Forbes, J. D., et al. (2017). A randomized trial to determine the impact of a digestion resistant starch composition on the gut microbiome in older and mid-age adults. *Clin. Nutr.* doi: 10.1016/j.clnu.2017.03.025 [Epub ahead of print].
- Allard, M. W., Strain, E., Melka, D., Bunning, K., Musser, S. M., Brown, E. W., et al. (2016). Practical value of food pathogen traceability through building a whole-genome sequencing network and database. *J. Clin. Microbiol.* 54, 1975–1983. doi: 10.1128/JCM.00081-16
- Almonacid, D. E., Kraal, L., Ossandon, F. J., Budovskaya, Y. V., Cardenas, J. P., Richman, J., et al. (2016). 16S rRNA gene sequencing as a clinical diagnostic aid for gastrointestinal-related conditions. *bioRxiv*. doi: 10.1101/084657
- Andersen, H., Connolly, N., Bangar, H., Staat, M., Mortensen, J., Deburger, B., et al. (2016). Use of shotgun metagenome sequencing to detect fecal colonization with multidrug-resistant bacteria in children. *J. Clin. Microbiol.* 54, 1804–1813. doi: 10.1128/JCM.02638-15

## CONCLUSION

The prospective use of diagnostic metagenomics and comparable techniques offers an assumption-free workflow (though biases are present with each methodology; **Table 1**) thus creating the ability to detect any and all pathogens (bacteria, virus, parasite, and others) from various biological specimens or food products. Through forthcoming improvements related to required technical expertise, throughput and cost-effectiveness of sequencing combined with enhanced and streamlined laboratory and bioinformatics pipelines, metagenomics will likely have a predominant role in the diagnostic and public health laboratory. We expect an automated metagenomics pipeline will complement and may even replace several methods currently employed in the diagnostic laboratory while concurrently providing additional information such as AMR, virulence, and genomic epidemiology. It is apparent that the functionality of diagnostic metagenomics has been established in research settings and further, in detecting etiological culprits of unidentified illnesses and outbreaks. We anticipate that within the next decade, detection and characterization of pathogens via metagenomics-based workflows will be implemented in routine usage in diagnostic and public health laboratories.

## AUTHOR CONTRIBUTIONS

NK and AR developed the concept for the manuscript. JF, NK, JR, FP, and AR wrote the manuscript. All authors read and approved the final version of the manuscript.

## ACKNOWLEDGMENTS

The authors would like to thank Dr. Celine Nadon for her expert review of the manuscript. A figure in this paper used icons made by Freepik, from [www.flaticon.com](http://www.flaticon.com).

- Anthony, R. M., Brown, T. J., and French, G. L. (2000). Rapid diagnosis of bacteremia by universal amplification of 23S ribosomal DNA followed by hybridization to an oligonucleotide array. *J. Clin. Microbiol.* 38, 781–788.
- Anthony, S. J., Epstein, J. H., Murray, K. A., Navarrete-Macias, I., Zambrana-Torrel, C. M., Solovyov, A., et al. (2013). A strategy to estimate unknown viral diversity in mammals. *mBio* 4:e00598–13. doi: 10.1128/mBio.00598-13
- Argenio, V. D., Casaburi, G., Colicchio, R., Sarnataro, D., Discepolo, V., Kim, S. M., et al. (2016). Mucosal gut microbiome is associated with celiac disease-specific microbiome alteration in adult patients. *Am. J. Gastroenterol.* 111, 1659–1661. doi: 10.1038/ajg.2016.227
- Ashton, P. M., Nair, S., Peters, T. M., Bale, J. A., Powell, D. G., Painsent, A., et al. (2016). Identification of *Salmonella* for public health surveillance using whole genome sequencing. *PeerJ* 4:e1752. doi: 10.7717/peerj.1752
- Aw, T. G., Wengert, S., and Rose, J. B. (2016). Metagenomic analysis of viruses associated with field-grown and retail lettuce identifies human and animal viruses. *Int. J. Food Microbiol.* 223, 50–56. doi: 10.1016/j.ijfoodmicro.2016.02.008
- Barratt, J. L. N., Harkness, J., Marriott, D., Ellis, J. T., and Stark, D. (2011). A review of *Dientamoeba fragilis* carriage in humans: several reasons why this organism

- should be considered in the diagnosis of gastrointestinal illness. *Gut Microbes* 2, 3–12. doi: 10.4161/gmic.2.1.14755
- Bavelaar, H. H., Rahamat-Langendoen, J., Niesters, H. G., Zoll, J., and Melchers, W. J. (2015). Whole genome sequencing of fecal samples as a tool for the diagnosis and genetic characterization of norovirus. *J. Clin. Virol.* 72, 122–125. doi: 10.1016/j.jcv.2015.10.003
- Bergholz, T. M., Moreno Switt, A. I., and Wiedmann, M. (2014). Omics approaches in food safety: fulfilling the promise? *Trends Microbiol.* 22, 275–281. doi: 10.1016/j.tim.2014.01.006
- Brown, J. R., Roy, S., Ruis, C., Yara Romero, E., Shah, D., and Williams, R. (2016). Norovirus whole-genome sequencing by SureSelect target enrichment: a robust and sensitive method. *J. Clin. Microbiol.* 54, 2530–2537. doi: 10.1128/JCM.01052-16
- Buchan, B. W., Olson, W. J., Pezewski, M., Marcon, M. J., Novicki, T., Uphoff, T. S., et al. (2013). Clinical evaluation of a real-time PCR assay for identification of *Salmonella*, *Shigella*, *Campylobacter* (*Campylobacter jejuni* and *C. coli*), and shiga toxin-producing *Escherichia coli* isolates in stool specimens. *J. Clin. Microbiol.* 51, 4001–4007. doi: 10.1128/JCM.02056-13
- Buss, S. N., Alter, R., Iwen, P. C., and Fey, P. D. (2013). Implications of culture-independent panel-based detection of *Cyclospora cayetanensis*. *J. Clin. Microbiol.* 51, 3909. doi: 10.1128/JCM.02238-13
- Buss, S. N., Leber, A., Chapin, K., Fey, P. D., Bankowski, M. J., Jones, M. K., et al. (2015). Multicenter evaluation of the BioFire FilmArray gastrointestinal panel for etiologic diagnosis of infectious gastroenteritis. *J. Clin. Microbiol.* 53, 915–925. doi: 10.1128/JCM.02674-14
- Carleton, H. A., and Gerner-Smidt, P. (2016). Whole-genome sequencing is taking over foodborne disease surveillance. *Microbe* 11, 311–317.
