

UCLA

UCLA Previously Published Works

Title

Metagenomic Sequencing of Positive Blood Culture Fluid for Accurate Bacterial and Fungal Species Identification: A Pilot Study.

Permalink

<https://escholarship.org/uc/item/3z03v4vt>

Journal

Microorganisms, 11(5)

ISSN

2076-2607

Authors

Kamau, Edwin
Yang, Shangxin

Publication Date

2023-05-01

DOI

10.3390/microorganisms11051259

Peer reviewed



Article

Metagenomic Sequencing of Positive Blood Culture Fluid for Accurate Bacterial and Fungal Species Identification: A Pilot Study

Edwin Kamau and Shangxin Yang ^{*,†}

Department of Pathology and Laboratory Medicine, UCLA David Geffen School of Medicine, Los Angeles, CA 90095, USA; edwin.kamau.mil@health.mil

* Correspondence: shangxinyang@mednet.ucla.edu

† Current address: UCLA Clinical Microbiology Laboratory, 11633 San Vicente Blvd, Los Angeles, CA 90049, USA.

Abstract: With blood stream infections (BSIs) representing a major cause of mortality and morbidity worldwide, blood cultures play a crucial role in diagnosis, but their clinical application is dampened by the long turn-around time and the detection of only culturable pathogens. In this study, we developed and validated a shotgun metagenomics next-generation sequencing (mNGS) test directly from positive blood culture fluid, allowing for the identification of fastidious or slow growing microorganisms more rapidly. The test was built based on previously validated next-generation sequencing tests, which rely on several key marker genes for bacterial and fungal identification. The new test utilizes an open-source metagenomics CZ-ID platform for the initial analysis to generate the most likely candidate species, which is then used as a reference genome for downstream, confirmatory analysis. This approach is innovative because it takes advantage of an open-source software's agnostic taxonomic calling capability while still relying on the more established and previously validated marker gene-based identification scheme, increasing the confidence in the final results. The test showed high accuracy (100%, 30/30) for both bacterial and fungal microorganisms. We further demonstrated its clinical utility especially for anaerobes and mycobacteria that are either fastidious, slow growing, or unusual. Although applicable in only limited settings, the Positive Blood Culture mNGS test provides an incremental improvement in solving the unmet clinical needs for the diagnosis of challenging BSIs.

Keywords: blood culture; metagenomics; next-generation sequencing; microbial identification; bloodstream infections; mNGS



Citation: Kamau, E.; Yang, S. Metagenomic Sequencing of Positive Blood Culture Fluid for Accurate Bacterial and Fungal Species Identification: A Pilot Study. *Microorganisms* **2023**, *11*, 1259. <https://doi.org/10.3390/microorganisms11051259>

Academic Editors: Eva Torres-Sangiao, Nelson C Soares and Monica Cartelle Gestal

Received: 3 April 2023
Revised: 27 April 2023
Accepted: 9 May 2023
Published: 10 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Bloodstream infections (BSIs) are often associated with severe diseases with high morbidity and mortality [1–4], making the detection of disease-causing pathogens in BSIs one of the most important functions of the clinical microbiology laboratory [5]. The increased mortality of BSIs is associated with delayed, insufficient, or inappropriate anti-infective treatment [1,3,4]. The accurate and timely identification of BSIs is critical for patient management, infection control, and disease epidemiology [6].

Currently, blood cultures coupled with antimicrobial susceptibility testing (AST) are considered the gold standard for diagnosing BSIs [1,5]. The conventional workflow consists of first subculturing the primary positive blood culture medium to a solid agar plate to obtain pure isolates (colonies) after the blood culture vial is flagged positive, and then the subsequent microbial identification and AST. For many bacteria or yeasts, only 12–48 h of incubation is needed to grow enough microbes for subsequent workups. However, in some instances, bacteria or fungi may grow poorly, slowly, or not grow at all at the subculture step once the primary blood cultures turn positive. In addition, some unusual microbes grown

from the subculture of positive blood culture may not be identified by conventional methods such as biochemical tests or mass spectrum-based methods including the matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF). In such situations, the microbial result usually stops short at the descriptive identification based on Gram stain morphology, which is insufficient information for both diagnosis and the empirical treatment of BSIs. This is a clinical dilemma encountered especially in tertiary hospitals, treating many immunocompromised patients who can be infected with uncommon microbes. In such settings, there is an unmet clinical need that warrants a solution.

In our previous case study, we demonstrated a proof of concept of using a shotgun metagenomics approach to correctly identify an unusual anaerobe, *Catabacter hongkongensis*, directly in a blood culture fluid because the conventional microbial workups including MALDI–TOF failed to identify the bacteria [7].

In the current study, we went further to develop and validate a shotgun metagenomics (mNGS)-based clinical test that can directly identify bacteria and fungi in positive blood culture fluid. This test is intended to be used in situations when conventional laboratory methods for positive blood culture fail to provide an appropriate identification of the microorganism, or a faster turn-around time (TAT) is desired especially for mycobacterial or certain fungal organisms that are typically slow growing. This pilot study is aimed at assessing the accuracy and potential clinical utility of a cutting-edge NGS-based molecular diagnostic assay for BSIs.

2. Materials and Methods

2.1. Specimen and Conventional Microbiology

Blood or sterile body fluids were collected from patients at the UCLA hospitals as part of routine care and transported to the clinical microbiology laboratory within 4 h, and incubated in the automated BD BACTEC™ FX blood culture system (Becton Dickinson, Franklin Lakes, NJ, USA). These specimens were collected in the following media/bottle types: BACTEC™ Plus Aerobic/F; BACTEC™ Lytic/10 Anaerobic/F; BACTEC™ Peds Plus; BACTEC™ Myco/F Lytic. When vials were flagged positive, they were unloaded and worked up following standard laboratory procedures. Gram stain, fungal stain, modified Ziehl–Neelsen stain or Ziehl–Neelsen staining were performed on the positive blood culture fluid. Specimens with positive stain results were subcultured to the appropriate media and incubated appropriately. For specimens with negative stain results, a blind subculture was also set up, and an acridine orange stain was performed to rule out organisms without cell walls. Isolates were given taxonomic or descriptive identification based on the staining results and one or more of the following methods: biochemical tests (e.g., oxidase, catalase) and/or commercial ID methods including VITEK MS (bioMérieux, Hazelwood, MO, USA), VITEK 2 (bioMérieux), API (bioMérieux), RapID™ ANA II System (Thermo Fisher Scientific, Waltham, MA, USA), PathoDX (Remel, Lenexa, KS, USA), and AccuProbe (Hologic, San Diego, CA, USA). After the standard microbiological workups were performed, 200 µL of each positive blood culture fluid was used for mNGS. A total of 30 samples were included in this study, including 28 randomly selected unique positive blood culture fluid specimens and 2 additional samples from 2 case studies (described in Section 3) due to either failed species identification by the conventional methods, or urgent clinical needs to improve the TAT (Table 1).

