UiT The Arctic University of Norway

Faculty of Health Sciences
Department of Medical Biology

A Shotgun-metagenomics approach for laboratory diagnostics in clinical microbiology

—

Adriana Maria Sanabria Moreno

A dissertation for the degree of Philosophiae Doctor - August 2020

A dissertation for the degree of Philosophiae Doctor

A Shotgun-metagenomics approach for laboratory diagnostics in clinical microbiology

ADRIANA MARIA SANABRIA MORENO

August 2020

Research Group of Host-Microbe Interactions
Department of Medical Biology
Faculty of Health Sciences
UiT- The Artic University of Norway

To my mother, Cecilia Moreno

for her endless love, support and encouragement

To my husband Andres Gonzalez and our daughters,

Maria Paz and Emma Sofia,

with love

Acknowledgments

Numerous people have contributed directly or indirectly to the work presented in this thesis.

First of all, I would like to thank my supervisors, Anne-Merethe Hanssen, Johanna Ericson Sollid, Mona Johannesen, and Gunnar Skov Simonsen. Maybe you have not realized how important you were to me; therefore, I want to highlight your huge significance for me and for this work. I have learned a lot from all of you, and I feel extremely lucky and grateful to have the opportunity to work with this incredible supervision team. Thanks for showing interest in my work and progress, despite of your very busy schedule.

I would like to express my deepest gratitude to my main supervisor Anne-Merethe Hanssen for the continuous support/guidance she has given me during my Ph.D. Anne-Merethe, thanks for your patience, motivation, and guidance during the whole process. For sharing your knowledge with me, for having your door always open for my never-ending list of questions, and for all the interesting discussions we had during the last four years. I could not have imagined a better supervisor for my Ph.D. studies. This work would have never been completed without your help, support, and dedication. You are one of the most honest, kind, and big-hearted persons I know, and I am very proud to be your Ph.D. student.

My big gratitude goes to my co-supervisors Johanna, Mona, and Gunnar. Johanna, thank you for your insightful comments, and your hard and always interesting questions. Mona, thanks for your encouragement and support not just as a co-supervisor but also as the group leader. Gunnar, thank you for your guidance, advice and thoughtful feedback. The extent of your knowledge never ceases to amaze me. The valuable lessons I have learned from you are invaluable.

I would like to thank the closest collaborators and co-authors for their invaluable contribution to this work.

Sincere thanks to Jessin Janice and Erik Hjerde, for sharing your knowledge in bioinformatics with me. Thanks for collaborating with my projects despite your busy schedules, and for being always available and patient for my silly questions. Jessin, I think I have never told you, how much you mean to me. Thank you so much for being there on my side and supporting me all the way through.

I would like to thank also the Department of Microbiology and Infection Control at the University Hospital of North Norway (UNN), Tromsø, especially to Bodil Pedersen for assistance in the use of the BCB system. Silje Nysted Hagen and Maria Frost for coordinating the sample collection at the preanalytical laboratory. Lisbeth Engen Mortensen for her kindness to help me with the phenotypical assays, and to my dear Merethe Røkeberg, thank you so much for all your technical help and amazing friendship.

I would like to thank all past, and current members of the Host-Microbe Interactions Research Group for providing a wonderful working environment. My special thanks to Tove Eriksen, for her kindness in helping to solve all administrative questions. Thanks to Alena, Ahmed, Kjersti, and Buphender for their important work of constantly updating and maintaining the lab, and thanks for your friendship.

To the 9th floor "the lab gang" including those that have left. Thanks for providing a happy distraction to rest my mind outside of my research with international birthday lunches, coffee breaks with pastries, and a lot of fun. Having you in my life made this journey amazing, I will cherish all the memories we have made together. Jessin, your friendship has been one of the greatest treasures in my life, thanks for always believe in me, for your help with bioinformatics, for your delicious food, for proof-reading my thesis and manuscripts, for your scripts, and the most important for loving my family. You know that my girls and me love you, and that you and Sudhagar are our family in Tromsø. Dianita, mi Peruana favorita, you are one of those persons I would like to have you by my side everywhere. Thanks for always listening to all my stories, complaints, and dramas. For taking care of my girls and be always there when I need you, despite my moods. You are a great friend. Maria Pain, thanks for your friendship and all our interesting talks. I have learned a lot during our evening walks. You are a researcher with huge potential, and you know I appreciate you a lot. Merethe, your friendship has been invaluable during my time in Tromsø, not just at work but also in my personal life. Thanks for everything, my friend. Clement, our paths have cross twice and it has been a very nice coincidence. Thanks a lot for your friendship amigo, for your advice and support. Bishnu, thanks for your friendship. You are a nice soul and a researcher from the heart. I hope you never forget that. Connysita, I still miss you, I miss our Zumba classes, our talks, your hugs, and off course, I miss my partner in crime to go for "scones". Katya, you are that friend that everybody wishes to have, always kind, humble, and helpful. Theresa, probably you do not know, but I appreciate you very much, thanks for your generosity and for being always eager to help me.

Last but not least, I would like to acknowledge the never-ending love and support of my family. My mother Cecilia Moreno and my siblings, Nelcy, Pilar, and Leonardo and their respective families who always believe in me and made my life meaningful from the very beginning. Wherever I will be, you will always be with me in my heart. I love you all so much. My father, Alonso Sanabria, and my siblings Carlos Alonso, Juan Manuel, and Antonio Jose, for all the love you give to me and my family. I am very proud and happy that you are in our lives. I love you. Finally, my beloved husband Andres, and our girls Maria Paz and Emma Sofia. You give me a strong sense of belonging and put all my strivings into perspective. Andres, we started this adventure four years ago when we decided to move to the north, thanks my love for being my teammate, for dreaming with me and for all your love and dedication. Maria and Emma, there are not enough words to tell you how much I love you; you are the biggest blessings of my life.

Thank you, God, for all your blessings to me and my family, for the strength you give me each day and for all the people around who make my life more meaningful.

Table of Contents

Abbreviations

AMR	Antimicrobial Resistance
ARG	Antimicrobial Resistance Gene
AST	Antibiotic Susceptibility Testing
BCB	Blood Culture Bottle
BD	Becton Dickinson
CA-MRSA	Community-acquired MRSA
CCs	Clonal Complexes
CFU	Colony-Forming Unit
CgMLST	Core genome MLST
CoNS	Coagulase Negative Staphylococci
EBJIS	European bone and joint infection society
ESR	Erythrocyte Sedimentation Level
FDA	Food and Drug Administration
HA-MRSA	Healthcare-acquired MRSA
HGT	Horizontal Gene Transfer
IDSA	Infectious Diseases Society of America
LA-MRSA	Livestock-associated MRSA
MALDI-TOF	Matrix-assisted laser desorption/ionization-time of flight mass spectrometer
MGEs	Mobile Genetic Elements
MLST	Multilocus Sequence Typing
mNGS	Metagenomics-Next Generation Sequencing
MRSA	Methicillin resistant *Staphylococcus aureus*
MSIS	Musculoskeletal Infection Society
MSSA	Methicillin susceptible *Staphylococcus aureus*
NGS	Next Generation Sequencing
PCR	Polymerase Chain Reaction
PFGE	Pulsed-Field Gel Electrophoresis
RPM	Revolutions Per Minute
SaPIs	Staphylococcal Pathogenicity Islands
SCC	Staphylococcal Cassette Chromosome
SMg	Shotgun-Metagenomics
STs	Sequence Types
THA	Total Hip Arthroplasties
TKA	Total Knee Arthroplasty
TNF-α	Tumor Necrosis Alpha Factor
TTD	Time to Detection
VFDB	Virulence Factor Database
VFs	Virulence Factors
WBC	White Blood Cells
wgMLST	Whole-genome MLST
WGS	Whole Genome sequencing

List of papers

Paper I

Adriana Sanabria, Merethe E. O. Røkeberg, Mona Johannessen, Johanna Ericson Sollid, Gunnar Skov Simonsen, Anne-Merethe Hanssen (2019).

Culturing periprosthetic tissue in BacT/Alert® Virtuo blood culture system leads to improved and faster detection of prosthetic joint infections. *BMC Infect Dis*. 19: 607.

Paper II

Adriana Maria Sanabria, Erik Hjerde, Mona Johannessen, Johanna Ericson Sollid, Gunnar Skov Simonsen, Anne-Merethe Hanssen (2020).

Shotgun-metagenomics on positive blood culture bottles inoculated with prosthetic joint tissue: a proof of concept study. *Front. Microbiol*. 10:3389.

Paper III

Adriana Maria Sanabria, Jessin Janice, Erik Hjerde, Anne-Merethe Hanssen.

Typing and prediction of antibiotic resistance and virulence determinants in *S. aureus* using shotgun-metagenomics data from prosthetic joint tissue on blood culture bottles. Manuscript in preparation.

Summary

Shotgun-metagenomics (SMg) has a huge potential, particularly in areas where conventional diagnostic methods have limitations such as in prosthetic joint infection (PJI). However, the application of metagenomic sequencing on periprosthetic joint tissue (PJT) specimens is often challenged by the low bacterial load in addition to a high level of inhibitor and contaminant host DNA, limiting the potential that this emerging approach can offer in clinical diagnostics. The main objective of this Ph.D. project was to explore the use of SMg in clinical diagnostics of PJI directly from blood culture bottles (BCBs) inoculated with PJT specimens.

In **paper I**, we assessed the use of a BacT/Alert® Virtuo blood culture system for culturing periprosthetic tissue (PJT) specimens. We showed that the blood culture bottle (BCB) method was found to detect a wider range of bacteria more rapidly than the conventional microbiological method. Our findings show that the use of BCB is a convenient approach to be used in the routine in the clinics for diagnosis of the PJI cases.

In **paper II**, we developed a proof of concept study with the aim of evaluating the use of SMg on BCBs inoculated with PJT for pathogen identification for diagnosis of PJI. For this purpose, we developed a method for the preparation of high-quality bacterial DNA from PJT for downstream SMg, established a bioinformatics pipeline, and compared SMg results with conventional culture method results. Our DNA preparation method resulted in high quality microbial DNA from all samples, both with and without human DNA depletion allowing downstream SMg. All bacteria identified by the culture method were also identified by SMg. We obtained an acceptable high number of bacterial reads, genome coverage and genome sequencing depth for identification of PJI pathogens

In **paper III**, we explored further applications of SMg, beyond pathogen identification. SMg from BCBs inoculated with PJT, allowed the identification of potential PJI pathogens, and strain-level typing of *S. aureus*. We obtained *S. aureus* antimicrobial resistance (AMR) and virulence gene profiles from both monomicrobial and polymicrobial samples. However, the use of this approach for the detection of AMR to help guide clinical antibiotic usage needs to be further elucidated, due to some disagreement between the AMR phenotype and genotype. Precise AMR prediction is required for the mainstream adoption of SMg into the clinical microbiology laboratory.

In conclusion, these studies present an alternative approach for the application of SMg in PJI diagnosis, complementing the currently available tools. Our results might be useful in further validation and standardization for solving challenges presented in PJI diagnoses.

Introduction

Infectious diseases remain the leading causes of morbidity and mortality worldwide. In order to reduce the burden of infectious diseases, an efficient and accurate diagnosis is important. Research into improving all aspects of culture-based methods and new alternative strategies to replace or support current microbiological methods is needed, especially when the first-line investigation tools present technical limitations, are time consuming and labor intensive. Currently, a large number of tests are available. Microbiological methods differ across countries, and within countries and institutions using varying criteria and protocols. The most efficient diagnostic strategy is likely formed by a combination of conventional and new diagnostic strategies.

Basic scientific applied research and emerging genomic technologies have enabled us to start exploring the utility of Next-generation sequencing (NGS) into multiple areas in the field of diagnostics of infectious diseases. NGS is actively moving from research to the clinical setting and has been presented as a tool with the potential to complement or replace the current diagnostic methods. It has the potential to change the microbial diagnosis of infectious diseases, especially when detection and identification of pathogens can be challenging, such as prosthetic joint infection (PJI) diagnosis. The application of NGS and Shotgun metagenomics (SMg) on PJI specimens is often challenged by the low bacterial load in addition to high inhibitor and contaminant host DNA limiting pathogen recovery.

This project investigates if SMg performed directly from blood culture bottles (BCBs) inoculated with periprosthetic joint tissue (PJT) specimens may be a convenient approach for overcoming these obstacles. We first established a method for culturing PJT on BCBs, continuing with establishing and evaluating a method for the preparation of bacterial DNA suitable for SMg directly from BCBs inoculated with PJT specimens and assessed their potential for identifying potential causative pathogens. Additionally, the potential of SMg for prediction of virulence factors (VFs) and antimicrobial resistance (AMR) determinants in *Staphylococcus aureus* on BCBs with PJT was assessed.

A general introduction is presented as follows: (1) The application of NGS methods for the analysis of clinical specimens that can be used for the identification of microorganisms and

detection of AMR and VFs in the clinical microbiology laboratory. (2) The microbiological diagnosis of PJI since it is the infectious disease model selected for this study.

1 NGS technology in the clinical microbiology laboratory

Over the last two decades, advances in sequencing technologies have revolutionized the ability to sequence nucleic acids. NGS also known as high-throughput sequencing is the term used to describe a number of different modern sequencing technologies whereby billions of nucleic acid fragments can be simultaneously and independently sequenced[1]. These technologies allow for sequencing of nucleic acids much more quickly and cheaper than the previously used Sanger sequencing, which have led to introducing this emerging technology as a tool in the clinical settings with the potential of revolutionizing the clinical laboratory by simplifying and improving the clinical diagnostic testing[2].

NGS is rapidly finding a variety of applications in both clinical and research laboratories. NGS has previously been used in clinical settings as a research tool and currently, it has begun to be an affordable alternative to be used in the clinical microbiology laboratory for the diagnosis of infectious disease[3–5]. However, despite the remarkable progress of NGS for improving the diagnoses of infectious diseases, the translational gap between NGS and clinical implementation remains. There are several challenges to overcome before NGS can deliver its potential in the clinical laboratory for patients, clinicians, and society. Therefore, further efforts need to be addressed.

1.1 NGS technologies

There are several NGS technologies commercially available, which are continuously evolving and improving. They differ substantially in terms of protocols (engineering and sequencing chemistry), output (length of reads, number of sequences), accuracy, and cost[6]. The unique combination of specific technical details differentiates one technology from another and determines the type of data produced from each platform.

NGS includes second generation and third generation sequencing methods defined as short-read (50-400 bp) and long-read (1-100 kb) technologies, respectively[7]. The most used short reads platforms include Illumina (San Diego, CA) and Ion Torrent (Waltham, MA) while for long- reads platforms are the single molecule real time sequencing (Pacific Biosciences) and

Oxford nanopore sequencing (Oxford, United Kingdom)[6,8,9] (Table 1). Choosing among available methods depends on the sequencing objectives and involves tradeoffs in accuracy, efficiency, and cost [6]. Due to differences in methodology and technology between the NGS platforms each platform has advantages and disadvantages that should be considered when choosing the technology to use in specific sequencing projects and for analyzing sequence data, both own and publicly available data[10].

Table 1. Comparison of the most used NGS platforms (Modified from[6,11])

Sequencing generation	Platform	Instrument	Throughput range (Gb)	Read length (bp)
Second generation (short-reads)	Illumina	MiniSeq	1.7–7.5	75 SE/PE - 150 PE
		MiSeq	3.3–15	75 - 300 PE
		NextSeq	10–120	75 - 150 PE
		HiSeq (2500/3000/4000)	10–1500	50 SE/PE - 150 PE
		NovaSeq 5000/6000	134-3000	50 -150 PE
	IonTorrent	PGM	0.08–2	Up to 400
		S5	0.6–15	Up to 400
		Proton	10–15	Up to 200
Third generation (long-reads)	Pacific BioSciences	PacBio RSII	0.5–1	Up to 60 Kb
		Sequel	5–10	Up to 60 Kb
	Oxford Nanopore	MInION	0.1–1	Hundreds to thousands of Kb
		GridION x5	2.5	
		PromethION (Beta)	375	

Abbreviations: Gb, Gigabytes; bp, base pairs; SE, single-end reads; PE, Pair-end reads; PGM, Ion Personal Genome Machine.

1.1.1 Short-reads sequencing

In general, the short-read technologies produce billions of nucleotide sequences during each run, where each genome is sequenced multiple times in small random pieces to generate very large data sets[6]. Short read sequencing can be divided further into two categories, based on the sequencing methodology used, as sequencing by ligation and sequencing by synthesis.

Sequencing by ligation is a method that uses the mismatch sensitivity of the DNA ligase to determine the identity of nucleotides in a given DNA sequence[12]. Currently, the only sequencing by ligation method commercially available is the BGIseq platform, which has been developed by the Beijing Genomics Institute (BGI) and has been into the market since 2016[13]. This platform uses DNA nanoball technology by cloning single DNA molecules locally on a flow cell that produces clonal DNA nanoballs that then undergo sequencing by synthesis[14]. This

platform has already been used for clinical applications, in reference gene catalog–based human gut metagenome studies[15].

Sequencing by synthesis includes a group of methodologies that make use of a DNA polymerase enzyme to incorporate a single nucleotide or short oligonucleotides (provided either one at a time or fluorescently labeled), containing a reversible terminator[14]. It is the most widely adopted NGS technology, with Illumina dominating the sequencing market.

Illumina sequencing by synthesis

Illumina has several platforms in the market including iSeq, MiSeq, MiniSeq, NextSeq, HiSeq, and NovaSeq. Depending on the specific Illumina platform it may be partitioned into 1 (miSeq), 2 (HiSeq2500), or 8 (HiSeq2000, HiSeq2500) separate sequencing lanes[16]. All of them use a strategy of bridge amplification on the surface of a flow cell[17]. Illumina NGS workflows include three basic steps: Library preparation, cluster generation, and sequencing. Briefly explained, following library preparation (the process that generates a collection of DNA fragments of uniform size ready to be sequenced), the DNA templates are hybridized (attached) to a glass slide (flow cell) that has patterned clusters of complementary adaptors and then the fragments are PCR amplified locally from one end only (single-end read) or from both ends (paired-end reads), producing millions to billions of clusters of clonal template DNA fragments that can be sequenced simultaneously[1]. This is followed by sequencing by synthesis which as mentioned before, utilizes DNA polymerase to build the complementary DNA utilizing modified deoxyribose nucleoside triphosphates (dNTPs) actin as chain terminators (one nucleotide per cycle), and then a readout of fluorescently labeled nucleotides determines the nucleotide identity[1]. This process is repeated for the length of the read to generate the sequence output, where read lengths are now typically between 75 to 350 bp (see Table 1).

Overall, all second-generation sequencing platforms have the disadvantages inherent to short-read sequencing platforms. They tend to generate relatively fragmented genome assemblies, causing an inability to resolve e.g. repetitive regions and making some types of genetic variants challenging to identify. The need for amplification is time consuming and could introduce laboratory bias, such as over/underrepresenting certain regions. Therefore, longer reads are desired in order to generate closed reference genomes particularly useful when sequencing through complex genomic regions such as repeats and phages (very common in bacterial genomes). To meet this demand, the so-called third-generation sequencing platforms have been

introduced. Third-generation methods are under development and generally are not clinically available.

1.1.2 Long-reads sequencing technology

New sequencing instruments are currently available for research use, that provide longer sequencing reads and are able to read the sequence of a single molecule rather than amplified DNA as a template: the PacBio SMRT (single molecule real time) (Menlo Park, California) and the Oxford Nanopore (Oxford, United Kingdom).

Oxford nanopore sequencing

Oxford Nanopore Technologies (Oxford, United Kingdom) includes the MinION (1 flow cell), GridION (5 flow cell capacity), and PromethION (48 flow cell capacity)[18]. The Oxford Nanopore uses a protein pore inserted into a membrane. A current is applied and flows through the pore between the 2 sides of the membrane. As the structure (DNA or RNA strand) passes through the pore, the current changes and the degree of change correlates with the individual base (A, C, G, or T) and also correlates with the methylation status of C; therefore, methylation and hydroxymethylation can be detected. Notably, for infectious disease diagnostics, nanopore DNA sequencing is orders of magnitude faster than other strategies that use sequence-by-synthesis methods (used by Illumina). Nanopore sequencing also does not require prior PCR amplification, although often this is still performed due to the high baseline sample input requirement (>500 ng).

Real-time nanopore sequencing enables real-time analyses, which is particularly interesting for the diagnosis of acute infectious diseases since timely diagnostics is essential when treating patients. The nanopore approach currently has more sequencing errors, lower throughput, and higher per-read costs than other NGS platforms, which may limit its utility[1]. The clinical adoption of these devices has been limited by the rapid pace of improvements to the platform, which can delay clinical validation efforts requiring standardized instruments and locked-down protocols[19]. Overall, long-read approaches currently tend to have lower throughput and higher total costs, limiting their widespread implementation. For these reasons, short-read technologies are more frequently used.

1.2 NGS basic workflow

Regardless of the differences in NGS methods and in the application, all NGS technologies are based on the same general process. Conceptually, this approach involves a series of steps

including the work carried out in the microbiology laboratory (wet lab) and in the computational laboratory (dry lab) (Figure 1). Wet lab analyses involve clinical sample processing including DNA isolation (and/or RNA, followed by reverse transcription), library preparation, and sequencing steps. It is followed by the dry lab consisting of the bioinformatic data analysis, where the sequences are analyzed commonly in the context of a reference database, to determine which organisms are present and their potential phenotypes, based on their genetic content[1,20].

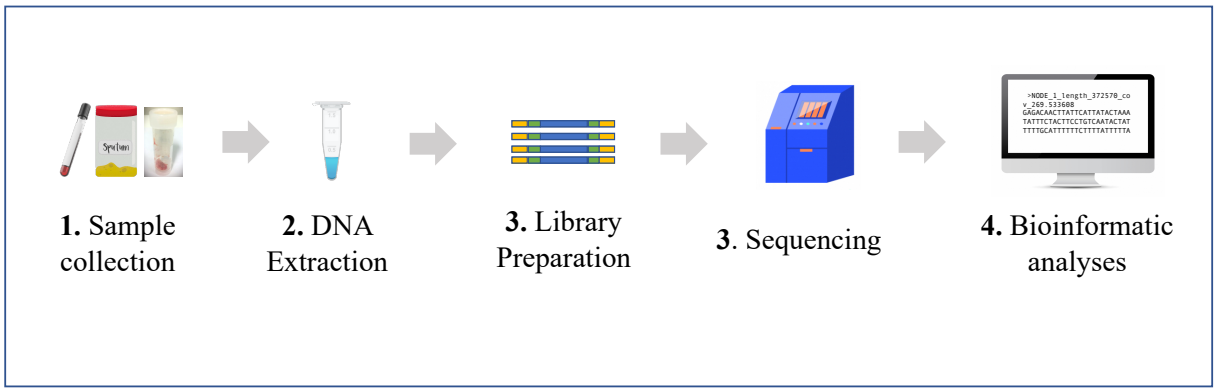


Figure 1. Next generation sequencing basic workflow.

1.2.1 Sample collection

Potential samples for NGS analysis include clinical specimens such as e.g. tissue, body fluids, and swabs[1]. Sample collection and preservation protocols can affect both the quality and accuracy of the NGS data. Protocols for sample collection and storage must be optimized and internally validated for different sample types. The length of time between sample collection and freezing for storage, as well as the freeze-thaw cycles used until DNA extraction, can affect the sample microbiological composition[21].

1.2.2 DNA extraction

DNA extraction is a crucial step during the NGS workflow to obtain reliable and quality results. The DNA extraction protocol used is highly dependent on the sample type[1]. DNA extraction is a procedure used to isolate DNA from cell membranes, proteins, and other cellular components by using physical and/or chemical methods[22]. The main features that are required in DNA extraction methods include good quantity and quality of DNA, removal of impurities and inhibitors, such as RNA and proteins, and high–throughput processing [23]. DNA extraction involves lysing the cells and solubilizing DNA, followed by chemical or enzymatic methods to remove macromolecules, lipids, RNA, or proteins. DNA extraction techniques can be classified

based on their purification strategies: organic extraction (phenol-chloroform method), nonorganic method (salting out and proteinase K treatment), solid-phase DNA extraction methods (silica-gel membrane), and ionic chelating resins (Chelex). Different strategies have been developed using some of these basic DNA isolation principles (or a combination of them) depending on the type of sample source. Currently, there are manual methods as well as commercially available kits that are used for DNA extraction[22].

Assessing the quality and yield of DNA

Estimating nucleic acid quality for library preparation is indispensable for the success of next-generation sequencing (NGS) applications. A broad variety of DNA sample types and extraction methods may introduce inhibitors that can negatively affect the sequencing process. Therefore, the quality and yield of DNA have to be assessed after the DNA extraction process, to avoid poor sequencing performance due to degradation, fragmentation, or contamination.

The origin of the sample, from which DNA is extracted, might influence the quantity and quality of the nucleic acid. Input DNA concentrations for SMg can be <100 pg or up to six orders of magnitude higher[1]. Commonly, the quality and yield of DNA are assessed by spectrophotometry or by gel electrophoresis. Illumina recommends UV spectrophotometry for purity assessment and fluorometric based methods such as Qubit or Pico/RiboGreen for nucleic acid quantitation. (http://www.illumina.com). Gel electrophoresis shows the DNA integrity, e.g. a ladder or smear below a band of interest may indicate sheared or degraded DNA. Furthermore, contaminating RNA can be detected on an agarose gel (since RNA has a lower molecular weight than genomic DNA). Other impurities, such as detergents or proteins, can be also be detected (smearing of DNA bands)[24].

