Gut Pathogens

RESEARCH Open Access

Metagenomics: a new frontier for routine pathology testing of gastrointestinal pathogens

Nicola Z. Angel[1*], Mitchell J. Sullivan[1], Areej Alsheikh-Hussain[1], Liang Fang[1], Samantha MacDonald[1], Alena Pribyl[1], Blake Wills[1], Gene W. Tyson[1,2], Philip Hugenholtz[1,3], Donovan H. Parks[1], Paul Griffin[1,4,5] and David L. A. Wood[1]

Abstract

Background Accurate and comprehensive identification of enteropathogens, causing infectious gastroenteritis, is essential for optimal patient treatment and effective isolation processes in health care systems. Traditional diagnostic techniques are well established and optimised in low-cost formats. However, thorough testing for a wider range of causal agents is time consuming and remains limited to a subset of pathogenic organisms. Metagenomic next-generation sequencing (mNGS) allows the identification of all pathogens in a sample in a single test, without a reliance on culture or introduction of target selection bias. This study aims to determine the ability to routinely apply mNGS testing, in comparison to traditional culture or polymerase chain reaction (PCR) based tests, for the identification of causal pathogens for gastrointestinal infections.

Results The performance of mNGS, PCR and microscopy, culture and sensitivity (MCS) assays was established using 2,619 prospectively collected faecal samples from patients with symptomology indicative of infectious gastroenteritiss. Commonly experienced pathogens including *Aeromonas* spp, *Campylobacter* spp, *Salmonella* spp and *Giardia* spp, in single and co-infected patients, were used to establish test outcomes. When testing for these organisms, using the combined result from either or both PCR and MCS testing as the comparator, the mNGS assay had clinically acceptable sensitivity (89.2–100%). Further, the mNGS assay detected 14 additional enteropathogens, that were either not detected or not tested, by initial PCR/MCS testing.

Conclusions The advantage of mNGS compared to other syndromic testing systems is the broad range of detectable targets and the ability to interrogate samples without clinician informed or assay specific bias. With the development of newer sequencing assays, it is now feasible to test for a wide range of target organisms in a sample using a single mNGS test. Overall, the mNGS based approach enabled pathogen detection that was comparable to conventional diagnostics and was shown to have the potential to be extended for the detection of many pathogens and genes of clinical interest. In conclusion, the mNGS assay offers an easy, sample to answer workflow with rapid detection of enteropathogens and has the potential to improve diagnosis, therapy and infection control precautions.

Keywords Infection, Pathogen, Gastrointestinal, Metagenomics, Faecal, Enteropathogens

*Correspondence:
Nicola Z. Angel
nicola.angel@microba.com
Full list of author information is available at the end of the article

Background

Gastrointestinal infections, characterised by vomiting and diarrhea, remain a leading cause of morbidity and mortality in both high and low-resource settings [1]. Persistent infection can lead to significant clinical outcomes including malabsorption of nutrients, food intolerances, allergies, diabetes, and diseases of the intestinal tract, such as ulcerative colitis and Crohn's disease, often requiring hospitalisation [2]. Transmission may occur by ingestion of contaminated food and water or person-to-person contact. Timely detection, identification and characterisation of causal microorganisms allows appropriate treatment plans to be formed by clinicians [3]. These causal microorganisms encompass a wide range of viral, bacterial, and eukaryotic species. Traditional diagnostic and molecular testing methods, which include microscopy and culture-based assays, fail to detect a pathogenic organism in up to 90% of cases [4]. Addition of multiplexed PCR based tests, to complement traditional microbiological tests, have improved diagnosis. PCR assays enable the identification of an increased number of pathogens with greater speed and sensitivity (Table 1) [5, 6]. The introduction of these improved tests was shown to result in a significant reduction in endoscopic procedures, antibiotic usage, length of hospital stays and abdominal radiology; leading to appreciable savings to health care systems [7]. However, limitations to testing are still encountered as the clinician must formulate a potential diagnosis on the basis of clinical presentation and patient history, in a disease that has overlapping infectious clinical manifestations and non-infectious causes [8]. Given the large range of pathogens, the construction of initially available diagnostic PCR panels was a compromise between providing less comprehensive diagnostic panels or less than desirable assay performance [9]. Current PCR assays are more comprehensive, however generally a physician-directed testing schema requires serial laboratory testing, which increases time to diagnosis and costs, and most often does not address rare pathogens. This ultimately impacts treatment and patient care that will vary with different infectious origins (bacterial, viral, eukaryotic), and delays the identification of both nosocomial and foodborne outbreaks [10].

The use of metagenomic next-generation sequencing (mNGS) in a diagnostic setting has been proposed because a broad range of causal organisms can be identified in a single assay with improved taxonomic resolution [11]. High throughput mNGS assays are fundamentally hypothesis free and allow detection and identification of uncultured microorganisms. The method has been applied extensively in research studies, associating specific human microbiome features or organisms to a wide range of health conditions, including gastrointestinal complaints [12], metabolic diseases [13], allergies [14], and neurological conditions [15]. The rapidly decreasing costs of sequencing and data processing, together with improved bioinformatic techniques, suggest that mNGS could be viable for routine clinical use, including diagnosis of human pathogens. In routine clinical diagnostic assays of gastrointestinal infections from faeces, the application of mNGS has been limited to microbial strain typing in epidemiological studies, although it is understood that implementation of the technique to microbial communities would lead to actionable clinical outcomes [16, 17].

Examples of the pioneering diagnostic application of mNGS include successful treatment of a critically ill child with meningoencephalitis [18] and investigation of *Escherichia coli* strain 0104:H4 outbreaks in public heath settings [19–21]. Several groups have now validated the use of mNGS in certified clinical laboratories to diagnose infections in different sample types including serum [22], sputum [23], and spinal fluid [24]. These established causal agents for a wide range of diseases including meningitis, encephalitis [18, 24], sepsis [25], and pneumonia [26, 27]. However, to our knowledge, there are no certified mNGS based tests for faecal samples. The application of this technology to faecal samples is challenging, as the causal agent is hidden within endogenous microorganisms native to the gastrointestinal tract [28], which is bioinformatically challenging to achieve acceptable diagnostic performance due to high frequency of similar genomic regions between pathogens and commensals. Preliminary research has identified potential pathogens in patients with acute cholecystitis [29] and *Clostridioides difficile* infections [30]. Additionally, a comprehensive description of the faecal microbiome has been demonstrated to provide valuable information by allowing the identification of individual microorganisms or sets of microorganisms that may be used as biomarkers for grading severity or prognosis of disease states [31, 32].

In this study, we demonstrate that mNGS has the necessary diagnostic performance for clinical application in infectious disease testing of faecal samples. The use of mNGS, by direct assay of stool samples, was compared to the outcome of conventional diagnostic testing to identify causal agents of infectious gastrointestinal complaints for a selected common pathogen panel.

Methods

Study design

To establish the performance of mNGS as a diagnostic assay, a retrospective study was designed which established test outcomes from samples that were submitted for routine pathology testing for GI pathogens and from

Table 1 Comparison of Existing Infectious Disease Diagnostic Tests

Test	Advantages	Disadvantages
Microscopy and staining	Rapid inexpensive	Requires staff interpretation of result low sensitivity unless high disease burden limited detection of subset of organisms low specificity
Culture	High throughput formats available inexpensive well documented historically	Limited sensitivity limited application for viral pathogens time consuming-long time to results
Targeted PCR	Simple rapid inexpensive potentially quantitative	Hypothesis driven requires known target reference material poor specificity (targets a small portion of the genome) variable reliability
Multiplexed PCR	Rapid multiple organisms detected high throughput format available	Low specificity (usually at genus level) false positives for many organisms targets small portion of genome
Targeted universal multiplex PCR NGS- 16S, ITS	Multiple organisms detected potentially quantitative Ability to multiplex samples in processing high throughput format available	Requires primer selection expensive time consuming targets small portion of genome
Targeted NGS Capture and enrichment panels	Multiple organisms detected increased sensitivity for detection	Complex sample processing targets specific known genome content expensive Time consuming
Metagenomic NGS	Hypothesis-free, unbiased assay includes novel and unexpected organisms multiple organisms detected detects entire genome covers all pathogen types including viral high throughput format available	Potential for human host background expensive time consuming requires comprehensive reference database

samples that were assumed to be pathogen free (workflow outlined in Fig. 1).

Clinical samples

A total of 2619 stool specimens submitted to a clinical laboratory for routine diagnostic testing of gastrointestinal pathogens were obtained as part of an ethically approved clinical trial to assess the performance of the mNGS assay (Mater Research Clinical Governance Approval; Project 46934 Ref AM/MML/46934 v4) within a local Australian population. Eligible samples were obtained from patients who submitted stool samples for testing as clinically indicated by the requesting practitioner. Direct sampling of the originally submitted sample was made to compare the diagnostic performance of the mNGS assay to the requested pathology testing. Aliquots of the original samples were made and stored at − 20 °C. Within this sample set, 87% of sample reports were negative for the pathogens tested with routine pathology testing. This allowed a retrospective panel of 329 samples to be selected after processing at the Mater Pathology (MP) clinical laboratory. Selection of samples was made to include reported gastrointestinal pathogen targets and samples failing to identify a pathogen target (negative) as indicated by routine analysis in the supplying laboratory. 238 samples were positive for a single targeted organism, 23 reported multiple infectious organisms, and 66 reported negative test results for all tested organisms. Two samples were removed from the testing panel due to high relative abundance of human reads, which resulted in insufficient microbial reads required for mNGS analysis. Complete test reporting is included in Supplemental File 1.

Non-Clinical samples

A set of 200 non-clinical samples were obtained as part of an ethically approved research program (Bellberry Approved Project No.: 2018-05-400-A-2) that was similarly formed from a local Australian population. This sample set was established from participants over the age of 18 that passed self-reported exclusion criteria that classified them as presumptively healthy. These criteria included no major medical conditions, no infections in the previous six months, no gastrointestinal disease indicators, a BMI less than 30, no antibiotic/antifungal use in the previous six months, no use of medications known to strongly impact the gut microbiome, not pregnant or breastfeeding, no excessive alcohol intake, and no smoking. These samples could be assumed to be free of gastrointestinal pathogens in the clinical range. The sample set could not be confirmed to be free of potential participants that may be asymptomatic.

