



# Prospective Comparison Between Shotgun Metagenomics and Sanger Sequencing of the 16S rRNA Gene for the Etiological Diagnosis of Infections

## OPEN ACCESS

### Edited by:

Silvia Turrone,  
University of Bologna, Italy

### Reviewed by:

Marko Verce,  
Catholic University of Louvain,  
Belgium  
Laure Flurin,  
Mayo Clinic, United States  
Erley Ferlope Lizarazo,  
University Medical Center Groningen,  
Netherlands

### \*Correspondence:

Paul-Louis Woerther  
paul-louis.woerther@aphp.fr

† These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Infectious Agents and Disease,  
a section of the journal  
Frontiers in Microbiology

Received: 20 August 2021

Accepted: 14 March 2022

Published: 06 April 2022

### Citation:

Lamoureux C, Surgers L,  
Fihman V, Gricourt G, Demontant V,  
Trawinski E, N'Debi M, Gomart C,  
Royer G, Launay N, Le Glaunec J-M,  
Wemmer C, La Martire G, Rossi G,  
Lepeule R, Pawlotsky J-M,  
Rodriguez C and Woerther P-L (2022)  
Prospective Comparison Between  
Shotgun Metagenomics and Sanger  
Sequencing of the 16S rRNA Gene  
for the Etiological Diagnosis  
of Infections.  
Front. Microbiol. 13:761873.  
doi: 10.3389/fmicb.2022.761873

Claudie Lamoureux<sup>1,2,3</sup>, Laure Surgers<sup>4,5</sup>, Vincent Fihman<sup>1,6</sup>, Guillaume Gricourt<sup>7</sup>,  
Vanessa Demontant<sup>7</sup>, Elisabeth Trawinski<sup>7</sup>, Melissa N'Debi<sup>7</sup>, Camille Gomart<sup>1</sup>,  
Guilhem Royer<sup>1</sup>, Nathalie Launay<sup>1</sup>, Jeanne-Marie Le Glaunec<sup>1</sup>, Charlotte Wemmer<sup>8</sup>,  
Giulia La Martire<sup>8</sup>, Geoffrey Rossi<sup>8</sup>, Raphaël Lepeule<sup>8</sup>, Jean-Michel Pawlotsky<sup>1,5</sup>,  
Christophe Rodriguez<sup>1,5,7†</sup> and Paul-Louis Woerther<sup>1,6,8\*†</sup>

<sup>1</sup> Microbiology Unit, Department of Diagnostic, Prevention and Treatment of Infections, Henri Mondor Hospital, AP-HP, University of Paris-Est-Créteil, Créteil, France, <sup>2</sup> Department of Bacteriology, Virology, Hospital Hygiene, and Parasitology-Mycology, Brest University Hospital, Brest, France, <sup>3</sup> Univ Brest, INSERM, EFS, UMR 1078, GGB, Brest, France, <sup>4</sup> GHU AP-HP Sorbonne Université, Service des Maladies Infectieuses et Tropicales, Hôpital Saint-Antoine, Paris, France, <sup>5</sup> INSERM U955, IMRB Institute, University of Paris-Est Créteil, Créteil, France, <sup>6</sup> EA 7380 Dynamyc, EnVA, UPEC, University of Paris-Est Créteil, Créteil, France, <sup>7</sup> NGS Platform, Henri Mondor Hospital, AP-HP, and IMRB Institute, University of Paris-Est-Créteil, Créteil, France, <sup>8</sup> Antimicrobial Stewardship Unit, Diagnostic, Prevention and Treatment of Infections Department, Henri Mondor Hospital, AP-HP, University of Paris-Est-Créteil, Créteil, France

Bacteriological diagnosis is traditionally based on culture. However, this method may be limited by the difficulty of cultivating certain species or by prior exposure to antibiotics, which justifies the resort to molecular methods, such as Sanger sequencing of the 16S rRNA gene (Sanger 16S). Recently, shotgun metagenomics (SMg) has emerged as a powerful tool to identify a wide range of pathogenic microorganisms in numerous clinical contexts. In this study, we compared the performance of SMg to Sanger 16S for bacterial detection and identification. All patients' samples for which Sanger 16S was requested between November 2019 and April 2020 in our institution were prospectively included. The corresponding samples were tested with a commercial 16S semi-automated method and a semi-quantitative pan-microorganism DNA- and RNA-based SMg method. Sixty-seven samples from 64 patients were analyzed. Overall, SMg was able to identify a bacterial etiology in 46.3% of cases (31/67) vs. 38.8% (26/67) with Sanger 16S. This difference reached significance when only the results obtained at the species level were compared (28/67 vs. 13/67). This study provides one of the first evidence of a significantly better performance of SMg than Sanger 16S for bacterial detection at the species level in patients with infectious diseases for whom culture-based methods have failed. This technology has the potential to replace Sanger 16S in routine practice for infectious disease diagnosis.

**Keywords:** shotgun metagenomics, molecular diagnostic, pathogen identification, Sanger sequencing of the 16S rRNA gene, microbial documentation

## INTRODUCTION

Molecular biology-based approaches, including next-generation sequencing (NGS), have changed the face of biological diagnosis in many medical areas (Couto et al., 2018; Simner et al., 2018). However, culture-based methods are still considered as the “gold standard” in Bacteriology, not only to identify pathogenic microorganisms, but also to determine their susceptibility to antimicrobial agents. In case of prior administration of antimicrobial drugs or when fastidious microorganisms are involved, culture-based approaches can yield falsely negative results. In that case, DNA or RNA detection, which could be less impacted, would be more appropriate. Amplification of the universal bacterial 16S rRNA gene and Sanger sequencing of the PCR products (Sanger 16S) is by far the most widely used molecular method for bacterial detection in clinical microbiology laboratories. However, the Sanger 16S approach has biases and limitations (Church et al., 2020). First this method is poorly adapted to detect more than one bacterial species per primers pair, which may be limiting in case of polymicrobial infections (Kommedal et al., 2009). Second, like for any targeted methods, the choice of the primers used for the PCR reaction can significantly impact the sensitivity and specificity of the amplification step. Third, discrimination at the species level is difficult for certain bacterial genera, according to the length of the DNA amplicon (e.g., *Staphylococci*, *Enterococci*) (Church et al., 2020), despite the fact that this information is essential for infectious gateway documentation and adequate antibiotic treatment prescription. Last, the prediction of antibiotic resistance is impossible (Abayasekara et al., 2017; Schriefer et al., 2018). These limitations make it necessary to evaluate alternative methods. Among them, NGS-based approaches, as shotgun metagenomics (SMg) or amplification of the universal bacterial 16S rRNA gene and NGS sequencing of the PCR products (NGS 16S), appear to be particularly promising. Focusing on SMg, recent studies have demonstrated its utility for the etiological diagnosis of infections (Miao et al., 2018; Filkins et al., 2020; Gu et al., 2021; Miller and Chiu, 2021), including meningitidis/encephalitis (Wilson et al., 2019; Rodino et al., 2020), necrotizing soft-tissue infections (Rodriguez et al., 2020c), pneumonia (Hilton et al., 2016; Zhou et al., 2021), bloodstream infections (Grumaz et al., 2016), bone and joint infections (Street et al., 2017; Thoendel et al., 2018; Gamie et al., 2021; Goswami et al., 2021) or fever with unknown origin (Fu et al., 2022). Moreover, because SMg has the capacity to generate the full-length sequence of bacterial genomes (unlike 16S approaches), this method could be used for the study of antimicrobial susceptibility and molecular typing. For all these reasons, Sanger 16S could be replaced by SMg, with the expectation that cost will continue to decrease with the advent of new technological solutions (Forbes et al., 2017).

