

# Identification of pathogen(s) in infectious diseases using shotgun metagenomic sequencing and conventional culture: a comparative study

Huan Chen<sup>1</sup>, Jun Li<sup>2</sup>, Shanshan Yan<sup>3</sup>, Hui Sun<sup>4</sup>, Chuyi Tan<sup>4</sup>, Meidong Liu<sup>4</sup>, Ke Liu<sup>4</sup>, Huali Zhang<sup>4</sup>, Mingxiang Zou<sup>2</sup> and Xianzhong Xiao<sup>4</sup>

<sup>1</sup> Postdoctoral Research Station of Clinical Medicine & Department of Hematology, Third Xiangya Hospital, Central South University, Changsha, China

<sup>2</sup> Department of Clinical Laboratory, Xiangya Hospital, Central South University, Changsha, China

<sup>3</sup> Department of Intensive Medicine, Third Xiangya Hospital, Central South University, Changsha, China

<sup>4</sup> Sepsis Translational Medicine Key Lab of Hunan Province, Department of Pathophysiology, School of Basic Medicine Science, Central South University, Changsha, China

## ABSTRACT

**Background:** Early and accurate diagnosis of microorganism(s) is important to optimize antimicrobial therapy. Shotgun metagenomic sequencing technology, an unbiased and comprehensive method for pathogen identification, seems to potentially assist or even replace conventional microbiological methodology in the diagnosis of infectious diseases. However, evidence in clinical application of this platform is relatively limited.

**Methods:** To evaluate the capability of shotgun metagenomic sequencing technology in clinical practice, both shotgun metagenomic sequencing and conventional culture were performed in the PCR-positive body fluid specimens of 20 patients with suspected infection. The sequenced data were then analyzed for taxonomic identification of microbes and antibiotic resistance gene prediction using bioinformatics pipeline.

**Results:** Shotgun metagenomic sequencing results showed a concordance of 17/20 compared with culture results in bacterial detection, and a concordance of 20/20 compared with culture results in fungal detection. Besides, drug-resistant types annotated from antibiotic resistance genes showed much similarity with antibiotic classes identified by susceptibility tests, and more than half of the specimens had consistent drug types between shotgun metagenomic sequencing and culture results.

**Conclusions:** Pathogen identification and antibiotic resistance gene prediction by shotgun metagenomic sequencing identification had the potential to diagnose microorganisms in infectious diseases, and it was especially helpful for multiple microbial co-infections and for the cases where standard culture approached failed to identify microorganisms.

Submitted 10 February 2021

Accepted 8 June 2021

Published 29 June 2021

Corresponding authors

Huali Zhang, zhanghuali@csu.edu.cn

Mingxiang Zou,

zoumingxiang@126.com

Academic editor

Hector Mora-Montes

Additional Information and  
Declarations can be found on  
page 15

DOI 10.7717/peerj.11699

© Copyright

2021 Chen et al.

Distributed under

Creative Commons CC-BY 4.0

**OPEN ACCESS**

**Subjects** Microbiology, Molecular Biology, Mycology, Infectious Diseases

**Keywords** Shotgun metagenomic sequencing, Culture, Pathogen identification, Infectious disease, Antibiotic resistance

## INTRODUCTION

Infectious diseases remain a major causes of death throughout the world (*World Health Organization, 2020*), and pathogen identification is essential for its diagnosis and treatment (*Rhodes et al., 2017*). Currently, culture-based approaches are still considered the gold standard for diagnosis of bacterial or fungal infection (*Vincent et al., 2015*), and culture-associated drug sensitive test provides effective information for antimicrobial therapy. However, some microorganisms may not easy to identify by routinely culture methods used in clinical microbiology department under the influence of many factors, such as non-viable bacteria, fastidious growth requirements, inhibition by other organisms, etc. (*Grumaz et al., 2016; Gu et al., 2021*). These factors make accurate diagnosis and treatment of infections a challenge, which might lead to unnecessary use of broad-spectrum antibiotics and increase the risk of microbial resistance and expense of hospitalization.

Shotgun metagenomic sequencing approach, as an unbiased, sensitive and high-throughput approach, has become more attractive to clinical microbiology laboratories for its fast detection of possible pathogens directly in a clinical sample without nucleonic acids enrichment or prior region selection and is especially suitable for unknown, atypical, and emerging etiologies of complex infectious diseases (*Miao et al., 2018; Thoendel et al., 2018; Zhou et al., 2019*). In addition, antibiotic resistance genes (ARGs) prediction by shotgun metagenomic sequencing shows the potential to determine appropriate antimicrobial therapy and guide clinical management (*Crofts, Gasparrini & Dantas, 2017; Wang et al., 2017*). Furthermore, shotgun metagenomic sequencing seems to be less influenced by prior application of antibiotics (*Miao et al., 2018; Zhang et al., 2019*). In view of these giant advantages, shotgun metagenomic sequencing technology might probably become a routine method in the near future, partly taking the place of conventional culture (*Goldberg et al., 2015*). However, literature relevant to clinical applications has mostly focused on its partial function of rapid detection in certain infectious diseases (e.g., central nervous system infections, respiratory infections, bone and joint infections, etc. (*Fang et al., 2020; Huang et al., 2020; Wilson et al., 2019; Yan et al., 2020; Zhang et al., 2020*)). On the other hand, to comprehensively interpret and evaluate the full-function of shotgun metagenomic sequencing (including pathogen identification and ARGs prediction) remains challenging, which needs more investigation and validation. For these purposes, the present study has been performed to expand shotgun metagenomic sequencing detection in real-world clinical practice while completely estimating its performance in both detecting causative microbes and providing resistance information.

## MATERIALS & METHODS

### Ethics statement

The Ethics Committee of the Xiangya Hospital of Central South University approved the study (No. 201703567), and written informed consents were obtained from all patients. All procedures were carried out in accordance with the Declaration of Helsinki and all experiments were performed by the relevant guidelines and the institutional regulations.

### **Patient recruitment**

Patients (18 years or older), admitted to hospital, with acute febrile illness were recruited at the Xiangya Hospital in Changsha, China, between March 2017 and May 2017. Patients were considered for inclusion with a suspected diagnosis of bacterial or fungal infection. Infection was defined as an axillary temperature of more than 38 °C or less than 35 °C; increased or decreased white blood cell count (>12 or <4 cells/mL<sup>3</sup>); and/or clinical signs associated with local or systemic infection, such as abscess, respiratory distress, prostration, altered consciousness, convulsions, clinical jaundice, signs of shock, etc.; and/or other infection associated radiographic results. Using our inclusion criteria, 89 patients were included to collect samples to perform bacterial/fungal PCR amplification tests to get candidate samples for the shotgun metagenomic sequencing and routine culture in a pairwise manner. After primarily screening and exclusion, specimens from 20 patients were finally subjected to the follow-up testing and enrolled into the comparative study between shotgun metagenomic sequencing and traditional methodology (see [Figs. 1A](#) and [2](#)). Detailed information of participated patients was collected and listed in [Table 1](#).