Routine pathology testing

Conventional testing, following the routine diagnostic methodology of the testing laboratory (Mater Pathology) was performed on all samples. Assays included microscopy, culture and sensitivity (MCS) testing, which comprised culture on selective media for *Salmonella* spp., *Shigella* spp., *Campylobacter* spp., *Aeromonas hydrophila, Yersinia* spp., *Vibrio parahaemolyticus/cholerae* and Shiga toxin-producing *Escherichia coli*. Alternatively, or additionally, testing was performed with PCR for *Salmonella* spp., *Campylobacter* spp., *Shigella* spp., *Yersinia enterocolitica, Rotavirus A, Adenovirus, Sapovirus, Norovirus 1, Giardia lamblia/intestinalis, Cryptosporidium parvum/hominis* and *Entamoeba histolytica*. Tests performed were on the direction of the submitting clinician and following the routine test selection for the laboratory.

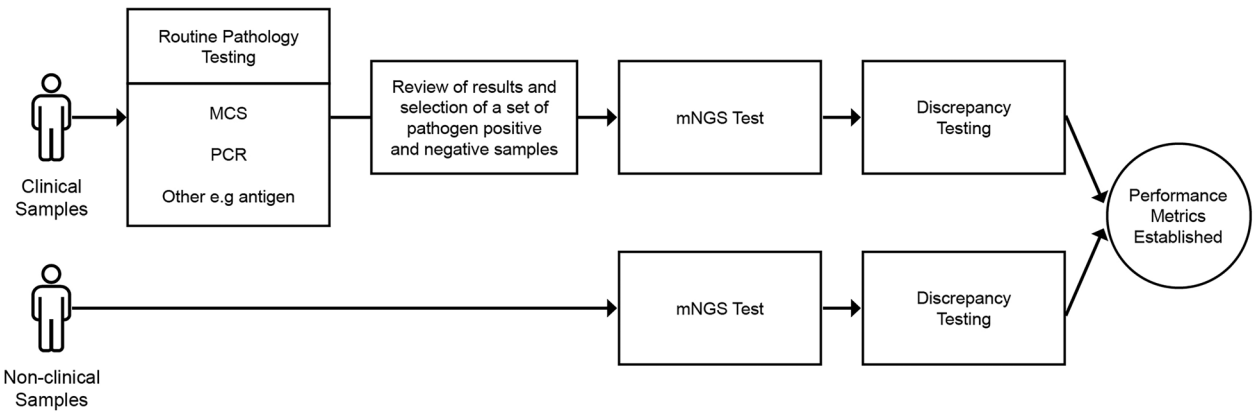


Fig. 1 Outline of Retrospective Study to Assess mNGS Assay Performance

mNGS testing

Faecal samples were collected in a routine pathology stool collection container (Sardstedt) and duplicate samples collected with a flocked active drying swab (Copan). Specimens were stored at − 20 °C until extracted with the QIAcube HT Automated DNA Extraction System, using the DNeasy 96 PowerSoil Pro extraction Kit (Cat. No. 47021), in accordance with the manufacturer's instructions. The quantity and purity of the extracted genomic DNA was evaluated with fluorometric assay (QuantIT High Sensitivity, ThermoFisher) and stored at 4 °C until used.

Samples were processed with secured version-controlled protocols, analytical pipelines and reference databases at Microba Laboratories (ISO15189:2022; NATA certified) following the routine diagnostic methodology of the laboratory. The mNGS workflow is outlined in Fig. 2. 1 ng of extracted genomic DNA was used in library construction using the DNA Tagmentation Kit (Illumina, CA, USA) according to the manufacturer's protocol, with slight modification to allow for high throughput processing of larger sample numbers in smaller reaction volumes. Sequencing was performed on the NovaSeq 6000 (Illumina, CA, USA) in 2×150 bp format, generating a minimum of 16 million read pairs per sample. Primary sequence output was demultiplexed using manufacturer supplied processing using Basespace (Illumina) and FASTQ sequence files generated using default parameters. The resulting reads were quality trimmed and subsequently filtered if shorter than 100 bases by Trimmomatic [33]. All sequences from metagenomic sequencing have been deposited at SRA (BioProject PRJNA1156595), after removal of reads containing human DNA. Quality-controlled read pairs were mapped to genomes in Microba's proprietary reference genome databases. Taxonomic profiling was performed using a proprietary genome-based bioinformatic pipeline utilizing the Microba Community Profiler (MCP) [37] and additional proprietary bioinformatics methods. This bioinformatic pipeline was parameterized using in silico communities to achieve accurate detection of target pathogens. Minimum requirements for the quality of sequencing and control sample results were met for each processing run. The clinical report lists detected organisms meeting target-specific reporting criteria, resulting in reporting of 'detected' or 'not detected' within the resulting report.

Control samples

Control samples were included in each mNGS processing run. These were designed to confirm the appropriate

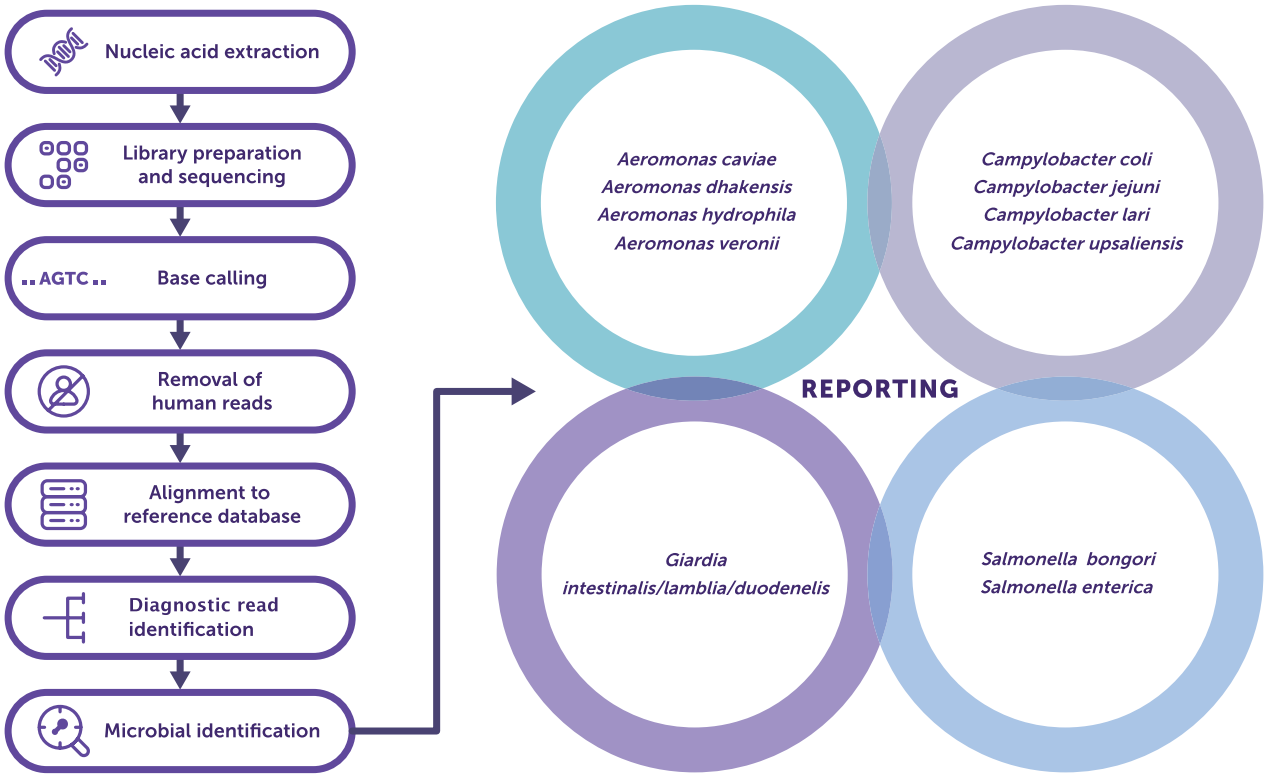


Fig. 2 Outline of Workflow for Clinical Metagenomics

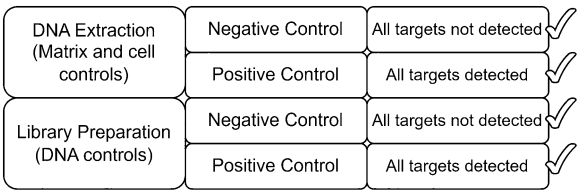


Fig. 3 Outline of required controls for sample processing to proceed to report generation

performance of the assay in each step of the assay and covered both positive and negative controls. To facilitate the creation of these controls a lot-controlled reference material was established from a faecal sample that had been previously screened to be free of the nominated gastrointestinal pathogens for the assay, by PCR and mNGS analysis. The control process applied is outlined in Fig. 3 for reference.

Positive control

The established faecal reference material was used as a base to which a spike in of mixed target species at a clinically reportable range was made. Target species included in the control material were isolated within the laboratory by culture directed approaches for *Aeromonas* spp., *Campylobacter* spp., *Salmonella* spp. and *Giardia* spp. Resulting axenic cultures were sequenced using DNA Tagmentation Library Preparation (Illumina), following the manufacturer's protocol for direct input of microbial material and validated genomically from 150 bp paired-end data generated on the NovaSeq 6000. Resulting data was trimmed and quality controlled using Trimmomatic v0.36 (ILLUMINACLIP:adapters_NexteraPE-PE_TruSeq3-PE.fa:2:30:10, LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, CROP:150 HEADCROP:0, MINLEN:100) [33], and phiX reads were removed using bbduk from bbmap v38.68 [34]. Reads were merged using bbmerge from bbmap and assembled using Spades v3.13.0 [35]. Resulting genome assemblies were evaluated with CheckM v1.0.18 [36] for genomic completeness and potential contamination of the reference material.