1.2.3 Library preparation

After DNA extraction, the library preparation is the first step of the sequencing process. This prepares the DNA to be sequenced into a form that is compatible with the sequencing system to be used, using specific protocols. Libraries allow optimal use of the output capacity of high-throughput sequencers and increase the number of samples that can be analyzed in a single sequencing run by adding an additional molecular barcode sequence to the adapters (multiplexing)[20]. There are several library preparation kits available commercially from different quantities of starting materials[25].

A library is a collection of DNA fragments of uniform size ready to be sequenced. Library preparation is performed by tagmentation (DNA recombination) of sequencing adapters to DNA (e.g., Illumina's Nextera preparation)[26] or ligation of the adapter to sheared DNA[27] (used for the rest of Illumina preparations different to Nextera)[1]. Libraries are tagged with a library-specific DNA sequence (barcode) allowing multiple libraries to be pooled together and sequenced simultaneously. The library preparation process usually includes DNA shearing into fragments of ~500bp or less, either mechanically or enzymatically, followed by ligation of DNA fragments to platform specific oligonucleotide adapters and barcodes/indexes, and library amplification[1,6]. Depending on the platform, the amplification method can vary. There are two amplification methods: emulsion PCR (EmPCR) or bridge amplification[6] (Figure 2).

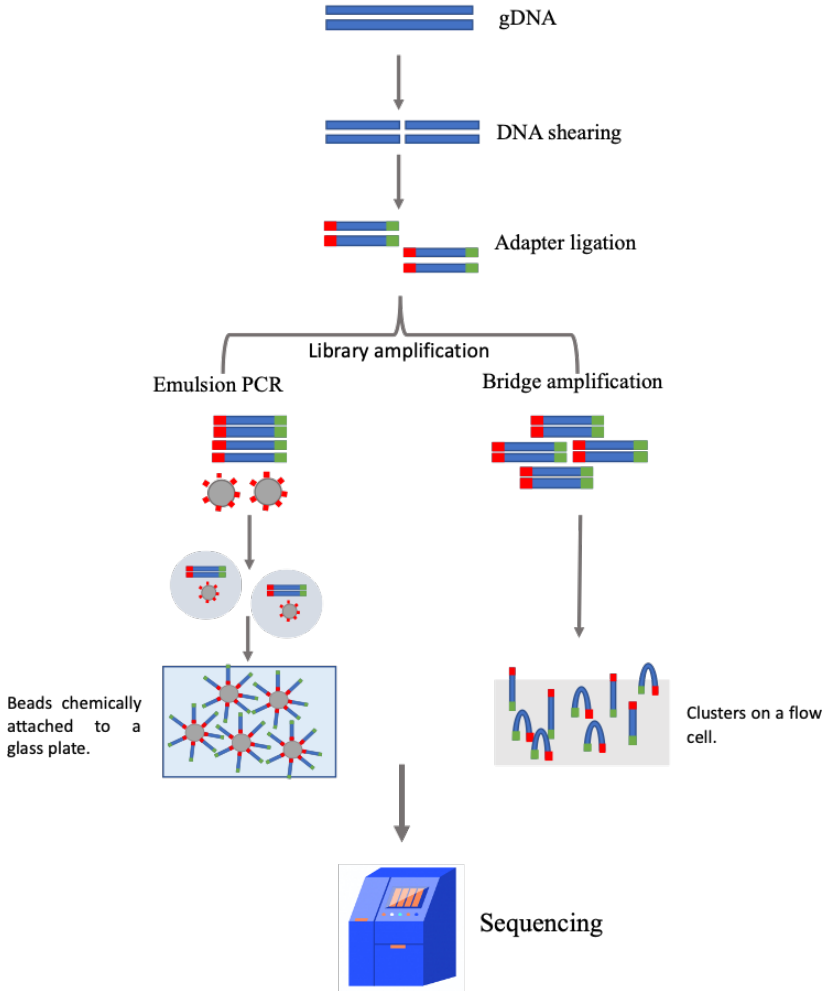


Figure 2. Schematic presentation of the workflow of library preparation.

The basic principle in EmPCR is dilution and compartmentalization of template molecules in a reaction mixture consisting of oil-aqueous droplets. PCR is performed within these droplets to

create beads with copies of the same template sequence. The beads are attached to a glass slide[28]. This type of amplification method is used by Ion Torrent platforms. On the other hand, Illumina platforms rely on a unique isothermal bridge amplification reaction, that occurs on the surface of the flow cell (Immobilized template) with immediately adjacent primers to form clusters (Figure 2).

1.2.4 Sequencing

After the library preparation, the sequencing follows, using one of the sequencing technologies described above. Then the reads obtained can be sorted by barcodes to get read data for each sample individually. Illumina reads are commonly 25-250 nucleotides long. Reads can be separated, known as single reads (SE), which involves sequencing DNA from only one end or reads can be paired (PE), allowing sequencing of both ends of a DNA fragment (Figure 3).

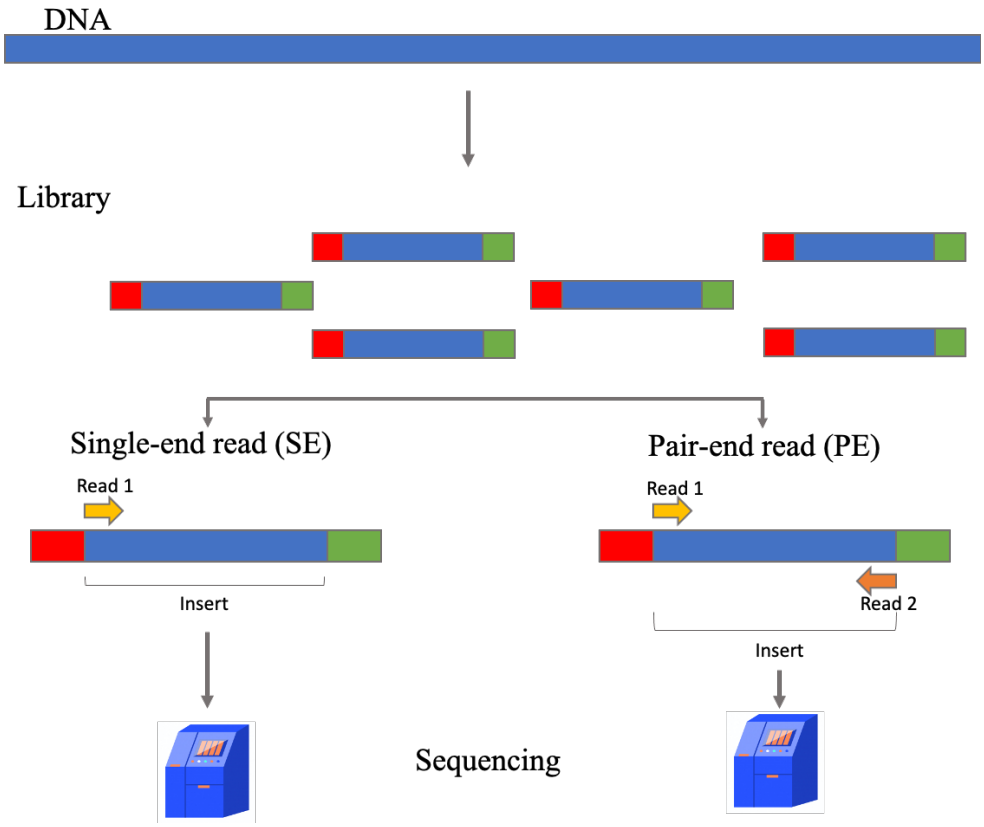


Figure 3. Single-end and pair-end reads.

1.2.5 Bioinformatic analyses

Bioinformatic analyses focus on the application of concepts, methods, tools, and software to analyze, integrate, and interpret biological data. Regardless of the NGS platform and the

application, the overall goal of each analysis is the same, therefore there are a few common steps for all the NGS approaches. However, each platform has its particularities and specificities[29]. Overall, all the bioinformatic workflows start with the preprocessing and quality assessment of the raw reads which is common for all the NGS applications, followed by the data analysis and interpretation, which can be directed onto many different paths depending on the needs of the user. The most common tools used for NGS analyses are listed in Table 2.

Table 2. Bioinformatic tools commonly used for NGS analyses in the microbiology laboratory.

Application	Tool	Link
Data quality Check		
Quality of the raw sequencing reads is evaluated, to remove, correct or trim the reads not meeting the standards.	FastQC NGS QC QC-Chain	www.bioinformatics.babraham.ac.uk http://www.nipgr.res.in/ngsqctoolkit.html http://www.computationalbioenergy.org/qc-chain.html
Trimming		
Adapter sequences should be removed from reads because they interfer with downstream analysis.	Trimmomatic Cutadapt BBDuk AdapterRemova	http://www.usadellab.org/cms/index.php?page=trimmomatic http://code.google.com/p/cutadapt https://jgi.doe.gov/data-and-tools/bbtools https://github.com/MikkelSchubert/adapterremoval/
Alignment and read mapping		
Read mapping is the process of aligning reads to a reference genome.	Deconseq Bowtie /Bowtie2 BWA Mauve Prokka	https://hpc.nih.gov/apps/DeconSeq.html http://bowtie-bio.sourceforge.net/bowtie2/ https://github.com/lh5/bwa http://darlinglab.org/mauve/mauve.html www.vicbioinformatics.com
Annotation		
Genome annotation includes identification of open reading frames and matching the identified segment to a database of known	Rast Spades	www.vicbioinformatics.com http://rast.nmpdr.org
Assembly		
Consist on using computer algorithms to align WGS reads to form longer DNA sequences known as contigs, and order the contigs into longer sequences called	Megahit Velvet OMEGA RAxML	http://bioinf.spbau.ru/spades www.ebi.ac.uk/~zerbino/velvet http://bioinf.spbau.ru/spades https://cme.h-its.org/exelixis/web/software/raxml/
Phylogeny		
Phylogenetic trees can be used to analyze and visualize SNP differences between	FastTree PhyML	http://sco.h-its.org/exelixis/software.html http://atgc_montpellier.fr/phyml/
Resistance		
Prediction of resistance genes in WGS data.	ARDB CARD ResFinder Abricate	https://ardb.cbcb.umd.edu https://ardb.cbcb.umd.edu https://card.mcmaster.ca https://github.com/tseemann/abricate
Virulence		
Prediction of virulence factors in bacteria.	VFDB VirulenceFinder	www.genomicepidemiology.org www.mgc.ac.cn/VFs
Typing		
Identification of bacteria at the species or clonal level.	CLC MLST BIGSdb	www.genomicepidemiology.org http://cge.cbs.dtu.dk/services/MLST/ http://bigsdb.readthedocs.io
Plasmids		
Prediction of plasmid sequences.	PlasmidFinder PlasmidSPAdes	https://cge.cbs.dtu.dk//services/PlasmidFinder/ http://cab.spbu.ru/software/plasmid-spades/

1.3 Application of NGS in clinical microbiology

The field of clinical microbiology comprises both diagnostic microbiology (identification of pathogens from clinical samples to species level and identification of their AMR patterns) and public health microbiology (surveillance and monitoring of infectious disease outbreaks in the community)[1]. NGS is rapidly finding a variety of applications in both clinical laboratories and research and it has been transitioning from research tools to diagnostic methods becoming more integrated into clinical microbiology laboratories.

The most common NGS approaches used in clinical microbiology include: whole-genome sequencing (WGS), targeted NGS methods (including amplification or probe hybridization), and metagenomic next generation sequencing (mNGS)[30,31]. The application of NGS in clinical settings is tremendous. They have successfully been used in a numerous range of clinical applications such as the diagnosis of infectious diseases[4], outbreak tracking and management[32–35], characterization and surveillance of pathogens[36,37], rapid identification of bacteria using the 16S-23S rRNA region, taxonomy[11,38,39], pathogen discovery[40], tracking the transmission of zoonotic microorganisms, microbiome studies, and metagenomics approaches on clinical samples among many others[11].

Whole genome sequencing analyses from cultured isolates have been extensively used for the characterization of pathogens, including species, strain type, antibiotic resistance, virulence, and other information for outbreak and case management[41]. However, this culture-isolate dependent approach is out of the scope of this study. Several reviews describe the different applications of the NGS in clinical microbiology in more detail[11,42].

1.3.1 Metagenomic approaches

Metagenomic approaches characterize all nucleic acid (DNA/RNA) present in a sample, which may contain mixed populations of microorganisms, and assigning these to their reference genomes to understand which microbes are present and in what proportions. The ability to sequence and identify nucleic acids from multiple different taxa for metagenomic analysis makes this a powerful new platform that can simultaneously identify genetic material from entirely different kingdoms of organisms. This approach has been used for characterizing several niches, from the environment to the microbiome. In the clinics, the study of metagenomics primarily was incorporated into the study of infectious disease diagnostics,

microbiome analysis, and oncological applications[19]. Nowadays, in clinical microbiology laboratories, these methods have a huge potential especially for the identification of microorganisms.

There are two general classes of metagenomic approaches that can be used for the identification of microorganisms in the clinical microbiology laboratory: *targeted amplicon sequencing*, usually based on amplification and sequencing of a phylogenetic marker and *shotgun metagenomics* (SMg), sequencing the genetic material present within a sample directly without amplification[43]. When applied to clinical samples, these approaches have been referred to as "metagenomic sequencing" or Metagenomic Next Generation sequencing (mNGS)[9,44]. However, not all sources define the amplicon sequencing approach as metagenomics, as many researchers prefer to use these terms for the shotgun sequencing approaches[44] since they consider that even universal or broad-range PCR methods are not sufficiently broad to be considered metagenomic. A general overview of both methods is shown in Figure 4.

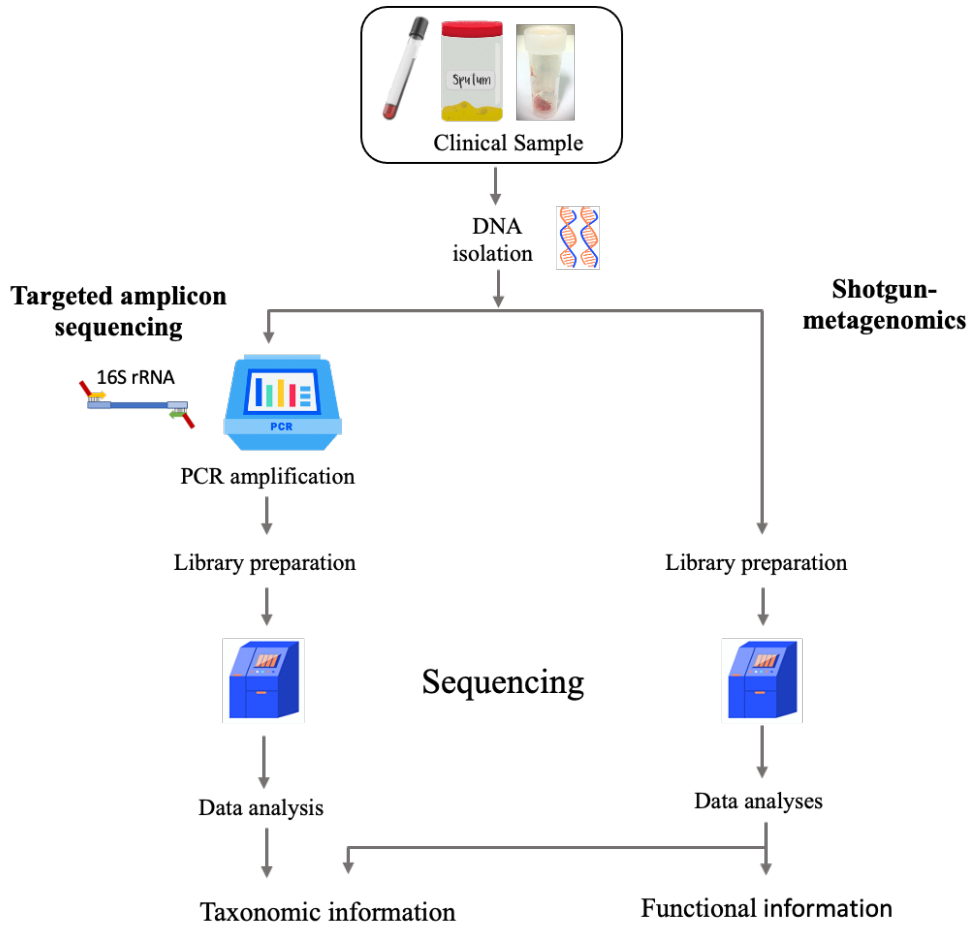


Figure 4. General overview of amplicon targeted (16S rRNA gene) and shotgun metagenomics.

Targeted amplicon sequencing is a widely used approach that provides only taxonomical classification for a broad range of pathogens. DNA is extracted from a clinical sample and subjected to PCR amplification using an appropriate set of PCR primers that targets a taxonomically informative gene such as the ribosomal RNA (rRNA) genes that are universally conserved among bacteria (16S or 23S rRNA) or fungi and parasites (18S rRNA, 28S rRNA or internal transcribed spacer (ITS). The other alternative is the use of other sets of primers that can be designed to target a defined set of pathogens and/or genes and used for multiplex reverse transcription PCR or PCR (multiplexed amplicon PCR). Amplicons from separate samples are then given molecular barcodes, pooled together, and sequenced. Following sequencing, raw data is analyzed with bioinformatics tools which include trimming, error correction, and comparison to a reference database. Subsequently, the reads are assigned to a phylogenetic rank, a taxonomy profile can be generated, by the alignment of consensus sequences to an appropriate reference database[45], enable to determine which microorganisms are present in the sample and their relative abundance allowing that pathogen(s) may be identified to the genus or species level.

The 16S rRNA gene is the most widely used taxonomically informative gene used in NGS methods. This gene is approximately 1500 base pairs (bp) in size and its genetic structure comprises nine highly conserved and nine hypervariable regions (V1–V9). The conserved regions can serve as universal primer binding sites for the PCR amplification of gene fragments, whereas the hypervariable regions contain considerable sequence diversity, useful for prokaryotic identification purpose[46]. By comparing these hypervariable regions to 16S rRNA gene sequences of designated prokaryotic type strains available on large public databases, researchers can generate accurate identification of the prokaryotic taxa present within clinical samples.

The benefits of targeted methods over SMg include: (i) increased sensitivity for the microorganism of interest since it has inherent an amplification step[9]. (ii) Decreased cost due to the possibility of extensive multiplexing of samples, which allows researchers to process hundreds of samples and analyze millions of PCR amplicons in a single NGS-run. (ii) Since sequencing is restricted only to the region of interest of a single gene, lower sequencing depths are required and the computational analyses are simpler[3,44], and (iv) It is faster when compared to SMg[11].

Among the limitations for using targeted amplicon metagenomics are: (i) it does not provide genetic information beyond species identification, such as the presence of AMR genes or VF genes unless these regions are specifically included in the targeted sequencing reaction[3,44]. (ii) This approach also requires a hypothesis about which organism group (bacterial or fungal) is suspected to ensure that appropriate amplification targets are chosen[44]. (iii) Universal primers used to target the 16S rRNA gene for polymicrobial may also be a problem. If polymicrobial populations are present when using 16S sequencing, multiple base-calls will be made per nucleotide, producing a mixed nucleotide chromatogram that cannot be interpreted. (iv) PCR amplification bias can cause preferential amplification of some targets, which may affect the proportions of the taxa identified in the clinical specimen[44].

1.3.2 SMg in clinical microbiology

The application of WGS on infectious agents in clinical samples to obtain the complete genome sequences are emerging and are currently known as Shotgun-Metagenomics (SMg)[47]. When SMg is directly applied to the clinical specimens it is known as "clinical metagenomics"[47].
It is an alternative approach to characterizing microbial communities, where all nucleic acids present in a specimen, including those derived from host and from any microorganism(s), are extracted and sheared into small fragments that are independently sequenced in parallel[47,48], producing relative abundance information for all genes detected. This method may not only identify microorganisms but also provides information on the types of genes presents within a clinical sample, eventually inferring functional characterization.

Currently, there are four main potential applications of SMg for pathogen characterization in the diagnostic microbiology laboratory: identification, molecular typing, antimicrobial resistance detection, and virulence gene detection. SMg has the possibility to be used as a single and *rapid comprehensive test* for potential pathogens[49]; has the advantage of not being limited to certain pathogens[50]. It is a promising approach to opening huge opportunities for detecting, identifying, and characterizing all potential pathogens in a sample, even if the sample present more than one bacterium (polymicrobial) [50,51].

Kumeren *et al.*[52] describe a systematic review and meta-analysis including studies that use metagenomic sequencing as a clinical diagnostic tool for infectious diseases. To date, several studies have provided evidence of the potential and successful applications of SMg as a

diagnostic tool for infectious diseases in a variety of clinical specimens, including urine[4,5,53], blood samples[54–56], respiratory secretions, spiked and clinical samples for bacterial (e.g. *Mycobacterium tuberculosis* and viral pathogens[57–60], Cerebrospinal fluid[59,61], intraocular fluid [62,63], cerebrospinal fluid[59,61,64], orthopedic device, sonication fluid[65–71] and bone and joint[65–72]. Several reviews have summarized the advances, limitations, and challenges in the field[19,31,73,74].

1.3.2.1 Detection of AMR and virulence-associated genes

The sequencing data provided by SMg may provide clinically useful information beyond the identification of a potential pathogen[75]. If the SMg depth of coverage of the microorganism is sufficiently high, valuable functional information such as antibiotic resistance and virulence-associated genes can be revealed[63]. SMg can identify previously described and potentially novel antimicrobial resistance determinants, although determinants on mobile genetic elements including plasmids are considered a greater challenge.

SMg-based detection of AMR can be further applied to predict phenotypic antimicrobial susceptibility testing (AST). There are two prediction models used to correlate AMR detection with AST phenotype: (i) rule-based approaches using databases of AMR loci and (ii) model based approaches using machine learning and/or statistical models[76]. The rule-based approach is based on the detection of antimicrobial resistance genes or k-mers (small portions of genes). Detection can be done using either raw reads or assembled genomes. It is done using software that searches databases of AMR. Resistance profiles from raw reads obviate the need for assembly and can, therefore, reduce the time to results. However, because of sequencing errors present in individual reads or DNA contamination from other organisms, false positives may be introduced[77]. In the model-based approaches, the classifier is trained based on a set of genomes with known phenotypes without linking individual AMR genes. The models can use k-mers, raw reads, contigs, or assemblies to predict phenotypic results. Most of the studies predicting AST from NGS have used Illumina platforms and the rule-based approach[76].

The other main potential use of SMg for organism characterization is the detection of genetic markers of virulence[78]. The pathogenesis of bacterial pathogens is determined by the virulence factors (VFs) that enable them to cause infection. The potential horizontal transfer of VFs between different strains or species of bacteria makes the emergence of new pathotypes of bacteria almost inevitable. Comprehensive characterization of the VFs carried by the new

pathotypes of emerging bacterial pathogens is critical for the effective prevention and control of infectious diseases.

Databases for ARGs and VFs play an essential role in studying the resistome and the virulome from (a) microorganism(s) in clinical samples. A critical first step in many such approaches is the construction and curation of databases. There are several AMR databases that have been developed, among them: ResFinder[79], the Comprehensive Antimicrobial Resistance Database (CARD)[80,81], Basic Local Alignment Search Tool (BLAST), Antibiotic Resistance Gene-ANNOTation, short read sequencing typing, and National Center for Biotechnology Information's (NCBI) National Database of Antibiotic Resistant Organisms[82]. However, they were not developed for clinical purposes and therefore they have a certain degree of restriction in clinical diagnosis. Similarly, there are comprehensive databases available to predict the VFs present in bacteria, such as VirulentPred [83], Virulent-GO[84], and the virulence factor database (VFDB)[85,86], which provide knowledge of VFs and serve as a comprehensive repository of bacterial pathogenesis knowledge for the scientific community[86].

1.3.2.2 Implementation of SMg in the microbiology laboratory

The use of any diagnostic test in the clinical laboratory requires analytical and clinical validation, as well as the careful monitoring and documentation of quality control and proficiency testing[87]. Currently, the FDA has provided general guidelines for clinical validation of NGS infectious disease testing[19]. However, so far, there are no FDA-approved SMg approaches for diagnosing infectious diseases. For SMg to be extensively implemented in the clinical microbiology laboratory, further work is needed in terms of quality control and validation. The base is to apply what is learned from the application of NGS in other areas and adapt that knowledge to the microbiological diagnosis, which present unique challenges, such as highly variable specimen complexity and quality, and the broad genetic diversity of microorganisms, among others.

Among the requirements suggested for NGS test establishment in clinical laboratory[87], the following aspects are included: test validation, quality control procedures (to assure and maintain accurate test results), independent assessment of test performance through proficiency testing or alternative approaches, and reference materials[88]. To date, most published studies are retrospective analyses of curated samples normally in sample sizes that are not optimal for test validation (case studies). They have been done using individualized laboratory-developed

workflows that are highly dependent on the type of infection and the type of sample[74]. Workflows vary regarding methods, tools, and interpretation, making it difficult to compare and assess the variability of the results. Therefore, rigorous and extensive validation must be performed.