2.2. Metagenomic Sequencing

Metagenomic sequencing was performed directly from the blood culture fluid as previously described [6]. Briefly, 200 µL of blood culture material was obtained and DNA was extracted using Qiagen EZ1 Blood and Tissue Kit on the EZ1 Advanced XL instrument per the manufacturer's recommendations (Qiagen, Valencia, CA, USA). Mycobacterial specimens were first heat-inactivated (100 °C for 30 min) and an additional bead-beating step was performed for the mechanical disruption of the cell wall. Bead-beating step was also performed for fungal specimens. Libraries were prepared and sequenced on the Illumina

MiSeq System (Illumina, San Diego, CA, USA) using a 2×250 bp standard protocol per the manufacturer's recommendations. Detailed steps regarding library preparation and sequencing have previously been described in our study on the validation and implementation of whole genome sequence-based bacterial identification in the clinical microbiology laboratory [8].

Table 1. Microbial organisms identified by Positive Blood Culture mNGS and conventional methods.

Sample ID #	Identification	Positive Blood Culture mNGS					Identification by Conventional Methods	Conventional Methods Used for Final Identification
		16S rRNA	<i>rpoB</i>	<i>groEL</i>	ITS Region	28S rRNA		
* UCLA_467-473, 562, 558	<i>Staphylococcus epidermidis</i>	100.00%	100.00%	100.00%			<i>S. epidermidis</i>	MALDI-TOF
UCLA_478	<i>Slackia exigua</i>	99.30%	n/a	n/a			<i>S. exigua</i>	MALDI-TOF
UCLA_499	<i>Desulfovibrio diazotrophicus</i>	99.7%	100.0%	99.9%			<i>Desulfovibrio</i> spp.	MALDI-TOF
UCLA_501	<i>Bacteroides dorei</i>	99.94%	99.79%	99.94%			<i>B. dorei</i>	MALDI-TOF
UCLA_503	<i>Bacteroides thetaiotaomicron</i>	99.96%	99.97%	100.00%			<i>B. thetaiotaomicron</i>	MALDI-TOF
UCLA_505	<i>Mycobacterium iranicum</i>	99.87%	97.86%	98.52%			Yellow pigmented rapid-growing <i>Mycobacterium</i>	Morphology and Stain
UCLA_510	<i>Proteus mirabilis</i>	100.00%	100.00%	100.00%			<i>P. mirabilis</i>	MALDI-TOF
UCLA_517	<i>Escherichia coli</i>	99.10%	99.98%	100.00%			<i>E. coli</i>	MALDI-TOF
UCLA_518	<i>Staphylococcus lugdunensis</i>	99.94%	100.00%	100.00%			<i>S. lugdunensis</i>	MALDI-TOF
UCLA_519	<i>Streptococcus intermedius</i>	99.94%	99.05%	98.34%			<i>S. intermedius</i>	MALDI-TOF
UCLA_520	<i>Clostridium butyricum</i>	100.00%	99.97%	99.88%			<i>C. butyricum</i>	MALDI-TOF
UCLA_524	<i>Mycobacterium intracellulare</i> subsp. <i>chimaera</i>	100.00%	100.00%	100.00%			<i>M. avium</i> complex	DNA Probe
UCLA_1007	<i>Mycobacterium avium</i> subsp. <i>hominissuis</i>	100.00%	100.00%	100.00%			<i>M. avium</i> complex	DNA Probe
UCLA_559	<i>Enterococcus faecium</i>	99.94%	100.00%	100.00%			<i>E. faecium</i>	MALDI-TOF
UCLA_560	<i>Fusobacterium nucleatum</i>	100.00%	98.96%	99.88%			<i>F. nucleatum</i>	MALDI-TOF
UCLA_561	<i>Staphylococcus haemolyticus</i>	99.81%	99.47%	98.40%			<i>S. haemolyticus</i>	MALDI-TOF
UCLA_496	<i>Candida tropicalis</i>				100%	n/p	<i>C. tropicalis</i>	MALDI-TOF
UCLA_497	<i>Candida glabrata</i>				100%	n/p	<i>C. glabrata</i>	MALDI-TOF
UCLA_498	<i>Fusarium proliferatum</i>				100%	n/p	<i>F. proliferatum</i>	MALDI-TOF
UCLA_500	<i>Clavispora lusitaniae</i>				100%	n/p	<i>C. lusitaniae</i>	MALDI-TOF
UCLA_502	<i>Candida albicans</i>				100%	n/p	<i>C. albicans</i>	MALDI-TOF
UCLA_596	<i>Aspergillus flavus-oryzae</i> group				100%	100%	<i>A. flavus-oryzae</i> group	MALDI-TOF

* Sample IDs were assigned sequentially to the blood culture bottles in no particular order. There were nine blood bottles that contained *S. epidermidis* with assigned sample IDs UCLA_467 to UCLA_473, UCLA_562, and UCLA_558. n/a: reference sequence not available in the NCBI nt database. n/p: analysis not performed.