Negative control

Target free wells were included in all stages of processing, from DNA extraction to sequence analysis, on each processing run. These were used to monitor for reagent/consumable contamination and any potential cross contamination between samples during processing.

Required control results

Each metagenomic assay required positive identification of all nominated organisms within the positive control and absence of identification of any organisms within the negative control for analysis of the samples to proceed to reporting.

Discrepant analysis

In the case of discrepant results, confirmatory PCR screening was performed using a validated clinical assay as per the manufacturer's instructions for the identification of *Campylobacter, Aeromonas* and *Salmonella* species (LightMix Modular Gastro Bacteria Multiplex Testing (TIB Molbiol)) or *Giardia* spp. (Gastro Parasite Multiplex Testing (TIB Molbiol)). Each discordant sample was tested in a minimum of two independent PCR tests to enable consensus of the sample content to be reported at the threshold of 3 of 4 concordant results. The result of this orthogonal confirmatory test was considered accurate and used to resolve the discrepancy. Six samples were unable to be resolved due to an inability to reach a consensus from multiple rounds of testing in which they returned a mix of detected/not detected results, presumably being at the limit of detection for the testing applied.

Statistical analysis

Statistical analysis was performed using Pearson's chi square test to obtain a correlation coefficient between DETECTED/NOT DETECTED results of pathogen detection by mNGS in comparison to the combined results of the validated routine testing of PCR and MCS. Statistical testing was performed with Stata 17 Data Analysis and Statistical Software and GraphPad Prism Software. A p-value of less than or equal to 0.05 was considered statistically significant.

Extended target testing

A subset of 32 samples from the original clinical sample set were also tested for the presence of *Adenovirus* spp. following the same laboratory processing and confirmatory PCR approach. Sample contents were evaluated with the Microba Community Profiler [37] as a proof of concept that the validated assay could be extended for additional targets.

Results

Limit of detection

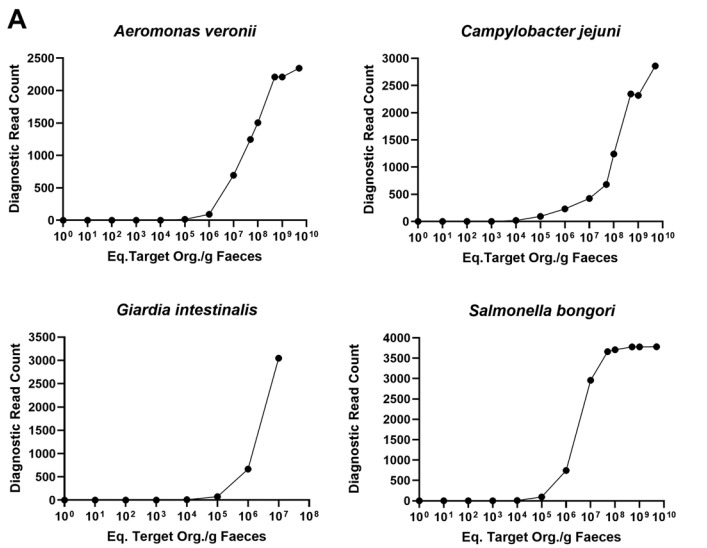
To determine if mNGS is viable for clinical applications, we first determined the limit of detection for clinically relevant species of *Campylobacter*, *Aeromonas*, *Salmonella* and *Giardia* that could be compared to standard qPCR and culture-based assays. This formed a set of 11 organisms, noted as reported in Fig. 2. A base faecal matrix was pre-screened for these potentially pathogenic organisms and was determined to have 2×10^8 organisms/g,

which was noted to be within the temporal variation experienced within faeces from healthy and unhealthy donors [38]. The limit of detection (LoD), defined as the lowest concentration at which 95% of positive samples were detected, was determined for each of the 11 organisms on the mNGS reporting panel at a standardised read sequencing depth. Serial dilutions from 5×10^9 genome equivalents/gram faeces were made across a minimum 4-log dilution range for each organism, and each sample assayed in triplicate by both PCR and mNGS. As the established value for all species within each genus was the same, an example of the read count correlation with increased number of organisms/g of faeces for one species within each genus has been included in Fig. 4a. The LoD ranged from 1×10^4 to 1×10^5 organisms/g for the mNGS assay. In comparison, the PCR assay LoD ranged from 1×10^4 to 1×10^6 organisms/g. This was confirmed using reference material for each of the 11 species and is presented as a collated concordant outcome for all species within each genus (Fig. 4b). All samples reported as negative by mNGS were also negative by PCR testing, except for *Aeromonas* spp., which required tenfold greater target organism levels to reliably return a positive result with mNGS testing. Conversely, the mNGS assay identified a positive result at tenfold less target than PCR for both *Giardia* and *Salmonella* spp. Testing of closely related species from the same genus, that were not clinically relevant, reported negative test results in both mNGS and PCR testing (Supplementary Table 2).

Clinical performance

To assess the validity of mNGS as a diagnostic for a range of infectious pathogens, we compared the results of the mNGS faecal assay to pathogens identified by standard testing of 329 faecal samples from selected patients in the Mater Hospital System, Brisbane, Australia. The required clinical presentation of these patients was gastrointestinal symptomology indicative of infectious disease.

Standard testing of faecal samples uses a combination of MCS, PCR and antigen testing. This testing was not uniform across the sample set due to technical limitations and individual clinician choice of test. However, each sample assessed contained at least one test outcome from the standard tests offered within the initial pathology service. These initial tests were all weighted equally in being considered a test outcome. The approach of the mNGS assay was to report only identified organisms that were clearly established as gastrointestinal pathogens, rather than listing all microbes present in the sample. Using this approach, from the 329 faecal samples, preliminary test results indicated that 18 were positive for more than one pathogen, 170 were positive for one pathogen and 150 were negative for all tested pathogens. Four pathogens of clinical interest typically identified at the genus level by standard testing were used for validation testing of the mNGS assay, *Campylobacter* (87/338), *Salmonella* (50/338), *Giardia* (28/338) and *Aeromonas* (15/338) [39]. Although, these may not be the most important organisms clinically, they were justified



B

Target organism CFU/g Faeces (eq. org per g faeces)	*Aeromonas spp*						*Campylobacter spp*					
	mNGS			PCR			mNGS			PCR		
	Replicate 1	Replicate 2	Replicate 3	Replicate 1	Replicate 2	Replicate 3	Replicate 1	Replicate 2	Replicate 3	Replicate 1	Replicate 2	Replicate 3
5×10^9	+	+	+	+	+	+	+	+	+	+	+	+
1×10^9	+	+	+	+	+	+	+	+	+	+	+	+
5×10^8	+	+	+	+	+	+	+	+	+	+	+	+
5×10^7	+	+	+	+	+	+	+	+	+	+	+	+
1×10^7	+	+	+	+	+	+	+	+	+	+	+	+
1×10^6	+	+	+	+	+	+	+	+	+	+	+	+
1×10^5	+	+	+	+	+	+	+	+	+	+	+	+
1×10^4	-	-	-	+	-	+	+	+	+	+	+	+
1×10^3	-	-	-	-	-	-	-	-	-	-	-	-
1×10^2	-	-	-	-	-	-	-	-	-	-	-	-
1×10^1	-	-	-	-	-	-	-	-	-	-	-	-
1×10^0	-	-	-	-	-	-	-	-	-	-	-	-
0	-	-	-	-	-	-	-	-	-	-	-	-

Target organism (eq. org per g faeces)	*Giardia spp*						*Salmonella spp*					
	mNGS			PCR			mNGS			PCR		
	Replicate 1	Replicate 2	Replicate 3	Replicate 1	Replicate 2	Replicate 3	Replicate 1	Replicate 2	Replicate 3	Replicate 1	Replicate 2	Replicate 3
5×10^9	+	+	+	+	+	+	+	+	+	+	+	+
1×10^9	+	+	+	+	+	+	+	+	+	+	+	+
5×10^8	+	+	+	+	+	+	+	+	+	+	+	+
5×10^7	+	+	+	+	+	+	+	+	+	+	+	+
1×10^7	+	+	+	+	+	+	+	+	+	+	+	+
1×10^6	+	+	+	+	+	+	+	+	+	+	+	+
1×10^5	+	+	+	+	+	+	+	+	+	-	-	-
1×10^4	+	+	+	-	-	-	-	+	-	-	-	-
1×10^3	-	-	-	-	-	-	-	-	-	-	-	-
1×10^2	-	-	-	-	-	-	-	-	-	-	-	-
1×10^1	-	-	-	-	-	-	-	-	-	-	-	-
1×10^0	-	-	-	-	-	-	-	-	-	-	-	-
0	-	-	-	-	-	-	-	-	-	-	-	-

Fig. 4 Established Limit of Detection for Representative Species used to Demonstrate Clinical Performance of Metagenomic Assays. The limit of detection was established using known isolates as a spike in to pre-established screened negative faecal samples at known numbers of pathogenic organisms per gram of faeces. A Each organism was graphed to map the number of target organisms to the resulting diagnostic reads recorded. An example of one species within each of the genus evaluated is included in section A. B Tabulated summary of the analysis of the spike in samples over the range of 0 to 5 x 10^9 target organisms per gram of screened negative faeces, tested in 3 replicates with both mNGS and PCR and reported as Detected (+) or Not Detected (-)