Prior to clinical implementation, all laboratory test protocols must be optimized and validated. Assay validation is the procedure to establish analytical performance specifications for certain performance characteristics, to ensure the analytical validity of test results prior to clinical testing. Analytical performance specifications include accuracy, precision, analytical sensitivity and specificity, reproducibility, and reportable range (limit of detection)[89].
So far, it has been reported just one developed and validated mNGS assay for clinical diagnosis of infectious causes of meningitis and encephalitis from cerebrospinal fluid (CSF) in a microbiology laboratory[59,61]. In fact, it is expected, that the application range of this approach will be limited to small targeted assays that are validated[90].

1.3.2.3 Quality assessment

Standardization of protocols and quality metrics for both the wet and dry laboratory (including sample preparations, reagent selection, sequencing process, bioinformatic pipeline, interpretation, and report) is critical to ensure the validity of the test result[91]. Quality control metrics specific for NGS should be established, including depth of coverage and quality scores as part of standard operating procedures. Additionally, it is important to establish clinical validity, to ensure that the test is appropriate for the intended clinical use[88,90].

The development of reference materials is important to support mNGS assay validation, performance assessment, and quality control. Reference standards materials are also important to be used as standard metrics suitable for comparison among laboratories[92]. Among them, characterized nucleic acids, whole microbes, spiked matrices (mock community), and data set, can be used as controls for assay validation and quality control[90].

Quality control is essential for reliable and valid metagenomic results and will be essential for regulatory approval for clinical applications. Both positive and negative controls must be run with the clinical samples through the entire process. Positive and internal controls mitigate the risk of a false negative result. Positive controls in mNGS assays consist of a negative matrix with a spiked-in microorganism(s) that are known[74]. External controls can be developed using

well characterized mock microbial communities to establish detection limits for mNGS[19]. Positive controls enable the detection of performance failures in several steps in the workflow. On the other hand, negative controls are required to detect contamination and avoid false positive results. Negative controls are samples that do not contain any possible pathogens. Negative controls enable the detection of contamination from the reagents or other sources, or it may indicate a problem in any of the workflow steps[74]. Previous studies have used elution buffer from the nucleic acid extraction kit[72], DNA extractions (blank) control[63], transport medium[93], water, and non-infectious samples as negative controls[94].

In addition to positive and negative controls, there are other controls during the process that must be considered for quality assurance for mNGS. Among them: (i) Process control and checkpoints that can ensure the quality of material before going to the next step in the workflow (e.g. quantification of DNA before library preparation). (ii) Contamination control since contamination may be introduced in every step of the workflow, for instance from laboratory reagents, cross-contamination from other samples. (iii) Database quality control, because databases may not be curated or updated, affecting sensitivity. Therefore, databases may require re-validation and version control, as they impact the accuracy of the results, and (iv) bioinformatics quality control, when for instance changes in the pipeline includes software updates, input parameters, algorithms, or databases. All of these changes must be documented and added to the standard operational procedures[1].

Proficiency testing is another important component of the quality assurance assessment that is required to verify the accuracy and reliability of laboratory testing (like in any clinical laboratory test). However, for mNGS, it is very challenging since there is not a well characterized reference material on the market. Therefore, the design of suitable alternative testing materials should be used, e.g. leftovers from clinical samples, previously analyzed and confirmed using traditional methods, that then may be compared to evaluate the consistency of the results[74].

2 Prosthetic joint infection (PJI)

2.1 Prosthetic joint replacement

Joint replacement surgery is the major procedure to alleviate pain, improve mobility in patients with destructed joints[95]. It is one of the most successful surgeries performed currently with the

significant patient and health economics benefits. Currently, there are arthroplasties available for hips, knees, shoulder, elbow, and ankle; where hip and knee arthroplasties are among the most commonly performed procedures[96]. Joint replacements are referred to as either primary or revision arthroplasty, according to the number of times that a given joint is replaced. Primary arthroplasty is the first time that a native joint is replaced; revision arthroplasty is a second or subsequent surgical procedure performed when a joint replacement fails and some or all parts of the original prosthesis need to be removed or replaced.

Globally, the number of joint arthroplasty surgeries performed each year is increasing and it is expected that as the need for primary arthroplasties increases so will the need for revision arthroplasties. In the United States, until 2014, approximately 4.7 and 2.5 million have undergone total knee arthroplasty (TKA) and total hip arthroplasties (THA) respectively[97]. The number of hip and knee replacements in most European countries has also increased in recent years, although rates between countries vary considerably. Data collected from 24 European joint replacement registries (most of them of national coverage), showed that until 2018, over 3.1 and 2.5 millions of Europeans have undergone primary THA and TKA respectively[98]. According to the Norwegian Arthroplasty register[99], 233 142 and 97 022 THA and TKA cases during the period between 1994 – 2018 were reported, respectively, with approximately 9000 primary hip arthroplasties and 6 900 primary knee operations performed in 2018 in Norwegian hospitals.

While the majority of joint arthroplasties are generally successful, postoperative complications commonly occur and will require additional surgery at some point during the life of the device. Among the complications of prosthetic joint surgeries are instability, fracture, dislocation, aseptic loosening, and infections of the prosthesis[100,101]. Aseptic loosening is nowadays the major cause of arthroplasty failure[102]. However, infections related to joint prostheses occur less frequently than aseptic failures but represent the most devasting complication[103].

2.2 PJI definition

Prosthetic joint infection (PJI), also referred to as periprosthetic infection, is defined as an infection involving the joint implant and adjacent tissue[104]. It is a devastating complication of total joint arthroplasty which appears in 0.5-4 % of all the cases[103,105] , and may exceed 10 % in revision surgery or patients with specific risk factors e.g. diabetes, systemic infections or immunocompromise[106]. PJI is associated with high morbidity, need for complex treatment

including the need for prolonged hospitalization, repeat surgery, antibiotic exposure, and societal costs (recent estimated costs of 20,000–40,000 dollars per infection[107]). PJI can lead to an unsatisfactory functional results or even permanent disability, including arthrodesis or leg amputation.

There is no standard definition of PJI. However, there are several consensus guidelines, such as those produced by the Infectious Diseases Society of American (IDSA)[108] and the Musculoskeletal infection society (MSIS)[109], that have attempted to provide a uniform, evidence-informed approach to the management of PJI. Among the most recently proposed, there is the one that has been presented at the annual meeting of the European Bone and Joint Infection Society (EBJIS) 2018 in Helsinki, Finland (Table 3)[110], which has been shown to offer better sensitivity for diagnosing PJI[111,112].

Table 3. Definition of PJI proposed by EBJIS, 2018[110].

Organization	Definition
The European Bone and Joint Infection Society (EBJIS)	***PJI is present when at least one of the 4 criteria are fulfilled***
	Sinus tract or visible purulence
	Acute inflammation in periprosthetic tissue
	Elevated synovial leukocyte count (>2000/μL)
	Elevated synovial granulocytes (>70%)
	Microbiology
	Synovial fluid or tissue samples or sonication fluid (≥ 50 CFU/mL)
	Microbial culture in ≥ 2 Positive samples
	For highly virulent organisms already one positive sample confirms infection.
	Under antibiotics and for anaerobes, <50 colony-forming unit (CFU)/mL can be significant

Abbreviations: CFU, Colony Forming Unit.

2.3 Diagnosis of PJI

2.3.1 Diagnostic criteria

The conventional diagnostic strategy to diagnose PJI is usually simple. Infection is suspected based on clinical findings (local pain, erythema, edema, fever and wound secretion) then confirmed by examination of samples obtained by simple aspiration (synovial fluid) or biopsy (periprosthetic tissue), which identifies the causative organism(s), thus defining the pharmacological and surgical treatment strategy[113]. However, the clinical presentation of prosthetic joint infection is variable and sometimes difficult to distinguish from that of aseptic implant failure[114] in addition to several interconnected issues that make the diagnosis of PJIs a challenging aspect of management for these infections, among them: the lack of a standard

definition of PJI[109,115], the lack of a perfect diagnostic reference standard, the presence of biofilm and challenges differentiating pathogens from contaminants.

Currently, there is an array of laboratory methods available that aim to improve the diagnosis of PJI (Table 4). However, none of these diagnostic methods has achieved satisfactory specificity or sensitivity in common practice and it is currently recognized that the complex management of these cases implies a multidisciplinary approach. One key part is the microbiology, especially due to the importance of diagnosis[107,113,116–118].

Table 4. Conventional PJI diagnosis test. Modified from [116].

Test type	Description of the diagnostic test
Pre-operative assessment	
Serological tests	White blood cells (WBC) count and differential; Erythrocyte sedimentation rate (ESR); Serum C-reactive protein level (CRP); Serum biomarkers for detecting PJI: Serum IL-6, procalcitonin, TNF-α, among others.
Synovial fluid aspiration	WBC count and differential; Gram stain and culture; PCR; leucocyte esterase (LE); Synovial CRP; synovial biomarkers: α-defensin, Interleukin (IL)-1β, IL-6, IL-8, IL17, TNF-α, among others.
Radiographic imaging	Plain radiographs Computer tomography.
Radionuclide bone scanning	Scintigraphy using a number of isotopes including: Technetium 99m, Gallium-67 citrate, or Indium-111-labeled leukocyte or immunoglobulin scan.
Positron emission tomography (FDP-PET)	Fluorine-18 fluorodeoxyglucose (F-18 FDG) positron emission tomography
Intra-operative assessment	
Periprosthetic tissue	Histopathology; Gram stain and culture
Explanted prosthesis	Culture

Abbreviations: WBC, White Blood Cells; ESR, Erythrocyte Sedimentation Level; CRP, C-reactive protein level; IL, Interleukin; TNF-α, Tumor Necrosis Alpha Factor; LE, Leukocyte esterase.

Microbiological diagnosis of PJI

Microbiological assessment is considered an important criterion for diagnosis of PJI playing a central role in infection confirmation, and it is also critical for determining antimicrobial susceptibility of the pathogen(s) to guide antimicrobial treatment. Although multiple definitions of PJI have been proposed, all of them include identification of pathogen(s) as a key component. The microbiological diagnosis of infection generally depends upon the isolation of a pathogen from a clinical specimen by conventional inoculation onto solid and/or liquid media under aerobic and anaerobic conditions[119]. Guidelines from different bodies have been published that

recommend strategies for identifying the organism(s) causing PJI from both pre-operative and at the time of surgery[108,120].

Among the microbiological definitions of PJI, is the isolation of two or more identical isolates from multiple specimens, or the isolation of one highly virulent organism from a single sample. Notably, single positive tissue or synovial fluid cultures, especially for organisms that may be contaminants (e.g. coagulase-negative staphylococci (CoNS), *Cutibacterium acnes*), should not be considered evidence of definite PJI[120]. The specimens commonly recommended for diagnosis of PJI are: synovial fluid, periprosthetic joint tissue (PJT) and sonication fluid for removed implant[101]. Two or more intraoperative cultures or a combination of preoperative aspiration and intraoperative cultures that yield the same organism is considered definite evidence of PJI.

The "usual suspects'' causing PJI

A wide range of microorganisms can cause PJI. They include Gram-positive bacteria, Gram-negative bacteria and, although less common also fungi. The majority of PJIs reported are caused by a single microorganism (~85%), but polymicrobial PJI is also possible[121]. Frequency of microorganisms causing PJI can vary significantly by time after implantation and the site of the joint prosthesis, among others[96]. Table 5 contains a list of the most common causes of prosthetic joint infection and the range reported in several studies from different countries and a study made in Norway in total hip arthroplasties (THAs)[122]. In Norway, in a nationwide study, based on 278 revisions of infected THA, staphylococci were the most common bacteria in THA revision for infection. S. aureus was more common in acute postoperative infections and CoNS were more common in early, delayed and late infections[122].

Table 5. Common causes of PJI.

Microorganism	Frecuency (%)	
	International data	Norwegian data
Aerobic Gram-positive cocci	>50	60
Staphylococcus aureus	12-28	19
Coagulase-negative Staphylococci (CoNS)	13-40	41
Staphylococcus epidermidis	23	16
Streptococcus sp.	8-10	11
Streptococcus agalactiae	2.8	-
Enterococcus sp.	3-10	9
Aerobic Gram-negative bacilli	10-42	6
Enterobacteriaceae	20	-
Escherichia coli	9	-
Anaerobic bacteria	3-24	-
Cutibacterium acnes	0.9-19	-
Polymicrobial	10-15	10
Culture negative	5-42.1	-

Data from: [a]International data[96,121,131,132,123–130], [b]Norwegian data[122].

Polymicrobial PJI tends to occur more frequently after implantation than monomicrobial PJI. There has been little research about polymicrobial PJI, including the frequency of the different bacteria involved. Existing studies have reported a prevalence of polymicrobial PJI infections in the range between 10-15%. Differences in the distribution of bacterial species between monomicrobial and polymicrobial have been observed, probably due to the fact that certain types of microorganisms may be more likely to grow together than separately. In a recent study by Flurin *et al*. 2019[121], *Staphylococcus epidermidis* and other CoNS were commonly found in polymicrobial PJIs. S. epidermidis was found in 60% of polymicrobial PJI and it was found at a higher rate in polymicrobial than monomicrobial infections. The most common co-pathogen found with S. epidermidis was E. faecalis. In addition to CoNS and enterococci, *Corynebacterium* sp. and *Finegoldia magna* are common in polymicrobial infections[121,127]. *S. aureus*, aerobic gram-negative bacilli, including *Pseudomonas aeruginosa* are also within the most isolated bacteria. In Norway, the incidence of the different bacteria is mostly unchanged throughout the study period. There is a trend towards more polymicrobial infection and the combination of CoNS and *Corynebacterium* sp. is most common[122].

The prime suspect in PJI: *Staphylococcus aureus*

Staphylococcus aureus is regularly reported to be the most common pathogen in PJIs *(12-27%)*[96,123–125,128]. For most healthy individuals, colonization is unproblematic. However, if *S. aureus* contaminates a breach in the skin or mucous membranes, it can go on to infiltrate the underlying tissue causing an infection. There is evidence that support that *S. aureus* from nose and PJI isolates of *S. aureus* in arthroplasty patients are genetically indistinguishable, suggesting that commensal *S. aureus* clones are capable of causing PJI [133].

S. aureus taxonomy and general features

S. aureus was first described in 1880 by the Scottish surgeon Sir Alexander Ogston, from a pus sample originated from a surgical abscess in a knee joint. The *Staphylococcus* name comes from the Greek *staphyle* (bunch of grapes) and *kokkos* (berry), due to its spherical shape (cocci), which cluster together into colonies that look like grapes, when observed under the microscope. Then, in 1884, Friedrich Julius Rosenbach named the bacteria *S. aureus* due the color of the colonies, *S. aureus* from the Latin *aurum*, gold[134].

S. aureus belongs to the phylum *Firmicutes*, class Bacilli, order *Bacillales*, family *Staphylococcaceae*, genus *Staphylococcus*. Today, there are 54 and 22 validated species and subspecies of *Staphylococcus* respectively (according to (https://lpsn.dsmz.de/), accessed June 07/2020)[135,136]. *S. aureus* in contrast to other *Staphylococci* members, form fairly large yellowish pigmentation on their colonies due to the production of staphyloxantin[137]. By testing the coagulation of rabbit serum, *S. aureus* can be readily distinguished from other *Staphylococcus* species, which are coagulase negative. They are Gram-positive bacteria, facultative anaerobes, catalase positive, oxidase negative, coagulase positive, and non-spore forming organisms. They can cause hemolysis in blood agar plates due to the secretion of several toxins (α, β, δ, ϒ).

S. aureus colonizer and pathogen

S. aureus is considered both a commensal colonizing different human body parts and a pathogen that causes a wide range of clinical infections[138]. Humans are constantly exposed to *S. aureus* in the environment, and the bacteria are well adapted to colonize multiple body sites which provide several niches for this species. *S. aureus* colonizes skin and mucous membranes in humans but also in several animal species. The nares, throat, and perineum are the most prevalent sites for carriage in the general adult population. In the human body, *S. aureus* is considered part of the human microbiota[139]. It has been estimated that around 20-30 % of the

healthy adult population are asymptomatically carriers with 20 % and 60% being colonized permanently or temporarily, respectively[140]. Colonization is considered an important risk factor for the development of *S. aureus* infection and nosocomial infections[140].

S. aureus is an opportunistic pathogen that has the potential to cause many different diseases. Among these, both healthcare- and community-associated bloodstream infections, skin, and soft tissue infections, endocarditis, septic arthritis, osteomyelitis, and infections of prostheses. The success of *S. aureus* reflects an array of abilities, for instance, its capacity to adapt to different environments, which can vary widely within the host, such as pH, nutrient availability, or oxygen tension; inducing the bacterium to express a variety of different phenotypes[139].

S. aureus genome evolution and molecular typing

Bacteria retain most of their genetic information from generation to generation. However, they also need to develop strategies that allow them to acquire new genetic material in their genomes to adapt and survive in an environment that changes continually. Genomes of more closely related bacteria are more conserved, but the genome variability exists within different genera and among different isolates of a single bacterial species. The mains mechanisms that contribute to the plasticity of the bacterial genome are the acquisition of DNA (gene gain), and the loss of genetic information (gene loss)[141]. The sequencing of *S. aureus* has provided valuable information to the understanding of this pathogen. Approximately 554 complete genomes assembly has been registered in the NCBI database to this day. The *S. aureus* genomes are about 2.83686 Mb in size (median total length), with approx. 2800 coding sequences and a median G+C content about 32.7% (NCBI accessed 07/06/2020).

The bacterial pan-genome is divided into the core and accessory genome. The core genome refers to the stable regions with relatively low mutational capacity containing the genes present in all strains. The core comprises ~75% of the 2.8 Mb genome of *S. aureus* and is highly conserved among strains. The accessory genome refers to genes not present in all strains of a species and it is where more of the genetic diversity of pathogenic bacteria occurs. The accessory genome comprises ~25% of the total *S. aureus* genome[142]. Mediators of virulence, immune evasion, and antibiotic resistance are commonly found in the accessory genome. Including, mobile genetic elements (MGEs) such as *S. aureus* pathogenicity islands (SaPIs), bacteriophages, staphylococcal cassette chromosomes (SCC), transposons and plasmids, which are acquired by horizontal gene transfer (HGT)[143]. Therefore, the accessory genome provides a

considerable resource for bacterial flexibility to improve their fitness and, potentially, pathogenicity, and virulence [144]. The success of S. aureus as both a colonizer and a pathogen is mostly because it has the ability to adapt to different environments due to the acquisition of new DNA by HGT and to spread clonally.

Molecular characterization of *S. aureus* populations is an important component in the study of clonal relatedness, evolutionary pathways, the genetic diversity of the pathogen, and tracking the spread of S. aureus infections[145]. Various typing techniques have been used for discriminating bacterial isolates and characterizing *S. aureus* populations, e.g. Pulsed-Field gel electrophoresis (PFGE), multilocus sequence typing (MLST), and *spa* typing (Staphylococcal protein A)[146].

MLST directly measures the DNA sequence variations in a set of housekeeping genes and characterizes strains by their unique allelic profiles, with the resulting data made publicly available online. Seven housekeeping genes are used in *S. aureus* MLST typing including carbamate kinase (*arcC*), shikimate dehydrogenase (*aroE*), glycerol kinase (*glpF*), guanylate kinase (*gmk*), phosphate acetyltransferase (*pta*), triosephosphate isomerase (*tpi*) and acetyl coenzyme A acetyltransferase (*yqiL*), as specified by the MLST website (http://saureus.mlst.net)[147]. The analyzed isolates are then assigned to a specific sequence type (ST). The comparison of different isolates and their STs has allowed the comparison of different isolates. MLST has become an important and reliable technique for epidemiological analysis. For instance, it has been found that methicillin-resistant S. aureus (MRSA) strains, including healthcare-acquired MRSA (HA-MRSA), community-acquired MRSA (CA-MRSA), and livestock-associated MRSA (LA-MRSA) strains have produced different STs, which help to identify them[148].

Based on typing, STs can be clustered together into clonal complexes (CCs). A clonal complex (CC) is defined as groups of STs in which every ST shares at least six out of seven identical alleles. The associated clonal complex (CC) is calculated using the eBURST algorithm (http://saureus.mlst.net.eburst)[149], which is used to cluster related STs[149]. S. aureus has a highly clonal population structure with CCs comprising closely related, although not identical, genetic backgrounds[150].

Now there is an extended MLST, which analyses the alleles of several hundred or even thousands of genes (core genome MLST (CgMLST) or Whole-genome MLST (wgMLST)) and consequently has a high discriminative power allowing the comparison of genetic relatedness between bacteria even on a sub-species level[151]. In principle, wgMLST can provide a higher resolution as the distance matrix is computed on a larger set of loci. However, some studies have demonstrated that results derived from wgMLST and cgMLST approaches are often quite similar[152]. Studies have confirmed that WGS is a suitable strain characterization and typing method for *S. aureus*[153].

S. aureus antibiotic resistance and virulence factors

Many factors seem to contribute to the success of *S. aureus* as a pathogen, such as (i) its ability to persist as a commensal, (ii) its acquired resistance to multiple antimicrobial agents, and (iii) the diverse repertoire of virulence factors (VFs) that facilitate its ability to invade the host, cause disease, and evade host defense mechanisms (Table 6)[143].

Table 6. Factors contributing to the success of *S. aureus*. Modified from[143]

Category	Gene name	Gene product	Function	Location
Virulence factors[a]	*aur*	Aureolysin	Tissue destruction	Core
	capA and *capP*	Capsular polysaccharide biosynthesis proteins	Immune evasion	Core
	chp	Chemotaxis inhibitory protein	Immune evasion	Bacteriophage
	clfA and *clfB*	Fibrinogen binding proteins	Adhesion	Core
	coa	Staphylocoagulase	Coagulation	Core
	ebhA and *ebhaB*	Extracellular matrix-binding proteins	Adhesion	Core
	eta	Exfoliative toxin A	Scalded skin syndrome	Plasmid
	etb	Exfoliative toxin B	Scalded skin syndrome	Bacteriophage
	etd	Exfoliative toxin D	Scalded skin syndrome	vSAγ
	geh	Lipase	Lipid degradation	Core
	hld	δ-Haemolysin	Haemolysis	Core
	hlgA,*hlgB* and *hlgC*	γ-Haemolysin components	Haemolysis	Core
	hysA	Hyaluronidase	Tissue invasion	vSaβ
	lukD and *lukE*	Leukotoxins	Immune evasion	SaPI
	lukS-PV and *lukF-PV*	Panton–Valentine leukocidin	Leukotoxin	Bacteriophage
	sak	Staphylokinase (protease III)	Clot dissolution	Bacteriophage
	sea	Enterotoxin A superantigen	Food poisoning	Bacteriophage
	seb and *sec*	Enterotoxin B and enterotoxin C (superantigens)	Food poisoning	SaPI
	seq2 and *sek2*	Enteroxin and superantigen	Food poisoning	SaPI
	sep	Enteroxin P	Food poisoning	Bacteriophage
	spa	Immunoglobulin G-binding protein A	Immune evasion	Core
	sspA	Serine protease	Tissue destruction	Core
	sspB	Cysteine protease	Tissue destruction	Core
	tst	Toxic shock syndrome toxin 1	Superantigen	SaPI
Antibiotic resistance	*aacA –aphD*	Bifunctional AAC–APH protein	Aminoglycoside resistance	Transposon
	aadD	Aminoglycoside adenyltransferase	Aminoglycoside resistance	Plasmid
	ant (4')	*O*-nucleotidyltransferase(4')	Aminoglycoside resistance	SCC*mec* Plasmid
	ant (9)	*O*-nucleotidyltransferase(9)	Aminoglycoside resistance	Transposon
	blaZ	β-Lactamase	Penicillin resistance	Plasmid Transpososn
	bleO	Bleomycin binding protein	Bleomycin resistance	SCC*mec Plasmid*
	cat	Chloramphenicol acetyltransferase	Chloramphenicol resistance	Plasmid
	dfrA and *dfrK*	Dihydrofolate reductase	Trimethoprim resistance	Plasmid
	ermA	rRNA methylase	Macrolide resistance	Transposon
	ermC	rRNA methylase	Macrolide, lincosamide and streptogramin resistance	Plasmid
	ileS	Isoleucyl-tRNA synthetase	Mupirocin resistance	Plasmid
	mecA	Penicillin-binding protein 2	Methicillin resistance	SCC*mec*
	tetK	Tetracycline resistance protein	Tetracycline resistance	Plasmid
	tetM	Tetracycline resistance protein	Tetracycline resistance	Transposon

Abbreviations: rRNA, ribosomal RNA; SaPI, *Staphylococcus aureus* pathogenicity island. [a]Not a comprehensive list.

Methicillin-resistant S. aureus (MRSA) is a major public health concern[154]. In Europe, MRSA proportions were reported to be around16.4% in 2018 (EARSS Annual Report. 2018). From *S. aureus* causing PJI *(12-27%), around 50% are MRSA. They are considered one of the most challenging organisms causing PJI* and can compromise the success of surgical treatment and result in much higher costs. It has acquired resistance to all the antibiotics that have entered the clinic to date[155], and the World Health Organization defined it as a high-priority pathogen for research and development of new antibiotics[156].

Methicillin resistance is due to the acquisition of a new gene, mecA, that codes for a novel penicillin-binding protein (PBP), designated PBP2a that makes the strain resistant to all beta-lactam antibiotics, including penicillins, cephalosporins, and carbapenems. The mecA gene is carried on a mobile genetic element designated Staphylococcal cassette chromosome (SCC) mec that is chromosomally integrated. Acquisition of mecA initiated the successful spread of MRSA, one of the most important multidrug-resistant (MDR) nosocomial pathogens. Some MRSA clones are spread globally while others are restricted to specific geographical regions, that are usually dominated by a single subclone, with different subclones present in different regions[157]. MRSA infection trend to occur in waves of infection, characterized by the emergence of predominant strains[158].