### **Sample collection, culture, and susceptibility testing**

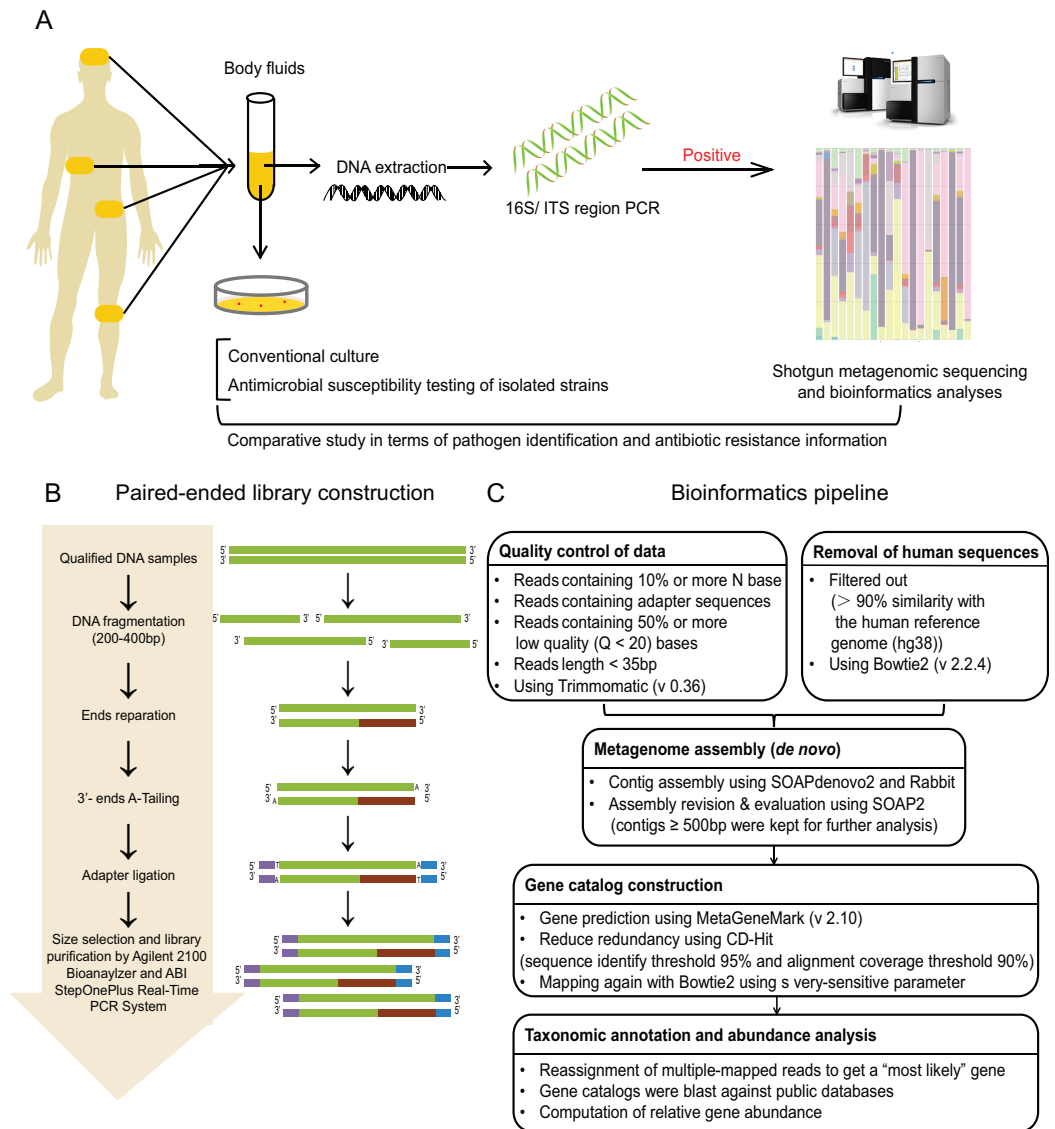
Sterile fluid specimens of suspected infection site were collected from the Microbiology department of the Xiangya Hospital within 48 h of infection, including cerebrospinal (CSF), pleural, peritoneal, or synovial fluid. Urine and sputum samples were not considered. The laboratory used standard aerobic culture methods for processing clinical specimens and they conformed to standard protocols published for clinical microbiology laboratories for detection of pathogens and interpretation of results ([Jorgensen et al., 2015](#)). Agar plates and BHI broth plates were used to culture samples of sterile sites, and incubated routinely for 5 days to observe bacterial growth. According to the Clinical and Laboratory Standards Institute (CLSI) guidelines ([CLSI, 2015](#)), antimicrobial susceptibility testing of isolated strains was carried out by Vitek 32 automated system (bioMérieux, Marcy-l'Étoile, France), and was identified by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) (Bruker Daltonik, GmbH, Berlin, Germany) with standard protocols.

### **DNA extraction**

The remaining specimens (that had been processed for culture) were stripped of patient identification and coded before experiments. The specimens were transferred to the laboratory for polymerase chain reaction (PCR). Sample DNA extraction was carried out using QIAamp DNA Mini kit (Qiagen, Hilden, Germany) according to manufacturer's instructions. The quality of DNA was analyzed using Qubit (Invitrogen, Carlsbad, CA, USA) and checked by 1% agarose gel electrophoresis. Extracted DNA was stored at -80 °C.

### **PCR amplification**

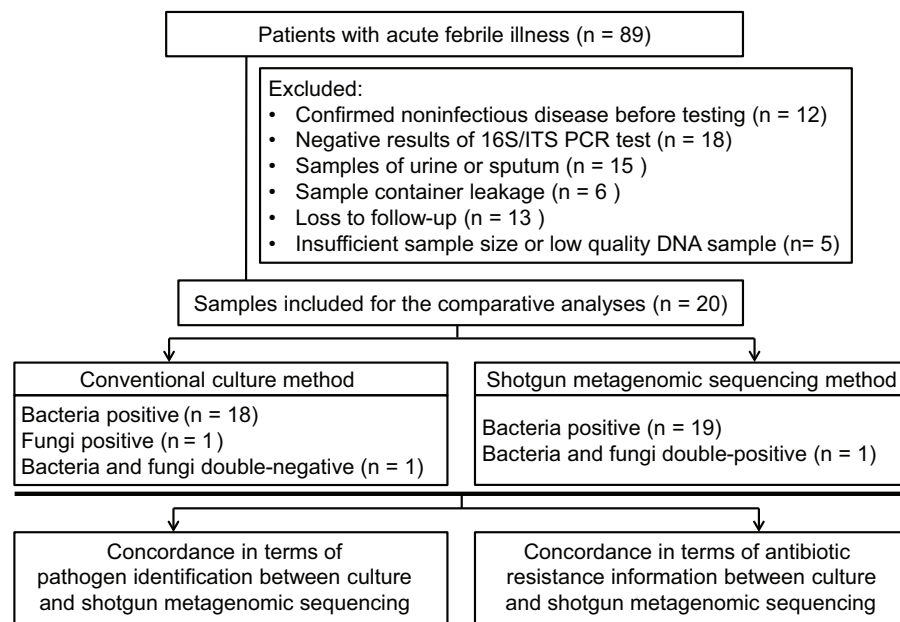
The presence of bacterial or fungal DNA (or both) was confirmed by using PCR amplification with the common primers of bacterial 16S gene region and fungal ITS1/4



**Figure 1 Study workflow.** (A) Schematic of comparative study workflow. Patients' samples of body fluids were collected for conventional culture and 16S/ITS region PCR test. PCR-positive specimens which passed the quality control of sequencing were selected to perform shotgun metagenomic sequencing and bioinformatics analyses, and further comparative study between culture results and sequencing results. (B) Schematic of paired-end library construction in accordance with Illumina's instruction. (C) Bioinformatics pipeline for shotgun metagenomic sequencing.

Full-size DOI: [10.7717/peerj.11699/fig-1](https://doi.org/10.7717/peerj.11699/fig-1)

gene region. To control the contamination, extracted DNA was tested in batches with positive and negative controls for bacteria and fungi, respectively. PCR products were run on a 2% agarose gel and visualized using nucleic acid dyestuffs. Strains of *E. coli* (ATCC: 25922) and *C. albicans* (ATCC: 10231) were used as positive controls (Abayasekara et al., 2017). Only specimens with positive PCR results were selected for and sequencing.



**Figure 2** Flowchart for enrollment.

Full-size DOI: 10.7717/peerj.11699/fig-2

**Table 1** Demographics of patients and sample characteristics.

Variables (n = 20)	Value
Patient demographics	
Age (years), median (IQR)	57 (44–62)
Gender, Male, n (%)	12 (60)
Days hospitalized, median (IQR)	17 (12–27)
Immunocompromised, n (%)	6 (30)
Empirical antibiotics use before sampling, n (%)	18 (90)
Temperature (°C), median (IQR)	38.5 (38–38.8)
WBC count (10 <sup>9</sup> /L), median (IQR)	15.4 (12.8–18.4)
Neutrophils (%), median (IQR)	89.15 (82.93–94.90)
Procalcitonin (ng/mL), median (IQR)	1.87 (0.49–55.92)
SOFA score, median (IQR)	2 (0–5)
Sample features	
Sample type, n (%)	
Abscess	5 (25)
Cerebrospinal fluid	5 (25)
Bile	4 (20)
Others*	6 (30)
Positive culture rate, n (%)	
Bacteria positive	18 (90)
Fungi positive	1 (5)
Time to final culture result (days), median (IQR)	4 (3–4)

**Note:**

\* Including ascites, bone marrow, joint aspirate (knee), and pleural fluid.