- Cartwright, E. J., Nguyen, T., Melluso, C., Ayers, T., Lane, C., Hodges, A., et al. (2016). A multistate investigation of antibiotic-resistant *Salmonella enterica* serotype I 4,[5], 12:i- infections as part of an international outbreak associated with frozen feeder rodents. *Zoonoses Public Health* 63, 62–71. doi: 10.1111/zph.12205
- Case, R. J., Boucher, Y., Dahllöf, L., Holmström, C., Doolittle, W. F., and Kjelleberg, S. (2007). Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Appl. Environ. Microbiol.* 73, 278–288. doi: 10.1128/AEM.01177-06
- Chattaway, M. A., Dallman, T. J., Gentle, A., Wright, M. J., Long, S. E., Ashton, P. M., et al. (2016). Whole genome sequencing for public health surveillance of shiga toxin-producing *Escherichia coli* other than serogroup O157. *Front. Microbiol.* 7:258. doi: 10.3389/fmicb.2016.00258
- Chrystoja, C. C., and Diamandis, E. P. (2014). Whole genome sequencing as a diagnostic test: challenges and opportunities. *Clin. Chem.* 60, 724–733. doi: 10.1373/clinchem.2013.209213
- Cinar, H. N., Gopinath, G., Jarvis, K., and Murphy, H. R. (2015). The complete mitochondrial genome of the foodborne parasitic pathogen *Cyclospora cayetanensis*. *PLoS ONE* 10:e0128645. doi: 10.1371/journal.pone.0128645
- Cleary, B., Brito, I. L., Huang, K., Gevers, D., Shea, T., Young, S., et al. (2015). Detection of low-abundance bacterial strains in metagenomic datasets by eigengene partitioning. *Nat. Biotechnol.* 33, 1053–1060. doi: 10.1038/nbt.3329
- Cohen, A., Bont, L., Engelhard, D., Moore, E., Fernández, D., Kreisberg-Greenblatt, R., et al. (2015). A multifaceted “omics” approach for addressing the challenge of antimicrobial resistance. *Fut. Microbiol.* 10, 365–376. doi: 10.2217/fmb.14.127
- Corrado, G. (2016). Advances in DNA typing in the agro-food supply chain. *Trends Food Sci. Technol.* 52, 80–89. doi: 10.1016/j.tifs.2016.04.003
- Cronquist, A. B., Mody, R. K., Atkinson, R., Besser, J., D’Angelo, M. T., Hurd, S., et al. (2012). Impacts of culture-independent diagnostic practices on public health surveillance for bacterial enteric pathogens. *Clin. Infect. Dis.* 54, S432–S439. doi: 10.1093/cid/cis267
- Davies, J., and Davies, D. (2010). Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.* 74, 417–433. doi: 10.1128/mmb.00016-10
- Demczuk, W., Lynch, T., Martin, I., Van Domselaar, G., Graham, M., Bharat, A., et al. (2015). Whole-genome phylogenomic heterogeneity of *Neisseria gonorrhoeae* isolates with decreased cephalosporin susceptibility collected in Canada between 1989 and 2013. *J. Clin. Microbiol.* 53, 191–200. doi: 10.1128/JCM.02589-14
- Denno, D. M., Shaikh, N., Stapp, J. R., Qin, X., Hutter, C. M., Hoffman, V., et al. (2012). Diarrhea etiology in a pediatric emergency department: a case control study. *Clin. Infect. Dis.* 55, 897–904. doi: 10.1093/cid/cis553
- Doggett, N. A., Mukundan, H., Lefkowitz, E. J., Slezak, T. R., Chain, P. S., Morse, S., et al. (2016). Culture-independent diagnostics for health security. *Health Secur.* 14, 122–142. doi: 10.1089/hs.2015.0074
- Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., et al. (2005). Diversity of the human intestinal microbial flora. *Science* 308, 1635–1638. doi: 10.1126/science.1110591
- Escobar-Zepeda, A., Sanchez-Flores, A., and Quirasco Baruch, M. (2016). Metagenomic analysis of a Mexican ripened cheese reveals a unique complex microbiota. *Food Microbiol.* 57, 116–127. doi: 10.1016/j.fm.2016.02.004
- Ettayebi, K., Crawford, S. E., Murakami, K., Broughman, J. R., Karandikar, U., Tenge, V. R., et al. (2016). Replication of human noroviruses in stem cell-derived human enteroids. *Science* 355, 1387–1393. doi: 10.1126/science.aaf5211
- Finkbeiner, S. R., Allred, A. F., Tarr, P. I., Klein, E. J., Kirkwood, C. D., and Wang, D. (2008). Metagenomic analysis of human diarrhea: viral detection and discovery. *PLoS Pathog.* 4:e1000011. doi: 10.1371/journal.ppat.1000011
- Finkbeiner, S. R., Li, Y., Ruone, S., Conrardy, C., Gregoricus, N., Toney, D., et al. (2009). Identification of a novel astrovirus (astrovirus VA1) associated with an outbreak of acute gastroenteritis. *J. Virol.* 83, 10836–10839. doi: 10.1128/JVI.00998-09