The prevalence of MRSA varies according to the geographical area. In some areas of the world, MRSA prevalence is very high, for example, in Latin American countries prevalence is estimated to be >80 % and in India proportions of 41– 80% were observed in 2008–2012[159]. In the United States and Canada, the prevalence of MRSA is 15 % to 45 % while it is generally lower in northern Europe and higher in the south and south-east Europe. In the Nordic countries, the incidence of MRSA is very low. In fact, Norway has one of the lowest prevalence of MRSA, with 0.8 % prevalence reported in the NORM surveillance system in 2018[160]. The most common spa types in 2018 were t002, t304 (CC6), t008 (CC8), t127 (CC1), t223 (CC22), t019 (CC30), t437 (CC59), t034 (CC398), t044 (CC80) and t105 (CC5)[160]. In a study, including PJI samples collected from different Swiss and French hospitals, *S. aureus* isolates were typed by agr (accessory gene regulator) group and the spa type. Results reported that the most common CCs were CC1, CC5, CC8, CC30, CC45, and CC59[161].

On the other hand, virulence factors contributing to *S. aureus* pathogenicity in PJI are involved in processes such as the invasion of host tissues, evasion of the immune system, adhesion to

surfaces, and biofilm formation[138]. Cell surface components, enzymes, and exotoxins are important virulence factors for invasion, evasion, and propagation [162]. By persisting in biofilm, bacteria evade neutrophil killing and display decreased susceptibility to antibiotics, which creates significant challenges, e.g. in the treatment of PJIs[133].

Clinical specimens for microbiological diagnosis of PJI

Needle aspiration of synovial fluid or a biopsy sample of intraoperative tissue provides the most accurate specimens for detecting the etiology of PJI[103]. Identifying the causative organism is crucial and requires the collection of high-quality clinical specimens from deep sites in contact with the prosthetic material and collected under aseptic conditions. The results of the culture of intraoperative specimens (fluids and/or tissue) are usually considered the gold standard for detecting the infecting organism(s)[163].

Analysis of synovial fluid culture is considered one of the main analyses to perform PJI diagnosis in patients with underlying inflammation, especially because the result from the test is independent of the presence of underlying inflammatory disease allowing the identification and susceptibility testing of microorganisms [164]. Needle aspiration can be performed routinely before surgery or samples are collected only during surgery [113]. Synovial fluid can be inoculated onto solid and/or liquid media. Culture sensitivity range between 82–94% and specificity of 94–97%. However, the use of antimicrobials prior to aspiration increases the number of false-negative in synovial fluid cultures [165]. As a solution for that, more recent studies have inoculated synovial fluid into blood culture bottles (BCB) [166–168]. It has been found that when combining both methods (agar plate vs BCB) while sampling joint fluid adds sensitivity (0.61 vs. 0.5) while reducing specificity (0.89 vs. 0.95) [169].

The microbiological culture of PJT is considered one of the most important criteria for the diagnosis of PJI. PJT provides the most accurate specimens for detecting the infecting microorganism(s)[103]. Previous studies have used various criteria for the microbiological diagnosis of infection, ranging from the isolation of organisms from one direct culture, or two or more broth enrichment cultures requiring even five of five specimens to yield the same organism. Multiple intraoperative tissue specimens should be submitted to the microbiology laboratory for aerobic and anaerobic culture to optimize culture yield, increase sensitivity, and because cultures may be falsely positive due to contamination introduced e.g. in the operating room, during specimen transport, or in the microbiology laboratory [120,170]. The Infectious

Diseases Society of America (IDSA) and the American Society of Microbiology recommendations suggest the use of at least three, but optimally six tissue specimens as the number of samples that should be obtained. A specification is done according to the culture method used for analyzing the samples. If the conventional plate and broth cultures are used, four samples are suggested, while if aerobic and anaerobic blood culture bottles are used, three samples are recommended as optimal numbers of tissue specimens for diagnosis of PJI[68,120]. Although the need to culture multiple samples is considered good practice in order to increase sensitivity and to overcome the problem of contamination, a standardized method has been neither defined nor validated.

Currently, there is a discussion about which specimen is superior for the PJI diagnosis. Sonication has been found in multiple studies to improve the yield of culture-based diagnostics[171–173]. Some studies suggest that sonication is more sensitive than tissue sample culture[172]. However, results vary and there are also several studies that have found the opposite[170,174].

Objectives of the study

The main objective was to explore the use of shotgun-metagenomics (SMg) in clinical diagnostics of prosthetic joint infection (PJI) for the identification of potential pathogens and prediction of antimicrobial resistance (AMR) and virulence determinants directly from blood culture bottles (BCBs) inoculated with periprosthetic tissue (PJT) specimens.

Paper I

Hypothesis: Use of the BacT/Alert® Virtuo blood culture system for culturing PJT specimens improves the microbiological diagnosis of PJI.

Specific objectives:

- Evaluate and optimize the use of the BacT/Alert® Virtuo BCB system for culturing PJT
- Compare the BCB method and conventional method for the identification of potential pathogen(s).

Paper II

Hypothesis: SMg can be performed directly on PJT from BCBs for pathogen identification.

Specific objectives:

- Establish a protocol for the preparation of high-quality gDNA from BCBs inoculated with PJT specimens for downstream SMg application.
- Optimize an SMg bioinformatics analysis pipeline for analyzing the SMg results for identification of potential pathogen(s)
- Compare the SMg results with the results from the conventional method and the BCB method for pathogen identification.

Paper III

Hypothesis: Typing and prediction of AMR and virulence determinants in *S. aureus*, can be performed with high accuracy by SMg from BCBs containing PJT.

Specific Objectives:

- Assess the use of SMg for typing, prediction of antimicrobial resistance genes (ARG) and virulence profiling of *S. aureus* from BCBs inoculated with PJT.
- Compare the genotypes predicted from the SMg data and AMR phenotypical results.

- Analyze and evaluate the sequencing data using reads-based and contigs-based approaches for taxonomic classification and prediction of ARD.

Choice of Methods

A detailed description of the methods is presented in papers I, II, and III. The following sections aim to present a brief description of the methods used in this study and the rationale for choosing them.

1 Sample collection

Currently, there are not well characterized reference materials intended for SMg assays or guidance for developing specific procedures to assess SMg[1]. Therefore, appropriate alternatives materials and assessment procedures were designed in this study. Three different types of BCB samples were obtained to get a comprehensive overview: BCB inoculated with PJT specimens from patients with suspicion of PJI, BCB inoculated with bacterial species common causes of PJI and a negative control which was prepared by inoculating sterilized tissue into a BCB. The choice was based on SMg workflows from different published studies, and they were adapted to the SMg approach from BCBs workflow. An overall description of the samples used in this study, discriminated by paper is listed in Table 7.

Table 7. Description of samples used in this study discriminated by paper.

Paper	No. of clinical specimens	No. of patients	Procedure	BCBs inoculated with tissue clinical specimens	No. of spiked BCBs (n=) and strains spiked-in	NC (n=)
I	158	62	Bact/Alert® Virtuo BCB system evaluation	158 FA 158 FN	24 FA 12 FN *Escherichia coli ATCC 25922* *Staphylococcus aureus ATCC 25923* *Enterococcus faecalis ATCC 29212* *Staphylococcus epidermidis ATCC 12228* *Bacteroides fragilis ATCC 25285* *Cutibacterium acnes (Clinical isolate)*	2
II	25	25	Assessmet of the DNA extraction methods	25 FA	3 FA *E. coli* ATCC 25922 *S. aureus* ATCC 25923 *C. acnes* + *S. aureus* ATCC 25923	1
	9	9	SMg sequencing for pathogen identification in PJI diagnosis	9 FA		
III (MNS)	19 (13 new samples + 6 from paper II)	18	SMg sequencing for *S. aureus* antimicrobial resistance prediction	20 FA	1 FA *S. aureus* ATCC 25923	

Abbreviations: MNS, manuscript; FA, Bact/Alert® BCB FA Plus for aerobic bacteria; FN, Bact/Alert® BCB FN Plus for anaerobic bacteria; NC, negative control.

Samples to be analyzed are obtained through collection techniques that vary in their level of invasiveness and risk to the patient[175]. In this project, the type of clinical samples used were all types of PJT specimens submitted on a routine basis to the Department of Microbiology and Infection Control (AMS) at the University Hospital in North Norway (UNN), comprising any type of PJT from e.g. hip, knee, elbow, and shoulder. Only residual PJT samples, also analyzed by the conventional routine method, were used as testing material in this study. Consequently, the samples used may be of lower volume and quality that would be the case if the method were used routinely. Samples were collected continuously over 28 months (August 2017–December 2019), anonymized, de-identified and stored at 4°C until processed.

PJT specimens used for paper I were randomly collected, while samples from paper II were selected based on the presence of bacteria (monomicrobial) reported as common cause of PJI. For paper III, samples were selected on the bases of being positive for *S. aureus* by the laboratory methods (conventional and BCB). Samples contained just *S. aureus* (monomicrobial, n= 17) or *S. aureus* and other bacterial species (polymicrobial, n=2). In total 22 BCBs inoculated with PJT samples were SMg sequenced (Paper II, n=9 and Paper III, n=19). All the BCB samples sequenced were taken from aerobic bottles, to facilitate the comparison between samples. Bacterial strains for spiking experiments were selected based on the criteria of covering a range of different bacteria (Gram-negative, Gram-positive, aerobic, and anaerobic bacteria) known as common causes of PJI.

2 Thesis workflow

The general workflow of the methods used in this study is summarized in Figure 5. The results from the conventional routine method were used for comparison with the BCB and/or the SMg methods, to validate the findings obtained.

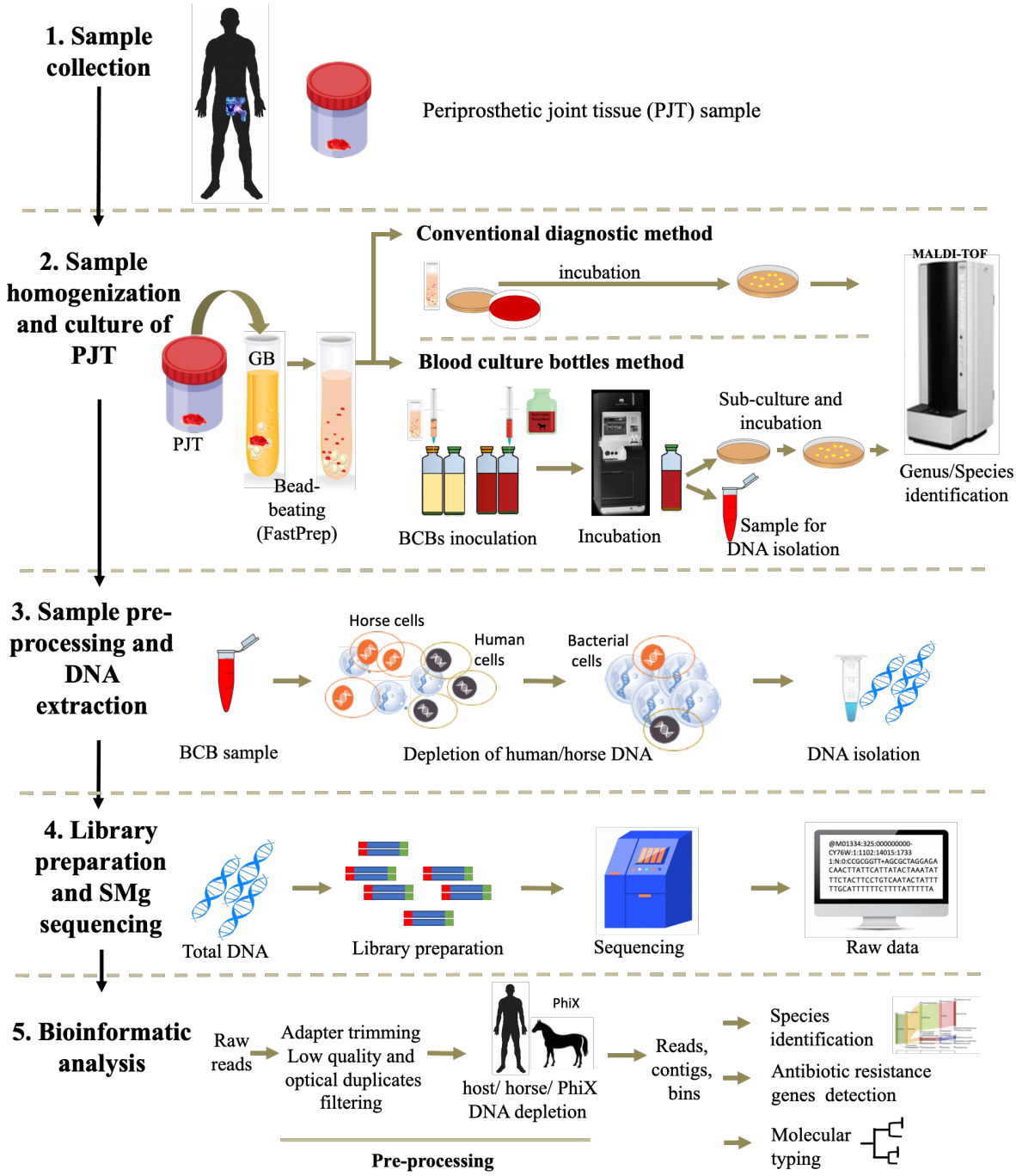


Figure 5. General workflow of methods used in this study.

2.1 Sample preparation

2.1.1 Tissue sample homogenization

Tissue samples need to be liquified to get access to the bacteria. Currently, mechanical disruption techniques are used on specimens to increase the sensitivity[113]. Bead mill processing using ceramic or glass beads and sonication of prosthetic material before inoculation onto solid

and/or liquid media are among the most used mechanical techniques in the clinical diagnosis of PJI[176]. In this project, sample homogenization included a bead-beating technique. It was done by adding an aliquot of the tissue samples (≈1 cm^3) to 15 ml tubes with glucose broth and ceramic beads. Then, the mixture was homogenized using a FastPrep-24 instrument (MP Biomedicals, France). This homogenized solution was used to inoculate the BCBs.

2.1.2 Sterilization of tissue material for spiking and negative control

The biological material used for mimicking infected tissue (spiked) and for the negative control were obtained from excess material of a native femoral head and surrounding tissue from an anonymous donor. The sample was crushed, sterilized, aliquoted, and frozen at -20˚C until use. Each time an aliquot was taken from the freezer, it was checked for contamination by plating on blood agar media.

Tissue sterilization was done using Gamma irradiation. Gamma radiation has several advantages and is the most suitable method for sterilization of biological tissues which has been used to assure adequate sterilization of tissue and bone allografts before transplantation procedures[177]. There are other sterilization techniques that may been use for tissue sterilization, among them: ethylene oxide gas, thermal treatment with moist heat, beta-propiolactone, chemical processing, and antibiotic soaks. Most current sterilization procedures have inherent disadvantages affecting biological properties of the tissue. Therefore, Gamma radiation offers a better alternative for sterilizing tissues[177]. Radiation sterilization at a dose of 25 kGy provides a high safety factor for sterility. Gamma radiation dose is measured in kilogray (kGy) units. One gray is the absorption of one joule of radiation energy by one kilogram of matter (one kGy = one joule/gram). 25 kGy was selected as the dose for sterilization as it is 40% above the minimum dose required to kill the microorganisms. Accordingly, 25 kGy is the minimum irradiation dose established for sterilization[177].

2.1.3 Horse blood as a BCB supplement

Defibrinated horse blood was used as enrichment supplement for the BCBs. This has previously been shown to produce high positivity rates and shortening of time to detection when testing simulated sterile body fluids other than blood models[178]. The use of horse blood does not significantly influence the performance of blood culture systems[178,179]. Horse blood is one of the most widely used animal blood products in culture media. BioMerieux has recommended the use of horse blood as a supplement in BacT/Alert FA blood culture bottles. In this study,

the BCB were inoculated with 4 mL of defibrinated horse blood (TCS Biosciences Ltd) as enrichment supplement.

2.1.4 Spiking tissue samples

Spiked samples were designed to evaluate the performance of the BacT/ ALERT® (BioMeriux) blood culture bottles system for culturing periprosthetic tissue samples (Paper I) and to create positive controls for the SMg experiments (Paper II and III), since positive controls mitigate the risk of false-positive results. A spiked sample is a negative matrix inoculated with a microorganism(s) that is known[74]. It is also known that positive controls are part of the quality control practices. It must be run with the clinical samples through the entire process to mitigate the risk of false-negative results and to enable the detection of performance failures in several steps in the workflow[74]. The design of this type of material is important to support the SMg assay validation, performance assessment, and quality control.

An important factor to consider when designing spiked samples is the bacterial inoculums concentration used to spike the negative matrix (in this project, the sterilized samples). The concentration of bacteria in spiked samples must be close to the inoculum of microorganisms needed to establish infection in real PJI samples. The size of the inoculum needed for bacteria to cause a PJI is a factor that depends in part on the type of microorganism that is causing the infection. Some microorganisms are highly pathogenic, and a relatively small number are capable of causing disease e.g. *S. aureus,* while others, e.g. *S. epidermidis*, require special settings or a relatively high bacterial inoculum, being more common in the delayed onset of PJI infections.

Commonly, in PJI, a low inoculum of microorganisms is needed to establish infection in the presence of the prosthetic material[96]. In fact, a low inoculum of infection is a factor that plays a role in making the diagnosis of PJI difficult[180]. Commonly, PJI are initiated through the introduction of microorganisms at the time of surgery and the infections occurs within 1 year of surgery. In a rabbit model, it was reported that lower inoculum concentration was needed to stablish infection ($<10^2$) if inoculated at the time of hip arthroplasty, compared with when no implant is present (10^4 CFU)[181]. In this project, to simulate PJI samples, bacterial inoculums were prepared and added to the glucose broth containing a piece of the sterilized tissue (≈1 cm^3) to get an inoculum concentration of ≈500 CFU on the glucose broth.

2.1.5 Negative control

Negative controls (NCs) are samples that do not contain any possible microbes and that follow the same entire process as the clinical samples. They are an essential part of the method of quality control for reliable and valid results of assays that are aimed for clinical applications[52]. NCs must be included in SMg assays to enable discriminating true pathogens from contaminants (external, reagent, and cross-sample contamination) and avoid false-positive results[74]. SMg studies have used as a variety of negative controls among them: elution buffer from the nucleic acid extraction kit[72], DNA extractions (blank) control[63], transport medium[93], water, and non-infectious samples[94]. In this study, a negative control sample was designed by inoculating a BCB with sterilized infection-negative tissue from a donor with no suspicion of infection.

2.2 Blood culture bottles (BCB)

Optimization of culture techniques for rapid isolation and detection of the infectious agent(s) from clinical specimens is desired. There are a number of changes in modern microbiological diagnoses and automated systems have become an integral part of many clinical microbiological laboratories. One such optimized culture technique is the inoculation of BCBs with clinical specimens. It is often used in the clinical microbiology laboratories for the detection of microorganisms, mainly from the blood where it is regarded as the "gold standard", but also for other body fluids and clinical specimens.

A blood culture is a laboratory test in which clinical specimens are inoculated into sets of bottles containing culture media to determine whether infection-causing microorganisms (bacteria or fungi) are present in the patient's sample. A set includes an aerobic and an anaerobic bottle.
The most common BCB systems used worldwide are BD BACTEC (Becton Dickinson Instrument Systems, Sparks, MD) and BacT/Alert® (bioMérieux, Marcy l'Etoile, France)[182–184]**.** BioMerieux launched a new version of the blood culture system to the market in 2014, the BacT/Alert Virtuo system which comes with automatic loading and unloading, and improved detection algorithm and temperature stability[179,185]. With colorimetric detection, BacT/ALERT® culture media use advanced technology including specialized Liquid Emulsion Sensors (LES) at the bottom of each culture bottle that visibly change color when the pH changes due to the rise in CO_2 as it is produced by microorganisms. BacT/ALERT VirtuO instruments measure the color changes every ten minutes and analyze the

changes. Sophisticated algorithms, including a unique "threshold" algorithm, ensure early detection of microorganisms from various sample types. Furthermore, the BacT/ALERT® FA Plus and FN Plus BCB series contains absorbent antibiotic-binding polymeric beads that neutralize the antibiotic effect optimizing the recovery of microorganisms for samples from patients on prior antimicrobial therapy, in addition to adsorbing inhibitory compounds[183]**.** BacT/ALERT® FA Plus medium composition contains peptones, yeast extracts, supplements, and absorbent resins. BacT/ALERT® FN Plus (anaerobic) media differ in their composition as well as their relative contents of two polymeric resins designed for neutralizing antibiotics and additional inhibitory substances. The anaerobic formulation includes a complex amino acid component not contained in the aerobic formulation[186].

The traditional approach for identifying pathogen(s) from positive blood culture bottles starts with Gram staining to confirm the presence of microbes and to determine the morphotype. The Gram staining allows disclosure of the presence of polymicrobial infections. Then the pathogen identification can be achieved by **(i)** subsequent steps of subculture on agar plates which requires 18-48 hours of additional time for subculture until colonies are present on the agar plates. Among them, Matrix-assisted laser desorption/ionization-time of flight mass spectrometer (MALDI-TOF) which is a spectrophotometry system that performs protein analysis of microorganisms for the identification or the Vitek 2 XL (BioMerieux)[187]. **(ii)** using molecular methods directly from positive BCBs that significantly reduce the time to results. These methods may be based on MALDI-TOF MS (on positive blood-culture pellet), PCR, and fluorescent nucleic acid probes[188].

2.3 Evaluation of the BacT/Alert Virtuo b system for culturing periprosthetic tissue

In order to ensure laboratory quality practice in the clinical microbiology laboratory, test validation and verification are needed. The terms verification and validation are often used interchangeably. However, each refers to a separate step in assuring that a test is performing as expected. ***Verification*** is the one-time process completed before a new test (or a modification of an existing test) is used for patient testing. In contrast, ***validation*** is the ongoing documentation that a test, which has already been verified, is repeatedly giving the expected results as the test is performed over a period of time[189]. Validation, verification, and evaluation of the performance of microbiological diagnostic tests are challenging. It is mainly because living microorganisms represent an extra source of uncontrollable variation and because the

test methods are the sum of sequential and conditional sub-processes that can affect the results[189–191].

One of the most important aims when trying a new application of an existing test is determining whether the sample to be examined has any inherent interfering properties and whether the incubation and growth conditions can recover microorganism to an acceptable level[192]. In this study, we present an evaluation of the use of the BacT/Alert® Virtuo blood culture system for microbiological analyses of periprosthetic tissue specimens as a tool that can be beneficially used for improving and accelerating the diagnosis of PJI (Paper I). Our analyses included spiked and clinical samples, as well as a comparison of the BCB method with the local routine diagnostic method, for time to detection (TTD) and bacterium (monomicrobial) or bacteria (polymicrobial) identified. A workflow summarizing the BCB evaluation process is presented in Figure 6. The translation of this test system into the clinical application at the local level was the main goal (paper I).

The use of BCBs is currently the standard diagnostic method for the diagnosis of blood infections[193]. They are approved by the FDA and they have been validated by the manufacturers for the detection of microorganisms from blood and other normally sterile body fluids (e.g. cerebrospinal fluid, ascites)[193]. When users practice a sub-process that differs from guidelines and for which the level of evidence in the literature does not ensure enough confidence for safe patient management, this particular sub-process needs an additional validation[191].

Previous studies have demonstrated the potential of BCB systems for microbiological detection from PJI specimens, including PJT[174,193–195]. The method has already been implemented in some clinical laboratories replacing the conventional diagnostic method. Multiple studies have validated BCB systems for their use with tissue preparations[193]. However, most studies are mainly using the BD BACTEC BCB system.

There are procedures suggested for evaluating the ability of a BCB system to support the growth and detection of typical clinical pathogens, allowing the validation to be performed with seeded suspensions versus patient specimens, in addition to a parallel comparison with a reference method[189]. Seeded suspensions refer to the inoculation of BCB using spiked materials that mimic patient specimens and that contain a known number of bacteria, preferably using

reference strains to verify that the analytic process is under control by detecting proper bacterial growth, positive bottle detection and time to detection. The use of this approach has been suggested for certain local accreditation committees and suppliers to meet the ISO 15189 standards for medical laboratories with respect to quality controls[191]. The parallel comparison approach, on the other hand, is important since the BCB method intended to replace another method, invariably requiring a level of comparison[189]. However, BCBs represent a particular process with very few equivalents in laboratory tests in many aspects, as for instance the volume of sample inoculated into the BCB, and this sample volume critically determines disease detection. In this study, the BCB method was based on, and modified from, similar methods used in the previous studies[194,196,197], while the conventional method was an already validated in-house method.