PCR amplification was performed as mentioned below: 95 °C for 2 min, 30 cycles at 95 °C for 30 s, 57 °C for 30 s and 72 °C for 60 s, and one cycle at 72 °C for 10 min for 16S amplification; and 94 °C for 5 min, 30 cycles at 94 °C for 30 s, 55 °C for 30 s and 72 °C for 60 s, and one cycle at 72 °C for 7 min for ITS amplification, respectively (*Abayasekara et al., 2017*). Primers were shown as follows: 16S rDNA F: 5'-ACTCC TACGGGAGGCAGCAGT-3', R: 5'-TGACGGGCGGTGTGTACAAG-3'; ITS1: 5'-TCC GTAGGTGAACCTGCGG-3', ITS4:5'-TCCTCCGCTTATTGATATGC-3' (Sangon Biotech, Shanghai, China).

## Sequencing and bioinformatics analyses

The workflow of sequencing and bioinformatics analyses was performed as follows (see [Figs. 2B](#) and [2C](#)), including library construction, sequencing, quality control and host sequence removal, metagenome assembly, gene prediction and taxonomic classification, calculation of relative gene abundance, and prediction of ARGs.

### Library construction

Library construction was performed using Illumina library preparation kit according to the manufacture's instruction (Illumina, San Diego, CA, USA) (*Liu et al., 2017*). To be brief, qualified DNA samples were sheared into small fragments (about 200–400 bp) by Covaris. After ends reparation, 3'-ends-adenylation, and adapter ligation, adequate size of DNA fragments were selected by Ampure beads. Then, Agilent 2100 Bioanalyzer and ABI StepOnePlus Real-Time PCR System were used to qualify and quantify the sample libraries. The qualified libraries were sequenced via Illumina platform (BGI, Shenzhen, China).

### Sequencing process

All 20 samples were sequenced on Illumina HiSeq 2500 platform (paired end; insert size, 350 bp; read length, 150 bp) via Beijing Genomics Institute in accordance with the manufacturer's instructions (BGI, Shenzhen, China). To control the contamination, a negative control with only reagents was added in each run. All the raw reads were deposited in the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>) with accession number [PRJNA527188](#).

### Quality control and removal of host sequences

In order to avoid contamination and acquire reliable results, unqualified reads and short reads were removed using Trimmomatic (version 0.36) (*Bolger, Lohse & Usadel, 2014*), including: (i) reads containing 10% or more N base; (ii) reads containing adapter sequences; (iii) reads containing 50% or more low quality (Q < 20) bases; (iv) reads length < 35 bp. Then, the host-related reads were filtered out which mapped to the human reference genome (hg19) with more than 90% similarity using Bowtie2 (version 2.2.4). Thus, high quality reads for these 20 samples were obtained for the next step.

### De novo metagenome assembly

The filtered reads were assembled de novo with SOAPdenovo2 (*Grabherr et al., 2011*), and further assembled with Rabbit (*You et al., 2013*) to obtain longer contigs. These longer

reads were mapped back to SOAP2 (Li *et al.*, 2009), and only contigs no less than 500 bp (with the optimal *k-mer*) were kept as clean data depending on both contig *N50* and mapping rate, and applied for further functional analysis.

### **Taxonomic classification**

We used MetaGeneMark (version 2.10) (Zhu, Lomsadze & Borodovsky, 2010) for gene prediction, and CD-Hit (Li & Godzik, 2006) for clustering genes in the whole cohort. The high quality genes after filtering were merged to generate gene catalog using Bowtie2 (on average of 10× coverage depth). Then, these selected genes were blasted against public databases, performed by the MEGAN (version 4.6) to classified these genes into four microbial genome databases (bacteria, fungi, viruses, and parasites) using LCA algorithm (Huson *et al.*, 2007). The classification reference databases were downloaded from the National Center for Biotechnology Information (NCBI; <https://ncbi.nlm.nih.gov/genomes/>). Notably, reads who mapping to multiple genes were reassigned to obtain a most possible gene by Pathoscope (version 1.0) (Francis *et al.*, 2013), which could resubmit the reads to a Bayesian framework to examine the global similarity of sequence.

### **Computation of relative gene abundance**

After classification, the relative abundance of any sample *S* was calculated by the formula below (Qin *et al.*, 2012):

$$a_i = \frac{b_i}{\sum_j b_j} = \frac{\frac{X_i}{L_i}}{\sum_j \frac{X_j}{L_j}}$$

where  $a_i$  represents the relative abundance of gene *i* in sample *S*,  $b_i$  represents the copy number of gene *i* in the sequenced data from sample *S*,  $L_i$  represents the length of gene *i*, and  $X_i$  represents the times which gene *i* can be identified in sample *S* (the number of mapped contigs).

In our sequence-based profiling analysis, only alignments with followed features could be accepted: (i) an entire of a pair-end read could be mapped onto a gene with the correct insert-size; (ii) one end of the pair-end read could be mapped onto the end of a gene, only if the other end of read was mapped outside the region of gene. Thus, the mapped read was regarded as one copy.

Then, the relative gene abundance of each level from the same taxonomy was added together, and the total relative abundance was regarded as the content of this taxonomy in a certain sample to construct the taxonomy profile of relative abundance on this sample.

### **Criteria for a positive result of shotgun metagenomic sequencing**

The results should be considered positive by the following steps. First, at least three copies were mapped to the pathogens whose relative abundance should exceed their own threshold set up by our preliminary data were included into the next. The threshold relative abundance of each pathogen was identified by shotgun metagenomic sequencing in samples of healthy volunteers with the highest abundance in genus level in the preliminary experiments (Table S1). Then, the identified pathogens ranked top five for



bacteria, virus, parasite and top ten for fungi in relative abundance were screened out as candidate pathogens for further consideration. After primary analysis, if the candidate microbes were commonly reported in the certain disease of that patient or in accordance with the patient's clinical manifestations, they would be considered as causative pathogens (with the help of experienced laboratory doctors).

### **Prediction and analysis of ARGs**

To identify potential in the genomic sequence of specimens, each specimen's genomic sequence with relative abundance higher than 1% was screened (Couto *et al.*, 2018) and aligned with the protein sequences of antibiotic resistance genes in the Comprehensive Antibiotic Resistance Database (CARD; <https://card.mcmaster.ca>) (Jia *et al.*, 2017).

### **Statistical analysis**

Statistical analyses were performed using SPSS software (version 19.0). For characteristics of patients, continuous data was described by median and interquartile range (*IQR*); categorical data was described by number (*n*) and percentage (%), and compared by *Chi-square* test or *Fisher's* exact test. A two-tailed *p*-value of < 0.05 was taken as a cut-off for statistical significance.

## **RESULTS**

### **Criteria for evaluation**

To compare easily, specimen was cultured (for detection of both bacterial and fungal growth) and extracted DNA (for PCR amplification) at the same time if it met the inclusion criteria. Specimens were selected to perform shotgun metagenomic sequencing if expected bands were seen in the PCR test (>1,000 bp for bacteria, and >750 bp for fungi, respectively). For the purposes of the validation, culture results of PCR-positive specimen were reserved as a standard to compare against shotgun metagenomic sequencing results. Contaminated specimen(s) were identified by the presence of bands in the negative controls or PCR blank, and were excluded from sequencing analysis. Specimens that did not pass the quality control before bioinformatics analysis were considered conflicts and ruled out from the final analysis. Discrepancies might be seen between culture-negative results and shotgun metagenomic sequencing results because of slow growing microorganisms, anaerobic bacteria, or other fastidious organisms.

The results were defined as "matches" when the culture isolates were identified in the shotgun metagenomic sequencing testing. If the culture-positive results were not detected in the shotgun metagenomic sequencing analysis or culture-negative results were detected in the shotgun metagenomic sequencing analysis, they were designated as "conflicts". The species tested in the shotgun metagenomic sequencing analysis were finally confirmed by a laboratory doctor when combined with clinical manifestations.

### **General characteristics**

From the 89 specimens of body fluids prospectively included for this study, 25 specimens showed positive in PCR reaction, but five of them were excluded from the sequencing workflow because of insufficient sample size, resulting in a final comparative analysis of



20 specimens (one specimen per patient; see [Fig. 2](#)). The demographic features of these patients are listed in [Table 1](#). The sample types included five abscess specimens, five CSF specimens, four bile specimens, two bone marrow specimens, two joint aspirate specimens, one ascites specimen and one pleural fluid specimen. Since Xiangya Hospital is a superior referral hospital, only two patients did not receive empirical antibiotic treatment before sampling.