- Fitzgerald, C., Patrick, M., Gonzalez, A., Akin, J., Polage, C. R., Wymore, K., et al. (2016). Multicenter evaluation of clinical diagnostic methods for detection and isolation of *Campylobacter* spp. from stool. *J. Clin. Microbiol.* 54, 1209–1215. doi: 10.1128/JCM.01925-15.Editor
- Fitzpatrick, D., Walsh, F., and Walsh, F. (2016). Antibiotic resistance genes across a wide variety of metagenomes. *FEMS Microbiol. Ecol.* 92:fiv168. doi: 10.1093/femsec/fiv168
- Flygare, S., Simmon, K., Miller, C., Qiao, Y., Kennedy, B., Di Sera, T., et al. (2016). Taxonomer: an interactive metagenomics analysis portal for universal pathogen detection and host mRNA expression profiling. *Genome Biol.* 17, 111. doi: 10.1186/s13059-016-0969-1
- Forbes, J. D., Van Domselaar, G., and Bernstein, C. N. (2016). Microbiome survey of the inflamed and noninflamed gut at different compartments within the gastrointestinal tract of inflammatory bowel disease patients. *Inflamm. Bowel Dis.* 22, 817–825. doi: 10.1097/MIB.0000000000000684
- Forster, S. C., Anonye, B. O., Kumar, N., Neville, B. A., Stares, M. D., Goulding, D., et al. (2016). Culturing of “unculturable” human microbiota reveals novel taxa and extensive sporulation. *Nature* 533, 543–546. doi: 10.1038/nature17645
- Gambarin, P., Magnabosco, C., Losio, M. N., Pavoni, E., Gattuso, A., Arcangeli, G., et al. (2012). *Listeria monocytogenes* in ready-to-eat seafood and potential hazards for the consumers. *Int. J. Microbiol.* 2012:497635. doi: 10.1155/2012/497635
- Gao, D., Yu, Q., Wang, G., Wang, G., and Xiong, F. (2016). Diagnosis of a malayan filariasis case using a shotgun diagnostic metagenomics assay. *Parasit. Vectors* 9, 86. doi: 10.1186/s13071-016-1363-2
- Gentry-Shields, J., and Jaykus, L.-A. (2015). Comparison of process control viruses for use in extraction and detection of human norovirus from food matrices. *Food Res. Int.* 77, 320–325. doi: 10.1016/j.foodres.2015.05.027
- Gibson, M. K., Forsberg, K. J., and Dantas, G. (2014). Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J.* 9, 207–216. doi: 10.1038/ismej.2014.106
- Gieraltowski, L., Higa, J., Peralta, V., Green, A., Schwensohn, C., Rosen, H., et al. (2016). National outbreak of multidrug resistant *Salmonella* heidelberg infections linked to a single poultry company. *PLoS ONE* 11:e0162369. doi: 10.1371/journal.pone.0162369
- Gill, A. (2017). The importance of bacterial culture to food microbiology in the age of genomics. *Front. Microbiol.* 8:777. doi: 10.3389/fmicb.2017.00777
- Gilmour, M. W., Graham, M., Van Domselaar, G., Tyler, S., Kent, H., Troutyakov, K. M., et al. (2010). High-throughput genome sequencing of two *Listeria monocytogenes* clinical isolates during a large foodborne outbreak. *BMC Genomics* 11:120. doi: 10.1186/1471-2164-11-120
- Gorski, L. (2012). Selective enrichment media bias the types of *Salmonella enterica* strains isolated from mixed strain cultures and complex enrichment broths. *PLoS ONE* 7:e34722. doi: 10.1371/journal.pone.0034722
- Hahn, B., Maldonado, Y., and Schreiber, E. (2003). Subtyping of foodborne and environmental isolates of *Escherichia coli* by multiplex-PCR, rep-PCR, PFGE,

- ribotyping and AFLP. *J. Microbiol. Methods* 53, 387–399. doi: 10.1016/S0167-7012(02)00259-2
- Hanson, K. E., and Couturier, M. R. (2016). Multiplexed molecular diagnostics for respiratory, gastrointestinal, and central nervous system infections. *Clin. Infect. Dis.* 63, 1361–1367. doi: 10.1093/cid/ciw494
- Harrington, S. M., Buchan, B. W., Doern, C., Fader, R., Ferraro, M. J., Pillai, D. R., et al. (2015). Multicenter evaluation of the BD max enteric bacterial panel PCR assay for rapid detection of *Salmonella* spp., *Shigella* spp., *Campylobacter* spp. (*C. jejuni* and *C. coli*), and shiga toxin 1 and 2 genes. *J. Clin. Microbiol.* 53, 1639–1647. doi: 10.1128/JCM.03480-14
- Hasan, M. R., Rawat, A., Tang, P., Jithesh, P. V., Thomas, E., Tan, R., et al. (2016). Depletion of human DNA in spiked clinical specimens to improve the sensitivity of pathogen detection by next generation sequencing. *J. Clin. Microbiol.* 54, 919–927. doi: 10.1128/JCM.03050-15