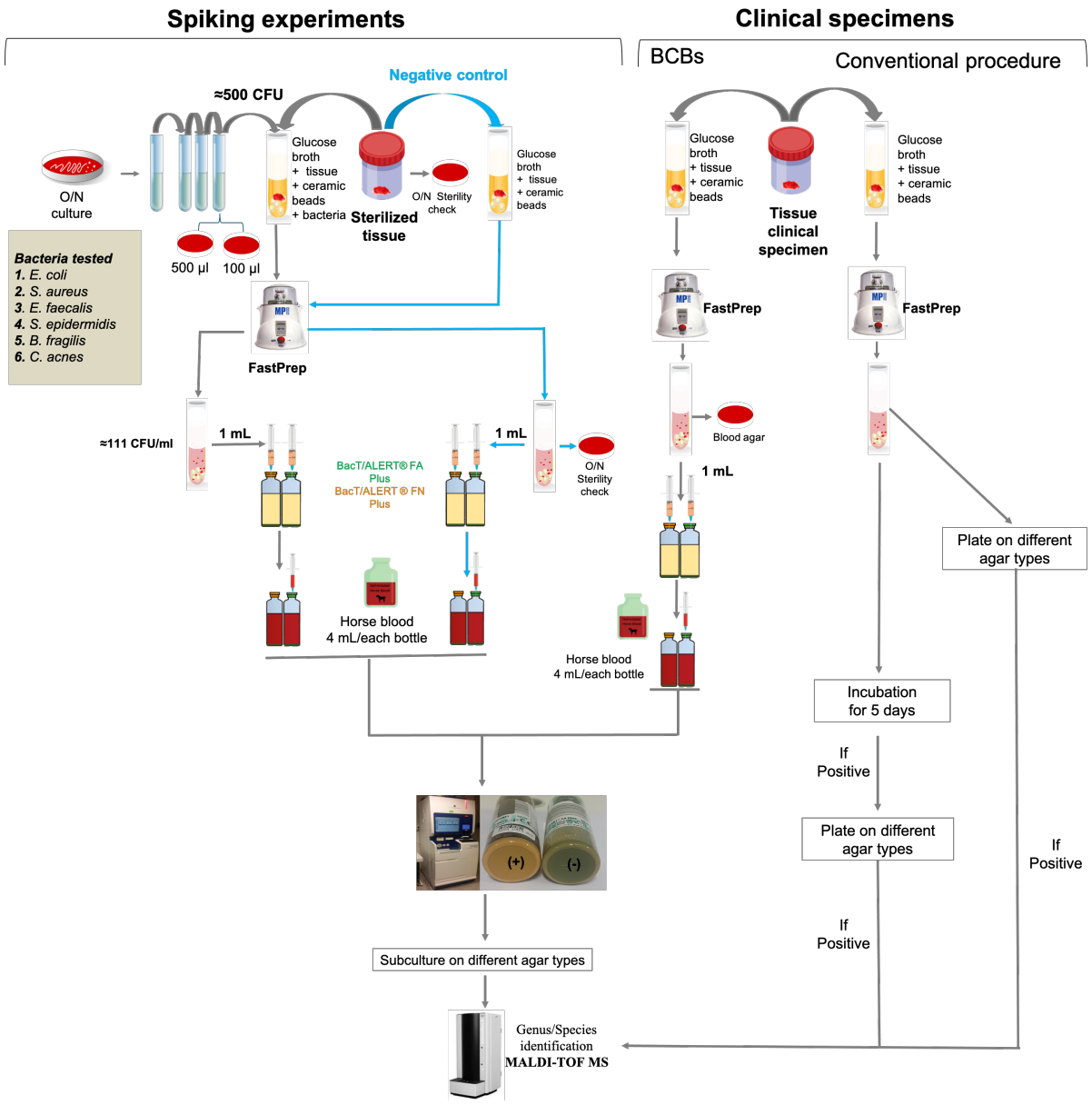


Figure 6. Overview of the process followed for evaluating the use of the BacT/Alert® Virtuo blood culture system for culturing periprosthetic tissue specimens. Abbreviations: O/N, overnight culture; CFU, Colony forming unit; MALDI-TOF MS, Matrix-assisted laser desorption/ionization-time of flight mass spectrometer; BCB, blood culture bottle.

2.4 SMg from positive BCBs

The workflow typically includes DNA extraction, library preparation, sequencing libraries, assembly of contigs, binning, and in silico analysis of the metadata.

2.4.1 DNA Sample preparation for SMg from BCB

The success of SMg is highly dependent on the quality and quantity of DNA extracted from a given specimen[59]. Therefore, for using SMg on clinical specimens in the microbiology laboratory, protocols must be tested, adapted, and validated according to the type of specimen.

Each specimen present unique and specific challenges reflecting their matrix and concentrations of the target pathogen and resident microflora. E.g. bacterial load will be not the same in specimens from normally sterile sites such as joint spaces or periprosthetic tissue, compared to specimens taken from sites with colonizing flora, such as the throat or surgical wounds [113,198].

DNA amplification inhibitors commonly found in BCBs include blood-derived heme compounds, heparin, EDTA, polyacetal sulfonate contained in the culture media, in addition to human DNA. Since the extraction method is from BCB media inoculated with PJT, a feasible sample preparation method must involve not only the removal of potent inhibitors but also efficient cell wall disruption and subsequent recovery of microbial DNA, of a diverse range of microorganisms that potentially may cause PJI (Table 8)[199].

Currently, it is possible to find products that deplete host DNA and simultaneously extracts enriched microbial DNA in one kit e.g. the Ultra-Deep Microbiome Prep or the Complete5 DNA extraction kit (Molzym GmbH, Bremen, Germany). There are also kits for the depletion of host DNA and microbial DNA enrichment that can be combined with any microbial DNA isolation kit. E.g. MolYsis Basic5 kit (Molzym, Bremen, Germany). Host DNA depletion and microbial DNA enrichment using Molzym kits are performed before DNA isolation. However, e.g. NEBNext Microbiome DNA Enrichment Kit (New England BioLabs, Ipswich, MA) do it after DNA isolation. Other methods have also been reported but are not available as commercial products, such as host cell lysis with detergents[200] or ox bile[201], among others[65].

Separation of bacterial DNA from host DNA in clinical samples may have an important impact on downstream applications, involving microbial diagnostic systems. The host genome interferes with the detection and diagnosis of pathogens resulting in decreased sensitivity[19]. The majority of sequencing reads produced during SMg are identified as human, which is on average 1000 times larger than the average bacterial genome. While bioinformatics tools can remove human reads, greater sequencing depth (which leads to an increase in cost) is required to obtain enough pathogen reads for the identification of a causative agent and obtain information regarding resistance or strain type.

Methodologies focusing on depleting host cells from the original specimen by exploiting differences in cell surface structure between human cells and bacteria for selective lysis of host

cells have been developed. One of these methodologies is the MolYsis basic kit from Molzym. The MolYsis kit works based on the principle that human cells are first selectively lysed by the use of chaotropic reagents. Then the released host DNA is degraded by a MolDNase (which is active in the presence of chaotropes in the lysis buffer) prior to the extraction of DNA from microorganisms. Bacteria present in the specimen are then sedimented, washed, and subjected to DNA extraction (Figure 7).

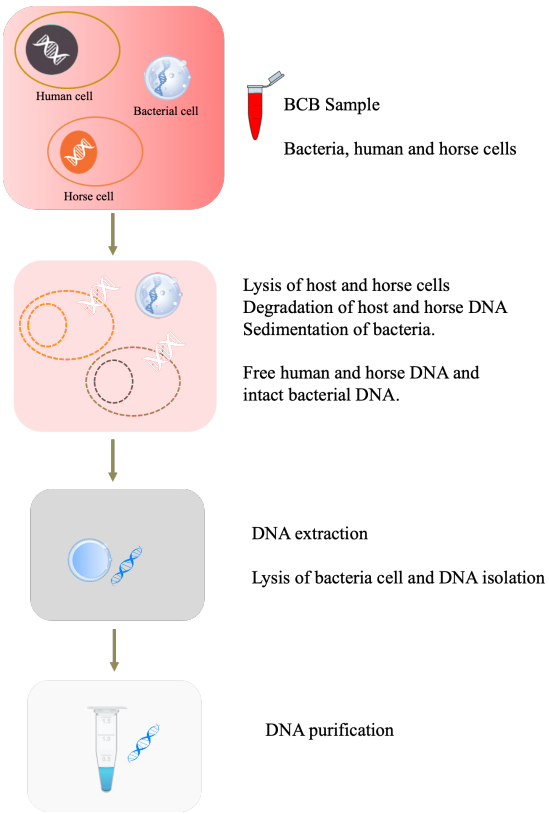


Figure 7. MolYsis kit Protocol for the depletion of host and horse DNA followed by DNA extraction. Modified from the source:(https://www.molzym.com/technology).

The MolYsis5 kit has been successfully used for direct sequencing from different types of specimens such as subgingival paque[202], sonicate fluid[65], and urine[5], among others. However, it is important to be aware of potential limitations that should be taken into consideration when using MolYsis in clinical diagnostics. For this study, total gDNA was extracted from the BCBs using the QIAamp BiOstic Bacteremia DNA Kit (Qiagen, Hilden, Germany). Samples were pre-treated using the MolYsisTM Basic5 kit (Molzym, Bremen, Germany) to deplete human DNA from the samples before DNA extraction. In order to find the most suitable procedure for extracting DNA from BCBs, DNA was extracted with or without pretreatment with the MolYsis5 kit (Figure 8).

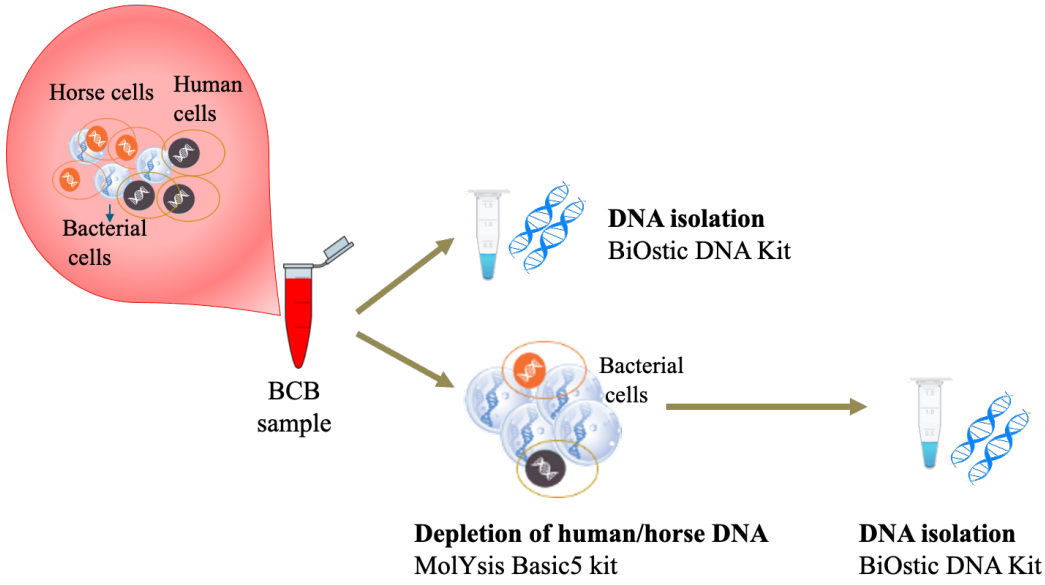


Figure 8. BCB DNA sample preparation methods tested in this project

2.4.2 Library preparation

The main objective when preparing a sequencing library is to create as little bias as possible[25]. Library preparation in SMg is possibly the most unbiased approach because the library covers all the genomes in the sample according to their frequency. However, the sequencing depth is minimized[1]. There are a variety of library preparation kits and sequencing platforms available that can be used for SMg. The selection depends on the cost, the number of samples, the availability of materials, and services[203]. The main concern about library preparation for SMg is related to amplification. The sequencing process requires a representative source of the genome(s) under investigation[25] and certain types of samples yield small amounts of DNA, sometimes requiring amplification during library preparation which may generate representation biases in the original DNA content[1].

In our study, DNA was prepared as a standard Illumina library using the ThruPLEX DNA-seq DNA-sequencing Kit (Rubicon). This library preparation kit has high multiplexing capability with improved performance for Illumina Next Generation Sequencing (NGS) Platforms. ThruPLEX DNA-seq DNA-sequencing Kit from Rubicon Genomics is compatible with the Illumina platforms for whole genome sequencing or for target enrichment workflows[204]. ThruPLEX™ kit enables library preparation from 0.05 to 50 ng of DNA distinct to other NGS library preparation kits, which are based on ligation of Y-adapters, ThruPLEX uses stem-loop adapters to construct high quality libraries in a fast (two hours) and efficient workflow. In fact, ThruPLEX is considered the fastest and most sensitive library prep kit available[205], with lower

duplication rates when compared with the Nextera, with less bias in the coverage (i.e fewer dropout regions and fewer regions with unnecessary deep coverage), easy for controlling the size of the library inserts, which has benefits for sample-sample yield and also increase the chances of success with the run. All of these features make ThruPLEX especially useful for clinical samples.

2.4.3 SMg Sequencing

One of the most commonly used methods in SMg studies is Illumina sequencing-by-synthesis (San Diego, CA), due to its high sequence throughput, low error rate, and low sequencing cost per base. It is important to note that this technology has the disadvantage of barcode index switching, in which high-frequency indices, that are designed to uniquely identify multiplexed samples may be misassigned during scanning of the flow cell. For SMg, this can lead to microbial reads from one sample containing a high titer pathogen cross-contaminating other samples on the same run, thus generating false-positive detections. This problem is intensified in the higher throughput HiSeq and NovaSeq sequencers due to the new chemistry techniques[1].

In this project, sequencing was performed using a MiSeq sequencer (Illumina) with v2 chemistry and 500 cycles for 250 bp paired-end sequencing. The criteria selection for using this platform was based on the number of samples sequenced (small sample size), previous reports in clinical metagenomic studies, and the availability of the sequencing center.

2.4.4 Bioinformatics Pipeline

The bioinformatic pipeline is a fundamental component of the SMg workflow. Particularly for clinical diagnosis, SMg bioinformatic pipelines use a number of different algorithms and tools developed for research on microbiome or environmental metagenome. Therefore, it is important to make custom optimization and modifications for the tools used but also to the reference databases used during the analyses for the purpose of clinical validation[206]. An SMg bioinformatic pipeline consists of a series of steps where the raw data (FASTQ files) generated from the sequencer can be processed through several different computational analysis tools. The typical NGS bioinformatic pipeline consists of three steps: reads preprocessing, data analyses, and interpretation. The workflow and the bioinformatic tools used during this study are summarized in Figure 9. In paper II, the analysis ends with the taxonomical assignment of the reads, while in paper III, the whole pipeline was followed.

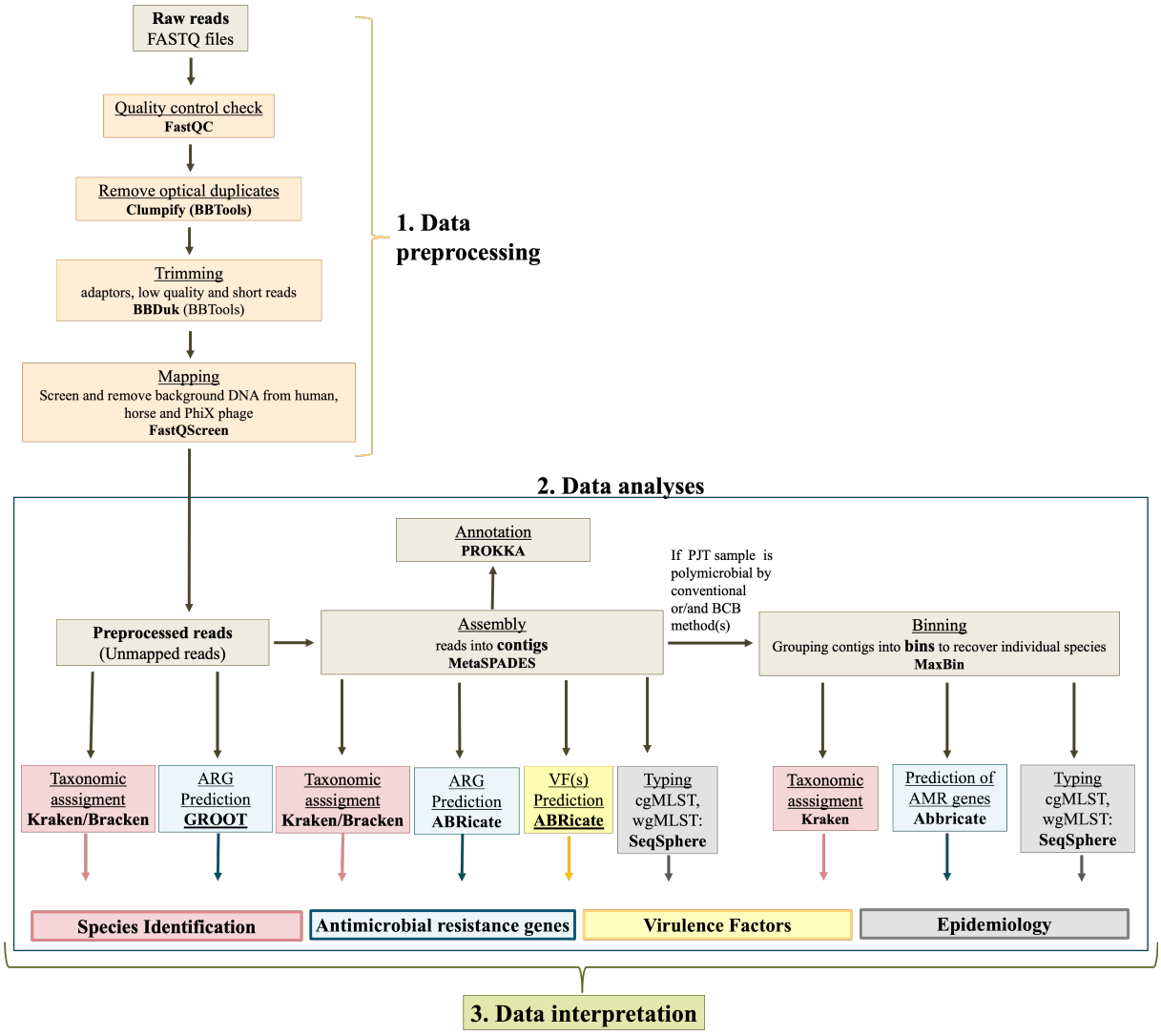


Figure 8. Bioinformatic pipeline followed during this project. Abbreviations: AMR, antimicrobial resistance; cgMLST, core genome multi-locus sequence typing; wgMLST, whole genome multi-locus sequence typing.

2.4.5 Data preprocessing and quality assessment

Sequencing data is often provided as raw reads which are preprocessed prior to data analysis. Reads preprocessing consists of the trimming of adaptors, low quality reads, and complex sequences filtering, followed by the data subtraction of reads mapped to background genomes, e.g. human genome, the PhiX phage and the horse genomes in the case of this project.

Errors such as base calling errors, poor quality reads, etc. occur during sequencing[29]. Therefore, it is important to assess the quality of the reads before starting to analyze the data generated. In this step, the quality of the raw reads is evaluated, to remove the reads that are not meeting the standards. Commonly, raw data are provided in FastQ files. They are an ASCII-based format

containing all the raw sequencing reads, the filenames and their ASCII-coded quality score for every single base[207]. Quality scores are based on the Phred score, a logarithmic error probability. In order to assess the quality of the raw sequencing reads, several bioinformatic tools have been developed, such as FastQC, NGS QC toolkit, and QC-Chain. One of the most used preprocessing procedures is read trimming. This procedure removes low quality portions of the reads while preserving the longest high-quality part of an NGS read. Then, adapter sequences used for the library preparation should be removed from the reads because they interfere with downstream analyses. Trimming is performed at the ends of each read to remove adapter sequences, which contain the sequencing primer binding sites, the index sequences, and the sites that allow library fragments to attach to the flow cell. The trimming step, although reducing the overall number and the length of reads, raises quality to acceptable levels. Several tools have been developed to perform trimming of Illumina data, such as Trimmonatic, Cutadapt, BBtools, and AdapterRemoval, among others[208]. The choice of the tool is highly dependent on the dataset, downstream analysis, and parameters used.

In the diagnostics of infectious diseases, it is important to detect and resolve DNA contamination from known sources. Common sources of contamination with foreign DNA include e.g. carry-over from samples previously loaded onto a sequencing machine or libraries from genomes of other species added intentionally for quality control or the host DNA[209]. For instance, Illumina offers PhiX libraries from the PhiX bacteriophage genome to provide quality control. If PhiX is used during library preparation, these libraries should be removed to avoid integration into the target genome(s)[210]. The raw reads produced need to be aligned to the reference genomes of the contamination sources in order to be detected and then filtered. The aim of aligning a read to the reference genome (mapping) is to find the chromosomal position the read is most likely coming from. If only perfect matching positions are sought for a reference sequence, no variation can be found. Several tools are currently available to map reads against pre-specified genomes, among them: BWA[211], Deconseq[212], Bowtie[213], BBMap[214], and fastqScreen[209].

2.4.6 Data analyses

Data analyses are considered one of the biggest challenges concerning the introduction of NGS in the clinical microbiology laboratory[215]. Depending on the analytic approach, the remaining reads can go directly to taxonomic classification and assignment of appropriate phylogenetic groups by comparing them to a genomic database, or they can be assembled into longer

sequences called contigs. Contigs can be grouped by species into discrete units, referred to as bins for taxonomic identification to the lowest taxonomical level and downstream analyses for pathogen identification[19,20].

Though it is possible to analyze sequence data without assembly, most analyses can be improved by constructing longer contiguous sequences (contigs)[216] through an assembly process. Metagenomic assembly is a computational process with the aim of reconstructing genes and genomes from metagenomic samples by combining the short reads to build larger stretches of DNA[217]. An assembly algorithm is then implemented, to compile reads into contigs. Several assembly tools have been published for short-read metagenomics, as Illumina reads. Among the assemblers using paired-end reads are MEGAHIT[218], SPAdes[219], and Omega[220], among others[221].

Metagenomics classifiers tools match sequences, typically reads or assembled contigs against a database of microbial genomes to identify the taxon of each sequence[222]. Some programs return an assignment of every read, while others only provide the overall composition of the sample. Methods for the analysis of metagenomic datasets can be classified into three classes: First, kmer-based read classifiers. This class includes tools like Kraken. Second, alignment-based methods, for complete genomes or marker genes only or translating the DNA and aligning to protein sequences. This category includes tools like MetaPhlan[223] and MG-RAST. Third, Bayesian or EM-based estimators. This class includes Bracken[224].

Kraken[225] is a fast taxonomic classifier using in-memory k-mer search of metagenomics reads against a database built from multiple genomes by using an algorithm that relies on exact k-mer matches, replacing alignment with a simple table lookup. Kraken constructs a database that stores, with every k-mer in every genome, the species identifier (taxonomy ID) for that k-mer. When a k-mer is found in two or more taxa, Kraken stores the lowest-common ancestor (LCA) of those taxa with that k-mer. This works on reads or contigs of a metagenomics data set. Kraken maps reads to the taxonomic tree, not to a specific level such as species or genus. **Bracken**[226] is an extension of Kraken that estimates species- or genus-level abundance based on a Bayesian probability algorithm[222].

Resistance genes and virulence factors databases play an essential role in studying the resistome and the virulome from microorganism(s) in clinical samples. A critical first step in many such

approaches is the construction and curation of databases. There are several AMR databases that have been developed, including CARD[80,81], ARDB[82], ARG-annot, and ResFinder[79]. However, they were not developed for clinical purposes and therefore they have a certain degree of restriction in clinical diagnosis. On the other hand, the virulence factor database (VFDB)[85,86] provides knowledge of VFs from various bacterial pathogens and serves as a comprehensive warehouse of bacterial pathogenesis knowledge for the scientific community[86].

2.4.7 Molecular typing

The genetic relationship between isolates can be investigated. There are two approaches using a multi-locus sequencing typing (MLST) approach that can be applied to SMg data: Core genome multi-locus sequence typing (cgMLST) and the whole genome multi-locus sequence typing (wgMLST)[11].

By cgMLST, the conserved core genome is analysed via Sanger sequencing, unlike the conventional MLST, which only looks at seven gene regions. cgMLST schemes interrogate thousands of gene regions or alleles. The allelic pattern (i.e. the identity of the sequence at each allele) is compared across isolates and an allelic distance number is calculated (i.e. the number of alleles that are different between 2 isolates). This number is used to define the relatedness of the isolates[227]. On the other hand, wgMLST, which includes a set of variable accessory genes. Several software packages, such as SeqSphere (Ridom) and BioNumerics (Applied Maths, Biomérieux), or online tools, such as EnteroBase and BIGSdb (Bacterial Isolate Genome Sequence Database, can be used for this approach[228]. Each approach has a source of bias and associated limitations[227].

Summary of Results

PAPER I: Culturing periprosthetic tissue in BacT/Alert® Virtuo blood culture system leads to improved and faster detection of prosthetic joint infections

- We showed that the BacT/Alert® Virtuo blood culture system effectively detected relevant bacteria from PJT specimens in both spiked and clinical samples, and the method was reproducible
- The BCB method was found to be as sensitive (79 %) as the conventional method (76 %) ($p = 0.844$) during analyses of clinical samples.
- The BCB method yielded positive results much faster than the conventional method: 89% versus 27% detection within 24 h, respectively.
- More than 80% of the bacteria were detected in less than 20 h using BCBs, compared with the five days needed to obtain a similar percentage by the conventional method.
- The median detection time was 11.1h for the BCB method (12 h and 11 h for aerobic and anaerobic BCBs, respectively.