The aim of this work was to compare the performance of SMg and Sanger 16S for bacterial detection in routine practice. We conducted a prospective study comparing two molecular methods for bacterial identification: a broadly used semi-automated targeted approach based on amplification of the 16S rRNA gene and Sanger sequencing of the PCR products,

and a semi-quantitative pan-microorganism DNA- and RNA-based SMg method.

## MATERIALS AND METHODS

### Study Population and Samples

All patients' samples, negative in culture at the time of inclusion, for which Sanger 16S analysis was requested and sent to our laboratory between November 1, 2019 and April 30, 2020, were included in this prospective, comparative, observational and non-inferiority study. Culture negativity of the sample was verified at the time of inclusion. The following sociodemographic and clinical parameters were recorded: age, sex, and antibiotic administration before or at the time of sample processing. The samples were prospectively processed with Sanger 16S and SMg approaches, both of which are available for routine care. The final diagnosis of infection was made by clinicians according to the clinical and biological arguments.

This study protocol was approved by the research ethics committee of the Henri Mondor University Hospital (registration number 00011558). According to French law, complete information about the study was given to each patient. This observational study was conducted under the “Méthodologie de Référence 004” (registration number 3312300420) and was declared to the “Commission Nationale de l'Informatique et des Libertés.” The Sanger 16S approach was performed first on the samples. Samples were stored at +4°C before Sanger 16S then stored at -20°C before SMg.

### Amplification of the 16S rRNA Gene and Sanger Sequencing of the PCR Products

The semi-automated method UMD-SelectNA CE-IVD kit (Molzym GmbH, Germany) was used for the 16S approach, according to the manufacturer's recommendations. Briefly, 1 mL of sample was pre-treated with protease and a chaotropic buffer lysing human cells, then subjected to DNase to degrade human nucleic acids. Bacterial DNA was subsequently extracted after a proteinase K treatment by means of a magnetic beads-driven procedure on the Arrow instrument (Norddiag, Oslo, Norway). A control (DNA fragment, sequence not provided by the manufacturer) was added to all samples during DNA extraction to detect the presence of potential PCR inhibitors. Extracts were stored at +4°C until real-time PCR analysis (storage at -20°C if real-time PCR not performed immediately after extraction). PCR amplifications targeting the hypervariable region V3-V4 of the bacterial 16S rRNA gene (primer's sequences not provided by the manufacturer) and the control (primers' sequences not provided by the manufacturer) were performed for each sample. The reaction mix comprised 8.0 µL of primers targeting 16S rRNA gene (2.5x concentrated solution), 2.0 µL of DNA staining solution, 6.0 µL of free DNA water, 0.8 µL of Taq DNA polymerase (MolTaq 16S, Molzym GmbH, Germany). A final reaction volume of 20 µL (16 µL of mastermix and 4 µL of DNA extract) was used for the PCR reaction. Real-time PCR was performed with a LightCycler 480 instrument (Roche, Basel, Switzerland) under the following conditions: 95°C for

1 min, 40 cycles at 95°C for 5 s, 55°C for 5 s, and 72°C for 25 s, followed by a melting curve analysis (65–95°C). Samples with a melting temperature value between 86°C (and 90°C were considered positive. PCR products from the positive samples were purified with the MinElute PCR purification kit (Qiagen, Hilden, Germany). Sequencing primers included in the UMD-SelectNA CE-IVD kit (one primer pair for Gram-positive and one for Gram-negative; primer's sequences not provided by the manufacturer) were used for Sanger sequencing on an ABI 3130xl instrument (Thermo Fisher Scientific, Waltham, United States). Sequences were initially aligned to the SepsitTest BLAST database<sup>1</sup> and subsequently, if no bacterial species were identified, to the Quick BioInformatic Phylogeny of Prokaryotes database.<sup>2</sup> Sequences with  $\geq 97$ –99% or  $\geq 99$ % identity to the database were assigned to the genus or species level, respectively, according to the manufacturer's instructions (Wellinghausen et al., 2009). In certain cases, when the database assigned more than one species belonging to the same species group with  $\geq 99$ % identity, the bacterium was identified at the group level. When assignment was  $< 97$ %, no result was provided.

## Shotgun Metagenomics

The SMg strategy used in this study consists of an in-house semi-quantitative pan-microorganism DNA- and RNA-based method, called MetaMIC, that has been previously described (Rodriguez et al., 2020c). This approach has received ISO 15189 certification and is routinely used in our laboratory for the etiological diagnosis of complex infectious syndromes (Rodriguez et al., 2020a,b; Courbin et al., 2022). Briefly, 400  $\mu$ L of samples were pre-treated with a combination of chemical cell disruption and bead homogenization. Nucleic acids were then extracted using the automated QIASymphony instrument with the DSP DNA Mini kit (Qiagen, Hilden, Germany). Extracts were stored at +4°C until libraries preparation (storage at –20°C if libraries not performed immediately after extraction). DNA libraries were prepared using 5  $\mu$ L extract at 0.2 ng/ $\mu$ L and Nextera XT DNA kit (Illumina, San Diego, United States). RNA libraries were prepared in parallel using 10  $\mu$ L extract at 10 ng/ $\mu$ L and RNA Human RiboZero TruSeq Stranded Total RNA Library Prep Kit (Illumina, San Diego, United States). The quality and quantity of each library were evaluated by means of a D1000 ScreenTape on a TapeStation (Agilent, Santa Clara, United States) and the Quant-it dsDNA Assay kit on Varioskan LUX instrument (Thermo Fisher Scientific, Waltham, United States), respectively. Then, DNA and RNA libraries were normalized to equal concentrations (final concentration: 2.1 pM) before pooling, denaturation and paired-end sequencing using the NextSeq 500/550 High Output Kit v2.5, 2  $\times$  150 bp on a NextSeq500 instrument (Illumina, San Diego, United States). An environmental control (molecular grade water) and a positive control including Gram-positive and -negative bacteria and fungi (ZymoBIOMICS Microbial Community Standards D6300, Zymo Research, Irvine, United States) were included in each SMg run and processed through the entire protocol to evaluate the performance of SMg. A minimal number of reads (10 million)