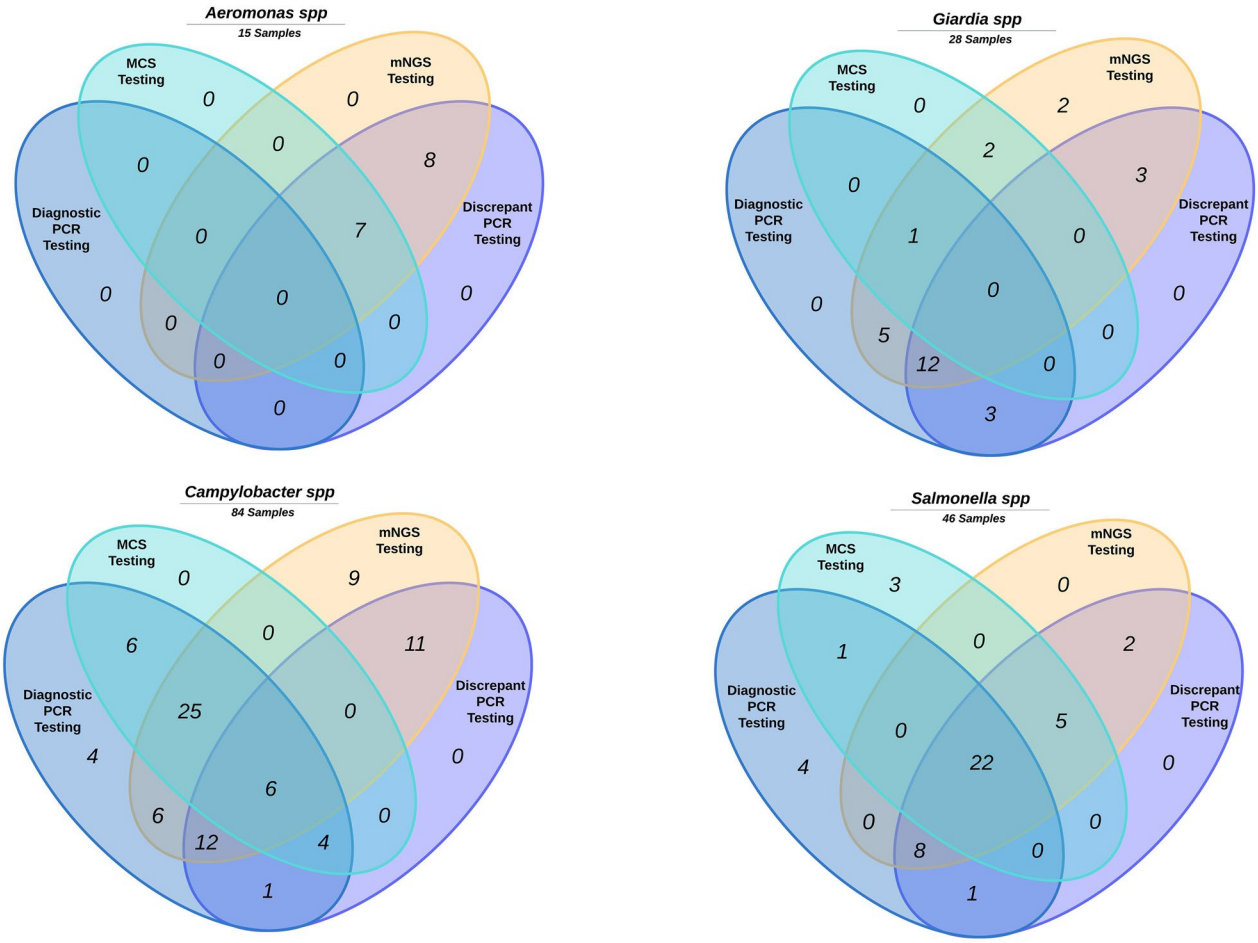


Fig. 5 Test Outcomes for MHC Clinical Sample Set Tested by PCR, MCS, mNGS and Discrepant PCR assay. A clinical sample set (MHC) was established from samples submitted for routine pathology testing, being symptomatically indicative of the presence of a gastrointestinal pathogen. The outcome from testing of these samples was determined for the initial pathology service (diagnostic PCR testing and MCS testing), and for the subsequent mNGS testing, and discrepant PCR testing, as required. The results were collated at a genus level for Aeromonas, Camplyobacter, Salmonella and Giardia species, given that PCR testing lacked resolution to the species level

in inclusion as they were contained on the reports issued by the primary pathology service and were noted to be present as targets in routinely offered diagnostic PCR panels. The selection of these organisms was also based on obtaining enough positive samples to generate statistically robust performance measures. As such, the study should be considered a proof of concept providing a basis for identification of other pathogenic species. Initially mNGS had modest concordance with standard testing, ranging from 46.7% (*Aeromonas*) to 75.9% (*Campylobacter*) sensitivity, requiring discordancy testing. Discordant test results were resolved according to the criteria outlined in the Methods (Supplementary Table 1). The notable difference in *Aeromonas* spp. testing outcomes was due to this genus not being included in the PCR assay used by the original testing laboratory. Therefore, identification of this pathogen was solely reliant on MCS testing, explaining the low concordance as MCS is known to be less reliable than molecular-based techniques [40, 41]. Overall, MCS failed to detect a pathogen where PCR did in 21 of 176 cases. However, in 3 cases, PCR failed to detect a pathogen where MCS did, demonstrating the limitations of current testing (Fig. 5).

After resolution of discordant test results with independent PCR testing, a composite result was made from the standard validated tests to establish pathogen presence in each sample, resulting in 152 positive samples. Using the composite results as ground truth, the sensitivity and specificity for detection of the 4 pathogens with mNGS was determined (Fig. 6). mNGS had a clinically acceptable sensitivity for all 4 pathogens; 89.2% for *Salmonella* spp., 88.5% for *Giardia* spp., 89.5% for *Campylobacter* spp. and 100% for *Aeromonas* spp. Analogous to PCR testing outcomes (130 of 152), the identification of

pathogens was higher with mNGS (137 of 152) than with MCS testing (83 of 152). Notably, mNGS testing resolves targets to the species level unlike MCS or PCR testing. Resolution of the positive *Aeromonas* samples at the species level revealed that each sample was dominated by a single species: *A. caviae* (4/15), *A. dhakensis* (5/15), *A. hydrophila* (5/15), and *A. veronii* (1/15). Our results are consistent with previous studies in which these four species are the most pervasive of the 19 known *Aeromonas* species considered potential pathogens [42].

The outcome of testing for *Campylobacter* species was mostly concordant (89.5%), with discordant outcomes being due to disagreements as to which species should be included as pathogenic in a diagnostic assay. For example, five of the samples were scored as positive for *Campylobacter* (Supplementary Table 2) from MCS and PCR testing but were scored as negative in the mNGS assay, because at the species level they were identified as either *C. hominis* or *C. concisus,* which we and others considered to be non-pathogenic due to their prevalence in healthy individuals [43, 44]. Further studies will be necessary to confirm the inference of non-pathogenicity applied in the mNGS test.

A number of samples were close to the PCR detection threshold (Ct > 35) for *Giardia, Salmonella* and *Campylobacter* spp., which presented as discordant results with mNGS. For example, in two cases duplicate PCR tests recorded alternating negative and positive results, indicating the inherent variability in assessing samples at the edge of clinical reporting ranges. Figure 7 demonstrates this trend, with the concordance of PCR test results when reported as units from a maximum of 40 cycles (assumed to be negative) to the number of informative sequencing reads identified in the mNGS assay for each species. Previous studies have concurred with this finding, with discrepant results containing Ct values of 35 or higher considered to be of questionable clinical relevance [6] and potentially incorrectly inflating the discrepancies reported (Supplementary Table 1). If these borderline results are removed, discrepancies between mNGS and standard testing are resolved in 92% of samples.

	Prevalence (%)	TP	TN	FP	FN
Aeromonas spp	0.7	15	315	0	0
Campylobacter spp	3.0	68	252	0	8
Giardia lambia	1.0	23	302	2	3
Salmonella spp	1.6	33	325	0	4
		Diagnostic Sensitivity		**Diagnostic Specificity**	
		%	**95% CI**	**%**	**95% CI**
Aeromonas spp		100.00	78.20%-100.00%	100.00	98.84%-100.00%
Campylobacter spp		89.47	80.31%-95.34%	100.00	98.55%-100.00%
Giardia lambia		88.46	69.85%-97.55%	99.34	97.64%-99.92%
Salmonella spp		89.19	74.58%-96.97%	100.00	98.87%-100.00%
		Diagnostic Accuracy		**Limit of Detection (95%CI) (CFU/g faeces)**	
		%	**95% CI**		
Aeromonas spp		100.00	98.89%-100.00%	10^5	
Campylobacter spp		99.68	98.29%-99.99%	10^4	
Giardia lambia		99.23	97.58%-99.87%	10^4	
Salmonella spp		99.83	98.65%-100.00%	10^5	
		Negative Predictive Value		**Positive Predictive Value**	
		%	**95% CI**	**%**	**95% CI**
Aeromonas spp		100.00	N/A	100.00	N/A
Campylobacter spp		99.68	99.38%-99.83%	100.00	N/A
Giardia lambia		99.88	99.66%-99.96%	57.59	25.31%-84.48%
Salmonella spp		99.82	99.56%-99.93%	100.00	N/A

Fig. 6 Clinical Performance Metrics for mNGS when Applied as a Routine Pathology Test. Standard performance metrics for the mNGS assay were established and are outlined for diagnostic sensitivity, specificity, accuracy, limit of detection, negative predictive value and positive predictive value