2.3. Bioinformatics Analysis

The data analysis pipeline is illustrated in Figure 1, with 3 major steps including (1) using Chan Zuckerberg ID (CZ-ID; previously known as IDseq) platform (<https://czid.org/> (accessed on 1 December 2022)), an open source cloud-based metagenomics pipeline for the initial assessment of the microbial content and selection of candidate species identification; (2) using CLC Genomic Workbench version 12 software (Qiagen, Valencia, CA, USA) to generate consensus sequences for the applicable marker genes including *16S rRNA*, *rpoB*, *groEL*, *ITS*, and *28S rRNA* based on the reference genome of the candidate species; and (3) using the NCBI BLAST tool (<https://blast.ncbi.nlm.nih.gov/Blast.cgi> (accessed on 1 December 2022)) and the Westerdijk Fungal Biodiversity Institute Pairwise Alignment tool (https://wi.knaw.nl/page/Pairwise_alignment (accessed on 1 December 2022)) for bacterial and fungal species level identification, respectively. The sequence (fastq) files generated on MiSeq were uploaded on CZ-ID, which performs host and quality filtration steps, and then executes an assembly-based alignment pipeline which results in the assignment of reads and contigs to taxonomic categories [9]. Based on the outcome of the analysis, the top taxonomies with the highest abundance were chosen as the candidate microorganisms present in the positive blood culture fluid. The subsequent bioinformatics were performed using the validated UCLA NGS bacterial or fungal workflows as previously described [8]. Briefly, for the identification of bacterial species, a reference genome suggested by the CZ-ID results was downloaded directly onto the CLC Genomics Workbench 12 (also referred to as CLCbio), and its three target genes (i.e., *16S rRNA*, *rpoB*, *groEL[hsp65]*) were identified, extracted, and concatenated. These genes have been extensively used for bacterial species level identification. The raw sequence reads of the sample were then mapped to the concatenated reference sequence. Various QC metrics were recorded including the total reads, the average coverage, the percentage with at least 5× and 10× coverage, and the percentage of ambiguous nucleotides. The consensus sequence of each target gene was queried using the BLAST NT and/or NR databases. The identification of the bacteria to species or genus level was based on results with at least 99% full-gene coverage, and at least 99% pairwise identity, which calculates the percentage of identical nucleotides in alignment between two sequences. High pairwise identity percentage provides increased confidence that the query sequence is identical to the reference sequences. A cascading taxonomic calling scheme starting from *16S rRNA* first, if not resolved (e.g., two species with the same pairwise identity), then to *rpoB*, and last to *groEL* was used for the species identification (Figure 1). A more detailed algorithm of the scheme was described previously [8]. For the identification of fungal organisms, raw sequences from each sample were processed using the same workflow except that the internal transcribed spacer (ITS) region (encompassing ITS1, 5.8S rDNA, and ITS2) and *28S rRNA* gene were used as marker genes [10]. These markers are widely used for the species identification of fungal organisms. Figure 1 provides an overview of the workflow used in the identification of the microorganism in this study.

The organism identification was carried out using the CZ-ID portal pipeline, then confirmed using UCLA NGS bacterial or fungal workflows, which are hereon now referred to as Positive Blood Culture mNGS. The clinical utility of the test is demonstrated by case studies presented here after the implementation of the validated test in our laboratory.

2.4. Clinical Utility Assessment and Study Ethics

A retrospective chart review was performed on patients for whom Positive Blood Culture mNGS was performed to assess the clinical utility of the NGS results. This study was reviewed by the UCLA Human Research Protection Program and received an IRB exemption.

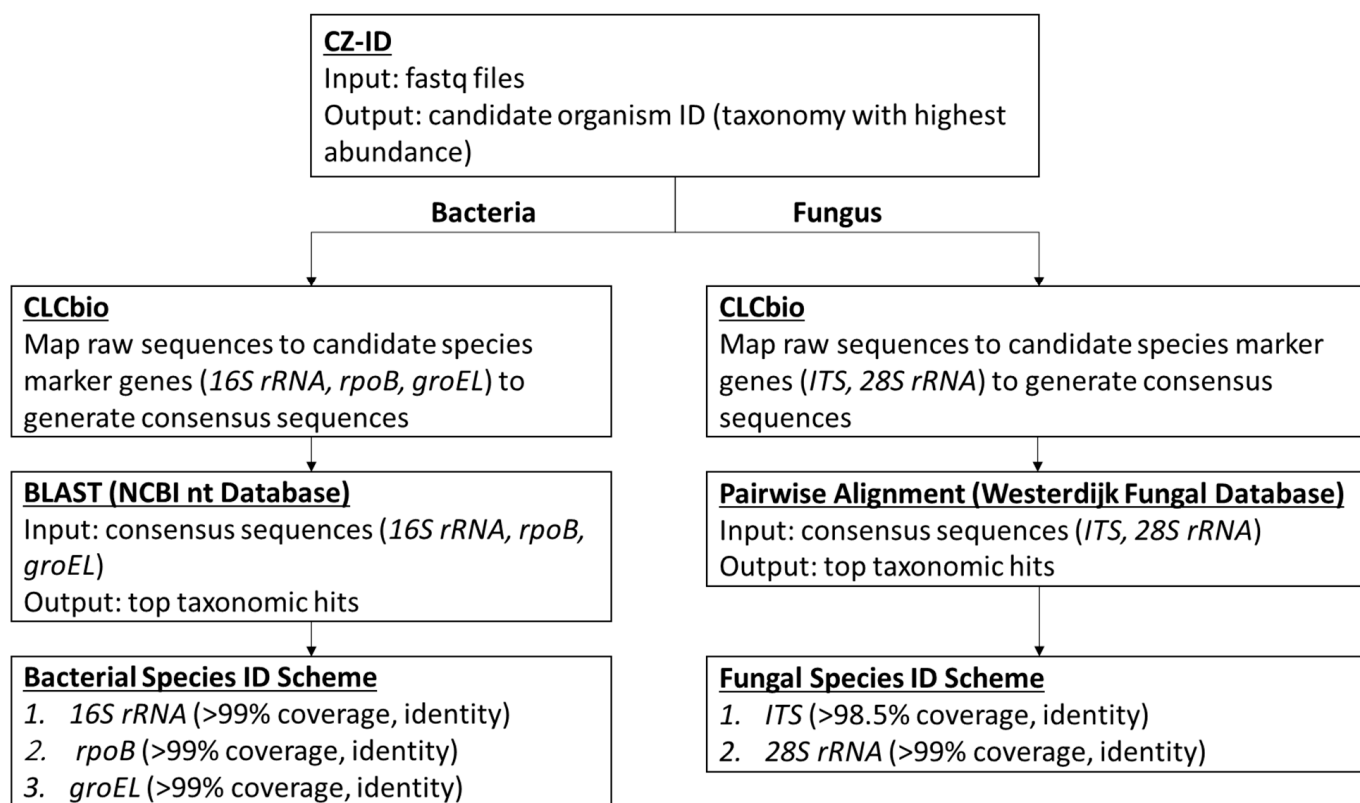


Figure 1. The Bioinformatics Workflow of Positive Blood Culture mNGS.

3. Results

3.1. Quality Matrices and Bioinformatics Performance

Overall, a range of 1,215,048–12,781,338 sequence reads was acquired per sample (Tables 2 and 3). For the CZ-ID portal pipeline, various metrics were obtained with QC criteria established on (1) Phred Quality Score, (2) reads per million (rPM), and (3) percent identity (%id). The Phred Quality Score represents confidence in the assignment of each base call by the sequencer, with a score of 60 representing an accuracy of 99.9999%; all the 28 samples passed the QC cutoff (60%). The rPM assesses the relative abundance of nucleic acid associated with the taxon present in the sample; a high rPM was acquired for all samples (mean= 280,248, range 7534.8–601,308), and thus 5000 rPM was set as the QC cutoff. The %id is the average percent identity of the reads and contigs that were aligned to the taxon in the NCBI NT/NR database; high %id (mean = 98.8, range 83.7–100) was achieved, indicating generally a high confidence in taxonomic calling. Overall, the CZ-ID provided very helpful preliminary microbial species identifications for the downstream analysis.