### **Overview of comparative analysis**

Of the 20 PCR-positive specimens, one specimen was reported as no pathogen growth (both bacterial-negative and fungal-negative), one specimen was reported as fungal growth (fungal-positive but bacterial-negative), and 18 specimens were bacterial-positive in conventional culture methods. Meanwhile, all 20 specimens were detected bacterial infection in the shotgun metagenomic sequencing analysis, including one case of bacterial and fungal co-infection. Thus, the match rate between shotgun metagenomic sequencing and culture is 17/20 in bacterial detection, and 20/20 in fungal detection. The comparison of culture and shotgun metagenomic sequencing results is shown in [Fig. 3A](#), and a comprehensive list with culture and sequencing results of each specimen and the relative abundance of each specimen are available in [Tables S2](#) and [S3](#).

### **Analysis of culture positive results**

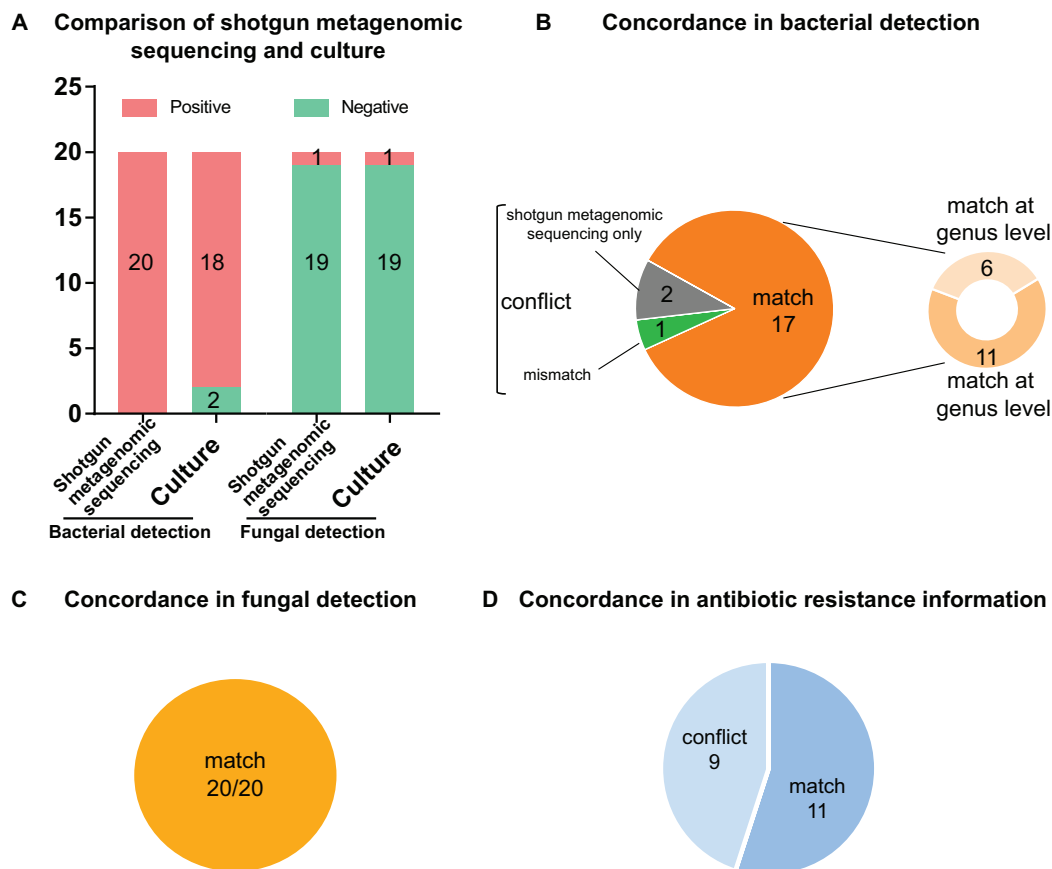
Of the 18 bacterial culture positive specimens, 6 specimens at genus level and 11 specimens at species level showed consistent pathogen results between culture and shotgun metagenomic sequencing analysis, making a match rate of 17/18 (see [Fig. 3B](#)). In addition, the fungal culture positive specimen was positive for both bacterial and fungal identification by shotgun metagenomic sequencing workflow, with a match of fungal result but conflict of bacterial result. Therefore, the match rates of fungal culture positive results were 1/1 (see [Fig. 3C](#)). Specific information of the conflicting results is shown in [Table 2](#). For conciseness, in spite of huge amounts of species identified from metagenomics analysis, only species of the highest abundance and/or clinical relevance are displayed.

### **Analysis of culture negative result**

The two bacterial culture negative specimens processed by shotgun metagenomic sequencing procedure were reported bacterial growth, which meant it conflicted with culture result (0/2) (see [Fig. 3B](#)). All 19 fungal culture negative specimens were not identified fungal growth in the shotgun metagenomic sequencing workflow (19/19) (see [Fig. 3C](#)). Details of these sequencing positive, culture negative specimens are revealed in [Table 2](#). Just species of the highest abundance and/or clinical relevance are displayed, although a large number of species are detected in the metagenomics process.

### **Comparative analysis of drug resistance information**

In order to analyze drug-resistance-related information, we compared ARGs obtained from sequencing data with clinical antimicrobial susceptibility tests. The ARGs of each specimen were aligned and annotated in the CARD database, and the top 10 genes with



**Figure 3** Comparison and concordance analysis between shotgun metagenomic sequencing and culture in pathogen detection and drug resistance information. Comparison and concordance analysis between shotgun metagenomic sequencing and culture in pathogen detection and drug resistance information. (A) The number of positive specimens (y-axis) for pairwise shotgun metagenomic sequencing and culture is plotted against bacterial detection and fungal detection (x-axis) ( $n = 20$ ). (B, C and D) Pie chart demonstrating the positivity distribution of shotgun metagenomic sequencing and culture for all specimens from bacterial detection (B), fungal detection (C), and antibiotic resistance information (D).  
Full-size [DOI: 10.7717/peerj.11699/fig-3](https://doi.org/10.7717/peerj.11699/fig-3)

**Table 2** The details of conflicting results between culture and shotgun metagenomic sequencing approach.

Specimen type	Culture results		Shotgun metagenomic sequencing results		
	Bacterial results	Fungal results	Species level of bacterial results	Genus level of bacterial results	Fungal results
Ascites (P5)	(-)	<i>Candida albicans</i>	<i>Pseudomonas aeruginosa</i> (19.104%)*, <i>Klebsiella pneumoniae</i> (9.090%)	<i>Pseudomonas</i> (39.559%)*	<i>Candida albicans</i> (78.152%)
Necrotic tissue (P13)	<i>Pseudomonas putida</i>	(-)	<i>Escherichia coli</i> (9.627%)*, <i>Acinetobacter baumannii</i> (6.225%)	<i>Acinetobacter</i> (14.583%)*	(-)
Necrotic tissue (P20)	(-)	(-)	<i>Acinetobacter baumannii</i> (7.263%)*	<i>Acinetobacter</i> (24.197%)*	(-)

**Notes:**

\* Species of highest abundance.

Species listed here are depended on species matching with highest abundance, and most clinical relevance.

relative abundance higher than 1% were screened and used for analysis. Since annotated information of genes was antibiotic types and drugs of susceptibility testing were more specific, antibiotic classes were used as the basis for evaluating the consistency of drug resistance.

If more than half of the antibiotic classes of susceptibility tests were the same as the annotated antibiotic classes of ARGs, the results were considered “consistent”; otherwise, the results were considered “conflict”.

As shown in Fig. 3D, 11 specimens showed consistent resistant antibiotic classes between sequencing results and susceptibility testing results, while nine specimens showed conflict between phenotype and genotype. Detailed information on the conflicting drug resistance conditions is shown in Table 3, and a comprehensive list with susceptibility test and sequencing prediction of each specimen is available in Table S4.

## DISCUSSION

Accurate and early identification of the pathogen driving infections is vitally important for guiding the diagnosis of disease, the decision of nursing management, and the application of proper antimicrobial therapy. Although culture-based approach regards as “gold standard” of microbes detection currently, it remains to be limited by many factors, which leads to the low positive rate or undetectable result (Abayasekara *et al.*, 2017; Decuyper *et al.*, 2016; Kirn & Weinstein, 2013; Vincent *et al.*, 2015). Moreover, false negative results may often emerge in culturing anaerobic bacteria or *M. Tuberculosis*, which require special culture media or long-time culture (Olaru *et al.*, 2018). Thus, a need of optimizing culture methods and seeking for better approaches is put on the agenda.

Recently, more and more studies were published on the use of shotgun metagenomic sequencing technology for the identification of pathogens in clinical samples (Kohl *et al.*, 2015; Naccache *et al.*, 2014; Wilson *et al.*, 2014), making it as a potential tool to speed pathogen discovery. Several case reports and studies showed potential of shotgun metagenomic sequencing approach to enhance diagnosis of difficult-to-detect pathogens in various conditions. To these aims, this study intends to explore the advantages and disadvantages of shotgun metagenomic sequencing technology compared with the conventional culture method in pathogen identification and resistance information, and expand evidence of its potential application in real-life clinical settings.