- Hasman, H., Saputra, D., Sicheritz-Ponten, T., Lund, O., Svendsen, C. A., Frimodt-Møller, N., et al. (2014). Rapid whole genome sequencing for the detection and characterization of microorganisms directly from clinical samples. *J. Clin. Microbiol.* 52, 139–146. doi: 10.1128/JCM.02452-13
- Hong, C., Manimaran, S., Shen, Y., Perez-Rogers, J. F., Byrd, A. L., Castro-Nallar, E., et al. (2014). PathoScope 2.0: a complete computational framework for strain identification in environmental or clinical sequencing samples. *Microbiome* 2, 33. doi: 10.1186/2049-2618-2-33
- Huang, J. Y., Henao, O. L., Griffin, P. M., Vugia, D. J., Cronquist, A. B., Hurd, S., et al. (2016). Infection with pathogens transmitted commonly through food and the effect of increasing use of culture-independent diagnostic tests on surveillance – foodborne diseases active surveillance network, 10 U.S. sites, 2012–2015. *Morb. Mortal. Wkly. Rep.* 65, 368–371. doi: 10.15585/mmwr.mm6154a2
- Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., et al. (2012). Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214. doi: 10.1038/nature11234
- Jackson, B. R., Tarr, C., Strain, E., Jackson, K. A., Conrad, A., Carleton, H., et al. (2016). Implementation of nationwide real-time whole-genome sequencing to enhance listeriosis outbreak detection and investigation. *Clin. Infect. Dis.* 63, 380–386. doi: 10.1093/cid/ciw242
- Jandaa, J. M., and Abbott, S. A. (2014). Culture-independent diagnostic testing: have we opened Pandora's box for good? *Diagn. Microbiol. Infect. Dis.* 80, 171–176. doi: 10.1016/j.diagmicrobio.2014.08.001
- Jarvis, K. G., White, J. R., Grim, C. J., Ewing, L., Ottesen, A. R., Beaubrun, J. J.-G., et al. (2015). Cilantro microbiome before and after nonselective pre-enrichment for *Salmonella* using 16S rRNA and metagenomic sequencing. *BMC Microbiol.* 15:160. doi: 10.1186/s12866-015-0497-2
- Jia, B., Raphenya, A. R., Alcock, B., Waglegchner, N., Guo, P., Tsang, K. K., et al. (2017). CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 45, D566–D573. doi: 10.1093/nar/gkw1004
- Jones, M. K., Watanabe, M., Zhu, S., Graves, C. L., Keyes, L. R., Grau, K. R., et al. (2014). Enteric bacteria promote human and mouse norovirus infection of B cells. *Science* 346, 755–759. doi: 10.1126/science.1257147
- Joseph, S. J., Li, B., Petit, R. A. III, Qin, Z. S., Darrow, L., and Read, T. D. (2016). The single-species metagenome: subtyping *Staphylococcus aureus* core genome sequences from shotgun metagenomic data. *PeerJ* 4, e2571. doi: 10.7717/peerj.2571
- Jung, J. Y., Lee, S. H., Kim, J. M., Park, M. S., Bae, J., Hahn, Y., et al. (2011). Metagenomic analysis of kimchi, a traditional Korean fermented food. *Appl. Environ. Microbiol.* 77, 2264–2274. doi: 10.1128/AEM.02157-10
- Kataoka, T., Yamaguchi, H., Sato, M., Watanabe, T., Taniuchi, Y., Kuwata, A., et al. (2016). Seasonal and geographical distribution of near-surface small photosynthetic-eukaryotes in the western North Pacific determined by pyrosequencing of 18S rDNA. *FEMS Microbiol. Ecol.* 93, fiw229. doi: 10.1093/femsec/fiw229
- Kawai, T., Sekizuka, T., Yahata, Y., Kuroda, M., Kumeda, Y., Iijima, Y., et al. (2012). Identification of *Kudoa septempunctata* as the causative agent of novel food poisoning outbreaks in Japan by consumption of *Paralichthys olivaceus* in raw fish. *Clin. Infect. Dis.* 54, 1046–1052. doi: 10.1093/cid/cir1040
- Kawakami, V. M., Bottichio, L., Angelo, K., Linton, N., Kissler, B., Basler, C., et al. (2016). Notes from the field: outbreak of multidrug-resistant *Salmonella* infections linked to pork – Washington, 2015. *Morb. Mortal. Wkly. Rep.* 65, 379–381. doi: 10.15585/mmwr.mm651444
- Kleiner, M., Hooper, L. V., and Duerkop, B. A. (2015). Evaluation of methods to purify virus-like particles for metagenomic sequencing of intestinal viromes. *BMC Genomics* 16:7. doi: 10.1186/s12864-014-1207-4
- Korsak, N., Taminiau, B., Leclercq, M., Nezer, C., Crevecoeur, S., Ferauche, C., et al. (2015). Short communication: evaluation of the microbiota of kefir samples using metagenetic analysis targeting the 16S and 26S ribosomal DNA fragments. *J. Dairy Sci.* 98, 3684–3689. doi: 10.3168/jds.2014-9065