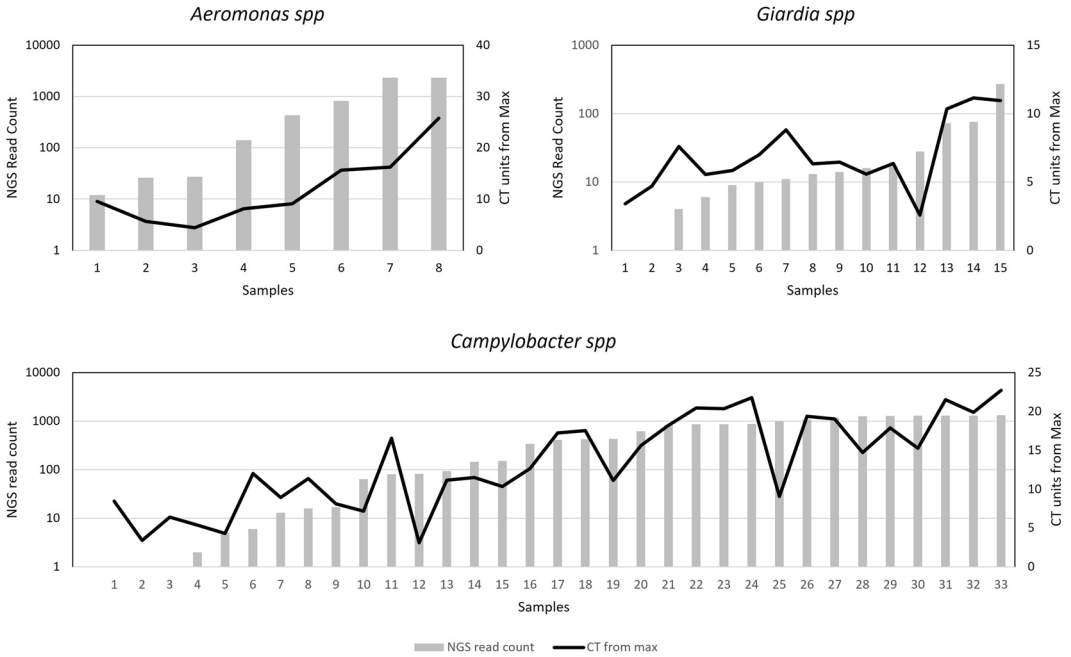


Fig. 7 Correlation of Diagnostic Read Identification with mNGS testing to Ct Value with PCR Testing of Clinical Samples. Comparison was made from samples within the MHC clinical sample set of the number of diagnostic reads (NGS read count) and PCR Ct value, displayed as the value below the upper Ct value cut off established by the manufacturer (CT units from max)

In total, mNGS testing identified 14 additional potential microbial pathogens in the sample set belonging to the genera *Aeromonas*, *Campylobacter*, *Giardia* and *Salmonella*, equating to a diagnostic result for an additional 4.14% of the originally tested sample set.

The evaluation of the selected pathogen panel indicates that mNGS can be applied as a standard testing approach with equal sensitivity and increased taxonomic resolution in comparison to routine pathology testing. While not having enough samples to perform the appropriate statistical tests, preliminary results suggest that the assay could equally be applied to a wider range of targets, including *Adenovirus, Cryptoporidium*, *Pleisomonas*, *Yersinia* and others. For example, *Adenovirus* could be resolved to subtypes A-H, using mNGS, providing clinically useful information as different subtypes have different clinical presentations and treatments e.g. A and F [45] which are associated with either diarrhea or cryptic enteric infection [45]. Other subtypes that are not reported by standard testing, such as C and D, are detected by mNGS providing a more complete picture (Fig. 8).

In the rare instances of a confirmed co-infection in the complete data set (12 of 2713 samples), mNGS identified common sets of co-occurring pathogens (Table 2). The most prevalent pathogens in co-infections were *Salmonella* (6/12), *Adenovirus* (5/12) and *Campylobacter* (5/12), which have been reported previously [46, 47].

Beyond pathogen detection, mNGS can detect many other features of potential clinical interest due to the untargeted nature of the technology in comparison to PCR. While PCR panels require selection and validation of each target primer set to be reported, mNGS covers all genomic material in the sample, relying on a selection of targets within the database to which the sequence data is compared. This means that the resulting data can be reanalysed for additional features of interest once the data is in hand. For example, antimicrobial resistance gene presence was recorded in 212 of 388 samples and toxin genes were detected in 110 of 388 samples (Supplementary Table 1) indicating that further clinical value can be developed for mNGS testing with refinement of the assay.

A limitation of mNGS noted in our study was host contamination that can mask microbial detection [48]. Human DNA content ranged from 0–5% in the healthy control samples and between 1–95% in the clinical sample set but was sufficiently low in the majority of samples within both populations to be generally appropriate for metagenomic based analysis. Results were unable to be generated from two samples due to high levels of human DNA (up to 99% of total DNA). These samples were from patients with active *C. difficile* infection that generally had higher human DNA concentrations. Methods to reduce host contamination, with commercially available solutions such as differential lysis or CpG methylated

Sample Number	Original PCR Result	mNGS Result	Confirmatory PCR Result	Consensus
BBF0325	Negative	*Adenovirus C*	Negative	**Negative**
BBF0660	Negative	*Adenovirus D*	Negative	**Negative**
BBF1911	Negative	*Adenovirus D*	Negative	**Negative**
BBF0354	Negative	Negative	Negative	**Negative**
BBF0375	Negative	Negative	Negative	**Negative**
BBF0526	Negative	Negative	Negative	**Negative**
BBF0358	*Adenovirus*	Negative	Negative	**Negative**
BBF0443	*Adenovirus*	Negative	Negative	**Negative**
BBF7394	*Adenovirus*	Negative	Negative	**Negative**
BBF8881	*Adenovirus*	Negative	Negative	**Negative**
BBF8885	*Adenovirus*	Negative	Negative	**Negative**
BBG6868	*Adenovirus*	Negative	Negative	**Negative**
BBG6876	*Adenovirus*	Negative	Negative	**Negative**
BBG6926	*Adenovirus*	Negative	Negative	**Negative**
BBG6927	*Adenovirus*	Negative	Negative	**Negative**
BBF0352	Negative	*Adenovirus F*	*Adenovirus*	**Positive**
BBF8866	Negative	*Adenovirus F*	*Adenovirus*	**Positive**
BBF8902	Negative	*Adenovirus F*	*Adenovirus*	**Positive**
BBG6869	*Adenovirus*	Negative	*Adenovirus*	**Positive**
BBG6872	*Adenovirus*	Negative	*Adenovirus*	**Positve**
BBG6874	*Adenovirus*	*Adenovirus A*	Negative	**Positive**
BBG6875	*Adenovirus*	*Adenovirus A*	Negative	**Positive**
BBG6873	*Adenovirus*	*Adenovirus A*	*Adenovirus*	**Positive**
BBF0350	*Adenovirus*	*Adenovirus F*	*Adenovirus*	**Positive**
BBF0449	*Adenovirus*	*Adenovirus F*	*Adenovirus*	**Positive**
BBF0525	*Adenovirus*	*Adenovirus F*	*Adenovirus*	**Positive**
BBF8901	*Adenovirus*	*Adenovirus F*	*Adenovirus*	**Positive**
BBG6870	*Adenovirus*	*Adenovirus F*	*Adenovirus*	**Positive**
BBG6871	*Adenovirus*	*Adenovirus F*	*Adenovirus*	**Positive**
BBG6877	*Adenovirus*	*Adenovirus F*	*Adenovirus*	**Positive**

**Adenovirus A* is associated with cryptic enteric infection
Adenovirus F is associated with diarrhea
Adenovirus C and *D* are not associated with gastrointestingal disease

Fig. 8 Detection of Adenovirus with mNGS and PCR Testing. A Subset of the MHC Clinical Sample Set was used to apply analysis for less prevalent organisms including Adenovirus. Comparison was made between the initial laboratory PCR test, the mNGS test, and the confirmatory PCR result to come a consensus outcome

DNA depletion, prior to sequencing is one approach to mitigate this issue [49–51].

mNGS performance was also assessed using a negative asymptomatic donor population. Of the 200 asymptomatic samples, 98% had no detectable pathogens, using both PCR and mNGS testing. In 3 positive samples, a single pathogenic species was detected in a clinically relevant range, with both mNGS and PCR, namely *Campylobacter coli*, *Aeromonas caviae,* and *Giardia intestinalis*. mNGS also identified a single sample that was positive for *Salmonella enterica* that was not detected by PCR testing. This species was identified as *Salmonella enterica* subsp. *arizonae*, which is determined to be a less common pathogenic species. Analysis of the genetic region used for the PCR test indicated that this serovar had only 93.9% identity to the genomic region

Table 2 Co-infections Within the Clinical SampleSset Identified with mNGSTtesting

Sample number	Coinfection type	Organisms identified
BBF0525	Bacteria-virus	*Salmonella bongori, Adenovirus F*
BBF8901	Eukaryotic-virus	*Giardia intestinalis, Adenovirus F*
BBF8888	Eukaryotic-bacteria	*Aeromonas dhakensis, Giardia intestinalis*
BBF8889	Bacteria-bacteria	*Aeromonas caviae, Salmonella enterica*
BBF8902	Bacteria-virus	*Salmonella enterica, Adenovirus F*
BBF8872	Bacteria-bacteria	*Aeromonas veronii, Campylobacter jejuni*
BBG0794	Bacteria-bacteria	*Aeromonas veronii, Campylobacter jejuni*
BBG0805	Eukaryotic-bacteria	*Salmonella enterica, Giardia intestinalis*
BBG5065	Eukaryotic-bacteria	*Salmonella enterica, Giardia intestinalis*
BBG5055	Bacteria-bacteria	*Campylobacter jejuni, Salmonella enterica*
BBF0660	Bacteria-virus	*Campylobacter jejuni, Adenovirus D*
BBF1911	Bacteria-virus	*Campylobacter jejuni, Adenovirus D*

targeted by the PCR test, which accounts for the lack of detection in this assay format. These results in a presumptive healthy population were similar to preliminary studies in smaller studies, in which *Salmonella* and *Camplyobacter* spp. were detected [8] and can be attributed to asymptomatic carriers.

Test turnaround time (TAT) is an important variable for clinicians. The median TAT was 1 week, which is comparable to standard testing. The fastest time to report generation was 48 h indicating that mNGS could be adopted for critical care. However, given the current routine test time frame, which is similar to sequential pathology testing including culture-based assays, we suggest that initial use of mNGS testing would be best applied to patients that have longer term chronic gastrointestinal distress. It is likely that these clinical samples, which have not been resolved with routine testing for common pathogens, require investigation for a broader range of organisms, which mNGS facilitates.

Discussion

Advances in diagnostics for gastrointestinal infectious diseases are critical to improve both the individual care of patients and health care systems globally. Current gold standard testing (PCR, MCS, Antibody staining, Maldi-TOF) lack specificity and coverage [41]. Therefore, new assays should identify a wide range of infectious agents in parallel with high precision. Identified pathogens should have sufficient characterisation to guide appropriate therapy and avoid mistreatment or overuse of antibiotics, that leads to the emergence of anti-microbial resistance. New tests need to enable surveillance of resistance patterns and screening of patients colonised with resistant pathogens, which supports prevention and infection control measures. Tests should also monitor for emerging pathogens within patient populations. Current techniques fall short of providing these functions, even when applied in a sequential syndromic testing format [17, 52].