In the downstream in-house bioinformatics pipelines, high sequence coverage was achieved in 100% (24/24) of the bacterial samples with an average coverage of >20× in all three (16S, rpoB, and groEL) marker genes and <1% ambiguous nucleotide (Table 2). The downstream in-house analysis also identified 6 fungal organisms, with 83.3% (5/6) achieving a high sequence coverage of the ITS region (>20×). However, one *Aspergillus* (UCLA_596) had a lower coverage of 5.9× for ITS, which prompted using an additional marker gene (28S rRNA, 9.5× coverage) for species confirmation.

Table 2. Bacterial Identification Quality Control Metrics.

Metric	Range	Average (SD)	QC Criteria
CZ-ID:			
Phred Quality Score (%)	65.64–95.33	88.3 (8.5)	>60%
rPM (highest value)	7534.8–601,308	280,248 (144,559)	>5000
%id	83.7–100	98.8 (3.4)	>80%
CLC Mapping to Marker Genes:			
Total Reads	1,215,048–12,781,338	2,853,970 (2,342,278)	>1,000,000
Average Coverage (×)	18.7–352.8	109.0 (87.6)	>10×
5× Coverage (%)	99.2–100	99.9 (0.21)	>90%
16S rRNA:			
Length	1513–1573	1547.7 (13.5)	None
Average Coverage	20.8–3943.7	942.4 (901.0)	>20×
Ambiguous Nucleotide (%)	0.0–0.9	0.08 (0.21)	<1%
rpoB:			
Length	3552–4389	3717.7 (242.0)	None
Average Coverage	30.2–822.4	176.6 (160.7)	>20×
Ambiguous Nucleotide (%)	0.0–0.9	0.05 (0.20)	<1%
groEL:			
Length	1620–1647	1627.0 (9.6)	None
Average Coverage	25.05–717.6	157.6 (140.8)	>20×
Ambiguous Nucleotide (%)	0.0–1.1	0.05 (0.23)	<1%

Table 3. Fungal Identification Quality Control Metrics.

Metric	Range	Average (SD)	QC Criteria
CZ-ID:			
Phred Quality Score (%)	56.5–90.5	80.6 (13.6)	>50%
rPM (highest value)	2981.9–619,105.8	225,922.8 (226,785.8)	>2000
%id	96.4–99.8	98.8 (1.2)	>80%
CLC Mapping to Marker Genes:			
Total Reads	2,385,524–4,777,322	3,357,613.7 (779,513.2)	>1,000,000
Average Coverage (×)	5.9–1256.3	653.2 (556.9)	>5×
5× Coverage (%)	66.6–100	94.4 (13.7)	>60%
ITS:			
Length (nt)	341–2587	880.7 (844.7)	None
Average Coverage (×)	5.9–1256.3	727.7 (569.8)	>5×
28S (Aspergillus and Mucorales):			
Length (nt)	593	593	None
Average Coverage (×)	9.5	9.5	>5×

3.2. Accuracy

Positive Blood Culture mNGS achieved 100% (30/30) species-level identification, including 24 bacterial and 6 fungal organisms in the positive blood culture fluids (Tables 1 and S1). Notably, there was 100% (30/30) concordance between the initial CZ-ID species identifications (the taxonomy with highest abundance) and the final species identifications made by Positive Blood Culture mNGS. In the 24 bacteria, the species identification was all based on the 16S rRNA gene (>99% pairwise identity) and confirmed by either *rpoB* or *groEL*, except for sample UCLA_478, for which *Slackia exigua* was identified by the 16S rRNA alone as there were no corresponding *rpoB* and *groEL* reference sequences in the NCBI nt database. All the six fungal organisms were identified based on the ITS region (100% pairwise identity). In sample UCLA_596, in addition to using ITS as the primary marker gene, the species identification (*Aspergillus flavus-oryzae* group) was further confirmed using the 28S rRNA gene which achieved 100% overlap and similarity.

3.3. Case Studies Demonstrating Clinical Utility

Case study 1. A 52-year-old male with a history of rheumatoid arthritis, interstitial lung disease, bilateral orthotopic lung transplant, pulmonary Aspergillosis, *Mycobacterium avium* complex (MAC) infection, abdominal hernia, and pancytopenia had presented with two months of abdominal discomfort and persistent dyspnea that was worsening. The patient reported unintended weight loss and complained of worsening fatigue, brain fog, shortness of breath, poor appetite, and distention of the abdomen. On admission, he was pancytopenic with elevated alkaline phosphatase. Lung imaging studies revealed increased small right pleural effusion with contiguous rounded atelectasis of the right posterior lung base, and ground-glass opacity of the left mid-lung. A liver ultrasound revealed a shrunken nodular liver and moderate to large ascites. A paracentesis was performed with two liters of serous removed. A liver biopsy revealed nodular regenerative hyperplasia and numerous non-caseating granulomas, but without fibrosis or cirrhosis. Bone marrow biopsy also revealed the presence of small non-caseating granulomas. On hospital day 9, blood samples were collected and submitted for blood cultures, and plasma was submitted to Karius® (Redwood City, CA, USA) for a microbial cell-free DNA next-generation sequencing test (KT). The KT results revealed the presence of MAC (254 molecules per microliter). Our laboratory was contacted by the infection prevention team to further speciate the MAC as part of an on-going FDA investigation and continued monitoring of *Mycobacterium intracellulare* subsp. *chimaera* infections associated with water-based heater-cooler devices (HCD) used during cardiopulmonary bypass in surgeries [11]. The blood culture turned positive after 9 days of incubation, and a blood culture fluid sample (UCLA_1007) was obtained directly from a blood bottle and a Positive Blood Culture mNGS was performed as described. In 96 h, Positive Blood Culture mNGS revealed that the MAC in the patient's blood was *M. avium* subsp. *hominissuis*, thus not related to the HCD-associated *M. intracellulare* subsp. *chimaera* outbreak, providing a quick subspecies-level identification for the infection prevention and clinical teams to be able to make relevant decisions on the next course of action. The patient was treated and discharged from the hospital and continued to be followed up as an outpatient.