was required for DNA or RNA libraries, with the exception of low cell matrices (i.e., cerebrospinal fluid,...). After sequencing, fastq files from both libraries (DNA and RNA sequence data) were analyzed by means of our in-house and patented MetaMIC software (V2.2.1), composed of a mosaic of modules.<sup>3</sup> The patent has been registered at European Patent Office at PCT application number: PCT/FR2020/052193 (Rodriguez et al., 2020d). All reads with Phred scores lower than 20 for at least one nucleotide were removed and a Phred score of 30 (Q30) had to be reached for at least 75% of reads for one sample according to the manufacturer's recommendations. Human reads were removed using the hg19 database. Identification of non-human reads was performed by means of a database dedicated to MetaMIC (derived from the NCBI nt database, including bacterial, viral and fungal reads but cleaned of mammals, insects, plants, and synthetic plasmids reads). All reads were grouped by taxonomic number and their genetic distances from each other were checked to rule out aberrant reads. A run was considered valid if each microorganism of the positive control was detected. For each sample, DNA libraries were first used for bacterial detection. A sequence background noise was evaluated by calculating a limit of "blank" and a limit of detection (LOD), based on the results of the environmental control. Each identification whose count had a value greater than the LOD was compared to the count of the reads obtained in the environmental control. Only the identifications superior to the LOD and the environmental control were kept as positive. Results were considered positive if one (or more) bacterium compatible with the suspected infectious process was identified from sterile anatomical sites (i.e., bone and joint, cardiovascular or central nervous system samples). In case of samples from sites normally colonized by commensal bacteria (i.e., pulmonary, intra-abdominal, genito-urinary, and skin and soft tissue samples), only classical pathogens were reported. Strictly pathogenic bacteria (e.g., *Mycobacterium tuberculosis*) were considered regardless of the number of DNA reads. Possible pathogenic bacteria with a low bacterial load (e.g., *Staphylococcus aureus*) were considered if corresponding reads were also found in RNA libraries (**Supplementary Table 1**). In a second time, results were interpreted by Infectious Disease specialists as "probable sample contaminant" or "pathogen," based on international recommendations and the combination of clinical, radiological, biological, and therapeutic evidences. In case of positive identification, semi-quantification was determined using the ratio between bacterial and human reads. Bacterial reads are publicly available on GenBank (SAMN16686554–SAMN16686620).

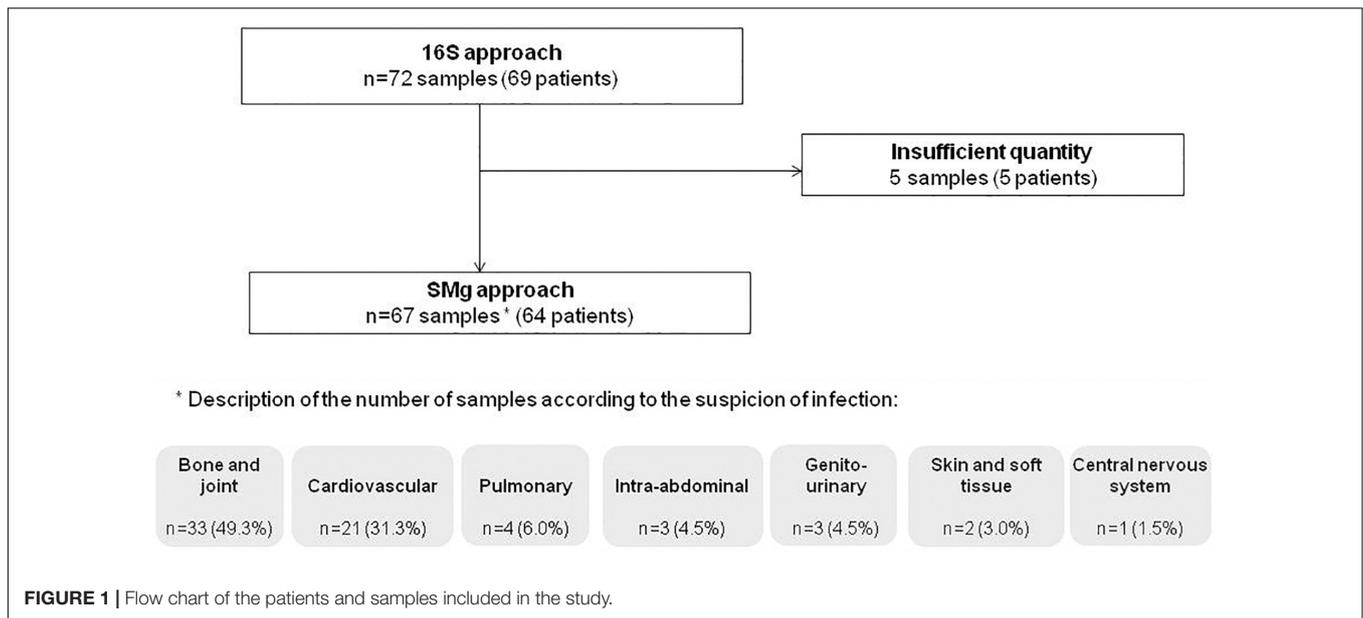
## Statistical Analysis

The Mc Nemar test was used to compare the Sanger 16S and SMg methods with an alpha risk considered significant when less than 0.05. As both methods are not considered as molecular "gold standard" for bacterial identification, sensitivities of the Sanger 16S and SMg approaches were evaluated by comparing the results provided by one single method with those obtained by the sum of the information generated by both methods.

<sup>1</sup>www.sepsitest-blast.net

<sup>2</sup>https://umr5558-proka.univ-lyon1.fr/lebibi/lebibi-legacy.cgi

<sup>3</sup>https://gitlab.com/mndebi/metamic



## RESULTS

### Patient and Sample Characteristics

After exclusion of five samples due to a lack of material, 67 samples from 64 patients were included in the study. The study population was composed of 57.8% of males (37/64) and the median age was 62.5 years (range: 19–96). Indications for Sanger 16S prescription were suspicion of bone and joint infection (49.3%; 33/67), cardiovascular infection (31.3%; 21/67), pulmonary infection (6.0%; 4/67), intra-abdominal infection (4.5%; 3/67), genito-urinary infection (4.5%; 3/67), skin and soft tissue infection (3.0%; 2/67), and central nervous system infection (1.5%; 1/67) (**Figure 1**). Previous exposure to antibiotics (1 month before and/or at the time of sampling) was reported in 62.5% of patients (40/59) [data not available for 7.8% of them (5/64)].

### Amplification of the 16S rRNA Gene and Sanger Sequencing of the PCR Products

A total of 38.8% of samples (26/67) were positive by Sanger 16S, of which 50.0% (13/26) were identified at the genus level and 50.0% (13/26) at the species level (**Table 1**). Among these 26 bacteria identified, 73.1% (19/26) were Gram-positive, 19.2% (5/26) were Gram-negative and 7.7% (2/26) were intracellular (**Figure 2**). No mycobacteria were found. The Sanger 16S approach was uninterpretable for three samples (control failure because of PCR inhibitors, despite dilution of the extracted DNA) (**Supplementary Table 1**).