In the present study, we comprehensively compared pathogen diagnosis and resistance information by shotgun metagenomic sequencing and culture in a pairwise manner, and found that shotgun metagenomic sequencing approach has advantages in several respects. First, shotgun metagenomic sequencing is known for its rapid and sensitive detection in diagnosing fungi and bacteria, and its short turn-around time might speed up early clinical decision and precise antimicrobial treatment. Then, shotgun metagenomic sequencing could provide ARGs information as well as resistance antibiotic types by bioinformatics pipeline, making it possible to apply proper antimicrobial therapy in time.

Of the 20 specimens in the study, fungi detection showed complete consistent results between these two methods. In terms of bacterial detection, due to the related technical complexity in taxonomic classification, both bacterial identification at genes level

**Table 3** The details of conflicting results between antibiotic resistance phenotypes and genotypes by susceptibility testing and shotgun metagenomic sequencing approach.

No	Resistance information in susceptibility tests		Resistance information in shotgun metagenomic sequencing	
	Susceptibility testing results*	Antibiotic classes of drugs	Resistance gene types from CARD#	Antibiotic classes of genes
P3	CZO, NIT	<b>Cephalosporin</b> , nitrofurantoin	<i>acrB</i> , <i>bl2be_shv2</i> , <i>tetA</i> , <i>sul2</i> , <i>mdtF</i> , <i>tetC</i> , <i>mdtO</i> , <i>mdtE</i> , <i>mdtP</i> , <i>aph(6)-Id</i>	Tetracycline, penam, <b>cephalosporin</b> , glycylcycline, rifamycin, chloramphenicol, fluoroquinolone, sulfonamide, macrolide, aminoglycoside
P4	(-)	(-)	<i>ermC</i> , <i>sul1</i> , <i>acrB</i> , <i>bacA</i> , <i>mexB</i> , <i>mexW</i> , <i>mexA</i> , <i>qnrB</i> , <i>mexI</i> , <i>mexD</i>	Macrolide, lincosamide, streptogramin, sulfonamide, diaminopyrimidine, tetracycline, penam, cephalosporin, glycylcycline, rifamycin, chloramphenicol, fluoroquinolone, monobactam, aminoglycoside
P5	(-)	(-)	<i>mexA</i> , <i>sul1</i> , <i>bl1_pao</i> , <i>acrB</i> , <i>mexB</i> , <i>bacA</i> , <i>mexD</i> , <i>qnrB</i> , <i>mexI</i> , <i>tetD</i>	Penam, sulfonamide, diaminopyrimidine, tetracycline, cephalosporin, monobactam, chloramphenicol, fluoroquinolone, macrolide, cephamycin, glycylcycline, rifamycin, aminoglycoside
P9	AMP, AMC, NIT	Penam, β-lactamase, nitrofurantoin	<i>bacA</i> , <i>adeA</i> , <i>adeB</i> , <i>aph(3')-Ia</i> , <i>ermC</i> , <i>tet39</i> , <i>lnuA</i> , <i>msrA</i>	Aminoglycoside, macrolide, lincosamide, streptogramin, tetracycline, glycylcycline
P12	AMK, AMP, AMC, TZP, CRO, FEP, FOX, ATM, IMP, TOB, GEN, CIP, SMZ-TMP, NIT	<b>Aminoglycoside</b> , penam, β-lactamase, cephalosporin, monobactam, fluoroquinolone, <b>sulfonamide</b> , <b>diaminopyrimidine</b> , nitrofurantoin	<i>aph(3')-Ia</i> , <i>aph(6)-Id</i> , <i>tetB</i> , <i>aph33ib</i> , <i>bacA</i> , <i>adeB</i> , <i>adeA</i> , <i>sul2</i> , <i>adeC</i> , <i>sul1</i>	<b>Aminoglycoside</b> , tetracycline, streptogramin, <b>sulfonamide</b> , <b>diaminopyrimidine</b> , glycylcycline
P17	CSL, SAM, CAZ, MEM, AMK, AMP, TZP, AMC, CRO, FEP, FOX, IMP, GEN, TOB, LVX, CIP, SMZ-TMP, NIT	β-lactamase, cephalosporin, penam, <b>aminoglycoside</b> , <b>aminoglycoside</b> , fluoroquinolone, <b>sulfonamide</b> , <b>diaminopyrimidine</b> , nitrofurantoin	<i>aph(3')-Ia</i> , <i>tetB</i> , <i>aph33ib</i> , <i>aph(6)-Id</i> , <i>bacA</i> , <i>sul1</i> , <i>ant(3'')-Ia</i> , <i>adeB</i> , <i>adeC</i> , <i>catB3</i>	<b>Aminoglycoside</b> , tetracycline, streptogramin, <b>sulfonamide</b> , <b>diaminopyrimidine</b> , glycylcycline, chloramphenicol
P18	CZO, FOX, SMZ-TMP, NIT	Cephalosporin, <b>sulfonamide</b> , <b>diaminopyrimidine</b> , nitrofurantoin	<i>tet39</i> , <i>bacA</i> , <i>aph(6)-Id</i> , <i>aph33ib</i> , <i>sul2</i>	Tetracycline, aminoglycoside, streptogramin, <b>sulfonamide</b> , <b>diaminopyrimidine</b>
P19	(-)	(-)	<i>tet39</i> , <i>lnuA</i> , <i>bacA</i>	Tetracycline, lincosamide
P20	(-)	(-)	<i>tet39</i> , <i>lnuA</i> , <i>ermC</i> , <i>blaZ</i> , <i>bacA</i> , <i>sul1</i> , <i>aph(3')-Ia</i> , <i>ermB</i> , <i>msrA</i> , <i>qacB</i>	Tetracycline, lincosamide, macrolide, streptogramin, penam, fluoroquinolone, aminoglycoside

**Notes:**

\* Only non-susceptibility is listed.

# Resistance genes listed here are the top 10 genes with relative abundance higher than 1%.

Reference number cells were highlighted in grey if resistance information of whole metagenome-shotgun sequencing results conflicted with susceptibility tests, where conflicts were defined by when the consistency of the antibiotic classes in susceptibility tests and annotated by CARD database was less than or equal to 50%. Consistent resistant classes between susceptibility tests and resistance genes were shown in bold. AMC, amoxicillin-clavulanic acid; AMK, amikacin; AMP, ampicillin; ATM, amikacin; CAZ, ceftazidime; CIP, ciprofloxacin; CLI, clindamycin; CRO, ceftriaxone; CSL, ceftiofur-sulbactam; CTT, cefotetan; CXM, cefuroxime; CZO, cefazolin; ERY, erythromycin; ESBLs, extended spectrum beta-lactamases; ETP, ertapenem; FEP, cefepime; FOX, ceftiofur; GAT, gatifloxacin; GENh, gentamicin high-level; IMP, imipenem; LVX, levofloxacin; MEM, meropenem; MFX, moxifloxacin; NIT, nitrofurantoin; OXA, oxacillin; PEN, penicillin; PIP, piperacillin; SAM, ampicillin-sulbactam; SMZ-TMP, sulfamethoxazole-trimethoprim; TCY, tetracycline; TOB, tobramycin; TZP, piperacillin-tazobactam.

and species level is of great value for comparison. We found that shotgun metagenomic sequencing identified extra pathogens in two specimens when conventional culture failed to get positive result, and reported 17 relatively consistent results in 18 bacterial positive

culture results (including 11/18 at species level and 6/18 at genus level). Due to the preliminary selection of PCR-positive specimens, shotgun metagenomic sequencing had a high negative predictive value than conventional culture, as is similar to previous studies ([Parize et al., 2017](#); [Thoendel et al., 2018](#); [Zhang et al., 2020](#)).

Except for those consistent results, three specimens (P5, P13, and P20) showed completely conflict between shotgun metagenomic sequencing and conventional culture. Besides, six specimens (P1, P3, P6, P10, P18, and P19) only matched at genus level, showing a widely rather than narrow consistence. Of note, unlike culture, more than one strain was identified in a specimen in the most of shotgun metagenomic sequencing workflow. The metagenomics results of culture negative specimens were further evaluated by the clinical microbiologist, which was finally concluded that the negative results of culture might be due to the application of antibiotics before admission and higher sensitivity of shotgun metagenomic sequencing workflow. Similar finding were also confirmed by previous studies, e.g., [Miao et al. \(2018\)](#) and [Zhang et al. \(2019\)](#) revealed shotgun metagenomic sequencing detected more potential microbes than culture in CNS infection cases with empirical antibiotic treatment.