- Köser, C. U., Ellington, M. J., Cartwright, E. J. P., Gillespie, S. H., Brown, N. M., Farrington, M., et al. (2012). Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. *PLoS Pathog.* 8:e1002824. doi: 10.1371/journal.ppat.1002824
- Kotton, C. N., Lankowski, A. J., and Hohmann, E. L. (2006). Comparison of rectal swabs with fecal cultures for detection of *Salmonella typhimurium* in adult volunteers. *Diagn. Microbiol. Infect. Dis.* 56, 123–126. doi: 10.1016/j.diagmicrobio.2006.04.003
- Kundu, S., Lockwood, J., Depledge, D. P., Chaudhry, Y., Aston, A., Rao, K., et al. (2013). Next-generation whole genome sequencing identifies the direction of norovirus transmission in linked patients. *Clin. Infect. Dis.* 57, 407–414. doi: 10.1093/cid/cit287
- Lakin, S. M., Dean, C., Noyes, N. R., Dettenwanger, A., Ross, A. S., Doster, E., et al. (2017). MEGARes: an antimicrobial resistance database for high throughput sequencing. *Nucleic Acids Res.* 45, D574–D580. doi: 10.1093/nar/gkw1009
- Lambert, D., Carrillo, C. D., Koziol, A. G., Manninger, P., and Blais, B. W. (2015). GeneSippr: a rapid whole-genome approach for the identification and characterization of foodborne pathogens such as priority Shiga toxinogenic *Escherichia coli*. *PLoS ONE* 10:e0122928. doi: 10.1371/journal.pone.0122928
- Langelier, C., Zinter, M., Kalantar, K., Yanik, G., Christenson, S., Odonovan, B., et al. (2017). Metagenomic next-generation sequencing detects pulmonary pathogens in hematopoietic cellular transplant patients with acute respiratory illnesses. *bioRxiv*. doi: 10.1101/102798
- Law, J. W., Mutalib, N. A., Chan, K. L., Lee, L., and Lee, L. (2015). An insight into the isolation, enumeration, and molecular detection of *Listeria monocytogenes* in food. *Front. Microbiol.* 6:1227. doi: 10.3389/fmicb.2015.01227
- Law, J. W. F., Mutalib, N. S., Chan, K. G., and Lee, L. H. (2014). Rapid methods for the detection of foodborne bacterial pathogens: principles, applications, advantages and limitations. *Front. Microbiol.* 5:770. doi: 10.3389/fmicb.2014.00770
- Leimena, M. M., Ramiro-garcia, J., Davids, M., Bogert, B., Van Den Smidt, H., Smid, E. J., et al. (2013). A comprehensive metatranscriptome analysis pipeline and its validation using human small intestine microbiota datasets. *BMC Genomics* 14:530. doi: 10.1186/1471-2164-14-530
- Leite, A. M. O., Mayo, B., Rachid, C. T. C. C., Peixoto, R. S., Silva, J. T., Paschoalin, V. M. F., et al. (2012). Assessment of the microbial diversity of Brazilian kefir grains by PCR-DGGE and pyrosequencing analysis. *Food Microbiol.* 31, 215–221. doi: 10.1016/j.fm.2012.03.011
- Leonard, S. R., Mammel, M. K., Lacher, D. W., and Elkins, C. A. (2015). Application of metagenomic sequencing to food safety: detection of shiga toxin-producing *Escherichia coli* on fresh bagged spinach. *Appl. Environ. Microbiol.* 81, 8183–8191. doi: 10.1128/AEM.02601-15
- Lesho, E., Clifford, R., Onmus-Leone, F., Appalla, L., Snedrud, E., Kwak, Y., et al. (2016). The challenges of implementing next generation sequencing across a large healthcare system, and the molecular epidemiology and antibiotic susceptibilities of carbapenemase-producing bacteria in the healthcare system of the U.S. Department of Defense. *PLoS ONE* 11:e0155770. doi: 10.1371/journal.pone.0155770
- Lienau, E. K., Strain, E., Wang, C., Zheng, J., Ottesen, A. R., Keys, C. E., et al. (2011). Identification of a salmonellosis outbreak by means of molecular sequencing. *N. Engl. J. Med.* 364, 981–982. doi: 10.1056/NEJMc1100443
- Liu, B., and Pop, M. (2009). ARDB - Antibiotic resistance genes database. *Nucleic Acids Res.* 37, 443–447. doi: 10.1093/nar/gkn656
- Loman, N. J., Constantinidou, C., Christner, M., Rohde, H., Chan, J. Z.-M., Quick, J., et al. (2013). A culture-independent sequence-based metagenomics approach to the investigation of an outbreak of shiga-toxicogenic *Escherichia coli* O104:H4. *J. Am. Med. Assoc.* 309, 1502–1510. doi: 10.1001/jama.2013.3231
- Mardis, E. R. (2017). DNA sequencing technologies: 2006–2016. *Nat. Protoc.* 12, 213–218. doi: 10.1038/nprot.2016.182