### Shotgun Metagenomics

Phred score (Q30) for all sample and positive control results for all experiments were in keeping with the specifications of the method for the SMg run validation. The median depth of SMg sequencing was 29,159,912 DNA reads (range:

435,424–159,677,808) and 19,426,688 RNA reads (range: 1,970–49,574,336) per sample. The total number of reads for DNA library was lower than 10 million for three samples (**Supplementary Table 1**). In total, 46.3% of the samples (31/67) were positive by SMg, of which 9.7% (3/31) were identified at the genus level and 90.3% (28/31) at the species level (**Table 1**). Among the 31 bacteria detected, 67.6% (21/31) were Gram-positive, 19.4% (6/31) were Gram-negative, 6.5% (2/31) were intracellular, and 6.5% (2/31) were mycobacteria (**Figure 2**). Semi-quantification by metagenomics provided high bacterial loads (ratio >  $10^{-2}$ ) for 3.3% (1/31) of cases, intermediate bacterial loads (ratio =  $10^{-2}$ – $10^{-5}$ ) for 41.9% (13/31) of cases, and low bacterial loads (ratio <  $10^{-5}$ ) for 54.8% (17/31) of cases (**Supplementary Table 1**). No viral or fungal DNA or RNA reads were detected in any of the samples.

### Comparison of Sanger Sequencing of the 16S rRNA Gene vs. Shotgun Metagenomics

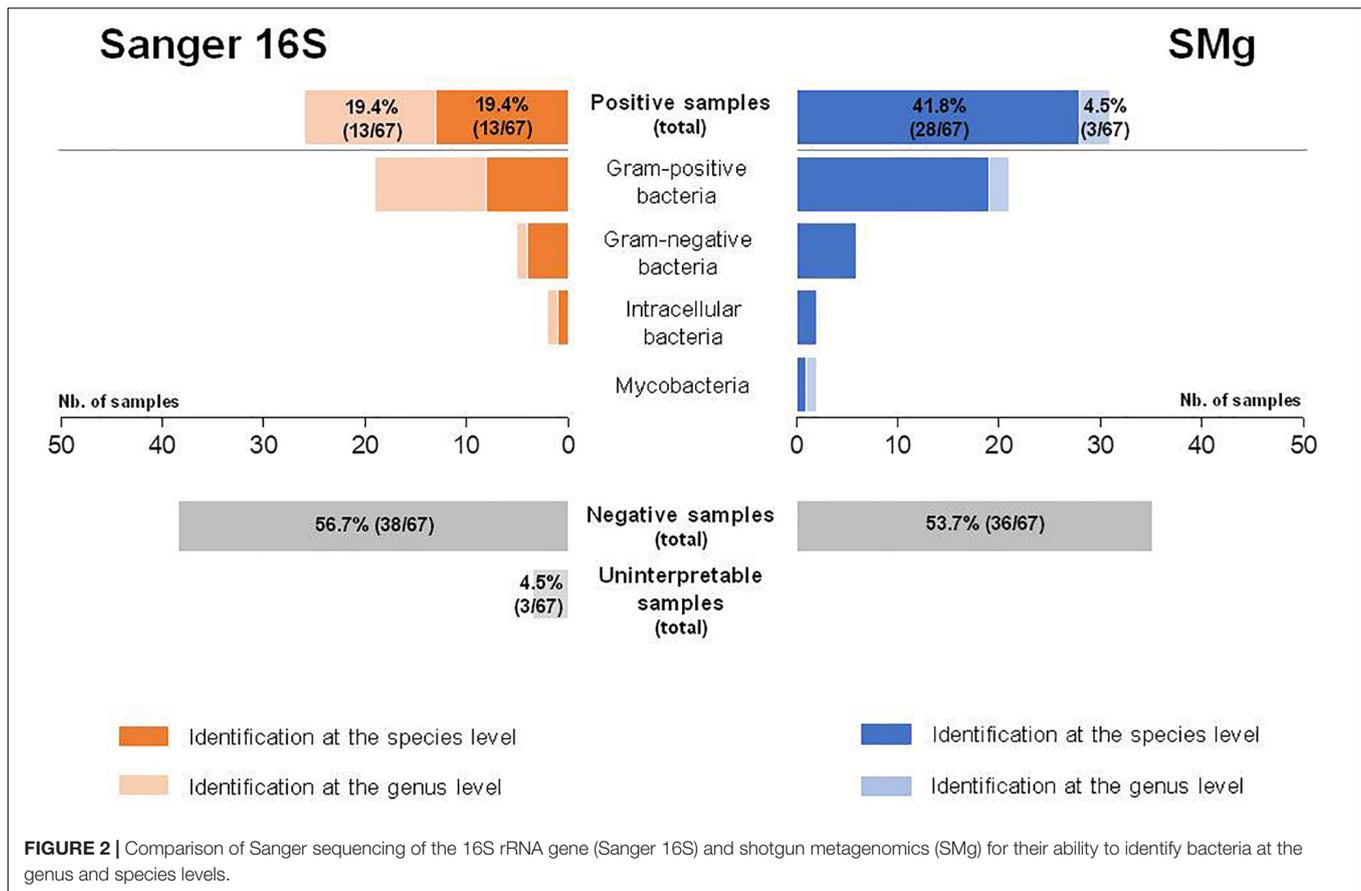
Bacterial identification was consistent in all the 23 samples found positive with both methods. However, 39.1% of them (9/23) could be identified at the species level by SMg, whereas Sanger 16S only provided the genus name in all of them. SMg detected bacteria in 88.5% (23/26) of Sanger 16S-positive samples. For the remaining three samples for which SMg was negative, the four bacteria not detected were: *Streptococcus* sp. from a biopsy; a mix of *Streptococcus* sp. and *Moraxella* sp. from a joint fluid and *Pseudomonas fluorescens* group from an abscess (**Table 1**). Except for *Streptococcus* sp., the others bacteria were considered as possible contaminants (human and environmental contaminants, respectively) and were not taken into account for patient care. Conversely, SMg identified eight bacteria in 19.5% (8/41) of samples negative with the Sanger 16S approach. Among these bacteria, 25.0% (2/8) were identified at the genus level and

**TABLE 1** | Clinical and biological data of the 34 positive samples detected positive with Sanger sequencing of the 16S rRNA gene (Sanger 16S) and/or shotgun metagenomics (SMg).