The inclusive nature of mNGS addresses many of these requirements by providing contemporaneous information on the presence of bacteria, fungi, eukaryotes, and viruses with up to strain level resolution if required [53]. In this study only a modest number of pathogens were analysed that can be replicated with multiplex PCR; however, mNGS has the potential to identify hundreds of targets and indeed more targets can be added as needed. In other studies, it was observed that the application of mNGS to prevent multiple sample testing and provide additional information in testing of immunocompromised neonates, transplant recipients, and critically ill intensive care patients significantly improved survival and quality of life [19, 54]. This was also associated with reduced patient isolation times, hospital stays and medical interventions, leading to an overall decreased cost to both the patient and the hospital [55]. Similar outcomes could be expected to be achieved when applying mNGS to testing of patients with persistent or severe gastrointestinal infections, and specifically hospitalised infants, elderly and immunocompromised patients, where nosocomial infectious gastroenteritis is a common complication, contributing to increased morbidity, mortality, length of hospital stay and hospital costs [5, 9, 56]. Further studies including economical calculations, including those associated with delayed diagnostics and increased morbidity and mortality, are called for.

Accurately determining the microbial species present in clinical samples enables understanding of microbial interactions that promote the expression of pathogenicity and virulence factors, and the transfer of AMR genes

[57–60]. Application of shotgun metagenomics for the detection of multi-drug resistant microbial colonization in high-risk clinical settings has shown promise in onco-hematological patients [61–63]. For these patients, gold standard testing lacks sensitivity and is restricted to detection of *Enterobacterales* species that carry extended spectrum beta-lactamase (ESBL)/carbapenemase resistance genes [64]. In this scenario, relapsing patients follow treatment plans that exert selective pressure, significantly increasing the risk of bacteraemia in patients colonised by ESBL containing organisms, even at relatively low abundances [65–67]. Under-detection of colonisation is also problematic in allowing the spread of resistant bacteria to other patients at risk, which has been comprehensively demonstrated in hospital-associated transmission events [68]. Thus, the application of mNGS in infectious disease testing holds great promise in addressing existing testing regime short comings.

In this study, we demonstrated the successful application of mNGS for routine infectious disease testing in faecal samples, in which the performance for detection of commonly encountered GI pathogens was equivalent to standard diagnostic testing but with improved resolution and inferred functional potential of antimicrobial resistance and virulence factors to guide clinical treatment. mNGS was able to detect pathogens in 137 of 152 samples with clinical sensitivity ranging between 88.5 to 100% and specificity ranging from 99.3 to 100% for the four target genera evaluated. This was achieved despite differences in diagnostic performance of testing methodologies and inclusion of cases only diagnosed with one test method, for example MCS. At least 14 of 16 additional organisms identified by mNGS are likely true positive detections which merit further investigation. As with any diagnostic assay, performance at the limit of detection is prone to unreliable reporting. This applied to 16 of the samples in the present study for which there was evidence of informative sequencing reads below the set target threshold or positive identification of closely related species to the PCR target organism being observed.

The four most consistently detected pathogens, by mNGS and PCR, are consistent with published studies globally, where *Campylobacter*, *Salmonella*, *Aeromonas* and *Giardia* organisms are commonly encountered pathogens [69–73]. This makes the selected targets used in the current study broadly applicable and translatable to global diagnostic testing. In total, mNGS identified 314 infectious agents compared to 209 with conventional testing alone. Four percent of samples were diagnosed solely on the basis of mNGS, in which testing had not been considered by the treating clinicians for the patient testing or the testing format did not include the reported pathogen. High host DNA background, which has been noted to be an issue for sensitivity in mNGS testing [74], was only seen in 2 samples related to active *C.difficile* infections. These findings suggest that gastrointestinal infections remain undiagnosed or misdiagnosed in a proportion of patients using conventional testing and demonstrate the potential usefulness of clinical mNGS testing. Additionally, eight of the samples testing positive for *Campylobacter* by PCR and MCS, were classified as the debatably pathogenic *C. hominis* or *C. conscisus* by mNGS [43]. This demonstrates the importance of using diagnostic testing that can report at the species or strain level.

While mNGS holds great promise for analyzing complex microbial populations in faecal samples, there are additional challenges that warrant consideration, which have been addressed in this study. Current limitations for the diagnostic use of mNGS broadly include technical restrictions, costs, standardisation, the complexity of bioinformatic analysis and turn-around-time (TAT). Technically, the use of any genomic technology cannot link to live organisms and is therefore limited in inferring clinical relevance. Furthermore, the absence of standardized workflows for sample preparation and data analysis can impact reproducibility across studies. Addressing these challenges is critical for maximizing the utility of mNGS in clinical and research applications.

With recent advances in sequencing technology, the costs associated with sample processing and sequencing have steadily declined, making this technology feasible to implement [75]. This is likely to continue, and the current cost of mNGS testing may be justifiable if it can make inroads into the ~ 60% of potential infections that remain undiagnosed using conventional testing [76]. As most countries do not have approved methods, instruments and/or databases for diagnostic mNGS, they are currently provided as in-house laboratory developed tests. These require extensive validation and demonstration of performance metrics before being able to be offered for clinical testing [76]. Encouragingly, networks are being established that are standardising both laboratory and bioinformatic procedures, required quality control processes and proficiency testing programs [77]. However, there is a scarcity of bioinformatic software capable for clinical diagnostic use. This is exacerbated by concerns over the quality and comprehensiveness of currently available reference databases. Notably, viral species are underrepresented [78] and reference genomes can harbour contamination with exogenous DNA [79] leading to high risk of erroneous alignments and false positive diagnoses. Although several open source metagenomic analysis pipelines are available, most have issues with false positive rates, precision and recall [37, 80, 81], making them unsuitable for formally accredited clinical application.

Accurate high-resolution classification, removal of host sequence data, and appropriate report visualisations are just some of the considerations required to make clinically robust software [82], for example FDA-ARGO [83], FDA RVDB [84]. The availability of proprietary clinical reporting services such as those described here, are therefore important, as they have been optimised for certified clinical diagnostics [36]. Use of such services also addresses previous concerns that TAT for mNGS assays is not within clinically actionable timeframes. Currently TAT, from specimen receipt to result report issue, has averaged 48 h for most mNGS platforms [10], with newly emerging sequencing technologies promising further reductions of TAT to as little as 6 h [85]. Given established time frames for conventional culture-based methods at 3 days for bacteria (but up to 45 days for some types such as mycobacteria) and 7 days for fungi, a 2 day TAT with mNGS is a clinically acceptable period [17]. The challenge for the field will be to achieve these outcomes in a commercially viable format.

Our mNGS approach did not allow for detection of all viral pathogens that had been found by conventional diagnostic testing, as most of the viral pathogens were RNA viruses, which are missed in standard DNA based metagenomics. However, the mNGS approach could be expanded to include RNA purification from stool samples that would benefit from the development of specific databases for virus diagnostics. Here we have validated the method for a small number of targets, however, the mNGS method can be greatly expanded to include more pathogens and enable much broader hypothesis free syndromic testing.

Conclusion

In summary, we demonstrate that mNGS is a promising approach for the certified diagnosis of infectious disease provided that laboratory and bioinformatic processes are standardised benchmarked and of sufficient diagnostic performance. Here we demonstrate this with a modest number of well-known clinical targets; however, mNGS has the potential to be scaled to hundreds of targets in a single validated assay, which would have greater clinical utility. This would justify the widespread adoption of mNGS by the medical community, particularly for gastrointestinal infections. Ultimately rapid and accurate identification of pathogens enables tailored treatments, reduces the abuse of broad-spectrum antibiotics and assists in patient recovery [3]. The emerging consensus amongst the medical fraternity is that metagenomics will play a role in clinical microbiology laboratories in the near future, in particular for hard to diagnose samples. Ultimately, we expect metagenomics to be an important method in gastrointestinal infectious disease testing.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13099-024-00673-1.

Supplementary materials 1: Compiled MHC sample set. Raw data from study collating the original laboratory result, confirmatory PCR test result, and mNGS test result

Supplementary materials 2: Specificity testing. Outcome from testing of closely related species in the mNGS test

Acknowledgements
The authors note the contribution of Quinn Cao for the review and assistance in creating publication quality illustrations.

Author contributions
N.A. prepared the initial draft and subsequent edits, literature review, and illustration designs; M.S., A.A. and D. P. analyzed and interpreted the patient data. L.F. and S.M. designed and performed the required laboratory-based sample processing; B.W. provided project management; P.H. and G.T. were major contributors in writing the manuscript; D.W. was a major contributor to the development, management and drafting relating to the clinical bioinformatic analysis performed; P.G. was a contributor to the study design, clinical review process and provided guidance; All authors read, edited and approved the final manuscript.

Funding
This study was funded by Microba Pty Ltd, a privately-owned company based in Australia.

Availability of data and materials
Metagenomic reads from patient samples in this study were depleted of human host sequences and have been submitted to NCBI under BioProject PRJNA1156595.

Declarations

Ethics approval and consent to participate
All participants provided informed consent within the approved research project by the national ethics committee of Bellberry Limited (Project No.: 2018–05-400-A-2) and local ethics committee, Mater Research Clinical Governance Office (Project 46934 Ref AM/MML/46934 v4).

Consent for publication
Not applicable.

Competing interests
All authors are shareholders and/or receive a salary or reimbursement from Microba Pty Ltd. It is noted that the reported outcomes in the paper are directly from validation of a metagenomics assay for the detection of infectious gastrointestinal organisms, which has been independently reviewed and assessed by an external certification body to ISO15189, as being true and ubiased account of the research performed (Microba Laboratories Accreditation No. 20563).