- Margot, H., Stephan, R., and Tasara, T. (2016). Mungo bean sprout microbiome and changes associated with culture based enrichment protocols used in detection of Gram-negative foodborne pathogens. *Microbiome* 4, 48. doi: 10.1186/s40168-016-0193-y
- Mason, O. U., Scott, N. M., Gonzalez, A., Robbins-Pianka, A., Bælum, J., Kimbrel, J., et al. (2014). Metagenomics reveals sediment microbial community response to Deepwater Horizon oil spill. *ISME J.* 8, 1464–1475. doi: 10.1038/ismej.2013.254
- McCollum, J. T., Cronquist, A. B., Silk, B. J., Jackson, K. A., O'Connor, K. A., Cosgrove, S., et al. (2013). Multistate outbreak of listeriosis associated with cantaloupe. *N. Engl. J. Med.* 369, 944–953. doi: 10.1056/NEJMoa1215837
- McDermott, P. F., Tyson, G. H., Kabera, C., Chen, Y., Li, C., Folster, J. P., et al. (2016). Whole-genome sequencing for detecting antimicrobial resistance in nontyphoidal *Salmonella*. *Antimicrob. Agents Chemother.* 60, 5515–5520. doi: 10.1128/AAC.01030-16.Address
- Mellmann, A., Andersen, P. S., Bletz, S., Friedrich, A. W., Kohl, T. A., Lilje, B., et al. (2017). High interlaboratory reproducibility and accuracy of next-generation-sequencing-based bacterial genotyping in a ring trial. *J. Clin. Microbiol.* 55, 908–913. doi: 10.1128/JCM.02242-16
- Mongkolrattanothai, K., Naccache, S. N., Bender, J. M., Samayoa, E., Pham, E., Yu, G., et al. (2017). Neurobrucellosis: unexpected answer from metagenomic next-generation sequencing. *J. Pediatric Infect. Dis. Soc.* doi: 10.1093/jpids/piw066 [Epub ahead of print].
- Moore, N. E., Wang, J., Hewitt, J., Croucher, D., Williamson, D. A., Paine, S., et al. (2015). Metagenomic analysis of viruses in feces from unsolved outbreaks of gastroenteritis in humans. *J. Clin. Microbiol.* 53, 15–21. doi: 10.1128/JCM.02029-14
- Morovic, W., Hibberd, A. A., Zabel, B., Rodolphe, B., and Sthal, B. (2016). Genotyping by PCR and high-throughput sequencing of commercial probiotic products reveals composition biases. *Front. Microbiol.* 7:1747. doi: 10.3389/fmicb.2016.01747
- Naccache, S. N., Federman, S., Veeraraghavan, N., Zaharia, M., Lee, D., Samayoa, E., et al. (2014). A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples. *Genome Res.* 24, 1180–1192. doi: 10.1101/gr.171934.113
- Nakamura, S., Maeda, N., Miron, I. M., Yoh, M., Izutsu, K., Kataoka, C., et al. (2008). Metagenomic diagnosis of bacterial infections. *Emerg. Infect. Dis.* 14, 1784–1786. doi: 10.3201/eid1411.080589
- Nasheri, N., Petronella, N., Ronholm, J., Bidawid, S., and Corneau, N. (2017). Characterization of the genomic diversity of norovirus in linked patients using a metagenomic deep sequencing approach. *Front. Microbiol.* 8:73. doi: 10.3389/fmicb.2017.00073
- Navidad, J. F., Griswold, D. J., Gradus, M. S., and Bhattacharyya, S. (2013). Evaluation of luminex xTAG gastrointestinal pathogen analyte-specific reagents for high-throughput, simultaneous detection of bacteria, viruses, and parasites of clinical and public health importance. *J. Clin. Microbiol.* 51, 3018–3024. doi: 10.1128/JCM.00896-13
- Onori, M., Coltella, L., Mancinelli, L., Argentieri, M., Menichella, D., Villani, A., et al. (2014). Evaluation of a multiplex PCR assay for simultaneous detection of bacterial and viral enteropathogens in stool samples of paediatric patients. *Diagn. Microbiol. Infect. Dis.* 79, 149–154. doi: 10.1016/j.diagmicrobio.2014.02.004
- Ottesen, A., Ramachandran, P., Reed, E., White, J. R., Hasan, N., Subramanian, P., et al. (2016). Enrichment dynamics of *Listeria monocytogenes* and the associated microbiome from naturally contaminated ice cream linked to a listeriosis outbreak. *BMC Microbiol.* 16:275. doi: 10.1186/s12866-016-0894-1
- Ottesen, A. R., Gonzalez, A., Bell, R., Arce, C., Rideout, S., Allard, M., et al. (2013). Co-enriching microflora associated with culture based methods to detect *Salmonella* from tomato phyllosphere. *PLoS ONE* 8:e73079. doi: 10.1371/journal.pone.0073079
- Pagotto, F., Hébert, K., and Farber, J. (2011a). “MFHPB-30. Isolation of *Listeria monocytogenes* and other *Listeria* spp. from foods and environmental samples,” in *Compendium of Analytical Methods*, Vol. 2, (Ottawa, ON: Health Products and Food Branch).