Suspected type of infection (%;n/N)*	Sample type	Antibiotic therapy**	Identification by SMg	Identification by Sanger 16S	Clinical diagnosis
Bone and joint (48.5%; 16/33)	Joint fluid	N	<i>Klebsiella pneumoniae</i>	Negative	Septic arthritis
	Joint fluid (Knee)	Y	<i>Mycoplasma hominis</i>	<i>Mycoplasma hominis</i>	Septic arthritis
	Joint fluid (Knee)	N	<i>Neisseria meningitidis</i>	<i>Neisseria meningitidis</i>	Septic arthritis
	Joint fluid	N	<i>Staphylococcus aureus</i>	<i>Staphylococcus aureus</i>	Septic arthritis
	Joint fluid	Y	<i>Streptococcus dysgalactiae</i>	<i>Streptococcus dysgalactiae</i>	Septic arthritis
	Joint fluid (Knee)	Y	Negative	<i>Streptococcus</i> sp. <sup>C</sup> <i>Moraxella</i> sp. <sup>C</sup>	Microcrystalline arthritis
	Biopsy (Tissue: sternum)	N	<i>Corynebacterium</i> sp. <sup>C</sup>	Negative	Non-documented infection
	Biopsy (Tissue: hip)	N	<i>Cutibacterium acnes</i>	Negative	Prosthetic joint infection
	Biopsy (Tissue: hip)	Y	<i>Escherichia coli</i>	<i>Escherichia coli</i>	Prosthetic joint infection
	Biopsy (Tissue: elbow)	Y	<i>Mycobacterium tuberculosis</i>	Negative	Osteoarthritis
	Biopsy (Tissue: foot)	Y	<i>Staphylococcus aureus</i>	<i>Staphylococcus aureus</i>	Osteoarthritis
	Biopsy (Tissue: spine)	N	<i>Staphylococcus aureus</i>	<i>Staphylococcus aureus</i>	Prosthetic joint infection
	Biopsy (Tissue)	Y	Negative	<i>Streptococcus</i> sp.	Septic arthritis
	Abscess	N	<i>Cutibacterium acnes</i>	<i>Cutibacterium acnes</i>	Prosthetic joint infection
	Abscess (Knee)	Y	Negative	<i>Pseudomonas fluorescens</i> group <sup>C</sup>	Non-documented septic arthritis
	Abscess (Psoas)	Y	<i>Staphylococcus aureus</i>	<i>Staphylococcus aureus</i>	Vertebral osteomyelitis
	Cardiovascular (61.9%; 13/21)	Biopsy (Tissue: aortic aneurysm)	Y	<i>Bacteroides fragilis</i>	<i>Bacteroides fragilis</i>
Biopsy (Tissue: vegetation)		N	<i>Bartonella quintana</i>	<i>Bartonella</i> sp.	Endocarditis
Biopsy (Tissue: vegetation)		Y	<i>Enterococcus faecalis</i>	<i>Enterococcus</i> sp.	Endocarditis
Biopsy (Tissue: vegetation)		Y	<i>Staphylococcus aureus</i>	<i>Staphylococcus</i> sp.	Endocarditis
Biopsy (Tissue: aortic valve)		Y	<i>Staphylococcus epidermidis</i>	Negative	Endocarditis
Biopsy (Tissue: valve prosthesis)		N	<i>Streptococcus anginosus</i>	<i>Streptococcus milleri</i> group	Endocarditis
Biopsy (Tissue: vegetation)		Y	<i>Streptococcus dysgalactiae</i>	β-haemolytic <i>Streptococcus</i>	Endocarditis
Biopsy (Tissue: mitral valve)		Y	<i>Streptococcus gordonii</i>	<i>Streptococcus gordonii</i>	Endocarditis
Biopsy (Tissue: vegetation)		Y	<i>Streptococcus mitis</i> group	<i>Streptococcus</i> sp.	Endocarditis
Biopsy (Tissue: carotid)		Y	<i>Streptococcus pneumoniae</i>	α-haemolytic <i>Streptococcus</i>	Aortitis
Biopsy (Tissue)		Y	<i>Streptococcus pyogenes</i>	<i>Streptococcus pyogenes</i>	Endocarditis
Biopsy (Tissue: aortic valve)		Y	<i>Streptococcus sanguinis</i>	<i>Streptococcus mitis</i> group	Endocarditis
Abscess (Mediastinum)		Y	<i>Cutibacterium acnes</i> (associated with cutaneous flora)	Negative	Mediastinitis
Intra-abdominal (100%; 3/3)	Abscess (Liver)	Y	<i>Fusobacterium nucleatum</i>	<i>Fusobacterium nucleatum</i>	Liver abscess
	Abscess (Liver)	Y	<i>Klebsiella pneumoniae</i>	Negative	Liver abscess
	Abscess (Intra-abdominal)	Y	<i>Staphylococcus lugdunensis</i>	<i>Staphylococcus</i> sp.	Abdominal abscess
Genito-urinary (33.3%; 1/3)	Abscess (Kidney)	Y	<i>Enterococcus faecalis</i>	<i>Enterococcus</i> sp.	Kidney graft abscess
Skin and soft tissue (50.0%; 1/2)	Granuloma	N	<i>Mycobacterium tuberculosis</i> complex	Negative	Post-BCG infection

\*Percentage of positive samples for each suspected type of infection in at least one method.

\*\*One month before and/or at the time of sampling; N, No; Y, Yes; <sup>C</sup>, Probable sample contaminant; SMg, Shotgun metagenomics; Sanger 16S, Amplification of the 16S rRNA gene and Sanger sequencing of the PCR products.



75.0% (6/8) at the species level (Figure 2). Seven of these bacteria not detected by Sanger 16S were considered relevant and taken into account for patient care.

Overall, SMg was able to identify a bacterial etiology in 31/67 samples vs. 26/67 with Sanger 16S ( $p = 0.25$ , Mc Nemar Test). Interestingly, this difference reached significance when only the results obtained at the species level were compared (28/67 vs. 13/67;  $p < 0.001$ , Mc Nemar Test) (Figure 2). Sensitivities for the identification at the genus and at the species level, respectively, were 91.2% (31/34) and 82.4% (28/34) for SMg and 76.5% (26/34) and 38.2% (13/34) for Sanger 16S. Considering only the clinically relevant results (i.e., exclusion of bacteria considered as probable sample contaminant), the performance for the identification at the genus and at the species level, respectively, were 96.8% (30/31) and 90.3% (28/31) for SMg and 77.4% (24/31) and 41.9% (13/31) for Sanger 16S.