Author details
[1]Microba Pty Ltd, Brisbane, Australia. [2]Centre for Microbiome Research, School of Biomedical Sciences, Queensland University of Technology (QUT), Translational Research Institute, Woolloongabba, QLD, Australia. [3]Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, University of Queensland, St Lucia, QLD, Australia. [4]Department of Medicine, The University of Queensland, Brisbane, QLD, Australia. [5]Mater Research Raymond Terrace, South Brisbane, Australia.

Received: 20 October 2024 Accepted: 24 December 2024
Published online: 18 January 2025

References

1. Wilber E, Baker JM, Rebolledo PA. Clinical implications of multiplex pathogen panels for the diagnosis of acute viral gastroenteritis. J Clin Microbiol. 2021. https://doi.org/10.1128/JCM.01513-19.
2. Ternhag A, Törner A, Svensson Å, Ekdahl K, Giesecke J. Short- and long-term effects of bacterial gastrointestinal infections. Emerg Infect Dis. 2008;14(1):143–8.
3. Brendish NJ, Beard KR, Malachira AK, Tanner AR, Sanga-Nyirongo L, Gwiggner M, et al. Clinical impact of syndromic molecular point-of-care testing for gastrointestinal pathogens in adults hospitalised with suspected gastroenteritis (GastroPOC): a pragmatic, open-label, randomised controlled trial. Lancet Infect Dis. 2023;23(8):945–55.
4. Axelrad JE, Freedberg DE, Whittier S, Greendyke W, Lebwohl B, Green DA. Impact of gastrointestinal panel implementation on health care utilization and outcomes. J Clin Microbiol. 2019;57(3):e01775-e1818.
5. Ozyurt OK, Saglik I, Ozhak B, Mutlu D, Levent B, Donmez L, et al. Evaluation of Diagnostic Performance of BD Max EBP Assay in Patients with Diarrheal Illness. In Review; 2021. https://www.researchsquare.com/article/rs-258584/v1. Accessed 6 Jan 2022.
6. Boers SA, Peters CJA, Wessels E, Melchers WJG, Claas ECJ. Performance of the QIAstat-Dx gastrointestinal panel for diagnosing infectious gastroenteritis. J Clin Microbiol. 2020;58(3):e01737-e1819.
7. Teh R, Tee WD, Tan E, Fan K, Koh CJ, Tambyah PA, et al. Review of the role of gastrointestinal multiplex polymerase chain reaction in the management of diarrheal illness. J Gastroenterol Hepatol. 2021;36(12):3286–97.
8. Joensen KG, Engsbro ALØ, Lukjancenko O, Kaas RS, Lund O, Westh H, et al. Evaluating next-generation sequencing for direct clinical diagnostics in diarrhoeal disease. Eur J Clin Microbiol Infect Dis. 2017;36(7):1325–38.
9. Maldonado-Garza HJ, Garza-González E, Bocanegra-Ibarias P, Flores-Treviño S. Diagnostic syndromic multiplex approaches for gastrointestinal infections. Expert Rev Gastroenterol Hepatol. 2021;15(7):743–57.
10. Miao Q, Ma Y, Wang Q, Pan J, Zhang Y, Jin W, et al. Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice. Clin Infect Dis. 2018;67:231–40.
11. Batool M, Galloway-Peña J. Clinical metagenomics—challenges and future prospects. Front Microbiol. 2023;28(14):1186424.
12. Aldars-García L, Chaparro M, Gisbert JP. Systematic review: the gut microbiome and its potential clinical application in inflammatory bowel disease. Microorganisms. 2021;9(5):977.
13. Metwaly A, Reitmeier S, Haller D. Microbiome risk profiles as biomarkers for inflammatory and metabolic disorders. Nat Rev Gastroenterol Hepatol. 2022;19(6):383–97.
14. De Filippis F, Paparo L, Nocerino R, Della Gatta G, Carucci L, Russo R, et al. Specific gut microbiome signatures and the associated pro-inflamatory functions are linked to pediatric allergy and acquisition of immune tolerance. Nat Commun. 2021;12(1):5958.
15. Wallen ZD, Demirkan A, Twa G, Cohen G, Dean MN, Standaert DG, et al. Metagenomics of Parkinson's disease implicates the gut microbiome in multiple disease mechanisms. Nat Commun. 2022;13(1):6958.
16. Maljkovic Berry I, Melendrez MC, Bishop-Lilly KA, Rutvisuttinunt W, Pollett S, Talundzic E, et al. Next generation sequencing and bioinformatics methodologies for infectious disease research and public health: approaches, applications, and considerations for development of laboratory capacity. J Infect Dis. 2020;221:292–307.
17. Schuele L, Cassidy H, Peker N, Rossen JWA, Couto N. Future potential of metagenomics in microbiology laboratories. Expert Rev Mol Diagn. 2021;21(12):1273–85.
18. Wilson MR, Naccache SN, Samayoa E, Biagtan M, Bashir H, Yu G, et al. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. N Engl J Med. 2014;370(25):2408–17.
19. Couto N, Rossen JW. Overview of Microbial NGS for Clinical and Public Health Microbiology. In: Moran-Gilad J, Yagel Y, editors. Application and Integration of Omics-powered Diagnostics in Clinical and Public Health Microbiology. Cham: Springer International Publishing; 2021. p. 9–24.
20. Deurenberg RH, Bathoorn E, Chlebowicz MA, Couto N, Ferdous M, García-Cobos S, et al. Application of next generation sequencing in clinical microbiology and infection prevention. J Biotechnol. 2017;243:16–24.
21. Loman NJ, Constantinidou C, Christner M, Rohde H, Chan JZM, Quick J, et al. A Culture-independent sequence-based metagenomics approach to the investigation of an outbreak of shiga-toxigenic escherichia coli o104:h4outbreak of shiga-toxigenic escherichia coli. JAMA. 2013;309(14):1502–10.
22. Williams A, Webster WZ, Cai C, Milgrom A, Al-Hasan M, Bookstaver PB. Evaluation of the diagnostic utility of metagenomic next-generation sequencing testing for pathogen identification in infected hosts: a retrospective cohort study. Ther Adv Infect Dis. 2024;11:20499361241232856.
23. Mu S, Hu L, Zhang Y, Liu Y, Cui X, Zou X, et al. Prospective evaluation of a rapid clinical metagenomics test for bacterial pneumonia. Front Cell Infect Microbiol. 2021;19(11): 684965.
24. Miller S, Naccache SN, Samayoa E, Messacar K, Arevalo S, Federman S, et al. Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. Genome Res. 2019;29(5):831–42.
25. Blauwkamp TA, Thair S, Rosen MJ, Blair L, Lindner MS, Vilfan ID, et al. Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. Nat Microbiol. 2019;4(4):663–74.
26. Diao Z, Han D, Zhang R, Li J. Metagenomics next-generation sequencing tests take the stage in the diagnosis of lower respiratory tract infections. J Adv Res. 2021 https://www.sciencedirect.com/science/article/pii/S2090123221001934. Accessed 7 Jan 2022.
27. Stojanovic Z, Gonçalves-Carvalho F, Marín A, Abad Capa J, Domínguez J, Latorre I, et al. Advances in diagnostic tools for respiratory tract infections: from tuberculosis to COVID-19—changing paradigms? ERJ Open Res. 2022;8(3):00113–2022.
28. Cummings LA, Hoogestraat DR, Rassoulian-Barrett SL, Rosenthal CA, Salipante SJ, Cookson BT, et al. Comprehensive evaluation of complex polymicrobial specimens using next generation sequencing and standard microbiological culture. Sci Rep. 2020;10(1):5446.
29. Kujiraoka M, Kuroda M, Asai K, Sekizuka T, Kato K, Watanabe M, et al. Comprehensive diagnosis of bacterial infection associated with acute cholecystitis using metagenomic approach. Front Microbiol. 2017;20(8):685.
30. Imwattana K, Knight DR, Riley TV. Can sequencing improve the diagnosis and management of Clostridioides difficile infection? Expert Rev Mol Diagn. 2021;21(5):429–31.
31. Hughes ER, Winter MG, Duerkop BA, Spiga L, Furtado de Carvalho T, Zhu W, et al. Microbial respiration and formate oxidation as metabolic signatures of inflammation-associated dysbiosis. Cell Host Microbe. 2017;21(2):208–19.
32. Liu H, Chen X, Hu X, Niu H, Tian R, Wang H, et al. Alterations in the gut microbiome and metabolism with coronary artery disease severity. Microbiome. 2019;7(1):68.
33. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.
34. Bushnell B, Rood J, Singer E. BBMerge—Accurate paired shotgun read merging via overlap. PLoS ONE. 2017;12(10): e0185056.
35. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De novo assembler. Curr Protoc Bioinforma. 2020;70(1): e102.
36. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25(7):1043–55.
37. Parks DH, Rigato F, Vera-Wolf P, Krause L, Hugenholtz P, Tyson GW, et al. Evaluation of the microba community profiler for taxonomic profiling of metagenomic datasets from the human gut microbiome. Front Microbiol. 2021;20(12): 643682.
38. Vandeputte D, Tito RY, Vanleeuwen R, Falony G, Raes J. Practical considerations for large-scale gut microbiome studies. FEMS Microbiol Rev. 2017;41(1):154–67.
39. Lowe AM, Yansouni CP, Behr MA. Causality and gastrointestinal infections: Koch, Hill, and Crohn's. Lancet Infect Dis. 2008;8(11):720–6.
40. Subbaram K, Gatasheh MK, Al Azzam KM, Kannan H. Molecular identification of diarrheal Aeromonas using immuno magnetic polymerase chain reaction (IM-PCR) technique: a comparative study with conventional culture method. Afr Health Sci. 2019;19(2):2036.
41. Amjad M. An overview of the molecular methods in the diagnosis of gastrointestinal infectious diseases. Int J Microbiol. 2020;24(2020):1–13.