Furthermore, the inexistent organisms in the culture identified by shotgun metagenomic sequencing pipeline may be attributed to growth inhibition induced by antibiotics usage, growth restriction by inappropriate culture conditions, or microbial interaction ([Abayasekara et al., 2017](#); [Decuypere et al., 2016](#)). Our results indicated that the shotgun metagenomic sequencing technology was able to detect pathogens that could not be directly identified by conventional culture, either for technical or practical reasons. That is to say, in some cases shotgun metagenomic sequencing may be more sensitive than conventional culture in identifying pathogens. Consequently, it should be realized that culture-negative specimens do not actually mean pathogen-free, but it might reflect the low sensitivity of culture-based methods or non-culturable characteristics of some specimens (e.g., prior antibiotics treatments). So, one of the applications of shotgun metagenomic sequencing may be to detect the specimens remaining negative in standard culture either for technical reasons, or for unsuspected organisms so that it can be used as a potential diagnostic tool with an added value.

Since we insisted on aseptic operation in our pipeline from beginning to end ([Graf et al., 2016](#); [Street et al., 2017](#)), we could firmly believe the results of shotgun metagenomic sequencing were accurate and true. Based on the above premise, we found out that most of the results by shotgun metagenomic sequencing were multiple bacterial infections, and the culture results were usually one single strain, which was also described by other studies ([Yan et al., 2020](#); [Yin et al., 2017](#); [Zhang et al., 2020](#)). Emerging evidence also exists that multiple microbial co-infections often occurred in patients with severe infections ([Herberg et al., 2016](#); [Song, Xu & Shen, 2016](#); [Yin et al., 2017](#)), and bacterial toxins could cause serious consequences even in the absence of live bacteria ([de Punder & Pruijboom, 2015](#)). So, here come the questions. Do all these bacteria detected from shotgun metagenomic sequencing exist in specimens or infected sites? Are they all the sources of infection? Among these conflict results, are the culture results or shotgun metagenomic sequencing results more credible? However, because the existing technical

approaches cannot explain or verify these issues, we cannot blindly hold the idea that one of them is more credible to be believed, but can only judge them in combination with clinical manifestations.

*Miao et al. (2018)* recently performed a similar study using shotgun metagenomic sequencing retrospectively to diagnose infectious disease for all suspected patients. However, due to their wider inclusion of all patients with suspected infections, they revealed that shotgun metagenomic sequencing approach did not identify as much more positive results as culture could, suggesting that shotgun metagenomic sequencing for common bacterial detection might not be as advantageous as it for viral and fungal might, and it is recommended as a supplement to conventional culture rather than replacement.

In addition, resistance information of causative pathogen is an important part of microbiological diagnosis, however, fewer studies except for several case reports investigated the performance of it by shotgun metagenomic sequencing in infectious diseases. Thus, we compared the drug resistance state in each sample in parallel between sequencing test and susceptibility test, to evaluate the application value of this novel technology in clinical practice ulteriorly. In order to enhance comparability in the present study, antibiotic classes were used as the basis for evaluating the consistency of drug resistance between shotgun metagenomic sequencing and drug susceptibility test. We confirmed more than half specimens had resembling resistant drug types, especially in those concordant specimens. The conflicting results (P3, P4, P5, P9, P12, P17, P18, P19, and P20) between phenotypes (susceptibility tests) and genotypes (annotated ARGs) are probably results from the complexity of drug resistance mechanisms. For example, the mechanism of bacterial resistance is not only caused by gene mutation, also may be changes in point mutations, gene expression changes, and posttranslational modifications, etc. (*Grumaz et al., 2016*); the applicability of the CARD database for detecting ARGs is only for genes per se resistant to antibiotics. So the detection of ARGs cannot fully represent the actual drug resistance of microbes. Notably, a shortcoming of resistant genes' analysis is that they do not provide an accurate measure of antibiotics susceptibility; in particular, if the presence of ARGs or mutations confer inducible resistance, phenotypic testing will still be required (*Jenkins & Schuetz, 2012; Lefterova et al., 2015*). Meanwhile, a resistance gene phenotype cannot be predicted in case that it has not been characterized genetically or is not in an existing database. Therefore, shotgun metagenomic sequencing-based ARGs' detection seems to be feasible, but it is unlikely to replace antimicrobial susceptibility tests at present.

Overall, these two approaches have their own strengths and weakness in the process of clinical application. For the moment, the combination of these two approaches with clinical evidences is the "best practice" for the diagnosis of pathogenic microorganisms, especially for those infectious diseases or critical illness that are difficult to diagnose.

There are still several limitations to this study. Although costs per base are steadily dropping, sequencing-based approaches are still comparatively expensive. Therefore, a small size of samples in our study was enrolled to achieve a relatively reliable sequencing depth for the detection of causative microbes (*Thoendel et al., 2018*). Then, a variety of specimen types was contained. Considering bias could have happened if only one type of



specimen had been used, it was profitable to use different types of specimens to obtain unbiased analysis. Moreover, blood specimens were not contained in this study. Owing to the low positive rate of blood culture and large amount of host cells in samples, blood samples were not included in this study. Thus, the standards in our study should be thoroughly modified and validated in the expanded studies before applying in clinical practice.

## CONCLUSIONS

In conclusion, this study suggested the specificity of shotgun metagenomic sequencing for the identification of causative pathogens, and provided new information and experimental evidence for the application of shotgun metagenomic sequencing approach in pathogenic identification of infectious diseases, especially for the cases where standard culture approaches failed to identify microorganisms.

## ACKNOWLEDGEMENTS

We thank all patients who participated in the studies.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This work was supported by grants from National Natural Science Foundation of China (No. 81871610, No. 81771766, No. 81870071, No. 81770306). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:

National Natural Science Foundation of China: 81871610, 81771766, 81870071, 81770306.

### Competing Interests

The authors declare that they have no competing interests.

### Author Contributions

- Huan Chen conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Jun Li performed the experiments, prepared figures and/or tables, and approved the final draft.
- Shanshan Yan performed the experiments, prepared figures and/or tables, and approved the final draft.
- Hui Sun analyzed the data, prepared figures and/or tables, and approved the final draft.
- Chuyi Tan analyzed the data, prepared figures and/or tables, and approved the final draft.
- Meidong Liu performed the experiments, prepared figures and/or tables, and approved the final draft.



- Ke Liu performed the experiments, prepared figures and/or tables, and approved the final draft.
- Huali Zhang conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Mingxiang Zou conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Xianzhong Xiao conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

### Human Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

The Ethics Committee of the Xiangya Hospital of Central South University approved the study (No. 201703567).

### DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

The primers of 16S rDNA, ITS1 and ITS4 were purchased from Sangon Biotech, Shanghai, China.

### Data Availability

The following information was supplied regarding data availability:

Data are available in the National Center for Biotechnology Information (NCBI) repository: [PRJNA527188](https://www.ncbi.nlm.nih.gov/PRJNA527188).

### Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.11699#supplemental-information>.

## REFERENCES

- Abayasekara LM, Perera J, Chandrasekharan V, Gnanam VS, Udunuwara NA, Liyanage DS, Bulathsinhala NE, Adikary S, Aluthmuhandiram JVS, Thanaseelan CS, Tharmakulasingam DP, Karunakaran T, Ilango J. 2017. Detection of bacterial pathogens from clinical specimens using conventional microbial culture and 16S metagenomics: a comparative study. *BMC Infectious Diseases* 17(1):631 DOI 10.1186/s12879-017-2727-8.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120 DOI 10.1093/bioinformatics/btu170.
- CLSI. 2015. *Methods for antimicrobial dilution and disc susceptibility testing of infrequently isolated or fastidious bacteria, M45*. Third Edition. Wayne, PA: Clinical and Laboratory Standards Institute.
- Couto N, Schuele L, Raangs EC, Machado MP, Mendes CI, Jesus TF, Chlebowicz M, Rosema S, Ramirez M, Carrico JA, Autenrieth IB, Friedrich AW, Peter S, Rossen JW. 2018. Critical steps in clinical shotgun metagenomics for the concomitant detection and typing of microbial pathogens. *Scientific Reports* 8(1):13767 DOI 10.1038/s41598-018-31873-w.