## DISCUSSION

The high performance (sensitivity and specificity) of SMg has been recently highlighted for the etiological diagnosis of a number of infectious diseases (Miao et al., 2018; Filkins et al., 2020; Gu et al., 2021; Miller and Chiu, 2021), including bone and joint infections (Street et al., 2017; Thoendel et al., 2018; Gamie et al., 2021; Goswami et al., 2021), meningitis

and encephalitis (Wilson et al., 2019; Rodino et al., 2020), or necrotizing fasciitis (Rodriguez et al., 2020c). Indeed, by allowing the detection and quantification of an unlimited panel of microorganisms, including bacteria, viruses, fungi, or parasites without *a priori* orientation, SMg is capable to identify the etiology of any infectious disease, including those due to microorganisms that have never been previously described in human infections (Rodriguez et al., 2020b). However, SMg has been barely implemented in the routine practice of microbiology laboratories, because of its cost and the specialized skills required for bioinformatics, biological and clinical interpretation of sequencing results. Nowadays, Sanger 16S remains the most widely used molecular tool to detect and identify bacteria from clinical samples when culture-based methods have failed. Nevertheless, the intrinsic performance of Sanger 16S and SMg in detecting bacteria and identifying them at the species level have rarely been assessed. A further comparative study between both approaches (five patients) was in favor of the superiority of SMg for bacteriological diagnosis (Gu et al., 2021).

In the present study, SMg detected bacteria in 19.5% of samples (8/41) that were found negative by Sanger 16S. Among these samples, the bacterial identification allowed a documentation in line with the clinical presentation in seven cases: *Cutibacterium acnes* in a case of a prosthetic joint infection; *Cutibacterium acnes* in a case of nosocomial mediastinitis; *Klebsiella pneumoniae* in a case of septic arthritis and liver

abscess, respectively; *Mycobacterium tuberculosis* in a case of osteoarthritis; *Mycobacterium tuberculosis* complex in a case of post-BCG infection and *Staphylococcus epidermidis* in a case of endocarditis (Table 1). Conversely, *Corynebacterium* sp., identified in one sample (tissular biopsy) taken for suspected bone infection, was considered as a sample contaminant by cutaneous flora in absence of clinical argument of infection due to this bacterium. Importantly, two samples that were negative by Sanger 16S were found positive for the presence of mycobacteria by SMg. This result, in keeping with the ability of SMg to detect mycobacteria (Miao et al., 2018; Zou et al., 2022) whereas 16S may lack specificity (Patel et al., 2000), highlights a key input of the SMg approach, in particular in situations where mycobacteriosis was not initially suspected from the clinical presentation.

Four bacteria (*Streptococcus* sp., *Streptococcus* sp. associated with *Moraxella* sp., and *Pseudomonas fluorescens* group) were detected by Sanger 16S, but not by SMg, and identified only at the genus level. Whereas one was considered as pathogenic (*Streptococcus* sp. in a patient with septic arthritis), *Pseudomonas fluorescens* group and the mix of *Streptococcus* sp. and *Moraxella* sp. were considered as probable environmental and human contaminants, respectively (contamination at the time of the sampling or during sample handling in the laboratory).

In all the 23 samples found positive with both methods, bacterial identification was consistent but SMg was more efficient than Sanger 16S to provide identification at the species level in nine samples (39.1%) identified only at the genus level by Sanger 16S. In comparison to the V3–V4 region of the 16S rRNA gene, the information provided by SMg allows covering a broader portion of the bacterial genome, including additional genes which may provide critical information to determine the species with precision. Accordingly, certain species are particularly difficult to identify by 16S rRNA gene sequencing (e.g., some species of the genus *Streptococcus*; Picard et al., 2004), compared with SMg approaches that theoretically allows to cover the entire bacterial genome. The ability of SMg to identify bacteria at the species level is of major importance for clinical management, including the choice of antimicrobial molecules and the treatment duration, although this has not been evaluated in this work. This was the case when *Staphylococcus aureus* was identified instead of *Staphylococcus* sp. or when *Enterococcus faecalis* was identified instead of *Enterococcus* sp. in two cardiac valves from patients with infectious endocarditis.

SMg is able to provide much more information than targeted approaches by detecting all potential bacterial and non-bacterial pathogens (virus, fungi) in one test (Wilson et al., 2019; Rodriguez et al., 2020b). As SMg was compared to Sanger 16S in samples from patients with suspected bacterial infections, only bacteria were detected in this work. Although not retrieved in the present study due to the depth of sequencing (minimal number of 10 million of reads per library), information on strain characterization, detection of virulence genes or resistance profile analysis can be theoretically obtained from SMg data (Andersen et al., 2016; Graf et al., 2016; Schmidt et al., 2017).

Considering the technical specifications of both approaches, the timeline has to be taken into account. Currently, in our institution, we estimate that the delay from the start of the

analysis to the results is estimated at 2.5 days for Sanger 16S and 5 days for SMg, respectively. Moreover, it should be noted that SMg requires specific skills and expensive equipment; a summary description of the main steps of the SMg workflow has been recently published (d'Humières et al., 2021). Finally, the cost of both methods is a considerable factor in the decision to realize novel diagnostic approaches in routine practice. Current SMg technology costs about three times more than the Sanger 16S approach in our laboratory with a large NGS platform (estimated SMg price: 200–300 euro per sample, excluding labor). However, these estimations deeply depend on the rate of flow and the size of the platform, the material and the technology used, as well as many other parameters including the country or the period during which the data were processed. Furthermore, it should be noted that most of the cost of SMg (approximately 2/3) is covered by the sequencing step, which is constantly decreasing.

Our study has some limitations. First, the potential bias due to samples freezing-thawing could not be accounted for in the interpretation of the results obtained in this study. Second, in order to make our study useful for most microbiologists and clinicians, we used the semi-automated kit UMD-SelectNA CE-IVD to perform the comparison between Sanger 16S and an accredited SMg approach (ISO 15189), as laboratories need to work with technologies approved by the FDA and / or CE-IVD. However, others methods used in microbiology laboratories for Sanger 16S (“in-house” or targeting other hypervariable regions of the 16S rRNA gene than V3–V4), may have slightly different performance. Moreover, it is important to mention that alternatives to Sanger sequencing using NGS technologies for 16S approach are numerous as amplicon sequencing or targeted capture technologies (Boers et al., 2019; Rassoulian Barrett et al., 2020), but still rarely used in routine practice currently. NGS 16S approaches enable bacterial identification in case of polymicrobial infections (Culbreath et al., 2019). Third, we used the threshold of 97% recommended by the manufacturer for the assignment at the genus level with Sanger 16S (Wellinghausen et al., 2009). However it should be noted that lower thresholds have been reported in others publications (Bemer et al., 2014; Yarza et al., 2014). Fourth, the minimal number of 10 million reads for DNA library was not obtained for three samples (Supplementary Table 1), which may lead to a decrease in the sensitivity of the SMg approach. Fifth, the turnaround time and the impact of our results on clinical care and outcomes were not evaluated. Nevertheless, SMg results were delivered to the clinicians and only the results taken into account in their treatment decisions were considered as true positives, suggesting clinical relevance. Finally, the results delivered to the clinicians depended on the nature of the sample (sterile or possibly colonized sample) and on the bacterium identified (i.e., pathogenicity level). In absence of “gold standard” other than culture, the categorization of the results delivered as true positives depended on the final decision made by the clinicians to consider the result in the diagnosis.