42. Bertran X, Rubio M, Gómez L, Llovet T, Muñoz C, Navarro F, et al. Taxonomic identification of different species of the genus aeromonas by whole-genome sequencing and use of their species-specific β-lactamases as phylogenetic markers. Antibiotics. 2021;10(4):354.
43. Liu F, Ma R, Wang Y, Zhang L. The clinical importance of campylobacter concisus and other human hosted campylobacter species. Front Cell Infect Microbiol. 2018. https://doi.org/10.3389/fcimb.2018.00243.
44. Lawson AJ, On SL, Logan JM, Stanley J. Campylobacter hominis sp. nov., from the human gastrointestinal tract. Int J Syst Evol Microbiol. 2001;51(2):651–60.
45. MacNeil KM, Dodge MJ, Evans AM, Tessier TM, Weinberg JB, Mymryk JS. Adenoviruses in medicine: innocuous pathogen, predator, or partner. Trends Mol Med. 2023;29(1):4–19.
46. Beal SG, Tremblay EE, Toffel S, Velez L, Rand KH. A gastrointestinal PCR panel improves clinical management and lowers health care costs. J Clin Microbiol. 2018;56(1):e01457-e1517.
47. Makimaa H, Ingle H, Baldridge MT. Enteric viral co-infections: pathogenesis and perspective. Viruses. 2020;12(8):904.
48. Wensel CR, Pluznick JL, Salzberg SL, Sears CL. Next-generation sequencing: insights to advance clinical investigations of the microbiome. J Clin Invest. 2022;132(7): e154944.
49. Kok NA, Peker N, Schuele L, De Beer JL, Rossen JWA, Sinha B, et al. Host DNA depletion can increase the sensitivity of Mycobacterium spp. detection through shotgun metagenomics in sputum. Front Microbiol. 2022;13:949328.
50. Islam Sajib MS, Brunker K, Oravcova K, Everest P, Murphy ME, Forde T. Advances in host depletion and pathogen enrichment methods for rapid sequencing-based diagnosis of bloodstream infection. J Mol Diagn. 2024;26(9):741–53.
51. Hasan MR, Rawat A, Tang P, Jithesh PV, Thomas E, Tan R, et al. Depletion of human DNA in spiked clinical specimens for improvement of sensitivity of pathogen detection by next-generation sequencing. J Clin Microbiol. 2016;54(4):919–27.
52. Malik YS, Verma AK, Kumar N, Touil N, Karthik K, Tiwari R, et al. Advances in diagnostic approaches for viral etiologies of diarrhea: from the lab to the field. Front Microbiol. 2019;13(10):1957.
53. Lefterova MI, Suarez CJ, Banaei N, Pinsky BA. Next-generation sequencing for infectious disease diagnosis and management: a report of the association for molecular pathology. J Mol Diagn. 2015;17(6):623–34.
54. Schluckebier L, Caetano R, Garay OU, Montenegro GT, Custodio M, Aran V, et al. Cost-effectiveness analysis comparing companion diagnostic tests for EGFR, ALK, and ROS1 versus next-generation sequencing (NGS) in advanced adenocarcinoma lung cancer patients. BMC Cancer. 2020;20(1):875.
55. Mintzer V, Moran-Gilad J, Simon-Tuval T. Operational models and criteria for incorporating microbial whole genome sequencing in hospital microbiology—a systematic literature review. Clin Microbiol Infect. 2019;25(9):1086–95.
56. Ransom EM, Potter RF, Dantas G, Burnham CAD. Genomic prediction of antimicrobial resistance: ready or not, here it comes! Clin Chem. 2020;66(10):1278–89.
57. Forslund K, Sunagawa S, Kultima JR, Mende DR, Arumugam M, Typas A, et al. Country-specific antibiotic use practices impact the human gut resistome. Genome Res. 2013;23(7):1163–9.
58. Mac Aogáin M, Narayana JK, Tiew PY, Ali NABM, Yong VFL, Jaggi TK, et al. Integrative microbiomics in bronchiectasis exacerbations. Nat Med. 2021;27(4):688–99.
59. Stecher B, Denzler R, Maier L, Bernet F, Sanders MJ, Pickard DJ, et al. Gut inflammation can boost horizontal gene transfer between pathogenic and commensal *Enterobacteriaceae*. Proc Natl Acad Sci. 2012;109(4):1269–74.
60. Karami N, Martner A, Enne VI, Swerkersson S, Adlerberth I, Wold AE. Transfer of an ampicillin resistance gene between two Escherichia coli strains in the bowel microbiota of an infant treated with antibiotics. J Antimicrob Chemother. 2007;60(5):1142–5.
61. Lumbreras-Iglesias P, Sabater C, Fernández Moreno A, López De Ugarriza P, Fernández-Verdugo A, Margolles A, et al. Evaluation of a shotgun metagenomics approach for detection of ESBL- and/or carbapenemase-producing enterobacterales in culture negative patients recovered from acute leukemia. Microorganisms. 2023;11(2):402.
62. Gudiol C, Tubau F, Calatayud L, Garcia-Vidal C, Cisnal M, Sanchez-Ortega I, et al. Bacteraemia due to multidrug-resistant Gram-negative bacilli in cancer patients: risk factors, antibiotic therapy and outcomes. J Antimicrob Chemother. 2011;66(3):657–63.
63. Geladari A, Karampatakis T, Antachopoulos C, Iosifidis E, Tsiatsiou O, Politi L, et al. Epidemiological surveillance of multidrug-resistant gram-negative bacteria in a solid organ transplantation department. Transpl Infect Dis. 2017;19(3): e12686.
64. Anjum MF, Zankari E, Hasman H. Molecular methods for detection of antimicrobial resistance. Microbiol Spectr. 2017. https://doi.org/10.1128/microbiolspec.ARBA-0011-2017.
65. Tacconelli E, Cataldo MA, Dancer SJ, De Angelis G, Falcone M, Frank U, et al. ESCMID guidelines for the management of the infection control measures to reduce transmission of multidrug-resistant Gram-negative bacteria in hospitalized patients. Clin Microbiol Infect. 2014;20:1–55.
66. Kirkup B. Culture-independence for surveillance and epidemiology. Pathogens. 2013;2(3):556–70.
67. Garcia-Vidal C, Cardozo-Espinola C, Puerta-Alcalde P, Marco F, Tellez A, Agüero D, et al. Risk factors for mortality in patients with acute leukemia and bloodstream infections in the era of multiresistance. PLoS ONE. 2018;13(6): e0199531.
68. David S, Reuter S, Harris SR, Glasner C, Feltwell T, Argimon S, et al. Epidemic of carbapenem-resistant Klebsiella pneumoniae in Europe is driven by nosocomial spread. Nat Microbiol. 2019;4(11):1919–29.
69. Ruzante JM, Olin K, Munoz B, Nawrocki J, Selvarangan R, Meyers L. Real-time gastrointestinal infection surveillance through a cloud-based network of clinical laboratories. PLoS ONE. 2021;16(4): e0250767.
70. Costa D, Iraola G. Pathogenomics of emerging campylobacter species. Clin Microbiol Rev. 2019. https://doi.org/10.1128/CMR.00072-18.
71. Yuwono C, Wehrhahn MC, Liu F, Riordan SM, Zhang L. The isolation of aeromonas species and other common enteric bacterial pathogens from patients with gastroenteritis in an australian population. Microorganisms. 2021;9(7):1440.
72. Elorza A, Rodríguez-Lago I, Martínez P, Hidalgo A, Aguirre U, Cabriada JL. Infección gastrointestinal por Aeromonas: incidencia y su posible relación con la enfermedad inflamatoria intestinal. Gastroenterol Hepatol. 2020;43(10):614–9.
73. Burd EM, Hinrichs BH. Gastrointestinal infections. In: Leonard DGB, editor. Molecular Pathology in Clinical Practice. Cham: Springer International Publishing; 2016.
74. Cheng WY, Liu WX, Ding Y, Wang G, Shi Y, Chu ESH, et al. High sensitivity of shotgun metagenomic sequencing in colon tissue biopsy by Host DNA depletion. Genomics Proteomics Bioinformatics. 2023;21(6):1195–205.
75. Nafea AM, Wang Y, Wang D, Salama AM, Aziz MA, Xu S, et al. Application of next-generation sequencing to identify different pathogens. Front Microbiol. 2024;29(14):1329330.
76. Simner PJ, Miller S, Carroll KC. Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. Clin Infect Dis. 2018;66(5):778–88.
77. López-Labrador FX, Brown JR, Fischer N, Harvala H, Van Boheemen S, Cinek O, et al. Recommendations for the introduction of metagenomic high-throughput sequencing in clinical virology, part I: Wet lab procedure. J Clin Virol. 2021;134: 104691.
78. Miller S, Chiu C, Rodino KG, Miller MB. Point-counterpoint: should we be performing metagenomic next-generation sequencing for infectious disease diagnosis in the clinical laboratory? J Clin Microbiol. 2019. https://doi.org/10.1128/JCM.01739-19.
79. Steinegger M, Salzberg SL. Terminating contamination: large-scale search identifies more than 2,000,000 contaminated entries in GenBank. Genome Biol. 2020;21(1):115.
80. Bharti R, Grimm DG. Current challenges and best-practice protocols for microbiome analysis. Brief Bioinform. 2021;22(1):178–93.
81. Chen Y, Fan L, Chai Y, Xu J. Advantages and challenges of metagenomic sequencing for the diagnosis of pulmonary infectious diseases. Clin Respir J. 2022;16(10):646–56.
82. Qian M, Zhan Y, Wu D, Ji L, Chen H, Cheng Y. Clinical standardization of metagenomic next generation sequencing (mNGS) in the pathogen diagnosis. Clin Transl Discov. 2021;1(1): e12.
83. Sichtig H, Minogue T, Yan Y, Stefan C, Hall A, Tallon L, et al. FDA-ARGOS is a database with public quality-controlled reference genomes for diagnostic use and regulatory science. Nat Commun. 2019;10(1):3313.

84. Goodacre N, Aljanahi A, Nandakumar S, Mikailov M, Khan AS. A reference viral database (RVDB) to enhance bioinformatics analysis of high-throughput sequencing for novel virus detection. MSphere. 2018;3(2):e00069-e118.
85. Charalampous T, Kay GL, Richardson H, Aydin A, Baldan R, Jeanes C, et al. Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. Nat Biotechnol. 2019;37(7):783–92.

Publisher's Note