Case study 2. A 76-year-old woman with a history of polymyositis, coronary artery disease, and hypertrophic obstructive cardiomyopathy status post septal myomectomy and mitral valve replacement presented to our emergency department with substernal chest pain and shortness of breath for one day. She also experienced fatigue and generalized weakness for several days and a 25-pound unintentional weight loss over the past year. The preliminary infectious work-up was unrevealing. However, after four days of incubation, aerobic blood cultures collected from her Port-a-Cath grew beaded Gram-positive rods, which were positive for Ziehl–Neelsen acid-fast staining and were reported to be *Mycobacterium* species. Subsequent aerobic blood cultures collected from the Port-a-Cath which were suspected to be the source of the infection grew the same organism. However, MALDI–TOF MS analysis failed to identify these isolates, which prompted us to perform the Positive Blood Culture mNGS from the blood culture bottle (UCLA_505) and we identified the organism as *Mycolicibacterium iranicum*, an emerging rapid-growing mycobacterial pathogen found to infrequently cause opportunistic infections in immunocompromised patients [12–14]. With this information, the clinical team optimized treatment and the Port-a-Cath was removed. The patient was successfully treated and discharged.

4. Discussion

With BSIs representing a major cause of mortality and morbidity worldwide, blood cultures play a crucial role in diagnosis, but the clinical application is dampened by the long TAT and the detection of only culturable pathogens. In this study, we have developed and validated a shotgun metagenomics test that uses a sample directly from positive blood culture bottles, allowing for the identification of fastidious or slow-growing microorganisms. The test was built based on previously validated WGS tests, which rely on several key marker genes for bacterial and fungal identification. The new test utilizes an open-source

metagenomics CZ-ID platform for the initial analysis to generate the most likely candidate species, which is then used as a reference genome for downstream, confirmation analysis. This approach is innovative because it takes advantage of CZ-ID's agnostic taxonomic calling capability while still relying on the more established and previously validated marker gene-based identification scheme, increasing the confidence in the final results.

In the past decade, pathogen-specific molecular assays targeting the identification of pathogens and some resistance determinants have been developed and tested as alternatives for, or complementary to, blood cultures [15]. Currently, there are several FDA approved tests. Depending on the Gram stain results of positive blood culture, different molecular test panels can be used for the rapid identification of microorganisms [4]. These tests are designed to target a specific set of pathogens and resistance markers, with the number of targets changing and/or increasing as manufacturers improve their technology to meet consumers' needs. However, these technologies are only as good as the targets included in the tests. As an alternative to the rapid molecular diagnostic panels, metagenomics is now being used for the genomic analysis of specific or all genetic material presented in a clinical sample [4,16]. The sequencing of nucleic acids can be based on specific amplicons (amplicon-based metagenomics) usually restricted to the universal gene targets found in bacteria or fungi [17–19] or on the entire genomes (shotgun metagenomics), which allows for the untargeted and unrestricted sequencing of DNA in each sample [16]. Shotgun metagenomics also allows for the detection of resistance markers, virulence genes, and may provide important information on molecular epidemiology [20]. In this study, we have developed and validated a highly accurate shotgun metagenomics-based clinical test for the direct identification of bacteria and fungi from positive blood culture fluids.

Blood cultures are usually incubated for a maximum duration of 5–7 days, with the typical TAT of 48–72 h [21–23]. However, prolonged incubation times for fastidious pathogens, mycobacterium, or fungal organisms may be required. Further, blood cultures have reduced sensitivity in patients who have received antibiotics, where there is an inability to culture certain pathogens, and there are requirements for additional testing and wait times for discerning detected pathogens (strain, virulence factors, and antimicrobial resistance) [23,24]. Novel technologies such as syndromic multiplex polymerase chain reaction (PCR) panels, 16S ribosomal DNA (16S rDNA) Sanger sequencing, and MALDI-TOF MS have decreased turnaround times and subsequently have beneficially impacted patient care [25]. Syndromes including pneumonia, meningitis/encephalitis, and sepsis can be caused by a wide array of pathogens with indistinguishable clinical presentations, often necessitating testing with a large panel of diagnostic assays in an attempt to establish the diagnosis. Despite these efforts, the etiology of infectious diseases remains unknown in up to 60% of cases depending on the clinical syndrome [26,27], which results in delayed or ineffective treatments, increased mortality, and excessive healthcare costs. Agnostic shotgun mNGS is emerging as a single, universal pathogen detection method that is revolutionizing infectious disease diagnostics. mNGS allows for the identification and genomic characterization of microorganisms directly from clinical specimens without any a priori information or knowledge of the likely disease causative etiology [28]. Recent workflow improvements and advancements in mNGS technology and bioinformatics tools have reduced barriers such as high costs, long sequencing times, and sophisticated data analysis that have historically made it impractical to apply these technologies routinely in diagnostic laboratories [27,29]. In this study, we have demonstrated clinical utilization of the Positive Blood Culture mNGS test by providing actionable clinical microbiological data in a timely manner from a blood culture specimen that otherwise would not have been possible. We demonstrated the value of this test in an infection control application, enabling the early release of a patient who otherwise would have spent a considerable amount of time in the hospital before *M. chimaera* was ruled out. A recent study demonstrated superior clinical performance (in terms of TAT, sensitivity, and range of organisms identified) of nanopore targeted sequencing compared to blood culture and PCR followed by Sanger

sequencing [23]. The clinical utilization of NGS is likely to continue increasing as workflows improve and open-source pipelines such as CZ-ID become more widely available.

Since first being described in 2020, the CZ-ID platform has proven to be uniquely suited for the detection of divergent pathogens [24,30–32]. As an open-source cloud-based pipeline, CZ-ID reduces the barrier to entry for mNGS data analysis and enables bench scientists, clinicians, and bioinformaticians to gain insight from mNGS datasets for both known and novel pathogens [9,15]. However, there are two main inherent concerns for using CZ-ID, a cloud-based server for the analysis of sequences that may contain human genomic data. The first concern is uploading human genomic DNA to a cloud server. At the start of this project, we used CLC Genomic Workbench software (Version 12) to interrogate the amount of human genome sequences that may be present in samples sequenced directly from the blood culture bottles, and found that generally human sequence reads were less than 100,000 reads, or <25 million bp (less than 1% of the human genome). The negligible amount of human DNA is due to several reasons. First, a typical blood bottle contains 30 mL of media with 8–10 mL of blood added, which is diluted > 4-fold. Second, the blood bottle is incubated on average 48 h before the growth of bacteria is detected; during this time, blood cells are lysed, and human DNA is rapidly degraded. Third, our method does not sequence very deeply (mean output = 2.9 million reads, maximum output = 12.9 million reads), and thus not many human sequence reads are generated. Last, the CZ-ID pipeline filters out all host genome sequences that are immediately discarded, therefore making sure no human sequences are stored or available for other analyses. Further, all uploaded samples have been completely de-identified. With these factors combined, there is minimal risk that sequences uploaded in the CZ-ID cloud server could compromise a patient's genetic information. The second inherent concern for using CZ-ID is that this is a research use only tool, and to the best of our knowledge, it has not been validated for diagnosis in the clinical setting. It is important to point out that our assay only uses the CZ-ID as a preliminary assessment for microbial species, mainly to identify a suitable reference genome for the downstream analysis; the definitive identification is based on our laboratory-developed NGS bacterial and fungal bioinformatics workflows. Although these workflows can be used for the identification of microbials without the need for CZ-ID, the use of CZ-ID for the initial analysis simplifies the process and increases confidence in the results obtained, with minimal additional steps in the workflow.