PAPER II: Shotgun-Metagenomics on Positive Blood Culture Bottles Inoculated with Prosthetic Joint Tissue: A Proof of Concept Study

- Our DNA preparation method resulted in high quality microbial DNA from all samples, both with and without human DNA depletion, allowing downstream SMg.
- By SMg we were able to identify relevant PJI pathogens, and all bacteria identified by the culture methods were also identified through SMg.
- The use of MolYsis resulted in a higher proportion of bacterial reads compared with untreated samples (96 % versus 87 %).
- Only 0.07 % of the reads were classified as humans in the clinical sample, and in the negative control 1.4 % human reads.
- We obtained a high enough sequencing depth by using this SMg method indicating that it is possible to multiplex samples reducing costs considerably.
- We show a high sequencing quality, low human DNA content (<1 %), a high number of bacterial reads, and complete genome coverage of sufficient depth (PC1, *S. aureus*: 775x and PC2, *E. coli*: 209x).

PAPER III: Shotgun-metagenomics for typing and prediction of antibiotic resistance and virulence determinants in *S. aureus* from prosthetic joint tissue on blood culture bottles.

- We obtained a high sequence quality, low human DNA content, a high number of bacterial reads, and complete genome coverage of sufficient depth for identification and typing of *S. aureus* and prediction of virulence and AMR profiles.
- Reads-based data analysis was found to be a better approach for pathogen identification, while for the AMR profiling the contigs-based approach gave a more complete picture.
- AMR and virulence profiles for all the samples were generated. We detected 7 and 73 different resistance and virulence genes in *S. aureus*, respectively, in the 20 samples analyzed (including the spiked sample).
- The disagreement rate between the phenotype and the genotype was 25 % for *S. aureus* phenotypically resistance and genotypically susceptible to penicillin (false negatives). In the rest, phenotypes could be explained by the presence of AMR genes, e.g. *fusB or fusE,* for fusidic acid resistance, and *bla-* operon members for penicillin resistance. However, predicted resistance by genotype but susceptible phenotype (false positives) were found in some cases, e.g. presence of the *tet(38)* gene in all *S. aureus* isolates, but any expressed tetracycline resistance.
- We identified genes encoding virulence factors known or proposed to play a role in PJI, among them *aur, clfA, can, fnbA, hld, hlgA-C, sdrD-E, spa, sspA-C, hly/hla*.
- Typing from SMg data showed that *S. aureus* grouped into eight phylogenetically diverse MLST CCs.
- Our work demonstrated that it is highly possible to perform identification, typing, and AMR and virulence gene prediction by SMg on PJT from BCBs.

General discussion

SMg has emerged as a promising approach for diagnosis from clinical samples[50]. A single metagenomics analysis has the potential to detect all, common, rare and novel pathogens[9], and provides information on all microorganisms present in a sample, even if the sample contains a complex (polymicrobial) community[54,229]. SMg applied to clinical specimens has the potential for pathogen identification, outbreak investigation, and prediction of antimicrobial resistance (AMR) in clinical microbiology laboratories. In this thesis, we wanted to explore the use of this emerging approach for the microbiological diagnosis of PJI, since it is an interesting infectious model, recognized for being challenging, especially in the detection of the microorganism(s) causing the infection.

Diagnosis of PJI is complicated by the wide array of pathogens implicated, potentially undetected polymicrobial infection, and the fact that most of the methods conventionally used to make microbiological diagnosis have low sensitivity and specificity, leading to 10-45% false-negative results[96,196,230,231]. Culture of peri-operative samples can be falsely negative because there are various factors that affect the microbiological results, such as previous antimicrobial therapy, a low inoculum of microorganism, the presence of microbial biofilm(s), low number of tissue specimens, inappropriate culture medium, inadequate culture incubation time, or prolonged time to transport the specimen to the laboratory. Despite the known low sensitivity, the culture of PJT specimens on agars and in broths is routine. However, a rapid, sensitive, and specific method is needed.

In recent years, a variety of measures and tools have emerged to aid the diagnosis of PJI including the application of SMg, which despite its huge potential is still in its infancy. There are significant concerns regarding performance, validity, and clinical significance[75]. In fact, it is currently best positioned as a complementary technique to be used alongside culture and other traditional methods and may be of special interest, when both conventional and enhanced molecular testing fails[9]. These have encouraged us for further research on possible solutions to overcome some of the challenges in the use of SMg for PJI diagnosis.

BCB, the base

The use of BCB is one of the techniques proposed to improve the sensitivity of PJI sample cultures, including PJT[194,196,197,232]. There are several studies published about the inoculation of BCB with PJI specimens for microbiological diagnosis. Automated BCB systems have been used for culturing synovial fluid[167–169,233–238], sonication fluids from explanted prostheses[166,170–172,174,239–241] and periprosthetic tissue specimens[104,174,194–197,241–244]. The focus has been mainly on the comparison between the use of the BCB method and the conventional culture. In addition, there are studies about the use of different BCB systems (BD BACTEC or BacT/Alert®), for a specific PJI specimen type and using the same BCB system with different PJI specimens. However, this approach has not been widely adopted and has been evaluated in a limited way[196]. To the best of our knowledge, most studies on PJI to date have used the BD BACTEC BCB system. Therefore, an evaluation of the BacT/Alert® Virtuo blood culture system for PJT was missed, despite being one of the most common BCB systems used worldwide. In the **paper I**, we evaluated the performance of this BCB system for this type of specimen (PJT).

We found that the use of the BacT/Alert® Virtuo blood culture system was found to be as sensitive as the local conventional method (79 % vs 76 %, $p = 0.844$, respectively). Although other studies have reported higher sensitivities[196,197,244], a meta-analysis evaluating the diagnostic accuracy of PJT culture in BCB reported that the pooled sensitivity of tissue culture in BCB was 70 % and the specificity 97 %, suggesting that BCBs may be of great value in PJI diagnosis[239].

Another important aspect of PJI diagnosis is the rapid detection of pathogens, the earlier the diagnosis, the better the outcome of subsequent treatment[243]. We found that the use of the BacT/Alert® Virtuo blood culture system yielded positive results much faster than the conventional method (89% against 27% detection within 24 h)[245]. This has been supported also by the previous studies[194,196,245], e.g. Minassian *et al. s*howed that PJT cultured in BCB resulted in the growth of most microorganisms within 3 to 5 days, in aerobic and anaerobic cultures, respectively, with detection rates of 95 % and 96 %, respectively. In addition, Peel *et al*. reported that aerobic and anaerobic BCB were positive within the first day of incubation[194,196].

Besides the culture of PJT samples, other PJI specimens have been inoculated into BCBs. However, when comparing the different PJI specimens, the culture of PJT into BCB improves

the diagnostic yield of microbiological culture and may approach or be superior to sonication culture[170]. We recommend the implementation of the BCB method for PJT samples in the laboratory if resources allow it. BCBs provide a semiautomated method for culturing PJI specimens that, incorporated into the routine lab, may provide many different significant advantages, among them:

(i) BCBs may enhance the culture yield, allowing a sensitivity rate as compared with the conventional plate/broth culture system. This by neutralizing the effect of antibiotics (by diluting and binding resin)[246], improving culture conditions (agitation and rich media composition)[195,247], and because the sample volume inoculated into the bottle is larger than the one inoculated onto agar plates[196,241].

(ii) BCBs allow the identification of a broader spectrum of pathogens including slow-growing bacteria[248].

(iii) BCBs also provide partial automation of the workflow, mechanical loading and unloading of bottles, faster detection of microorganisms present in the samples, minimizing the risk of contamination, and reducing labor requirements in the clinical laboratory which minimizes human errors[104,194,196,245].

On the other hand, there are certain disadvantages using BCBs, for instance, the higher rate of false-positive results (in our study this was low 1.2 %), cost of BCBs, cost of waste disposal, and capacity problems in the laboratories[170,232]. Therefore, before implementation, laboratories adopting the protocol need to perform the appropriate validation, since BCB are not approved by the FDA for the culture of tissue specimens[170]. In fact, the translation of this test system into the clinical application at the local level was the main goal of this part of the project (**Paper I**).

The combination of BCB and SMg

First-line culture-based methods are powerful tools that usually are successful in establishing the diagnosis. However, they are looking for etiological agents that are associated with a specific disease and the results obtained may be incomplete when novel pathogens, uncultivable, strictly anaerobic, or very fastidious species are involved. When the first- line investigation tools fail to identify the causative pathogenic agent, it is important to have alternatives[113]. Currently, a large number of tests are available. Microbiological methods differ across institutions and have varying criteria and protocols. The most effective diagnostic strategy is likely formed by a combination of conventional and new diagnostic strategies for

PJI. Research into improving all aspects of culture-based methods and new alternative strategies to replace or support current microbiological methods is needed [249]. Further research is needed to integrate many of the new techniques in the clinical laboratory routine.

SMg may make a further contribution to the PJI diagnosis. To the best of our knowledge, there are no studies about SMg on BCBs inoculated with any specimens related to PJI. Although it is true that there are no real problems in identifying bacteria that grow in blood culture media, our study was conducted initially as a laboratory method development (**Paper II**), where the aim was to test the use of SMg directly on positive BCBs inoculated with PJT to try to solve the limitations observed when using SMg directly on PJI specimens (synovial fluid[48,71], sonication fluid[66,68,69,250], and tissue[72]). Those studies found that DNA extracted from PJI samples was contaminated with a high concentration of human DNA, while the bacterial DNA yield was very low. Special attention has to be paid to the sample preparation, background contamination, bioinformatic pipeline for the data analysis, and the data interpretation.

Host DNA and bacterial DNA yield

The presence of an overwhelming amount of host DNA is one of the most important problems to be addressed in SMg[200]. Host DNA removal should depend on appropriate host DNA removal methods in samples rather than bioinformatic analysis[250]. Reduction of host contamination may facilitate greater pathogen coverage and depth, especially when the bacterial cells in the PJI samples is low since the host DNA reduce the sequencing yield. Several methods have been developed for host DNA depletion and enrichment of bacterial DNA. They include two main different approaches: depletion of host DNA prior to DNA extraction, including centrifugation and the MolYsis kit[65,198,200,251,252], or depletion of human DNA post-extraction, e.g. NEBNext® Microbiome DNA Enrichment[65]. For this purpose, in **Paper II**, we used a method which includes a sample pretreatment step using MolYsis. The rationale for using MolYsis was based on a higher proportion of bacterial reads obtained, compared with untreated samples. Previous studies have demonstrated the efficacy of MolYsis for selective lysis of human or eukaryotic cells from PJI samples. In synovial fluid spiked with *S. aureus,* the use of MolYsis Kit produced a higher fold of microbial reads compared to the NEBNext® Microbiome DNA Enrichment kit[65]. However, it is important to be aware of potential limitations when using MolYsis in clinical diagnostics. MolYsis provides a solution for the removal of host DNA and enrichment of intact microorganisms[253], but this raises a question about which bacterial DNA is removed during the MolYsis treatment. It may create

decreased sensitivity for microorganisms without cell walls, such as *Mycoplasma* spp., *Ureaplasma*, *Chlamydia,* or parasites[19,254,255].

An effective sample preparation method is important for use of SMg for diagnosing infectious diseases since their quality will be reflected in the data produced, which will directly affect the accuracy of species and AMR prediction. A sufficient number of bacterial reads is needed from the sequencing process in order to give accurate results. Our results show that the sample preparation method used results in a high sequencing quality, low human DNA content (<1%), and a high number of bacterial reads (3,7 M, 98 %), in addition to a complete genome coverage of sufficient depth (775x), allowing us to explore SMg application further. These results show significant improvement since a current metadata study reports a median of 79 % of sequence data classified as human DNA in SMg studies, with a median of 89 % (range 76-98 %) when using samples related to PJI including sonication fluid and tissue samples[52].

Background contaminants or clinically significant species?

A significant limitation for the use of SMg assays for infectious disease diagnostic is the possibility of false-positives due to contamination[48,66,67,196]. The spectrum of organisms defined as reportable by SMg assays should be defined, and organisms determined to be background contaminants or clinically insignificant should be described[1]. This is very complicated since the common contaminants also are among the common causes of PJI, making it difficult to discern based on identification alone[48]. We are aware that defining a contaminant is not clear for BCB in the laboratory and this presents a challenge for SMg[256]. However, this issue is not restricted to SMg testing; in fact, this is a recognized problem in the microbiology laboratory, especially for culture and molecular assays[75]. BCBs can lead to contamination and it will vary from batch to batch. In fact, contamination may be present not just from the BCB media but also from reagents, kits, environment, or labware. This is a common problem due to the sensitivity of shotgun metagenomics[67,72], e.g. whole genome amplification kits are also sources of contamination.

There are some recommendations that have been suggested as good practice for quality assurance and validation of SMg for pathogen detection, among them: documenting lot numbers, replacing material, and including controls. Various strategies have been employed to identify the true presence of a microorganism versus contamination. Most studies look at the

reads found in the controls and compare with reads in other samples[66], others set a threshold[56], and some use the reads, DNA quantity, and relative importance of the bacteria[68].

We validated our findings by comparison with conventional culturing methods, where we already knew the outcome, i.e. complete microbiological data from culturing were available for all samples (**Paper II**). Positive and negative controls were included, and all of this information allowed us to assess the true species present versus contamination. In our data analysis, we include several of the various strategies that may help in the discrimination between species presence versus contamination, among them: the proportion of reads assigned to the species present, higher depth and breadth coverage with respect to other species detected by the taxonomical classifier, and comparison with respect to the species, the proportions of reads found in the negative control and to set an appropriate threshold. **In paper II,** we show how the threshold set for taxonomical classification influences the results. Our experience is that it is difficult to create sequencing libraries from negative controls since sometimes sufficient DNA yield is not reached for library preparation protocols[257]. Since only positive BCBs are included in this project, the threshold of detection was not determined.

SMg Sensitivity

The application of SMg for detecting causal pathogens has been the most studied so far. In fact, SMg has actually performed well in detecting the causal agents. Previous studies have demonstrated the potential of SMg in the diagnosis of PJI and that this can be used as an effective tool for the identification of pathogens[258,259]. The studies were using both, Illumina[48,66], and nanopore [69,71,72,260] sequencing platforms. In **paper II**, three negative samples by the conventional culture method reported as positive from the BCB method were included. All bacterial species identified from the BCB culture method were detected indicating that the BCB method and the SMg were consistent concerning bacterial species found and they were found to be the most abundant species present in the sample (mean rate, 97.9 %). In other previous studies, the sensitivity of species identification using SMg on PJI samples has been estimated to be 82 % (72 – 88 %)[52]. We detected 100 % of the species detected by BCB (**Paper II and III**).

Beyond Species identification

The use of SMg on clinical samples could potentially guide therapeutic treatment strategies. Currently, there are some examples of successful identification of drug resistance from a microorganism in samples analyzed by SMg[37,72]. So far, there are studies that attempt to predict drug susceptibility mainly by comparison with phenotype, WGS of isolates or no comparison is made[52]. We did not perform WGS on pure isolates for confirmation of the SMg results, but WGS data from the reference genome from the strain used in the positive control (Spiked sample) confirmed the AMR-profile, typing and virulence gene determinants from SMg. In a meta-analyses study by Kumeren *et al.*[52], it was shown that agreement between the genotype and phenotype was 83 % (68-92 %), very major rates (prediction sensitive, phenotype resistant) was 9% (2-27 %) and major error (prediction resistant, phenotype sensitive) was 1 % (0-20%). Correct antibiotic susceptibility could be predicted in 94 % and 76 % in monomicrobial and polymicrobial samples, respectively in samples of bone and joint infection[72].

In **Paper III**, AMR profiles were generated for *S. aureus* in all the samples containing this species, alone (monomicrobial) or accompanied by *S. agalactiae* (polymicrobial). Prediction of AMR by SMg showed that the *S. aureus* isolates in this study are not very resistant, which is in accordance with data from the NORM surveillance system in 2018[160], where AMR in *S. aureus* is dominated by only beta-lactamase (69.8 %) in blood cultures, while for the rest of the antimicrobials tested were <5 %, with 3 % of the isolates, phenotypically resistant to fusidic acid (n=1,445 isolates)[160]. No MRSA isolates were observed, neither by phenotypical nor genotypical testing. This result is not surprising, since the prevalence of MRSA varies according to the geographical area and Norway has a low prevalence of MRSA (0.8 % prevalence reported in the NORM surveillance system in 2018[160]).

When comparing genotype and phenotype, the disagreement rate was 25 % for *S. aureus* phenotypically resistance and genotypically susceptible to penicillin (false negatives). In the rest, phenotypes could be explained by the presence of ARGs, e.g. *fusB or fusE,* for fusidic acid resistance, and *bla*- operon members for penicillin resistance. However, predicted resistance by genotype but susceptible phenotype (false positives) were found in some cases, e.g. presence of the *tet(38)* gene in all *S. aureus* isolates, but any expressed tetracycline resistance. This confirms that detecting the presence of ARGs in the clinical specimen does not guarantee that a gene is expressed. One of the limitations when using short-reads sequencing technology is

that it is difficult to reliably determine if an ARG is derived from plasmids or chromosomes. Moreover, plasmid detection based on short-read sequences is error-prone and unreliable particularly when larger plasmids (>50 kb) are involved[261].

A factor that plays a vital role in identifying the ARGs from SMg, is the database used. Analysis of SMg data relies on the ability of bioinformatics pipelines to classify reads or contigs using established databases. In **papers II and III,** we have explored available free tools for data analyses and a pipeline was assessed. SMg requires bioinformatics expertise for results interpretation. Caution must be taken about the fact that different pipelines use different microbial databases, which in turn have different degrees of curation, accuracy, and completeness, resulting in varying sensitivity and specificity at this level. In **paper III**, we use the NCBI databases to detect ARGs. It is essential that those implementing SMg in the clinical laboratory use standardized database resources. However, there are many limitations in this regard. Among them, the fact is that no databases have been qualified for clinical use. NCBI is used only for detection of acquired resistance, and CARD is limited to only known resistance genes, without considering other resistance mechanisms such as point mutations, gene expression changes, and posttranslational modifications[262]. The choice of data analysis approach clearly influences the results. We consider that further research is needed to explore which approach gives the most suitable results.

To the best of our knowledge, this is the first study using SMg for the generation of virulence profiles from bacteria, directly on clinical samples (**Paper III**). The strategy for identifying virulence genes is very similar to the one aiming for identifying ARGs[73]. We have generated virulence profiles of *S. aureus* in our samples. We identified some VFs that are known or proposed to play a role in *S. aureus* PJI, e.g. genes involved in colonization and attachment of host tissues or implanted biomaterials such as the *sdrD*, coding for the serine aspartate repeat containing protein D (SdrD) that has been found to play an important role in adhesion to desmoglein in human cells[263] and the adhesins *clfA* and *fnbA* encoding clumping factor A and the fibronectin-binding protein A, respectively[264]. Several studies have indicated that fibronectin is the most important protein in this process. Comprehensive characterization of the VFs carried by the new pathotypes of emerging bacterial pathogens is critical for the effective prevention and control of infectious diseases.

In addition, cgMLST was used for strain-level typing. We used a contig-based approach, which means that we had sufficient depth for metagenomic assembly. cgMLST demonstrated that *S. aureus* in the samples belonged to several lineages, CC15, CC22, CC30, CC45, and CC5. These have also been reported before as common CCs assigned to the *S. aureus* collection from The Tromsø Staph and Skin Study (TSSS)[265]. Interestingly, in a recent study, it was found that commensal nasal isolates belonged to the same CCs as the PJI isolates, including the CCs assigned to our PJT samples, which also displayed the same genetic determinants as the PJI isolates exhibiting a great diversity within the groups[133].

Concluding remarks and future aspects

Despite the initial interest, SMg has many challenges to overcome before being implemented as a diagnostic tool in clinical laboratories. This thesis explored the use of SMg in clinical diagnostics of PJI for the identification of potential pathogens, AMR genes, and virulence determinants directly from BCBs inoculated with PJT specimens.

Overall, we have covered various potential applications of SMg for diagnosis and characterization of bacteria potentially causing PJI. However, we do not believe that SMg can replace conventional culturing today. Most likely, SMg can be a potential diagnostic tool in PJI under certain situations, for instance, to support conventional culture in cases when PJI diagnosis is challenging, e.g. when bacteria are not detected by the laboratory methods but there are still clinical signs of the presence of infection. It is still possible to extract DNA from a negative BCB and analyze if pathogenic bacteria are present. The results of SMg can also be valuable even when concordant with laboratory results, not only providing a guarantee that the laboratory diagnosis is correct, but also allowing extra information, e.g. detecting coinfections and/or predicting antimicrobial susceptibility. SMg will most likely be a good supplementary tool for complicated infection with AMR and long-term antibiotic treatment. It will certainly also play an important role in future clinical microbiology research.

Our results can be useful for further validation, and standardization for the use of SMg on BCBs inoculated with clinical samples for routine diagnostics. However, it is too early to define its place in the clinical microbiology laboratory and PJI diagnosis. We do not know how useful it is at this point, but SMg is in continuous development. There are still challenges that must be solved in order to be able to explore the full potential of SMg. Despite the high cost of SMg and the other current limitations, the role of SMg in the microbiology laboratories will increase during the next years. Our SMg approach presents an alternative tool in PJI diagnosis, complementing the currently available tools.

References

1. Gu, W., Miller, S. & Chiu, C. Y. Clinical Metagenomic Next-Generation Sequencing for Pathogen Detection. *Annu. Rev. Pathol. Mech. Dis.* **14**, (2019).
2. Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **17**, 333–351 (2016).
3. Petersen, L. M., Martin, I. W., Moschetti, W. E., Kershaw, C. M. & Tsongalis, G. J. Third-Generation Sequencing in the Clinical Laboratory: Exploring the Advantages and Challenges of Nanopore Sequencing. *J. Clin. Microbiol.* **58**, e01315-19 (2019).
4. Hasman, H. *et al.* Rapid Whole-Genome Sequencing for Detection and Characterization of Microorganisms Directly from Clinical Samples. *J. Clin. Microbiol.* **52**, 139 LP – 146 (2014).
5. Schmidt, K. *et al.* Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *J. Antimicrob. Chemother.* **72**, 104–114 (2017).
6. Besser, J., Carleton, H. A., Gerner-Smidt, P., Lindsey, R. L. & Trees, E. Next-generation sequencing technologies and their application to the study and control of bacterial infections. *Clin. Microbiol. Infect.* **24**, 335–341 (2018).
7. Schadt, E. E., Turner, S. & Kasarskis, A. A window into third-generation sequencing. *Hum. Mol. Genet.* **19**, R227–R240 (2010).
8. Motro, Y. & Moran-Gilad, J. Next-generation sequencing applications in clinical bacteriology. *Biomol. Detect. Quantif.* **14**, 1–6 (2017).
9. Miller, R. R., Montoya, V., Gardy, J. L., Patrick, D. M. & Tang, P. Metagenomics for pathogen detection in public health. *Genome Med.* **5**, 81 (2013).
10. Metzker, M. L. Sequencing technologies - the next generation. *Nat. Rev. Genet.* **11**, 31–46 (2010).
11. Deurenberg, R. H. *et al.* Application of next generation sequencing in clinical microbiology and infection prevention. *J. Biotechnol.* **243**, 16–24 (2017).
12. Ho, A., Murphy, M., Wilson, S., Atlas, S. R. & Edwards, J. S. Sequencing by ligation variation with endonuclease V digestion and deoxyinosine-containing query oligonucleotides. *BMC Genomics* **12**, 598 (2011).
13. Zhu, F.-Y. *et al.* Comparative performance of the BGISEQ-500 and Illumina HiSeq4000 sequencing platforms for transcriptome analysis in plants. *Plant Methods* **14**, 69 (2018).
14. Pereira, D. M., Fernandes, J. C., Valentão, P. & Andrade, P. B. "Omics" Technologies. in *Principles of Translational Science in Medicine* (ed. Wehling, M. B. T.-P. of T. S. in M. (Second E.) 25–39 (Elsevier, 2015). doi:10.1016/B978-0-12-800687-0.00003-7
15. Fang, C. *et al.* Assessment of the cPAS-based BGISEQ-500 platform for metagenomic sequencing. *Gigascience* **7**, (2018).
16. Buermans, H. P. J. & den Dunnen, J. T. Next generation sequencing technology: Advances and applications. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1842**, 1932–1941 (2014).
17. Dunne, W. M., Westblade, L. F. & Ford, B. Next-generation and whole-genome sequencing in the diagnostic clinical microbiology laboratory. *Eur. J. Clin. Microbiol. Infect. Dis.* **31**, 1719–1726 (2012).
18. Greninger, A. L. *et al.* Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med.* **7**, 99 (2015).
19. Chiu, C. Y. & Miller, S. A. Clinical metagenomics. *Nature Reviews Genetics* **20**, 341–355 (2019).