In conclusion, SMg showed a better contribution to bacterial documentation at the species level when compared to Sanger 16S for the etiological diagnosis of infectious syndromes in case of standard cultures negative. These results suggest that SMg may replace in the future Sanger 16S in routine practice

to diagnose bacterial infections. Future studies assessing the clinical benefit of SMg and the possible impact on prognosis are, however, still needed.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Henri Mondor Hospital Institutional Review Board, Créteil, France (approval number 00011558). The patients/participants provided their written informed consent to participate in this study.

## REFERENCES

- Abayasekara, L. M., Perera, J., Chandrasekharan, V., Gnanam, V. S., Udunuwara, N. A., Liyanage, D. S., et al. (2017). Detection of bacterial pathogens from clinical specimens using conventional microbial culture and 16S metagenomics: a comparative study. *BMC Infect. Dis.* 17:631. doi: 10.1186/s12879-017-2727-8
- Andersen, H., Connolly, N., Bangar, H., Staat, M., Mortensen, J., Deburger, B., et al. (2016). Use of shotgun metagenome sequencing to detect fecal colonization with multidrug-resistant bacteria in children. *J. Clin. Microbiol.* 54, 1804–1813. doi: 10.1128/JCM.02638-15
- Bemer, P., Plouzeau, C., Tande, D., Leger, J., Giraudeau, B., Valentin, A. S., et al. (2014). Evaluation of 16S rRNA gene PCR sensitivity and specificity for diagnosis of prosthetic joint infection: a prospective multicenter cross-sectional study. *J. Clin. Microbiol.* 52, 3583–3589. doi: 10.1128/JCM.01459-14
- Boers, S. A., Jansen, R., and Hays, J. P. (2019). Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. *Eur. J. Clin. Microbiol. Infect. Dis.* 38, 1059–1070. doi: 10.1007/s10096-019-03520-3
- Church, D. L., Cerutti, L., Gürtler, A., Griener, T., Zelazny, A., and Emler, S. (2020). Performance and application of 16S rRNA gene cycle sequencing for routine identification of bacteria in the clinical microbiology laboratory. *Clin. Microbiol. Rev.* 33:e00053–19. doi: 10.1128/CMR.00053-19
- Courbin, V., Riller, Q., Amegnizin, J.-L., Gricourt, G., Demontant, V., Fihman, V., et al. (2022). Case report: cerebral *Nocardiosis* caused by *Nocardia cyriacigeorgica* detected by metagenomics in an apparently immunocompetent patient. *Front. Immunol.* 13:719124. doi: 10.3389/fimmu.2022.719124
- Couto, N., Schuele, L., Raangs, E. C., Machado, M. P., Mendes, C. I., Jesus, T. F., et al. (2018). Critical steps in clinical shotgun metagenomics for the concomitant detection and typing of microbial pathogens. *Sci. Rep.* 8:13767. doi: 10.1038/s41598-018-31873-w
- Culbreath, K., Melanson, S., Gale, J., Baker, J., Li, F., Saebo, O., et al. (2019). Validation and retrospective clinical evaluation of a quantitative 16S rRNA gene metagenomic sequencing assay for bacterial pathogen detection in body fluids. *J. Mol. Diagn.* 21, 913–923. doi: 10.1016/j.jmoldx.2019.05.002
- d’Humières, C., Salmona, M., Dellièrre, S., Leo, S., Rodriguez, C., Angebault, C., et al. (2021). The potential role of clinical metagenomics in infectious diseases: therapeutic perspectives. *Drugs* 81, 1453–1466. doi: 10.1007/s40265-021-01572-4
- Filkins, L. M., Bryson, A. L., Miller, S. A., and Mitchell, S. L. (2020). Navigating clinical utilization of direct-from-specimen metagenomic pathogen detection: clinical applications, limitations, and testing recommendations. *Clin. Chem.* 66, 1381–1395. doi: 10.1093/clinchem/hvaa183

## AUTHOR CONTRIBUTIONS

CL, LS, VF, J-MP, CR, and P-LW contributed in every aspect of this research work including conception, study design, and data analysis and interpretation. VD, ET, NL, and J-ML performed the laboratory experiments. CG and MN'D performed bioinformatics analysis. CL, LS, J-MP, CR, and P-LW wrote the manuscript. LS, CW, GL, GRos, and RL participated in interpretation of data and conflicting results (adjudication committee). VF, CG, and GRoy participated in review of the manuscript. All authors contributed to manuscript revision, read and approved the submitted manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.761873/full#supplementary-material>

- Forbes, J. D., Knox, N. C., Ronholm, J., Pagotto, F., and Reimer, A. (2017). Metagenomics: the next culture-independent game changer. *Front. Microbiol.* 8:1069. doi: 10.3389/fmicb.2017.01069
- Fu, Z., Zhang, H., Zhang, Y., Cui, P., Zhou, Y., Wang, H., et al. (2022). Evaluations of clinical utilization of metagenomic next-generation sequencing in adults with fever of unknown origin. *Front. Cell. Infect. Microbiol.* 11:745156. doi: 10.3389/fcimb.2021.745156
- Gamie, Z., Karthikappallil, D., Gamie, E., Stamiris, S., Kenanidis, E., and Tsiridis, E. (2021). Molecular sequencing technologies in the diagnosis and management of prosthetic joint infections. *Expert Rev. Mol. Diagn.* doi: 10.1080/14737159.2021.1894929
- Goswami, K., Shope, A. J., Tokarev, V., Wright, J. R., Unverdorben, L. V., Ly, T., et al. (2021). Comparative meta-omics for identifying pathogens associated with prosthetic joint infection. *Sci. Rep.* 11:23749. doi: 10.1038/s41598-021-02505-7
- Graf, E. H., Simmon, K. E., Tardif, K. D., Hymas, W., Flygare, S., Eilbeck, K., et al. (2016). Unbiased detection of respiratory viruses by use of RNA sequencing-based metagenomics: a systematic comparison to a commercial PCR panel. *J. Clin. Microbiol.* 54, 1000–1007. doi: 10.1128/JCM.03060-15
- Grumaz, S., Stevens, P., Grumaz, C., Decker, S. O., Weigand, M. A., Hofer, S., et al. (2016). Next-generation sequencing diagnostics of bacteremia in septic patients. *Genome Med.* 8:73. doi: 10.1186/s13073-016-0326-8
- Gu, W., Deng, X., Lee, M., Sucu, Y. D., Arevalo, S., Stryker, D., et al. (2021). Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nat. Med.* 27, 115–124. doi: 10.1038/s41591-020-1105-z
- Hilton, S. K., Castro-Nallar, E., Pérez-Losada, M., Toma, I., McCaffrey, T. A., Hoffman, E. P., et al. (2016). Metataxonomic and metagenomic approaches vs. culture-based techniques for clinical pathology. *Front. Microbiol.* 7:484. doi: 10.3389/fmicb.2016.00484
- Kommedal, O., Kvello, K., Skjastad, R., Langeland, N., and Wiker, H. G. (2009). Direct 16S rRNA gene sequencing from clinical specimens, with special focus on polybacterial samples and interpretation of mixed DNA chromatograms. *J. Clin. Microbiol.* 47, 3562–3568. doi: 10.1128/JCM.00973-09
- Miao, Q., Ma, Y., Wang, Q., Pan, J., Zhang, Y., Jin, W., et al. (2018). Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice. *Clin. Infect. Dis.* 67, S231–S240. doi: 10.1093/cid/ciy693
- Miller, S., and Chiu, C. (2021). The role of metagenomics and next-generation sequencing in infectious disease diagnosis. *Clin. Chem.* 68, 115–124. doi: 10.1093/clinchem/hvab173
- Patel, J. B., Leonard, D. G., Pan, X., Musser, J. M., Berman, R. E., and Nachamkin, I. (2000). Sequence-based identification of *Mycobacterium* species using the