At our institution, we have launched this Positive Blood Culture mNGS test for the diagnosis of BSI. Currently, we perform one NGS run per week with a TAT of 5 to 7 days, which is suboptimal for BSI diagnosis. The main challenges of providing a faster TAT for any NGS test are the cost and labor required to perform it at a higher frequency (e.g., three times per week), which is practically prohibitive in most clinical microbiology laboratories. Significantly reduced cost and full automation of NGS workflow are two essential factors to allow this technology to become a primary identification tool instead of a secondary backup method for only rare scenarios of challenging samples. Furthermore, it should be noted that the Positive Blood Culture mNGS test described in this study has several major differences compared to the plasma cell-free DNA test (e.g., the Karius test). First, the plasma cell-free DNA test does not require blood culture and thus provides faster TAT; however, it also suffers from a poor specificity issue with many false or questionable positive results [33,34]. Conversely, the Positive Blood Culture mNGS test requires incubation time and thus has a slower TAT, but with much higher specificity since the blood culture enriches the microbes to a very high abundance which leads to highly-confident results. Second, due to the limitation of short-read sequencing, the taxonomic calling may not be accurate to the species or the subspecies level when insufficient sequence reads are acquired, which is the common scenario in shotgun metagenomics results. In contrast, this is not a problem for Positive Blood Culture mNGS since large amounts of sequence reads can be obtained to provide sufficient definition for species and even sub-species identification. Due to this, our Positive Blood Culture mNGS is especially useful for mycobacterium species and subspecies identification, which is often critical in the clinical setting as demonstrated in the two case

studies described here. In addition, there can be a potential for anti-microbial resistance (AMR) prediction, further justifying utilizing this approach especially for organisms with established AMR genetic markers (e.g., *Mycobacterium tuberculosis*). This is an area we are continuing to explore.

In addition to shotgun mNGS used in this study, pan-bacterial (e.g., based on 16S rRNA) or pan-fungal (e.g., based on ITS) targeted amplicon deep sequencing, also known as targeted NGS (tNGS), is another powerful tool for diagnostic testing [35,36], human microbiome research [37], and ecological microbiome surveillance [38,39]. Compared to shotgun mNGS, tNGS has narrower microbial detection spectrum, but is more cost-effective and less prone to host/background DNA interference, and therefore represents an attractive option for pathogen detection directly from clinical samples. The main limitation of tNGS is its lack of resolution to distinguish certain closely related species; however, clinical validation studies have shown that tNGS identifies most clinically significant pathogens with satisfactory performance [35,36,40].

This study has several limitations. First, the use of a privately funded web portal for the analysis of clinical specimens poses unique challenges: there is no guarantee the portal will always be available, and upgrades or changes may impact its performance. Second, the current TAT for this test is 5–7 days, which is not ideal as discussed above. Third, only a limited species and samples were validated, and no polymicrobial positive blood culture sample was tested in this study. Fourth, similar to blood culture results, caution should be taken to differentiate colonization/contamination from true infection. Importantly, extraneous sources of nucleic acid contamination in the blood bottle medium had been reported to cause false positive results in a commercial direct blood culture molecular identification panel assay (<https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfRes/res.cfm?ID=181223> (accessed on 21 April 2023); <https://asm.org/ASM/media/Policy-and-Advocacy/BICx-contaminating-DNA-FINAL.pdf> (accessed on 21 April 2023)). Therefore, stringent quality control and correlation with other laboratory results (e.g., Gram stain) and clinical presentations are needed to safeguard the accuracy of the test. Due to these limitations, a post-launch test validation is still ongoing.

In conclusion, we developed and validated an innovative NGS test for the accurate species identification directly from the fluid of a positive blood culture bottle, and demonstrated its clinical utility especially for anaerobes and mycobacteria that are either fastidious, slow growing, or unusual. Although applicable only in limited settings, the Positive Blood Culture mNGS test provides an incremental improvement in solving the unmet clinical needs for the diagnosis of challenging BSIs.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/microorganisms11051259/s1>, Table S1: Detailed Conventional Microbiological Results.

Author Contributions: Conceptualization, E.K. and S.Y.; Methodology, E.K. and S.Y.; Software, E.K.; Validation, E.K.; Formal analysis, E.K.; Investigation, E.K. and S.Y.; Resources, S.Y.; Data curation, E.K.; Writing—original draft, E.K.; Writing—review & editing, S.Y.; Visualization, E.K. and S.Y.; Supervision, S.Y.; Funding acquisition, S.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the UCLA Department of Pathology and Laboratory Medicine.

Data Availability Statement: The sequencing data of this study is available upon request.