20. Watts, G. S. & Hurwitz, B. L. Metagenomic Next-Generation Sequencing in Clinical Microbiology. *Clin. Microbiol. Newsl.* **42**, 53–59 (2020).
21. Probst, A. J., Weinmaier, T., DeSantis, T. Z., Santo Domingo, J. W. & Ashbolt, N. New Perspectives on Microbial Community Distortion after Whole-Genome Amplification. *PLoS One* **10**, e0124158 (2015).
22. Gupta, N. DNA Extraction and Polymerase Chain Reaction. *J. Cytol.* **36**, 116–117 (2019).
23. Alonso, A. DNA Extraction and Quantification. in *Encyclopedia of Forensic Sciences* 214–218 (Elsevier, 2013). doi:10.1016/B978-0-12-382165-2.00039-8
24. Glasel, J. A. Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios. *Biotechniques* **18**, 62–63 (1995).
25. Head, S. R. *et al.* Library construction for next-generation sequencing: overviews and challenges. *Biotechniques* **56**, 61-passim (2014).
26. Adey, A. *et al.* Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biol.* **11**, R119 (2010).
27. Bentley, D. R. *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53–59 (2008).
28. Bahassi, E. M. & Stambrook, P. J. Next-generation sequencing technologies: breaking the sound barrier of human genetics. *Mutagenesis* **29**, 303–310 (2014).
29. Pereira, R., Oliveira, J. & Sousa, M. Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics. *J. Clin. Med.* **9**, 132 (2020).
30. Mitchell, S. L. & Simner, P. J. Next-Generation Sequencing in Clinical Microbiology: Are We There Yet? *Clin. Lab. Med.* **39**, 405–418 (2019).
31. 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.* **66**, 778–788 (2018).
32. Bols, E., Smits, L. & Weijenberg, M. Healthy Living: The European Congress of Epidemiology, 2015. *Eur. J. Epidemiol.* **30**, 709–1001 (2015).
33. Zhou, K. *et al.* Characterization of a CTX-M-15 Producing Klebsiella Pneumoniae Outbreak Strain Assigned to a Novel Sequence Type (1427). *Front. Microbiol.* **6**, (2015).
34. Zhou, K. *et al.* Use of whole-genome sequencing to trace, control and characterize the regional expansion of extended-spectrum β-lactamase producing ST15 Klebsiella pneumoniae. *Sci. Rep.* **6**, 20840 (2016).
35. Maljkovic Berry, I. *et al.* Global Outbreaks and Origins of a Chikungunya Virus Variant Carrying Mutations Which May Increase Fitness for Aedes aegypti: Revelations from the 2016 Mandera, Kenya Outbreak. *Am. J. Trop. Med. Hyg.* **100**, 1249–1257 (2019).
36. Moran-Gilad, J. Whole genome sequencing (WGS) for food-borne pathogen surveillance and control – taking the pulse. *Eurosurveillance* **22**, 30547 (2017).
37. Aanensen, D. M. *et al.* Whole-Genome Sequencing for Routine Pathogen Surveillance in Public Health: a Population Snapshot of Invasive Staphylococcus aureus in Europe. *MBio* **7**, (2016).
38. Lasken, R. S. & McLean, J. S. Recent advances in genomic DNA sequencing of microbial species from single cells. *Nat. Rev. Genet.* **15**, 577–584 (2014).
39. Chun, J. & Rainey, F. A. Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea. *Int. J. Syst. Evol. Microbiol.* **64**, 316–324 (2014).
40. Tortoli, E. Microbiological Features and Clinical Relevance of New Species of the Genus Mycobacterium. *Clin. Microbiol. Rev.* **27**, 727–752 (2014).
41. Forbes, J. D., Knox, N. C., Ronholm, J., Pagotto, F. & Reimer, A. Metagenomics: The

Next Culture-Independent Game Changer. *Front. Microbiol.* **8**, (2017).
42. Ronholm, J., Nasheri, N., Petronella, N. & Pagotto, F. Navigating Microbiological Food Safety in the Era of Whole-Genome Sequencing. *Clin. Microbiol. Rev.* **29**, 837–857 (2016).
43. Boers, S. A., Jansen, R. & Hays, J. P. 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 (2019).
44. Dulanto Chiang, A. & Dekker, J. P. From the Pipeline to the Bedside: Advances and Challenges in Clinical Metagenomics. *J. Infect. Dis.* **221**, S331–S340 (2020).
45. Cummings, L. A. *et al.* Clinical Next Generation Sequencing Outperforms Standard Microbiological Culture for Characterizing Polymicrobial Samples. *Clin. Chem.* **62**, 1465–1473 (2016).
46. Van de Peer, Y., Chapelle, S. & De Wachter, R. A quantitative map of nucleotide substitution rates in bacterial rRNA. *Nucleic Acids Res.* **24**, 3381–91 (1996).
47. Forbes, J. D., Knox, N. C., Peterson, C. L. & Reimer, A. R. Highlighting Clinical Metagenomics for Enhanced Diagnostic Decision-making: A Step Towards Wider Implementation. *Computational and Structural Biotechnology Journal* **16**, 108–120 (2018).
48. Ivy, M. I. *et al.* Direct Detection and Identification of Prosthetic Joint Infection Pathogens in Synovial Fluid by Metagenomic Shotgun Sequencing. *J. Clin. Microbiol.* **56**, (2018).
49. Rawat, A., Engelthaler, D. M., Driebe, E. M., Keim, P. & Foster, J. T. MetaGeniE: Characterizing Human Clinical Samples Using Deep Metagenomic Sequencing. *PLoS One* **9**, e110915 (2014).
50. Couto, N. *et al.* Critical steps in clinical shotgun metagenomics for the concomitant detection and typing of microbial pathogens. *Sci. Rep.* **8**, (2018).
51. Wilson, M. R. *et al.* Clinical Metagenomic Sequencing for Diagnosis of Meningitis and Encephalitis. *N. Engl. J. Med.* **380**, 2327–2340 (2019).
52. Govender, K. N., Street, T. L., Sanderson, N. D. & Eyre, D. W. Metagenomic sequencing as a clinical diagnostic tool for infectious diseases: a systematic review and meta-analysis. *medRxiv* 2020.03.30.20043901 (2020). doi:10.1101/2020.03.30.20043901
53. Graham, R. M. A., Doyle, C. J. & Jennison, A. V. Epidemiological typing of Neisseria gonorrhoeae and detection of markers associated with antimicrobial resistance directly from urine samples using next generation sequencing. *Sex. Transm. Infect.* **93**, 65–67 (2017).
54. Gyarmati, P. *et al.* Metagenomic analysis of bloodstream infections in patients with acute leukemia and therapy-induced neutropenia. *Sci. Rep.* **6**, 23532 (2016).
55. Vijayvargiya, P. *et al.* Application of metagenomic shotgun sequencing to detect vector-borne pathogens in clinical blood samples. *PLoS One* **14**, e0222915–e0222915 (2019).
56. Anson, L. W. *et al.* DNA extraction from primary liquid blood cultures for bloodstream infection diagnosis using whole genome sequencing. *J. Med. Microbiol.* **67**, 347–357 (2018).
57. Nakamura, S. *et al.* Direct metagenomic detection of viral pathogens in nasal and fecal specimens using an unbiased high-throughput sequencing approach. *PLoS One* **4**, e4219 (2009).
58. Bogaert, D. *et al.* Variability and Diversity of Nasopharyngeal Microbiota in Children: A Metagenomic Analysis. *PLoS One* **6**, e17035 (2011).

59. Schlaberg, R., Chiu, C. Y., Miller, S., Procop, G. W. & Weinstock, G. Validation of Metagenomic Next-Generation Sequencing Tests for Universal Pathogen Detection. *Arch. Pathol. Lab. Med.* **141**, 776–786 (2017).
60. Langelier, C. *et al.* Metagenomic Sequencing Detects Respiratory Pathogens in Hematopoietic Cellular Transplant Patients. *American journal of respiratory and critical care medicine* **197**, 524–528 (2018).
61. Miller, S. *et al.* Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. *Genome Res.* **29**, 831–842 (2019).
62. Doan, T. *et al.* Illuminating uveitis: metagenomic deep sequencing identifies common and rare pathogens. *Genome Med.* **8**, 90 (2016).
63. Kirstahler, P. *et al.* Genomics-Based Identification of Microorganisms in Human Ocular Body Fluid. *Sci. Rep.* **8**, 4126 (2018).
64. Xing, X.-W. *et al.* Apparent performance of metagenomic next-generation sequencing in the diagnosis of cryptococcal meningitis: a descriptive study. *J. Med. Microbiol.* **68**, 1204–1210 (2019).
65. Thoendel, M. *et al.* Comparison of microbial DNA enrichment tools for metagenomic whole genome sequencing. *J. Microbiol. Methods* **127**, 141–145 (2016).
66. Schmidt, K. *et al.* Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *J. Antimicrob. Chemother.* **72**, 104–114 (2017).
67. Thoendel, M. *et al.* Impact of contaminating DNA in whole-genome amplification kits used for metagenomic shotgun sequencing for infection diagnosis. *J. Clin. Microbiol.* **55**, 1789–1801 (2017).
68. Street, T. L. *et al.* Molecular diagnosis of orthopedic-device-related infection directly from sonication fluid by metagenomic sequencing. *J. Clin. Microbiol.* **55**, 2334–2347 (2017).
69. Sanderson, N. D. *et al.* Real-time analysis of nanopore-based metagenomic sequencing from infected orthopaedic devices. *BMC Genomics* **19**, 714 (2018).
70. Yan, Q. *et al.* Evaluation of the CosmosID Bioinformatics Platform for Prosthetic Joint-Associated Sonicate Fluid Shotgun Metagenomic Data Analysis. *J. Clin. Microbiol.* **57**, (2018).
71. Wang, C. *et al.* Preliminary assessment of nanopore-based metagenomic sequencing for the diagnosis of prosthetic joint infection. *Int. J. Infect. Dis.* **97**, 54–59 (2020).
72. Ruppé, E. *et al.* Clinical metagenomics of bone and joint infections: a proof of concept study. *Sci. Rep.* **7**, 7718 (2017).
73. Padmanabhan, R., Mishra, A. K., Raoult, D. & Fournier, P.-E. Genomics and metagenomics in medical microbiology. *J. Microbiol. Methods* **95**, 415–424 (2013).
74. Han, D. *et al.* mNGS in clinical microbiology laboratories: on the road to maturity. *Crit. Rev. Microbiol.* **45**, 668–685 (2019).
75. Miller, S., Chiu, C., Rodino, K. G. & Miller, M. B. Point-Counterpoint: Should We Be Performing Metagenomic Next-Generation Sequencing for Infectious Disease Diagnosis in the Clinical Laboratory. *J. Clin. Microbiol.* **58**, e01739-19 (2020).
76. Yee, R. & Simner, P. J. Next-Generation Sequencing Approaches to Predicting Antimicrobial Susceptibility Testing Results. *Adv. Mol. Pathol.* **2**, 99–110 (2019).
77. Su, M., Satola, S. W. & Read, T. D. Genome-Based Prediction of Bacterial Antibiotic Resistance. *J. Clin. Microbiol.* **57**, (2018).
78. Kwong, J. C., Mccallum, N., Sintchenko, V. & Howden, B. P. Whole genome sequencing in clinical and public health microbiology. *Pathology* **47**, 199–210 (2015).
79. Zankari, E. *et al.* Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* **67**, 2640–2644 (2012).

80. Jia, B. *et al.* CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **45**, D566–D573 (2017).
81. Alcock, B. P. *et al.* CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **48**, D517–D525 (2020).
82. Liu, B. & Pop, M. ARDB--Antibiotic Resistance Genes Database. *Nucleic Acids Res.* **37**, D443–D447 (2009).
83. Garg, A. & Gupta, D. VirulentPred: a SVM based prediction method for virulent proteins in bacterial pathogens. *BMC Bioinformatics* **9**, 62 (2008).
84. Tsai, C. T., Huang, W. L., Ho, S. J., Shu, L. S. & Ho, S. Y. Virulent-GO: Prediction of Virulent Proteins in Bacterial Pathogens Utilizing Gene Ontology Terms. *Int. J. Biol. Life Sci.* (2009).
85. Liu, B., Zheng, D., Jin, Q., Chen, L. & Yang, J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* **47**, D687–D692 (2018).
86. Chen, L. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res.* **33**, D325–D328 (2004).
87. Lefterova, M. I., Suarez, C. J., Banaei, N. & Pinsky, B. A. Next-Generation Sequencing for Infectious Disease Diagnosis and Management. *J. Mol. Diagnostics* **17**, 623–634 (2015).
88. Gargis, A. S., Kalman, L. & Lubin, I. M. Assuring the Quality of Next-Generation Sequencing in Clinical Microbiology and Public Health Laboratories. *J. Clin. Microbiol.* **54**, 2857–2865 (2016).
89. The Clinical Laboratory Improvement Amendments of 1988. *Health Care Financ. Rev.* (1989).
90. Chelsie, G. *Applications of Clinical Microbial Next-Generation Sequencing*. *American society for microbiology* (2015). doi:10.1128/AAMCol.Apr.2015
91. Aziz, N. *et al.* College of American Pathologists' Laboratory Standards for Next-Generation Sequencing Clinical Tests. *Arch. Pathol. Lab. Med.* **139**, 481–493 (2014).
92. Gargis, A. S. *et al.* Assuring the quality of next-generation sequencing in clinical laboratory practice. *Nat. Biotechnol.* **30**, 1033–1036 (2012).
93. Bal, A. *et al.* Quality control implementation for universal characterization of DNA and RNA viruses in clinical respiratory samples using single metagenomic next-generation sequencing workflow. *BMC Infect. Dis.* **18**, 537 (2018).
94. Wilson, M. R. *et al.* Chronic Meningitis Investigated via Metagenomic Next-Generation Sequencing. *JAMA Neurol.* **75**, 947–955 (2018).
95. Renz, N. & Trampuz, A. Bone and Joint Infections. in *Evidence-Based Infectious Diseases* 23–36 (John Wiley & Sons, Ltd, 2018). doi:10.1002/9781119260363.ch3
96. Tande, A. J. & Patel, R. *Prosthetic Joint Infection*. *Clinical Microbiology Reviews* **27**, (American Society for Microbiology, 2014).
97. Maradit Kremers, H. *et al.* Prevalence of Total Hip and Knee Replacement in the United States. *J. Bone Joint Surg. Am.* **97**, 1386–1397 (2015).
98. Lübbeke, A., Silman, A. J., Barea, C., Prieto-Alhambra, D. & Carr, A. J. Mapping existing hip and knee replacement registries in Europe. *Health Policy (New. York).* **122**, 548–557 (2018).
99. *Norwegian National Advisory Unit on Arthroplasty and Hip Fractures - Annual report*. (2019).
100. Gemmel, F. *et al.* Prosthetic joint infections: radionuclide state-of-the-art imaging. *Eur. J. Nucl. Med. Mol. Imaging* **39**, 892–909 (2012).
101. Li, C., Renz, N. & Trampuz, A. Management of Periprosthetic Joint Infection. *Hip*

pelvis **30**, 138–146 (2018).
102. Sundfeldt, M., V Carlsson, L., B Johansson, C., Thomsen, P. & Gretzer, C. Aseptic loosening, not only a question of wear: A review of different theories. *Acta Orthop.* **77**, 177–197 (2006).
103. Trampuz, A. & Widmer, A. F. Infections associated with orthopedic implants. *Curr. Opin. Infect. Dis.* **19**, 349–356 (2006).
104. Peel, T. N. *et al.* Laboratory workflow analysis of culture of periprosthetic tissues in blood culture bottles. *J. Clin. Microbiol.* **55**, 2817–2826 (2017).
105. Zimmerli, W. trampuz a, ochsner PE. *Prosthetic-joint Infect. n engl J Med* **351**, 1645–1654 (2004).
106. George, D. A. *et al.* Predicting lower limb periprosthetic joint infections: A review of risk factors and their classification. *World J. Orthop.* **8**, 400–411 (2017).
107. Schwarz, E. M. *et al.* 2018 International Consensus Meeting on Musculoskeletal Infection: Research Priorities from the General Assembly Questions. *J. Orthop. Res.* **37**, 997–1006 (2019).
108. Osmon, D. R. *et al.* Diagnosis and Management of Prosthetic Joint Infection: Clinical Practice Guidelines by the Infectious Diseases Society of America. *Clin. Infect. Dis.* **56**, e1–e25 (2013).
109. Parvizi, J. *et al.* New definition for periprosthetic joint infection: From the workgroup of the musculoskeletal infection society. in *Clinical Orthopaedics and Related Research* **469**, 2992–2994 (2011).
110. Parvizi, J., Gehrke, T. & Chen, A. F. Proceedings of the International Consensus on Periprosthetic Joint Infection. *Bone Joint J.* **95-B**, 1450–1452 (2013).
111. Renz, N., Yermak, K., Perka, C. & Trampuz, A. Alpha Defensin Lateral Flow Test for Diagnosis of Periprosthetic Joint Infection. *J. Bone Jt. Surg.* **100**, 742–750 (2018).
112. Izakovicova, P., Borens, O. & Trampuz, A. Periprosthetic joint infection: current concepts and outlook. *EFORT open Rev.* **4**, 482–494 (2019).
113. Arvieux, C. & Common, H. New diagnostic tools for prosthetic joint infection. *Orthop. Traumatol. Surg. Res.* **105**, S23–S30 (2019).
114. Trampuz, A., Osmon, D. R., Hanssen, A. D., Steckelberg, J. M. & Patel, R. Molecular and Antibiofilm Approaches to Prosthetic Joint Infection. *Clin. Orthop. Relat. Res.* **414**, 69–88 (2003).
115. Della Valle, C. *et al.* American Academy of Orthopaedic Surgeons Clinical Practice Guideline on: The Diagnosis of Periprosthetic Joint Infections of the Hip and Knee. *JBJS* **93**, (2011).
116. Trampuz, A. *et al.* Advances in the laboratory diagnosis of prosthetic joint infection. *Reviews in Medical Microbiology* **14**, 1–14 (2003).
117. Parvizi, J. *et al.* The 2018 Definition of Periprosthetic Hip and Knee Infection: An Evidence-Based and Validated Criteria. *J. Arthroplasty* **33**, 1309-1314.e2 (2018).
118. Prieto-Borja, L., Pérez-Jorge, C. & Esteban, J. Chapter 9 - Microbiological Diagnosis of Prosthetic Joint Infections. in *Clinical Microbiology: Diagnosis, Treatments and Prophylaxis of Infections* (eds. Kon, K. & Rai Soft Tissue, Bone and Joint Infections, M. B. T.-T. M. of S.) **2**, 141–152 (Academic Press, 2017).
119. Atkins, B. L. *et al.* Prospective Evaluation of Criteria for Microbiological Diagnosis of Prosthetic-Joint Infection at Revision Arthroplasty. *J. Clin. Microbiol.* **36**, 2932–2939 (1998).
120. Miller, J. M. *et al.* A Guide to Utilization of the Microbiology Laboratory for Diagnosis of Infectious Diseases: 2018 Update by the Infectious Diseases Society of America and the American Society for Microbiologya. *Clin. Infect. Dis.* **67**, e1–e94 (2018).

121. Flurin, L., Greenwood-Quaintance, K. E. & Patel, R. Microbiology of polymicrobial prosthetic joint infection. *Diagn. Microbiol. Infect. Dis.* **94**, 255–259 (2019).
122. Langvatn, H. *et al.* Bacterial and Hematological Findings in Infected Total Hip Arthroplasties in Norway Assessment of 278 Revisions Due to Infection in the Norwegian Arthroplasty Register. *Open Orthop. J.* **9**, 445–449 (2015).
123. Li, Z.-L. *et al.* Identifying Common Pathogens in Periprosthetic Joint Infection and Testing Drug-resistance Rate for Different Antibiotics: A Prospective, Single Center Study in Beijing. *Orthop. Surg.* **10**, 235–240 (2018).
124. Tsai, J.-C., Sheng, W.-H., Lo, W.-Y., Jiang, C.-C. & Chang, S.-C. Clinical characteristics, microbiology, and outcomes of prosthetic joint infection in Taiwan. *J. Microbiol. Immunol. Infect.* **48**, 198–204 (2015).
125. Peel, T. N., Buising, K. L. & Choong, P. F. M. Prosthetic joint infection: challenges of diagnosis and treatment. *ANZ J. Surg.* **81**, 32–39 (2011).
126. Shah, N. B., Tande, A. J., Patel, R. & Berbari, E. F. Anaerobic prosthetic joint infection. *Anaerobe* **36**, 1–8 (2015).
127. Marculescu, C. E. & Cantey, J. R. Polymicrobial prosthetic joint infections: risk factors and outcome. *Clin. Orthop. Relat. Res.* **466**, 1397–1404 (2008).
128. Benito, N. *et al.* Time trends in the aetiology of prosthetic joint infections: a multicentre cohort study. *Clin. Microbiol. Infect.* **22**, 732.e1-732.e8 (2016).
129. Rieber, H. *et al.* Periprosthetic joint infection caused by anaerobes. Retrospective analysis reveals no need for prolonged cultivation time if sensitive supplemented growth media are used. *Anaerobe* **50**, 12–18 (2018).
130. Rasouli, M. R. *et al.* Low Rate of Infection Control in Enterococcal Periprosthetic Joint Infections. *Clin. Orthop. Relat. Res.* **470**, 2708–2716 (2012).
131. Fernandes, A. The Microbiological Profiles of Infected Prosthetic Implants with an Emphasis on the Organisms which Form Biofilms. *J. Clin. DIAGNOSTIC Res.* (2013). doi:10.7860/JCDR/2013/4533.2732
132. Yoon, H.-K. *et al.* A Review of the Literature on Culture-Negative Periprosthetic Joint Infection: Epidemiology, Diagnosis and Treatment. *Knee Surg. Relat. Res.* **29**, 155–164 (2017).
133. Wildeman, P. *et al.* Genomic characterization and outcome of prosthetic joint infections caused by Staphylococcus aureus. *Sci. Rep.* **10**, 5938 (2020).
134. Licitra, G. Etymologia: Staphylococcus. *Emerg. Infect. Dis.* **19**, 1553 (2013).
135. Parte, A. C. LPSN – List of Prokaryotic names with Standing in Nomenclature (bacterio.net), 20 years on. *Int. J. Syst. Evol. Microbiol.* **68**, 1825–1829 (2018).
136. Lpsn, S. J.P. Euzéby: List of Prokaryotic names with Standing in Nomenclature. *Int. J. Syst. Evol. Microbiol.* (2009).
137. Guentzel, M. N. *Escherichia, Klebsiella, Enterobacter, Serratia, Citrobacter, and Proteus*. *Medical Microbiology* (1996). doi:NBK8035 [bookaccession]
138. Ricciardi, B. F. *et al.* Staphylococcus aureus Evasion of Host Immunity in the Setting of Prosthetic Joint Infection: Biofilm and Beyond. *Curr. Rev. Musculoskelet. Med.* **11**, 389–400 (2018).
139. Dastgheyb, S. S. & Otto, M. Staphylococcal adaptation to diverse physiologic niches: an overview of transcriptomic and phenotypic changes in different biological environments. *Future Microbiol.* **10**, 1981–1995 (2015).
140. Peacock, S. J., de Silva, I. & Lowy, F. D. What determines nasal carriage of Staphylococcus aureus? *Trends Microbiol.* **9**, 605–610 (2001).
141. Dobrindt, U. Whole genome plasticity in pathogenic bacteria. *Curr. Opin. Microbiol.* **4**, 550–557 (2001).
142. Lindsay, J. A. & Holden, M. T. G. Staphylococcus aureus: superbug, super genome?

Trends Microbiol. **12**, 378–385 (2004).
143. Turner, N. A. *et al.* Methicillin-resistant Staphylococcus aureus: an overview of basic and clinical research. *Nat. Rev. Microbiol.* **17**, 203–218 (2019).
144. Jackson, R. W., Vinatzer, B., Arnold, D. L., Dorus, S. & Murillo, J. The influence of the accessory genome on bacterial pathogen evolution. *Mob. Genet. Elements* **1**, 55–65 (2011).
145. Sabat, A. J. *et al.* Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Euro Surveill.* **18**, 20380 (2013).
146. Chen, F.-J., Siu, L.-K. K., Lin, J.-C., Wang, C.-H. & Lu, P.-L. Molecular typing and characterization of nasal carriage and community-onset infection methicillin-susceptible Staphylococcus aureus isolates in two Taiwan medical centers. *BMC Infect. Dis.* **12**, 343 (2012).
147. Aanensen, D. M. & Spratt, B. G. The multilocus sequence typing network: mlst.net. *Nucleic Acids Res.* **33**, W728–W733 (2005).
148. Yang, X. *et al.* Multilocus Sequence Typing and Virulence-Associated Gene Profile Analysis of Staphylococcus aureus Isolates From Retail Ready-to-Eat Food in China. *Front. Microbiol.* **9**, (2018).
149. Cha, J. O. *et al.* Molecular analysis of Staphylococcus aureus isolates associated with staphylococcal food poisoning in South Korea. *J. Appl. Microbiol.* **101**, 864–871 (2006).
150. Lindsay, J. A. *et al.* Microarrays Reveal that Each of the Ten Dominant Lineages of Staphylococcus aureus Has a Unique Combination of Surface-Associated and Regulatory Genes. *J. Bacteriol.* **188**, 669–676 (2006).
151. Uelze, L. *et al.* Typing methods based on whole genome sequencing data. *One Heal. Outlook* **2**, 3 (2020).
152. Henri, C. *et al.* An Assessment of Different Genomic Approaches for Inferring Phylogeny of Listeria monocytogenes. *Front. Microbiol.* **8**, (2017).
153. Durand, G. *et al.* Routine Whole-Genome Sequencing for Outbreak Investigations of Staphylococcus aureus in a National Reference Center. *Front. Microbiol.* **9**, 511 (2018).
154. Goudarzi, M., Fazeli, M., Goudarzi, H., Azad, M. & Seyedjavadi, S. S. Spa Typing of Staphylococcus aureus Strains Isolated From Clinical Specimens of Patients With Nosocomial Infections in Tehran, Iran. *Jundishapur J. Microbiol.* **9**, (2016).
155. Monaco, M., de Araujo, F., Cruciani, M., Coccia, E. M. & Pantosti, A. Worldwide Epidemiology and Antibiotic Resistance of Staphylococcus aureus. in *Staphylococcus aureus: Microbiology, Pathology, Immunology, Therapy and Prophylaxis* (eds. Bagnoli, F., Rappuoli, R. & Grandi, G.) 21–56 (Springer International Publishing, 2017). doi:10.1007/82_2016_3
156. Manara, S. *et al.* Whole-genome epidemiology, characterisation, and phylogenetic reconstruction of Staphylococcus aureus strains in a paediatric hospital. *Genome Med.* **10**, 82 (2018).
157. Copin, R. *et al.* Sequential evolution of virulence and resistance during clonal spread of community-acquired methicillin-resistant Staphylococcus aureus. *Proc. Natl. Acad. Sci.* **116**, 1745–1754 (2019).
158. Boswihi, S. S., Udo, E. E. & Al-Sweih, N. Shifts in the Clonal Distribution of Methicillin-Resistant Staphylococcus aureus in Kuwait Hospitals: 1992-2010. *PLoS One* **11**, e0162744–e0162744 (2016).
159. Rossen, J. W. A., Friedrich, A. W. & Moran-Gilad, J. Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology. *Clin. Microbiol. Infect.* **24**, 355–360 (2018).