- Crofts TS, Gasparrini AJ, Dantas G. 2017. Next-generation approaches to understand and combat the antibiotic resistance. *Nature Reviews Microbiology* 15(7):422–434 DOI 10.1038/nrmicro.2017.28.
- de Punder K, Pruimboom L. 2015. Stress induces endotoxemia and low-grade inflammation by increasing barrier permeability. *Frontiers in Immunology* 6(17):223 DOI 10.3389/fimmu.2015.00223.
- Decuyper S, Meehan CJ, Van Puyvelde S, De Block T, Maltha J, Palpouguini L, Tahita M, Tinto H, Jacobs J, Deborggraeve S. 2016. Diagnosis of bacterial bloodstream infections: a 16S metagenomics approach. *PLOS Neglected Tropical Diseases* 10(2):e0004470 DOI 10.1371/journal.pntd.0004470.
- Fang X, Cai Y, Shi T, Huang Z, Zhang C, Li W, Zhang C, Yang B, Zhang W, Guan Z. 2020. Detecting the presence of bacteria in low-volume preoperative aspirated synovial fluid by metagenomic next-generation sequencing. *International Journal of Infectious Diseases* 99(1):108–116 DOI 10.1016/j.ijid.2020.07.039.
- Francis OE, Bendall M, Manimaran S, Hong C, Clement NL, Castro-Nallar E, Snell Q, Schaalje GB, Clement MJ, Crandall KA, Johnson WE. 2013. Pathoscope: species identification and strain attribution with unassembled sequencing data. *Genome Research* 23(10):1721–1729 DOI 10.1101/gr.150151.112.
- Goldberg B, Sichtig H, Geyer C, Ledebner N, Weinstock GM. 2015. Making the leap from research laboratory to clinic: challenges and opportunities for next-generation sequencing in infectious disease diagnostics. *MBio* 6(6):e0188801815 DOI 10.1128/mBio.01888-15.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29(7):644–652 DOI 10.1038/nbt.1883.
- Graf EH, Simmon KE, Tardif KD, Hymas W, Flygare S, Eilbeck K, Yandell M, Schlager R. 2016. Unbiased detection of respiratory viruses by use of RNA sequencing-based metagenomics: a systematic comparison to a commercial PCR panel. *Journal of Clinical Microbiology* 54(4):1000–1007 DOI 10.1128/JCM.03060-15.
- Grumaz S, Stevens P, Grumaz C, Decker SO, Weigand MA, Hofer S, Brenner T, von Haeseler A, Sohn K. 2016. Next-generation sequencing diagnostics of bacteremia in septic patients. *Genome Medicine* 8(1):73 DOI 10.1186/s13073-016-0326-8.
- Gu W, Deng X, Lee M, Sucu YD, Arevalo S, Stryke D, Federman S, Gopez A, Reyes K, Zorn K, Sample H, Yu G, Ishpuniani G, Briggs B, Chow ED, Berger A, Wilson MR, Wang C, Hsu E, Miller S, DeRisi JL, Chiu CY. 2021. Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nature Medicine* 27(1):115–124 DOI 10.1038/s41591-020-1105-z.
- Herberg JA, Kafrou M, Wright VJ, Shailes H, Eleftherohorinou H, Hoggart CJ, Cebey-Lopez M, Carter MJ, Janes VA, Gormley S, Shimizu C, Tremoulet AH, Barendregt AM, Salas A, Kanegaye J, Pollard AJ, Faust SN, Patel S, Kuijpers T, Martinon-Torres F, Burns JC, Coin LJ, Levin M. 2016. Diagnostic test accuracy of a 2-transcript host RNA signature for discriminating bacterial vs viral infection in febrile children. *JAMA* 316(8):835–845 DOI 10.1001/jama.2016.11236.
- Huang J, Jiang E, Yang D, Wei J, Zhao M, Feng J, Cao J. 2020. Metagenomic next-generation sequencing versus traditional pathogen detection in the diagnosis of peripheral pulmonary infectious lesions. *Infection and Drug Resistance* 13:567–576 DOI 10.2147/idr.s235182.

- Huson DH, Auch AF, Qi J, Schuster SC. 2007. MEGAN analysis of metagenomic data. *Genome Research* 17(3):377–386 DOI 10.1101/gr.5969107.
- Jenkins SG, Schuetz AN. 2012. Current concepts in laboratory testing to guide antimicrobial therapy. *Mayo Clinic Proceedings* 87(3):290–308 DOI 10.1016/j.mayocp.2012.01.007.
- Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, Lago BA, Dave BM, Pereira S, Sharma AN, Doshi S, Courtot M, Lo R, Williams LE, Frye JG, Elsayegh T, Sardar D, Westman EL, Pawlowski AC, Johnson TA, Brinkman FS, Wright GD, McArthur AG. 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Research* 45(D1):D566–D573 DOI 10.1093/nar/gkw1004.
- Jorgensen JPM, Carroll K, Funke G, Landry M, Richter S, Jorgensen JH, Pfaller MA, Carroll KC, Funke G, Landry ML, Richter SS, Warnock DW. 2015. *Manual of clinical microbiology*. 11th Edition. Hoboken, New Jersey: Wiley.
- Kirn TJ, Weinstein MP. 2013. Update on blood cultures: how to obtain, process, report, and interpret. *Clinical Microbiology and Infection* 19(6):513–520 DOI 10.1111/1469-0691.12180.
- Kohl C, Brinkmann A, Dabrowski PW, Radonic A, Nitsche A, Kurth A. 2015. Protocol for metagenomic virus detection in clinical specimens. *Emerging Infectious Diseases* 21(1):48–57 DOI 10.3201/eid2101.140766.
- Lefterova MI, Suarez CJ, Banaei N, Pinsky BA. 2015. Next-generation sequencing for infectious disease diagnosis and management: a report of the association for molecular pathology. *Journal of Molecular Diagnostics* 17(6):623–634 DOI 10.1016/j.jmoldx.2015.07.004.
- Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, Wang J. 2009. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25(15):1966–1967 DOI 10.1093/bioinformatics/btp336.
- Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22(13):1658–1659 DOI 10.1093/bioinformatics/btl158.
- Liu R, Hong J, Xu X, Feng Q, Zhang D, Gu Y, Shi J, Zhao S, Liu W, Wang X, Xia H, Liu Z, Cui B, Liang P, Xi L, Jin J, Ying X, Wang X, Zhao X, Li W, Jia H, Lan Z, Li F, Wang R, Sun Y, Yang M, Shen Y, Jie Z, Li J, Chen X, Zhong H, Xie H, Zhang Y, Gu W, Deng X, Shen B, Xu X, Yang H, Xu G, Bi Y, Lai S, Wang J, Qi L, Madsen L, Wang J, Ning G, Kristiansen K, Wang W. 2017. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nature Medicine* 23(7):859–868 DOI 10.1038/nm.4358.
- Miao Q, Ma Y, Wang Q, Pan J, Zhang Y, Jin W, Yao Y, Su Y, Huang Y, Wang M, Li B, Li H, Zhou C, Li C, Ye M, Xu X, Li Y, Hu B. 2018. Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice. *Clinical Infectious Diseases* 67(suppl\_2):S231–S240 DOI 10.1093/cid/ciy693.
- Naccache SN, Federman S, Veeraraghavan N, Zaharia M, Lee D, Samayoa E, Bouquet J, Greninger AL, Luk KC, Enge B, Wadford DA, Messenger SL, Genrich GL, Pellegrino K, Grard G, Leroy E, Schneider BS, Fair JN, Martinez MA, Isa P, Crump JA, DeRisi JL, Sittler T, Hackett J Jr., Miller S, Chiu CY. 2014. A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples. *Genome Research* 24(7):1180–1192 DOI 10.1101/gr.171934.113.
- Olaru ID, Patel H, Kranzer K, Perera N. 2018. Turnaround time of whole genome sequencing for mycobacterial identification and drug susceptibility testing in routine practice. *Clinical Microbiology and Infection* 24(6):659.e655–659.e657 DOI 10.1016/j.cmi.2017.10.001.
- Parize P, Muth E, Richaud C, Gratigny M, Pilmis B, Lamamy A, Mainardi JL, Cheval J, de Visser L, Jagorel F, Ben Yahia L, Bamba G, Dubois M, Join-Lambert O, Leruez-Ville M, Nassif X, Lefort A, Lanternier F, Suarez F, Lortholary O, Lecuit M, Eloit M. 2017. Untargeted