Acknowledgments: We thank the UCLA Clinical Microbiology NGS Team for their technical assistance.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Peker, N.; Couto, N.; Sinha, B.; Rossen, J. Diagnosis of bloodstream infections from positive blood cultures and directly from blood samples: Recent developments in molecular approaches. *Clin. Microbiol. Infect.* **2018**, *24*, 944–955. [[CrossRef](#)] [[PubMed](#)]
2. Rutanga, J.P.; Nyirahabimana, T. Clinical Significance of Molecular Diagnostic Tools for Bacterial Bloodstream Infections: A Systematic Review. *Interdiscip. Perspect. Infect. Dis.* **2016**, *2016*, 6412085. [[CrossRef](#)] [[PubMed](#)]
3. Opota, O.; Croxatto, A.; Prod'Hom, G.; Greub, G. Blood culture-based diagnosis of bacteraemia: State of the art. *Clin. Microbiol. Infect.* **2015**, *21*, 313–322. [[CrossRef](#)] [[PubMed](#)]
4. Loonen, A.J.M.; Wolffs, P.F.G.; Bruggeman, C.A.; Brule, A.J.C.V.D. Developments for improved diagnosis of bacterial bloodstream infections. *Eur. J. Clin. Microbiol. Infect. Dis.* **2014**, *33*, 1687–1702. [[CrossRef](#)]
5. Clinical and Laboratory Standards Institute (CLSI). *Principles and Procedures for Blood Cultures; Approved Guideline*; CLSI document M47-A; Clinical and Laboratory Standards Institute: Wayne, PA, USA, 2007.
6. Weinstein, M.P.; Towns, M.L.; Quartey, S.M.; Mirrett, S.; Reimer, L.G.; Parmigiani, G.; Reller, L.B. The Clinical Significance of Positive Blood Cultures in the 1990s: A Prospective Comprehensive Evaluation of the Microbiology, Epidemiology, and Outcome of Bacteremia and Fungemia in Adults. *Clin. Infect. Dis.* **1997**, *24*, 584–602. [[CrossRef](#)]
7. Kamau, E.; Maliksi, E.; Kwan, N.; Garner, O.B.; Yang, S. *Catabacter hongkongensis* bacteremia identified by direct metagenomic sequencing of positive blood culture fluid, first case report in the US. *Anaerobe* **2021**, *71*, 102421. [[CrossRef](#)]
8. Price, T.K.; Realegeno, S.; Mirasol, R.; Tsan, A.; Chandrasekaran, S.; Garner, O.B.; Yang, S. Validation, Implementation, and Clinical Utility of Whole Genome Sequence-Based Bacterial Identification in the Clinical Microbiology Laboratory. *J. Mol. Diagn.* **2021**, *23*, 1468–1477. [[CrossRef](#)]
9. Kalantar, K.L.; Carvalho, T.; De Bourcy, C.F.A.; Dimitrov, B.; Dingle, G.; Egger, R.; Han, J.; Holmes, O.B.; Juan, Y.-F.; King, R.; et al. IDseq—An open source cloud-based pipeline and analysis service for metagenomic pathogen detection and monitoring. *Gigascience* **2020**, *9*, giaa111. [[CrossRef](#)]
10. Salem-Bango, Z.; Price, T.K.; Chan, J.L.; Chandrasekaran, S.; Garner, O.B.; Yang, S. Fungal Whole-Genome Sequencing for Species Identification: From Test Development to Clinical Utilization. *J. Fungi* **2023**, *9*, 183. [[CrossRef](#)]
11. Xu, K.; Finn, L.E.; Geist, R.L.; Prestel, C.; Moulton-Meissner, H.; Kim, M.; Stacey, B.; McAllister, G.A.; Gable, P.; Kamali, T.; et al. *Mycobacterium chimaera* infections among cardiothoracic surgery patients associated with heater-cooler devices—Kansas and California, 2019. *Infect. Control. Hosp. Epidemiol.* **2022**, *43*, 1333–1338. [[CrossRef](#)]
12. Shojaei, H.; Daley, C.; Gitti, Z.; Hashemi, A.; Heidarieh, P.; Moore, E.R.B.; Naser, A.D.; Russo, C.; van Ingen, J.; Tortoli, E. *Mycobacterium iranicum* sp. nov., a rapidly growing scotochromogenic species isolated from clinical specimens on three different continents. *Int. J. Syst. Evol. Microbiol.* **2013**, *63 Pt 4*, 1383–1389. [[CrossRef](#)]
13. Lapierre, S.G.; Toro, A.; Drancourt, M. *Mycobacterium iranicum* bacteremia and hemophagocytic lymphohistiocytosis: A case report. *BMC Res. Notes* **2017**, *10*, 372. [[CrossRef](#)]
14. Ranson, E.L.; Tsevat, R.K.; von Bredow, B.; Kamau, E.; Yang, S.; Prabaker, K.K. Catheter-Related Bloodstream Infection Caused by *Mycobacterium iranicum*, California, USA. *Emerg. Infect. Dis.* **2023**, *29*, 217–219. [[CrossRef](#)]
15. Dubourg, G.; Raoult, D. Emerging methodologies for pathogen identification in positive blood culture testing. *Expert Rev. Mol. Diagn.* **2016**, *16*, 97–111. [[CrossRef](#)]
16. Gyarmati, P.; Kjellander, C.; Aust, C.; Song, Y.; Öhrmalm, L.; Giske, C.G. Metagenomic analysis of bloodstream infections in patients with acute leukemia and therapy-induced neutropenia. *Sci. Rep.* **2016**, *6*, 23532. [[CrossRef](#)]
17. Faria, M.; Conly, J.; Surette, M. The development and application of a molecular community profiling strategy to identify polymicrobial bacterial DNA in the whole blood of septic patients. *BMC Microbiol.* **2015**, *15*, 215. [[CrossRef](#)]
18. Decuypere, S.; Meehan, C.J.; Van Puyvelde, S.; De Block, T.; Maltha, J.; Palpouguini, L.; Tahita, M.; Tinto, H.; Jacobs, J.; Deborggraeve, S. Diagnosis of Bacterial Bloodstream Infections: A 16S Metagenomics Approach. *PLoS Negl. Trop. Dis.* **2016**, *10*, e0004470. [[CrossRef](#)]
19. Sabat, A.J.; van Zanten, E.; Akkerboom, V.; Wisselink, G.; van Slochteren, K.; de Boer, R.F.; Hendrix, R.; Friedrich, A.W.; Rossen, J.W.A.; Kooistra-Smid, A.M.D. Targeted next-generation sequencing of the 16S-23S rRNA region for culture-independent bacterial identification—Increased discrimination of closely related species. *Sci. Rep.* **2017**, *7*, 3434. [[CrossRef](#)]
20. Couto, N.; Schuele, L.; Raangs, E.C.; Machado, M.P.; Mendes, C.I.; Jesus, T.F.; Chlebowicz, M.; Rosema, S.; Ramirez, M.; Carriço, J.A.; et al. Critical steps in clinical shotgun metagenomics for the concomitant detection and typing of microbial pathogens. *Sci. Rep.* **2018**, *8*, 13767. [[CrossRef](#)]
21. Didelot, X.; Bowden, R.; Wilson, D.J.; Peto, T.E.A.; Crook, D.W. Transforming clinical microbiology with bacterial genome sequencing. *Nat. Rev. Genet.* **2012**, *13*, 601–612. [[CrossRef](#)]
22. Laupland, K.B.; Valiquette, L. The changing culture of the microbiology laboratory. *Can. J. Infect. Dis. Med. Microbiol.* **2013**, *24*, 125–128. [[CrossRef](#)] [[PubMed](#)]
23. Fu, Y.; Chen, Q.; Xiong, M.; Zhao, J.; Shen, S.; Chen, L.; Pan, Y.; Li, Z.; Li, Y. Clinical Performance of Nanopore Targeted Sequencing for Diagnosing Infectious Diseases. *Microbiol. Spectr.* **2022**, *10*, e0027022. [[CrossRef](#)] [[PubMed](#)]
24. Yang, S.; Rothman, R.E. PCR-based diagnostics for infectious diseases: Uses, limitations, and future applications in acute-care settings. *Lancet Infect. Dis.* **2004**, *4*, 337–348. [[CrossRef](#)] [[PubMed](#)]