- Pagotto, F., Trottier, Y.-L., Upham, J., and Iugovaz, I. (2011b). “MFLP-74. Enumeration of *Listeria monocytogenes* in foods,” in *Compendium of Analytical Methods*, Vol. 3, (Ottawa, ON: Health Products and Food Branch).
- Park, E., Chun, J., Cha, C., Park, W., Ok, C., and Bae, J. (2012). Bacterial community analysis during fermentation of ten representative kinds of kimchi with barcoded pyrosequencing. *Food Microbiol.* 30, 197–204. doi: 10.1016/j.fm.2011.10.011
- Perez-Llarena, F. J., and Bou, G. (2016). Proteomics as a tool for studying bacterial virulence and antimicrobial resistance. *Front. Microbiol.* 7:410. doi: 10.3389/fmicb.2016.00410
- Perry, J. A., and Wright, G. D. (2014). Forces shaping the antibiotic resistome. *Bioessays* 36, 1179–1184. doi: 10.1002/bies.201400128
- Ronholm, J., Nasheri, N., Petronella, N., and Pagotto, F. (2016). Navigating microbiological food safety in the era of Q12 whole genome sequencing. *Clin. Microbiol. Rev.* 29, 837–857. doi: 10.1128/CMR.00056-16
- Ruppé, E., Baud, D., Schicklin, S., Guigon, G., and Schrenzel, J. (2016). Clinical metagenomics for the management of hospital- and healthcare-acquired pneumonia. *Fut. Microbiol.* 11, 427–439. doi: 10.2217/fmb.15.144
- Sahl, J. W., Schupp, J. M., Rasko, D. A., Colman, R. E., Foster, J. T., and Keim, P. (2015). Phylogenetically typing bacterial strains from partial SNP genotypes observed from direct sequencing of clinical specimen metagenomic data. *Genome Med.* 7, 52. doi: 10.1186/s13073-015-0176-9
- Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., et al. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* 12:87. doi: 10.1186/s12915-014-0087-z
- Schellenberg, J. J., Jayaprakash, T. P., Gamage, N. W., Patterson, M. H., Vanechoutte, M., and Hill, J. E. (2016). *Gardnerella vaginalis* subgroups defined by *cpn60* sequencing and sialidase activity in isolates from Canada, Belgium and Kenya. *PLoS ONE* 11:e0146510. doi: 10.1371/journal.pone.0146510
- Schloss, P. D., and Handelsman, J. (2005). Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. *Genome Biol.* 6:229. doi: 10.1186/gb-2005-6-8-229
- Schneeberger, P. H. H., Becker, S. L., Pothier, J. F., Duffy, B., N'Goran, E. K., Beuret, C., et al. (2016). Metagenomic diagnostics for the simultaneous detection of multiple pathogens in human stool specimens from Cote d'Ivoire: a proof-of-concept study. *Infect. Genet. Evol.* 40, 389–397. doi: 10.1016/j.meegid.2015.08.044
- Segata, N., Waldron, L., Ballarini, A., Narasimhan, V., Jousson, O., and Huttenhower, C. (2012). Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* 9, 811–814. doi: 10.1038/nmeth.2066
- Sharpton, T. J. (2014). An introduction to the analysis of shotgun metagenomic data. *Front. Plant Sci.* 5:209. doi: 10.3389/fpls.2014.00209
- Shea, S., Kubota, K. A., Maguire, H., Gladbach, S., Woron, A., Atkinson-Dunn, R., et al. (2017). Clinical microbiology laboratories' adoption of culture independent diagnostic tests are a threat to food-borne disease surveillance in the United States. *J. Clin. Microbiol.* 55, 10–19. doi: 10.1128/JCM.01624-16
- Smits, S. L., Schapendonk, C. M. E., Beek, J., Van Vennema, H., Schürch, A. C., Schipper, D., et al. (2014). New viruses in the Netherlands. *Emerg. Infect. Dis.* 20, 1218–1222. doi: 10.3201/eid2007.140190
- Spletstoesser, W. D., Seibold, E., Zeman, E., Trebesius, K., and Podbielski, A. (2010). Rapid differentiation of *Francisella* species and subspecies by fluorescent in situ hybridization targeting the 23S rRNA. *BMC Microbiol.* 10:72. doi: 10.1186/1471-2180-10-72
- Steyer, A., Jevšnik, M., Petrovec, M., Pokorn, M., Grosek, Š., Fratnik Steyer, A., et al. (2016). Narrowing of the diagnostic gap of acute gastroenteritis in children 0–6 years of age using a combination of classical and molecular techniques, delivers challenges in syndromic approach diagnostics. *Pediatr. Infect. Dis. J.* 35, e262–e270. doi: 10.1097/INF.0000000000001208