- MicroSeq 500 16S rDNA bacterial identification system. *J. Clin. Microbiol.* 38, 246–251. doi: 10.1128/JCM.38.1.246-251.2000
- Picard, F. J., Ke, D., Boudreau, D. K., Boissinot, M., Huletsky, A., Richard, D., et al. (2004). Use of tuf sequences for genus-specific PCR detection and phylogenetic analysis of 28 Streptococcal species. *J. Clin. Microbiol.* 42, 3686–3695. doi: 10.1128/JCM.42.8.3686-3695.2004
- Rassouliani Barrett, S., Hoffman, N. G., Rosenthal, C., Bryan, A., Marshall, D. A., Lieberman, J., et al. (2020). Sensitive identification of bacterial DNA in clinical specimens by broad-range 16S rRNA gene enrichment. *J. Clin. Microbiol.* 58, e1605–e1620. doi: 10.1128/JCM.01605-20
- Rodino, K. G., Toledano, M., Norgan, A. P., Pritt, B. S., Binnicker, M. J., Yao, J. D., et al. (2020). Retrospective review of clinical utility of shotgun metagenomic sequencing testing of cerebrospinal fluid from a U.S. tertiary care medical center. *J. Clin. Microbiol.* 58:e01729–20. \*e01729-20 doi: 10.1128/JCM.01729-20
- Rodriguez, C., Jary, A., Hua, C., Woerther, P. L., Bosc, R., Desroches, M., et al. (2020c). Pathogen identification by shotgun metagenomics of patients with necrotizing soft-tissue infections. *Br. J. Dermatol.* 183, 105–113. doi: 10.1111/bjd.18611
- Rodriguez, C., Gouilh, M. A., Weiss, N., Stroer, S., Mokhtari, K., Seilhean, D., et al. (2020a). Fatal measles inclusion-body encephalitis in adult with untreated AIDS, France. *Emerg. Infect. Dis.* 26, 2231–2234. doi: 10.3201/eid2609.200366
- Rodriguez, C., Gricourt, G., Ndebi, M., Demontant, V., Poiteau, L., Burrel, S., et al. (2020b). Fatal encephalitis caused by *Cristoli* virus, an emerging *Orthobunya* virus, France. *Emerg. Infect. Dis.* 26, 1287–1290. doi: 10.3201/eid2606.191431
- Rodriguez, C., Pawlotsky, J.-M., and Demontant, V. (2020d). *Method For Identifying An Infectious Agent, PCT/FR2020/052193*. Paris: European Patent Office, 1–50.
- Schmidt, K., Mwaigwisya, S., Crossman, L. C., Doumith, M., Munroe, D., Pires, C., et al. (2017). Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *J. Antimicrob. Chemother.* 72, 104–114. doi: 10.1093/jac/dkw397
- Schriefer, A. E., Cliften, P. F., Hibberd, M. C., Sawyer, C., Brown-Kennerly, V., Burcea, L., et al. (2018). A multi-amplicon 16S rRNA sequencing and analysis method for improved taxonomic profiling of bacterial communities. *J. Microbiol. Methods* 154, 6–13. doi: 10.1016/j.mimet.2018.09.019
- Simner, P. J., Miller, S., and Carroll, K. C. (2018). Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. *Clin. Infect. Dis.* 66, 778–788. doi: 10.1093/cid/cix881
- Street, T. L., Sanderson, N. D., Atkins, B. L., Brent, A. J., Cole, K., Foster, D., et al. (2017). Molecular diagnosis of orthopedic-device-related infection directly from sonication fluid by metagenomic sequencing. *J. Clin. Microbiol.* 55, 2334–2347. doi: 10.1128/JCM.00462-17
- Thoendel, M. J., Jeraldo, P. R., Greenwood-Quaintance, K. E., Yao, J. Z., Chia, N., Hanssen, A. D., et al. (2018). Identification of prosthetic joint infection pathogens using a shotgun metagenomics approach. *Clin. Infect. Dis.* 67, 1333–1338. doi: 10.1093/cid/ciy303
- Wellinghausen, N., Kochem, A.-J., Disqué, C., Mühl, H., Gebert, S., Winter, J., et al. (2009). Diagnosis of bacteremia in whole-blood samples by use of a commercial universal 16S rRNA gene-based PCR and sequence analysis. *J. Clin. Microbiol.* 47, 2759–2765. doi: 10.1128/JCM.00567-09
- Wilson, M. R., Sample, H. A., Zorn, K. C., Arevalo, S., Yu, G., Neuhaus, J., et al. (2019). Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *N. Engl. J. Med.* 380, 2327–2340. doi: 10.1056/NEJMoa1803396
- Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F. O., Ludwig, W., Schleifer, K.-H., et al. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol.* 12, 635–645. doi: 10.1038/nrmicro3330
- Zhou, H., Larkin, P. M. K., Zhao, D., Ma, Q., Yao, Y., Wu, X., et al. (2021). Clinical impact of metagenomic next-generation sequencing of bronchoalveolar lavage in the diagnosis and management of pneumonia. *J. Mol. Diagn.* 23, 1259–1268. doi: 10.1016/j.jmoldx.2021.06.007
- Zou, J., Wang, T., Qiu, T., Chen, Z., Zhou, J., Ma, X., et al. (2022). Clinical characteristics of tuberculous infection following renal transplantation. *Transpl. Immunol.* 70:101523. doi: 10.1016/j.trim.2021.101523

**Conflict of Interest:** CR served as an advisor, and/or speaker for Illumina, and Vela Diagnostics. P-LW has served as speaker for MSD. J-MP served as an advisor, and/or speaker for Abbvie, Gilead, GlaxoSmithKline, Merck, Regulus, and Siemens Healthcare.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Lamoureux, Surgers, Fihman, Gricourt, Demontant, Trawinski, N'Debi, Gomart, Royer, Launay, Le Glaunec, Wemmert, La Martire, Rossi, Lepeule, Pawlotsky, Rodriguez and Woerther. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.