25. Tan, K.E.; Ellis, B.C.; Lee, R.; Stamper, P.D.; Zhang, S.X.; Carroll, K.C. Prospective Evaluation of a Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry System in a Hospital Clinical Microbiology Laboratory for Identification of Bacteria and Yeasts: A Bench-by-Bench Study for Assessing the Impact on Time to Identification and Cost-Effectiveness. *J. Clin. Microbiol.* **2012**, *50*, 3301–3308. [[CrossRef](#)]
26. Glaser, C.A.; Honarmand, S.; Anderson, L.J.; Schnurr, D.P.; Forghani, B.; Cossen, C.K.; Schuster, F.L.; Christie, L.J.; Tureen, J.H. Beyond Viruses: Clinical Profiles and Etiologies Associated with Encephalitis. *Clin. Infect. Dis.* **2006**, *43*, 1565–1577. [[CrossRef](#)]
27. Schlager, R.; Chiu, C.Y.; Miller, S.; Procop, G.W.; Weinstock, G.; the Professional Practice Committee and Committee on Laboratory Practices of the American Society for Microbiology; the Microbiology Resource Committee of the College of American Pathologists. Validation of Metagenomic Next-Generation Sequencing Tests for Universal Pathogen Detection. *Arch. Pathol. Lab. Med.* **2017**, *141*, 776–786. [[CrossRef](#)]
28. Simner, P.J.; Miller, S.; Carroll, K.C. Understanding the Promises and Hurdles of Metagenomic Next-Generation Sequencing as a Diagnostic Tool for Infectious Diseases. *Clin. Infect. Dis.* **2017**, *66*, 778–788. [[CrossRef](#)]
29. Flygare, S.; Simmon, K.; Miller, C.; Qiao, Y.; Kennedy, B.; Di Sera, T.; Graf, E.H.; Tardif, K.D.; Kapusta, A.; Rynearson, S.; et al. Taxonomer: An interactive metagenomics analysis portal for universal pathogen detection and host mRNA expression profiling. *Genome Biol.* **2016**, *17*, 111. [[CrossRef](#)]
30. Han, Y.; Jia, Z.; Shi, J.; Wang, W.; He, K. The active lung microbiota landscape of COVID-19 patients through the metatranscriptome data analysis. *Bioimpacts* **2022**, *12*, 139–146. [[CrossRef](#)]
31. Ramachandran, P.S.; Ramesh, A.; Creswell, F.V.; Wapniarski, A.; Narendra, R.; Quinn, C.M.; Tran, E.B.; Rutakingirwa, M.K.; Bangdiwala, A.S.; Kagimu, E.; et al. Integrating central nervous system metagenomics and host response for diagnosis of tuberculosis meningitis and its mimics. *Nat. Commun.* **2022**, *13*, 1675. [[CrossRef](#)]
32. Bohl, J.A.; Lay, S.; Chea, S.; Ah Yong, V.; Parker, D.M.; Gallagher, S.; Fintzi, J.; Man, S.; Ponce, A.; Sreng, S.; et al. Discovering disease-causing pathogens in resource-scarce Southeast Asia using a global metagenomic pathogen monitoring system. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2115285119. [[CrossRef](#)]
33. Hogan, C.A.; Yang, S.; Garner, O.B.; Green, D.A.; Gomez, C.A.; Bard, J.D.; Pinsky, B.A.; Banaei, N. Clinical Impact of Metagenomic Next-Generation Sequencing of Plasma Cell-Free DNA for the Diagnosis of Infectious Diseases: A Multicenter Retrospective Cohort Study. *Clin. Infect. Dis.* **2021**, *72*, 239–245. [[CrossRef](#)]
34. Babady, N.E. Clinical Metagenomics for Bloodstream Infections: Is the Juice Worth the Squeeze? *Clin. Infect. Dis.* **2021**, *72*, 246–248. [[CrossRef](#)]
35. Larkin, P.M.; Lawson, K.L.; Contreras, D.A.; Le, C.Q.; Trejo, M.; Realegeno, S.; Hilt, E.E.; Chandrasekaran, S.; Garner, O.B.; Fishbein, G.A.; et al. Amplicon-Based Next-Generation Sequencing for Detection of Fungi in Formalin-Fixed, Paraffin-Embedded Tissues: Correlation with Histopathology and Clinical Applications. *J. Mol. Diagn.* **2020**, *22*, 1287–1293. [[CrossRef](#)]
36. Culbreath, K.; Melanson, S.; Gale, J.; Baker, J.; Li, F.; Saebo, O.; Kommedal, O.; Contreras, D.; Garner, O.B.; Yang, S. Validation and Retrospective Clinical Evaluation of a Quantitative 16S rRNA Gene Metagenomic Sequencing Assay for Bacterial Pathogen Detection in Body Fluids. *J. Mol. Diagn.* **2019**, *21*, 913–923. [[CrossRef](#)]
37. Pannaraj, P.S.; Li, F.; Cerini, C.; Bender, J.M.; Yang, S.; Rollie, A.; Adisetiyo, H.; Zabih, S.; Lincez, P.J.; Bittinger, K.; et al. Association Between Breast Milk Bacterial Communities and Establishment and Development of the Infant Gut Microbiome. *JAMA Pediatr.* **2017**, *171*, 647–654. [[CrossRef](#)]
38. Tan, H.; Liu, T.; Yu, Y.; Tang, J.; Jiang, L.; Martin, F.M.; Peng, W. Morel Production Related to Soil Microbial Diversity and Evenness. *Microbiol. Spectr.* **2021**, *9*, e0022921. [[CrossRef](#)]
39. Yu, Y.; Liu, T.; Liu, L.; Chen, Y.; Tang, J.; Peng, W.; Tan, H. Application of the mushroom volatile 1-octen-3-ol to suppress a morel disease caused by *Paecilomyces penicillatus*. *Appl. Microbiol. Biotechnol.* **2022**, *106*, 4787–4799. [[CrossRef](#)]
40. Hilt, E.E.; Ferrieri, P. Next Generation and Other Sequencing Technologies in Diagnostic Microbiology and Infectious Diseases. *Genes* **2022**, *13*, 1566. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.