- next-generation sequencing-based first-line diagnosis of infection in immunocompromised adults: a multicentre, blinded, prospective study. *Clinical Microbiology and Infection* 23(8):574.e571–574.e576 DOI 10.1016/j.cmi.2017.02.006.
- Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, Liang S, Zhang W, Guan Y, Shen D, Peng Y, Zhang D, Jie Z, Wu W, Qin Y, Xue W, Li J, Han L, Lu D, Wu P, Dai Y, Sun X, Li Z, Tang A, Zhong S, Li X, Chen W, Xu R, Wang M, Feng Q, Gong M, Yu J, Zhang Y, Zhang M, Hansen T, Sanchez G, Raes J, Falony G, Okuda S, Almeida M, LeChatelier E, Renault P, Pons N, Batto JM, Zhang Z, Chen H, Yang R, Zheng W, Li S, Yang H, Wang J, Ehrlich SD, Nielsen R, Pedersen O, Kristiansen K, Wang J. 2012. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490(7418):55–60 DOI 10.1038/nature11450.
- Rhodes A, Evans LE, Alhazzani W, Levy MM, Antonelli M, Ferrer R, Kumar A, Sevransky JE, Sprung CL, Nunnally ME, Rochweg B, Rubinfeld GD, Angus DC, Annane D, Beale RJ, Bellingham GJ, Bernard GR, Chiche JD, Coopersmith C, De Backer DP, French CJ, Fujishima S, Gerlach H, Hidalgo JL, Hollenberg SM, Jones AE, Karnad DR, Kleinpell RM, Koh Y, Lisboa TC, Machado FR, Marini JJ, Marshall JC, Mazuski JE, McIntyre LA, McLean AS, Mehta S, Moreno RP, Myburgh J, Navalesi P, Nishida O, Osborn TM, Perner A, Plunkett CM, Ranieri M, Schorr CA, Seckel MA, Seymour CW, Shieh L, Shukri KA, Simpson SQ, Singer M, Thompson BT, Townsend SR, Van der Poll T, Vincent JL, Wiersinga WJ, Zimmerman JL, Dellinger RP. 2017. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock: 2016. *Intensive Care Medicine* 43(3):304–377 DOI 10.1007/s00134-017-4683-6.
- Song Q, Xu BP, Shen KL. 2016. Bacterial co-infection in hospitalized children with Mycoplasma pneumoniae pneumonia. *Indian Pediatrics* 53(10):879–882 DOI 10.1007/s13312-016-0951-8.
- Street TL, Sanderson ND, Atkins BL, Brent AJ, Cole K, Foster D, McNally MA, Oakley S, Peto L, Taylor A, Peto TEA, Crook DW, Eyre DW. 2017. Molecular diagnosis of orthopedic-device-related infection directly from sonication fluid by metagenomic sequencing. *Journal of Clinical Microbiology* 55(8):2334–2347 DOI 10.1128/JCM.00462-17.
- Thoendel MJ, Jeraldo PR, Greenwood-Quaintance KE, Yao JZ, Chia N, Hanssen AD, Abdel MP, Patel R. 2018. Identification of prosthetic joint infection pathogens using a shotgun metagenomics approach. *Clinical Infectious Diseases* 67(9):1333–1338 DOI 10.1093/cid/ciy303.
- Vincent JL, Brealey D, Libert N, Abidi NE, O'Dwyer M, Zacharowski K, Mikaszewska-Sokolewicz M, Schrenzel J, Simon F, Wilks M, Picard-Maureau M, Chalfin DB, Ecker DJ, Sampath R, Singer M. 2015. Rapid diagnosis of infection in the critically ill, a multicenter study of molecular detection in bloodstream infections, pneumonia, and sterile site infections. *Critical Care Medicine* 43(11):2283–2291 DOI 10.1097/CCM.0000000000001249.
- Wang H, Wang J, Yu P, Ge P, Jiang Y, Xu R, Chen R, Liu X. 2017. Identification of antibiotic resistance genes in the multidrug-resistant *Acinetobacter baumannii* strain, MDR-SHH02, using whole-genome sequencing. *International Journal of Molecular Medicine* 39(2):364–372 DOI 10.3892/ijmm.2016.2844.
- Wilson MR, Naccache SN, Samayoa E, Biagtan M, Bashir H, Yu G, Salamat SM, Somasekar S, Federman S, Miller S, Sokolic R, Garabedian E, Candotti F, Buckley RH, Reed KD, Meyer TL, Seroogy CM, Galloway R, Henderson SL, Gern JE, DeRisi JL, Chiu CY. 2014. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *New England Journal of Medicine* 370(25):2408–2417 DOI 10.1056/NEJMoa1401268.
- Wilson MR, Sample HA, Zorn KC, Arevalo S, Yu G, Neuhaus J, Federman S, Stryke D, Briggs B, Langelier C, Berger A, Douglas V, Josephson SA, Chow FC, Fulton BD, DeRisi JL, Gelfand JM, Naccache SN, Bender J, Dien Bard J, Murkey J, Carlson M, Vespa PM, Vijayan T, Allyn PR, Campeau S, Humphries RM, Klausner JD, Ganzon CD, Memar F,

- Ocampo NA, Zimmermann LL, Cohen SH, Polage CR, DeBiasi RL, Haller B, Dallas R, Maron G, Hayden R, Messacar K, Dominguez SR, Miller S, Chiu CY. 2019. Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *New England Journal of Medicine* **380**(24):2327–2340 DOI 10.1056/NEJMoa1803396.
- World Health Organization. 2020. The top 10 causes of death. Available at <http://www.who.int/zh/news-room/fact-sheets/detail/the-top-10-causes-of-death>.
- Yan L, Sun W, Lu Z, Fan L. 2020. Metagenomic next-generation sequencing (mNGS) in cerebrospinal fluid for rapid diagnosis of Tuberculosis meningitis in HIV-negative population. *International Journal of Infectious Diseases* **96**(1):270–275 DOI 10.1016/j.ijid.2020.04.048.
- Yin C, Yang W, Meng J, Lv Y, Wang J, Huang B. 2017. Co-infection of *Pseudomonas aeruginosa* and *Stenotrophomonas maltophilia* in hospitalised pneumonia patients has a synergic and significant impact on clinical outcomes. *European Journal of Clinical Microbiology & Infectious Diseases* **36**(11):2231–2235 DOI 10.1007/s10096-017-3050-4.
- You M, Yue Z, He W, Yang X, Yang G, Xie M, Zhan D, Baxter SW, Vasseur L, Gurr GM, Douglas CJ, Bai J, Wang P, Cui K, Huang S, Li X, Zhou Q, Wu Z, Chen Q, Liu C, Wang B, Li X, Xu X, Lu C, Hu M, Davey JW, Smith SM, Chen M, Xia X, Tang W, Ke F, Zheng D, Hu Y, Song F, You Y, Ma X, Peng L, Zheng Y, Liang Y, Chen Y, Yu L, Zhang Y, Liu Y, Li G, Fang L, Li J, Zhou X, Luo Y, Gou C, Wang J, Wang J, Yang H, Wang J. 2013. A heterozygous moth genome provides insights into herbivory and detoxification. *Nature Genetics* **45**(2):220–225 DOI 10.1038/ng.2524.
- Zhang XX, Guo LY, Liu LL, Shen A, Feng WY, Huang WH, Hu HL, Hu B, Guo X, Chen TM, Chen HY, Jiang YQ, Liu G. 2019. The diagnostic value of metagenomic next-generation sequencing for identifying *Streptococcus pneumoniae* in paediatric bacterial meningitis. *BMC Infectious Diseases* **19**(1):495 DOI 10.1186/s12879-019-4132-y.
- Zhang Y, Cui P, Zhang HC, Wu HL, Ye MZ, Zhu YM, Ai JW, Zhang WH. 2020. Clinical application and evaluation of metagenomic next-generation sequencing in suspected adult central nervous system infection. *Journal of Translational Medicine* **18**(1):199 DOI 10.1186/s12967-020-02360-6.
- Zhou X, Wu H, Ruan Q, Jiang N, Chen X, Shen Y, Zhu YM, Ying Y, Qian YY, Wang X, Ai JW, Zhang WH. 2019. Clinical evaluation of diagnosis efficacy of active mycobacterium tuberculosis complex infection via metagenomic next-generation sequencing of direct clinical samples. *Frontiers in Cellular and Infection Microbiology* **9**:351 DOI 10.3389/fcimb.2019.00351.
- Zhu W, Lomsadze A, Borodovsky M. 2010. Ab initio gene identification in metagenomic sequences. *Nucleic Acids Research* **38**(12):e132–e132 DOI 10.1093/nar/gkq275.