- Taylor, A. J., Lappi, V., Wolfgang, W. J., Lapierre, P., Palumbo, M. J., Medus, C., et al. (2015). Characterization of foodborne outbreaks of *Salmonella enterica* serovar Enteritidis with whole-genome sequencing single nucleotide polymorphism-based analysis for surveillance and outbreak detection. *J. Clin. Microbiol.* 53, 3334–3340. doi: 10.1128/JCM.01280-15
- The NIH HMP Working Group (2009). The NIH human microbiome project. *Genome Res.* 19, 2317–2323. doi: 10.1101/gr.096651.109
- Thomas, M. K., Murray, R., Flockhart, L., Pintar, K., Fazil, A., Nesbitt, A., et al. (2015). Estimates of foodborne illness-related hospitalizations and deaths in



- Canada for 30 specified pathogens and unspecified agents. *Foodborne Pathog. Dis.* 12, 820–827. doi: 10.1089/fpd.2015.1966
- Thurber, R. V., Haynes, M., Breitbart, M., Wegley, L., and Rohwer, F. (2009). Laboratory procedures to generate viral metagenomes. *Nat. Protoc.* 4, 470–483. doi: 10.1038/nprot.2009.10
- Todd, E. C. D., Greig, J. D., Bartleson, C. A., and Michaels, B. S. (2008). Outbreaks where food workers have been implicated in the spread of foodborne disease. Part 4. Infective doses and pathogen carriage. *J. Food Prot.* 71, 2339–2373. doi: 10.4315/0362-028X-71.11.2339
- Van Belkum, A., and Dunne, W. M. (2013). Next-generation antimicrobial susceptibility testing. *J. Clin. Microbiol.* 51, 2018–2024. doi: 10.1128/JCM.00313-13
- van Duin, D., and Paterson, D. L. (2016). Multidrug-resistant bacteria in the community: trends and lessons learned. *Infect. Dis. Clin. North Am.* 30, 377–390. doi: 10.1016/j.idc.2016.02.004
- Verhoef, L., Williams, K. P., Kroneman, A., Sobral, B., van Pelt, W., and Koopmans, M. (2012). Selection of a phylogenetically informative region of the norovirus genome for outbreak linkage. *Virus Genes* 44, 8–18. doi: 10.1007/s11262-011-0673-x
- Vinje, J. (2014). Advances in laboratory methods for detection and typing of norovirus. *J. Clin. Microbiol.* 53, 373–381. doi: 10.1128/JCM.01535-14
- Wilson, M. R., Naccache, S. N., Samayoa, E., Biagtan, M., Bashir, H., Yu, G., et al. (2014). Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *N. Engl. J. Med.* 370, 2408–2417. doi: 10.1056/NEJMoa1401268
- Won, Y. J., Park, J. W., Han, S. H., Cho, H. G., Kang, L. H., Lee, S. G., et al. (2013). Full-genomic analysis of a human norovirus recombinant GII.12/13 novel strain isolated from South Korea. *PLoS ONE* 8:e85063. doi: 10.1371/journal.pone.0085063
- Wood, D. E., and Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 15:R46. doi: 10.1186/gb-2014-15-3-r46
- Zagordi, O., Bhattacharya, A., Eriksson, N., and Beerenwinkel, N. (2011). ShoRAH: estimating the genetic diversity of a mixed sample from next-generation sequencing data. *BMC Bioinformatics* 12:119. doi: 10.1186/1471-2105-12-119
- Zhao, X., Lin, C.-W., Wang, J., and Oh, D. H. (2014). Advances in rapid detection methods for foodborne pathogens. *J. Microbiol. Biotechnol.* 24, 297–312. doi: 10.4014/jmb.1310.10013
- Zhou, Y., Wylie, K., Ei Feghaly, R. E., Mihindukulasuriya, K., Elward, A., Haslam, D. B., et al. (2016). Metagenomic approach for identification of the pathogens associated with diarrhea in stool specimens. *J. Clin. Microbiol.* 54, 368–375. doi: 10.1128/JCM.01965-15
- Zilelidou, E., Karmiri, C., Zoumpopoulou, G., Mavrogonatou, E., Kletsas, D., and Tsakalidou, E. (2016). *Listeria monocytogenes* strains underrepresented during selective enrichment with an iso method might dominate during passage through simulated gastric fluid and *in vitro* infection of caco-2 cells. *Appl. Environ. Microbiol.* 82, 6846–6858. doi: 10.1128/AEM.02120-16
- 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 © 2017 Forbes, Knox, Ronholm, Pagotto and Reimer. 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) or licensor 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.