160. 2018, N.-V. *NORM/NORM-VET 2018. Usage of Antimicrobial Agents and Occurrence of Antimicrobial Resistance in Norway*. (2018). doi:ISSN:1502-2307
161. Post, V. *et al.* Phenotypic and genotypic characterisation of Staphylococcus aureus causing musculoskeletal infections. *Int. J. Med. Microbiol.* **304**, 565–576 (2014).
162. Lenguerrand, E. *et al.* Risk factors associated with revision for prosthetic joint infection following knee replacement: an observational cohort study from England and Wales. *Lancet Infect. Dis.* **19**, 589–600 (2019).
163. Bauer, T. W., Parvizi, J., Kobayashi, N. & Krebs, V. Diagnosis of Periprosthetic Infection. *J. Bone Jt. Surg.* **88**, 869–882 (2006).
164. Esteban, J., Pérez-Jorge, C., Pérez-Tanoira, R. & Gómez-Barrena, E. Microbiological Diagnosis of Prosthetic Joint Infection BT - Infected Total Joint Arthroplasty: The Algorithmic Approach. in (ed. Trebše, R.) 165–179 (Springer London, 2012). doi:10.1007/978-1-4471-2482-5_17
165. Bar??, J., MacDonald, S. J. & Bourne, R. B. Preoperative Evaluations in Revision Total Knee Arthroplasty. *Clin. Orthop. Relat. Res.* **446**, 40–44 (2006).
166. Janz, V., Wassilew, G. I. & Perka, C. Incubation of sonicate fluid in blood culture bottles leads to an improved and quicker rate of bacterial isolation. *Orthop. Proc.* **98-B**, 71 (2016).
167. Font-Vizcarra, L., García, S., Martínez-Pastor, J. C., Sierra, J. M. & Soriano, A. Blood culture flasks for culturing synovial fluid in prosthetic joint infections. *Clin. Orthop. Relat. Res.* **468**, 2238–2243 (2010).
168. Hughes, J. G. *et al.* Culture with BACTEC Peds Plus/F bottle compared with conventional methods for detection of bacteria in synovial fluid. *J. Clin. Microbiol.* **39**, 4468–4471 (2001).
169. Cohen, D., Natshe, A., Ben Chetrit, E., Lebel, E. & Breuer, G. S. Synovial fluid culture: agar plates vs. blood culture bottles for microbiological identification. *Clin. Rheumatol.* **39**, 275–279 (2020).
170. Vasoo, S. Improving the Diagnosis of Orthopedic Implant-Associated Infections: Optimizing the Use of Tools Already in the Box. *J. Clin. Microbiol.* **56**, e01379-18 (2018).
171. Portillo, M. E. *et al.* Improved Diagnosis of Orthopedic Implant-Associated Infection by Inoculation of Sonication Fluid into Blood Culture Bottles. *J. Clin. Microbiol.* **53**, 1622 LP – 1627 (2015).
172. Trampuz, A. *et al.* Sonication of Removed Hip and Knee Prostheses for Diagnosis of Infection. *N. Engl. J. Med.* **357**, 654–663 (2007).
173. Zhai, Z. *et al.* Meta-Analysis of Sonication Fluid Samples from Prosthetic Components for Diagnosis of Infection after Total Joint Arthroplasty. *J. Clin. Microbiol.* **52**, 1730–1736 (2014).
174. Dudareva, M. *et al.* Sonication versus Tissue Sampling for Diagnosis of Prosthetic Joint and Other Orthopedic Device-Related Infections. *J. Clin. Microbiol.* **56**, (2018).
175. Mach, A. J., Adeyiga, O. B. & Di Carlo, D. Microfluidic sample preparation for diagnostic cytopathology. *Lab Chip* **13**, 1011–1026 (2013).
176. Roux, A.-L. *et al.* Diagnosis of prosthetic joint infection by beadmill processing of a periprosthetic specimen. *Clin. Microbiol. Infect.* **17**, 447–450 (2011).
177. Singh, R., Singh, D. & Singh, A. Radiation sterilization of tissue allografts: A review. *World J. Radiol.* **8**, 355–369 (2016).
178. Nylén, T., Saeedi, B., Borg, C., Ullberg, M. & Özenci, V. The performance of 4 different supplements and 5 blood culture bottles types in detection of bacteria and Candida spp. in simulated sterile body fluid cultures. *Diagn. Microbiol. Infect. Dis.* **77**, 1–4 (2013).

179. Altun, O. *et al.* Controlled Evaluation of the New BacT/Alert Virtuo Blood Culture System for Detection and Time to Detection of Bacteria and Yeasts. *J. Clin. Microbiol.* **54**, 1148 LP – 1151 (2016).
180. Berbari, E. F. *et al.* Culture-Negative Prosthetic Joint Infection. *Clin. Infect. Dis.* **45**, 1113–1119 (2007).
181. Southwood, R., Rice, J., McDonald, P., Hakendorf, P. & Rozenbilds, M. Infection in experimental hip arthroplasties. *J. Bone Joint Surg. Br.* **67-B**, 229–231 (1985).
182. Jacobs, M. R. *et al.* Multicenter Clinical Evaluation of BacT/Alert Virtuo Blood Culture System. *J. Clin. Microbiol.* **55**, 2413–2421 (2017).
183. Mueller-Premru, M., Jeverica, S., Papst, L. & Nagy, E. Performance of two blood culture systems to detect anaerobic bacteria. Is there any difference? *Anaerobe* **45**, 59–64 (2017).
184. Weinstein, M. P. Current Blood Culture Methods and Systems: Clinical Concepts, Technology, and Interpretation of Results. *Clin. Infect. Dis.* **23**, 40–46 (1996).
185. She, R. C. *et al.* Performance of the BacT/Alert Virtuo Microbial Detection System for the culture of sterile body fluids: prospective multicentre study. *Clin. Microbiol. Infect.* **24**, 992–996 (2018).
186. Ford, B. & Kallstrom, G. The bioMérieux BacT/Alert: Automation at Last in the Black Box. in *The Dark Art of Blood Cultures* (ed. Jr. W, B. C.) 85–111 (American Society of Microbiology, 2017). doi:10.1128/9781555819811.ch5
187. Broyer, P. *et al.* An Automated Sample Preparation Instrument to Accelerate Positive Blood Cultures Microbial Identification by MALDI-TOF Mass Spectrometry (Vitek®MS). *Front. Microbiol.* **9**, 911 (2018).
188. Opota, O., Croxatto, A., Prod'hom, G. & Greub, G. Blood culture-based diagnosis of bacteraemia: state of the art. *Clin. Microbiol. Infect.* **21**, 313–322 (2015).
189. Elder, B. L. Verification and validation of procedures in the clinical microbiology laboratory. *Clin. Microbiol. Newsl.* **19**, 153–156 (1997).
190. SEWLL & DL. Quality assurance: quality improvement, quality control, and test validation. *Man. Clin. Microbiol.* 55–66 (1995).
191. Lamy, B., Ferroni, A., Henning, C., Cattoen, C. & Laudat, P. How to: accreditation of blood cultures' proceedings. A clinical microbiology approach for adding value to patient care. *Clin. Microbiol. Infect.* **24**, 956–963 (2018).
192. Sandle, T. *Approaching Microbiological Method Validation*. *Journal of Validation Technology* (2015).
193. Schroeter, J., Wilkemeyer, I., Schiller, R. A. & Pruss, A. Validation of the microbiological testing of tissue preparations using the BACTEC blood culture system. *Transfus. Med. Hemotherapy* **39**, 387–390 (2012).
194. Minassian, A. M. *et al.* Use of an automated blood culture system (BD BACTEC™) for diagnosis of prosthetic joint infections: Easy and fast. *BMC Infect. Dis.* **14**, 233 (2014).
195. Velay, A., Schramm, F., Gaudias, J., Jaulhac, B. & Riegel, P. Culture with BACTEC Peds Plus bottle compared with conventional media for the detection of bacteria in tissue samples from orthopedic surgery. *Diagn. Microbiol. Infect. Dis.* **68**, 83–85 (2010).
196. Peel, T. N. *et al.* Improved Diagnosis of Prosthetic Joint Infection by Culturing Periprosthetic Tissue Specimens in Blood Culture Bottles. *MBio* **7**, e01776 (2016).
197. Hughes, H. C. *et al.* Microbiological diagnosis of prosthetic joint infections: a prospective evaluation of four bacterial culture media in the routine laboratory. *Clin. Microbiol. Infect.* **17**, 1528–1530 (2011).
198. Bachmann, N. L., Rockett, R. J., Timms, V. J. & Sintchenko, V. Advances in clinical

sample preparation for identification and characterization of bacterial pathogens using metagenomics. *Front. Public Heal.* **6**, (2018).
199. Mäki, M. Bacterial and Fungal DNA Extraction from Positive Blood Culture Bottles: A Manual and an Automated Protocol. in *Sepsis: Diagnostic Methods and Protocols* (ed. Mancini, N.) 57–64 (Springer New York, 2015). doi:10.1007/978-1-4939-1776-1_6
200. Hasan, M. R. *et al.* Depletion of human DNA in spiked clinical specimens for improvement of sensitivity of pathogen detection by next-generation sequencing. *J. Clin. Microbiol.* **54**, (2016).
201. Zhou, L. & Pollard, A. J. A novel method of selective removal of human DNA improves PCR sensitivity for detection of Salmonella Typhi in blood samples. *BMC Infect. Dis.* **12**, 164 (2012).
202. Duran-Pinedo, A. E. *et al.* Community-wide transcriptome of the oral microbiome in subjects with and without periodontitis. *ISME J.* **8**, 1659–1672 (2014).
203. Sczyrba, A. *et al.* Critical Assessment of Metagenome Interpretation—a benchmark of metagenomics software. *Nat. Methods* **14**, 1063–1071 (2017).
204. Bennett, C. W., Berchem, G., Kim, Y. J. & El-Khoury, V. Cell-free DNA and next-generation sequencing in the service of personalized medicine for lung cancer. *Oncotarget* **7**, 71013–71035 (2016).
205. Angevine, K. *et al.* Detect Rare Alleles in NGS with Molecular Tags. *Genet. Eng. Biotechnol. News* **36**, 24–25 (2016).
206. Roy, S. *et al.* Standards and Guidelines for Validating Next-Generation Sequencing Bioinformatics Pipelines: A Joint Recommendation of the Association for Molecular Pathology and the College of American Pathologists. *J. Mol. Diagnostics* **20**, 4–27 (2018).
207. Kulkarni, P. & Frommolt, P. Challenges in the Setup of Large-scale Next-Generation Sequencing Analysis Workflows. *Comput. Struct. Biotechnol. J.* **15**, 471–477 (2017).
208. Del Fabbro, C., Scalabrin, S., Morgante, M. & Giorgi, F. M. An Extensive Evaluation of Read Trimming Effects on Illumina NGS Data Analysis. *PLoS One* **8**, e85024 (2013).
209. Wingett, S. W. & Andrews, S. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research* **7**, 1338 (2018).
210. Wright, M. N., Gola, D. & Ziegler, A. Preprocessing and Quality Control for Whole-Genome Sequences from the Illumina HiSeq X Platform. in (ed. Elston, R. C.) 629–647 (Springer New York, 2017). doi:10.1007/978-1-4939-7274-6_30
211. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
212. Schmieder, R. & Edwards, R. Fast Identification and Removal of Sequence Contamination from Genomic and Metagenomic Datasets. *PLoS One* **6**, e17288 (2011).
213. Langmean, B. & Salzberg, S. L. Bowtie 2. *Nat. Methods* (2012). doi:10.1038/nmeth.2250.Digestion
214. Chaisson, M. J. & Tesler, G. Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. *BMC Bioinformatics* **13**, 238 (2012).
215. Deurenberg, R. H. *et al.* Reprint of "Application of next generation sequencing in clinical microbiology and infection prevention". *J. Biotechnol.* **250**, 2–10 (2017).
216. Sangwan, N., Xia, F. & Gilbert, J. A. Recovering complete and draft population genomes from metagenome datasets. *Microbiome* **4**, 8 (2016).
217. Olson, N. D. *et al.* Metagenomic assembly through the lens of validation: recent advances in assessing and improving the quality of genomes assembled from

metagenomes. *Brief. Bioinform.* **20**, 1140–1150 (2019).
218. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).
219. Bankevich, A. *et al.* SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
220. Haider, B. *et al.* Omega: an Overlap-graph de novo Assembler for Metagenomics. *Bioinformatics* **30**, 2717–2722 (2014).
221. Ayling, M., Clark, M. D. & Leggett, R. M. New approaches for metagenome assembly with short reads. *Brief. Bioinform.* **21**, 584–594 (2019).
222. Breitwieser, F. P., Lu, J. & Salzberg, S. L. A review of methods and databases for metagenomic classification and assembly. *Brief. Bioinform.* **20**, 1125—1136 (2019).
223. Segata, N. *et al.* Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* **9**, 811–814 (2012).
224. Dilthey, A. T., Jain, C., Koren, S. & Phillippy, A. M. Strain-level metagenomic assignment and compositional estimation for long reads with MetaMaps. *Nat. Commun.* **10**, 3066 (2019).
225. Venter, J. C. Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science (80-.).* **304**, 66–74 (2004).
226. Lu, J., Breitwieser, F. P., Thielen, P. & Salzberg, S. L. Bracken: estimating species abundance in metagenomics data. *PeerJ Comput. Sci.* **3**, e104 (2017).
227. Dougherty, C. E. & Graf, E. Next Generation Sequencing for Outbreak Investigation in the Clinical Microbiology Laboratory. *Am. Soc. Clin. Lab. Sci.* ascls.119.001750 (2019). doi:10.29074/ascls.119.001750
228. Jolley, K. A. & Maiden, M. C. J. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* **11**, 595 (2010).
229. Wilson, M. R. *et al.* Actionable Diagnosis of Neuroleptospirosis by Next-Generation Sequencing. *N. Engl. J. Med.* **370**, 2408–2417 (2014).
230. Tan, T. L. *et al.* Culture-Negative Periprosthetic Joint Infection: An Update on What to Expect. *JB JS open access* **3**, e0060–e0060 (2018).
231. Bejon, P. *et al.* Two-stage revision for prosthetic joint infection: predictors of outcome and the role of reimplantation microbiology. *J. Antimicrob. Chemother.* **65**, 569–575 (2010).
232. Peel, T. N. *et al.* Optimal Periprosthetic Tissue Specimen Number for Diagnosis of Prosthetic Joint Infection. *J. Clin. Microbiol.* **55**, 234–243 (2017).
233. Fink, B. *et al.* The value of synovial biopsy, joint aspiration and C-reactive protein in the diagnosis of late peri-prosthetic infection of total knee replacements. *J. Bone Joint Surg. Br.* **90-B**, 874–878 (2008).
234. Yagupsky, P. *et al.* High prevalence of Kingella kingae in joint fluid from children with septic arthritis revealed by the BACTEC blood culture system. *J. Clin. Microbiol.* **30**, 1278–1281 (1992).
235. von Essen, R. Culture of Joint Specimens in Bacterial Arthritis: Impact of blood culture bottle utilization. *Scand. J. Rheumatol.* **26**, 293–300 (1997).
236. Fuller, D. D. *et al.* Comparison of BACTEC Plus 26 and 27 media with and without fastidious organism supplement with conventional methods for culture of sterile body fluids. *J. Clin. Microbiol.* **32**, 1488–1491 (1994).
237. Fuller, D. D. & Davis, T. E. Comparison of BACTEC plus Aerobic/F, Anaerobic/F, Peds Plus/F, and Lytic/F media with and without fastidious organism supplement to conventional methods for culture of sterile body fluids. *Diagn. Microbiol. Infect. Dis.* **29**, 219–225 (1997).

238. von Essen, R. & Holtta, A. Improved method of isolating bacteria from joint fluids by the use of blood culture bottles. *Ann. Rheum. Dis.* **45**, 454–457 (1986).
239. Li, C., Renz, N., Thies, C. O. & Trampuz, A. Meta-analysis of sonicate fluid in blood culture bottles for diagnosing periprosthetic joint infection. *J. Bone Jt. Infect.* **3**, 273–279 (2018).
240. Shen, H., Tang, J., Wang, Q., Jiang, Y. & Zhang, X. Sonication of Explanted Prosthesis Combined with Incubation in BD Bactec Bottles for Pathogen-Based Diagnosis of Prosthetic Joint Infection. *J. Clin. Microbiol.* **53**, 777 LP – 781 (2015).
241. Yan, Q. *et al.* Comparison of Diagnostic Accuracy of Periprosthetic Tissue Culture in Blood Culture Bottles to That of Prosthesis Sonication Fluid Culture for Diagnosis of Prosthetic Joint Infection (PJI) by Use of Bayesian Latent Class Modeling and IDSA PJI Criteria for. *J. Clin. Microbiol.* **56**, e00319-18 (2018).
242. van den Bijllaardt, W. *et al.* Culturing periprosthetic tissue in blood culture bottles results in isolation of additional microorganisms. *Eur. J. Clin. Microbiol. Infect. Dis.* **38**, 245–252 (2019).
243. Li, C., Ojeda-Thies, C. & Trampuz, A. Culture of periprosthetic tissue in blood culture bottles for diagnosing periprosthetic joint infection. *BMC Musculoskelet. Disord.* **20**, 299 (2019).
244. Pohlig, F. *et al.* Diagnostic accuracy of arthroscopic biopsy in periprosthetic infections of the hip. *Eur. J. Med. Res.* **22**, 6 (2017).
245. Sanabria, A. *et al.* Culturing periprosthetic tissue in BacT/Alert® Virtuo blood culture system leads to improved and faster detection of prosthetic joint infections. *BMC Infect. Dis.* **19**, 607 (2019).
246. Flayhart, D., Borek, A. P., Wakefield, T., Dick, J. & Carroll, K. C. Comparison of BACTEC PLUS Blood Culture Media to BacT/Alert FA Blood Culture Media for Detection of Bacterial Pathogens in Samples Containing Therapeutic Levels of Antibiotics. *J. Clin. Microbiol.* **45**, 816 LP – 821 (2007).
247. Hawkins, B. L., Peterson, E. M. & de la Maza, L. M. Improvement of positive blood culture detection by agitation. *Diagn. Microbiol. Infect. Dis.* **5**, 207–213 (1986).
248. Simor, A. E., Scythes, K., Meaney, H. & Louie, M. Evaluation of the BacT/Alert® microbial detection system with FAN aerobic and FAN anaerobic bottles for culturing normally sterile body fluids other than blood. *Diagn. Microbiol. Infect. Dis.* **37**, 5–9 (2000).
249. Osmon, D. R. Microbiology and Antimicrobial Challenges of Prosthetic Joint Infection. *J. Am. Acad. Orthop. Surg.* **25**, S17–S19 (2017).
250. Zhang, C. *et al.* Value of mNGS in sonication fluid for the diagnosis of periprosthetic joint infection. *Arthroplasty* **1**, 9 (2019).
251. Ji, X.-C. *et al.* Reduction of Human DNA Contamination in Clinical Cerebrospinal Fluid Specimens Improves the Sensitivity of Metagenomic Next-Generation Sequencing. *J. Mol. Neurosci.* **70**, 659–666 (2020).
252. Nelson, M. T. *et al.* Human and Extracellular DNA Depletion for Metagenomic Analysis of Complex Clinical Infection Samples Yields Optimized Viable Microbiome Profiles. *Cell Rep.* **26**, 2227-2240.e5 (2019).
253. Votintseva, A. A. *et al.* Same-Day Diagnostic and Surveillance Data for Tuberculosis via Whole-Genome Sequencing of Direct Respiratory Samples. *J. Clin. Microbiol.* **55**, 1285 LP – 1298 (2017).
254. Rieber, H., Frontzek, A. & Fischer, M. Periprosthetic joint infection associated with Mycoplasma hominis after transurethral instrumentation in an immunocompetent patient. Unusual or underestimated? A case report and review of the literature. *Int. J. Infect. Dis.* **82**, 86–88 (2019).

255. Geipel, U. Pathogenic organisms in hip joint infections. *Int. J. Med. Sci.* **6**, 234–240 (2009).
256. Greninger, A. L. & Naccache, S. N. Metagenomics to Assist in the Diagnosis of Bloodstream Infection. *J. Appl. Lab. Med.* **3**, 643–653 (2019).
257. Greninger, A. L. The challenge of diagnostic metagenomics. *Expert Rev. Mol. Diagn.* **18**, 605–615 (2018).
258. Tarabichi, M. *et al.* Diagnosis of Periprosthetic Joint Infection: The Potential of Next-Generation Sequencing. *JBJS* **100**, (2018).
259. Huang, Z. *et al.* Metagenomic next-generation sequencing contribution in identifying prosthetic joint infection due to Parvimonas micra: a case report. *J. Bone Jt. Infect.* **4**, 50–55 (2019).
260. Helmersen, K. & Aamot, H. V. DNA extraction of microbial DNA directly from infected tissue: an optimized protocol for use in nanopore sequencing. *Sci. Rep.* **10**, 2985 (2020).
261. Peter, S. *et al.* Tracking of antibiotic resistance transfer and rapid plasmid evolution in a hospital setting by Nanopore sequencing. *bioRxiv* 639609 (2019). doi:10.1101/639609
262. Grumaz, S. *et al.* Next-generation sequencing diagnostics of bacteremia in septic patients. *Genome Med.* **8**, 73 (2016).
263. Askarian, F. *et al.* The interaction between Staphylococcus aureus SdrD and desmoglein 1 is important for adhesion to host cells. *Sci. Rep.* **6**, 22134 (2016).
264. Greene, C. *et al.* Adhesion properties of mutants of Staphylococcus aureus defective in fibronectin-binding proteins and studies on the expression of fnb genes. *Mol. Microbiol.* **17**, 1143–1152 (1995).
265. SANGVIK, M. *et al.* Prevalence and population structure of Staphylococcus aureus nasal carriage in healthcare workers in a general population. The Tromsø; Staph and Skin Study. *Epidemiol. Infect.* **141**, 143–152 (2013).

Paper I

BMC Infectious Diseases

RESEARCH ARTICLE Open Access

Culturing periprosthetic tissue in BacT/Alert® Virtuo blood culture system leads to improved and faster detection of prosthetic joint infections

Adriana Sanabria[1*], Merethe E. O. Røkeberg[2], Mona Johannessen[1], Johanna Ericson Sollid[1], Gunnar Skov Simonsen[1,2] and Anne-Merethe Hanssen[1]

Abstract

Background: Blood culture bottles (BCBs) provide a semiautomated method for culturing periprosthetic tissue specimens. A study evaluating BCBs for culturing clinical samples other than body fluids is needed before implementation into clinical practice. Our objective was to evaluate use of the BacT/Alert® Virtuo blood culture system for culturing periprosthetic tissue specimens.

Methods: The study was performed through the analysis of spiked ($n = 36$) and clinical ($n = 158$) periprosthetic tissue samples. Clinical samples were analyzed by the BCB method and the results were compared to the conventional microbiological culture-based method for time to detection and microorganisms identified.

Results: The BacT/Alert® Virtuo blood culture system detected relevant bacteria for prosthetic joint infection in both spiked and clinical samples. The BCB method was found to be as sensitive (79%) as the conventional method (76%) ($p = 0.844$) during the analyses of clinical samples. The BCB method yielded positive results much faster than the conventional method: 89% against 27% detection within 24 h, respectively. The median detection time was 11.1 h for the BCB method (12 h and 11 h for the aerobic and the anaerobic BCBs, correspondingly).

Conclusion: We recommend using the BacT/Alert® Virtuo blood culture system for analyzing prosthetic joint tissue, since this detect efficiently and more rapidly a wider range of bacteria than the conventional microbiological method.

Keywords: BacT/Alert® Virtuo blood culture system, Blood culture bottle, Prosthetic joint infection, Periprosthetic tissue specimens

Background

Prosthetic joint infection (PJI) is one of the most serious complications in joint implantation, and if untreated, it may lead to severe pain, persistent dislocation and death [1]. Approximately 2 % of all patients undergoing joint replacement worldwide are affected by this complication [2–4] and this number is expected to increase with the increasing incidence of arthroplasty surgery [3, 5–7]. The early diagnosis of PJI plays a key role in successful treatment, however, the condition is difficult to diagnose [8, 9].

The diagnosis of PJI is not standardized [9]. The scheme currently in use combines clinical findings and laboratory results [10] where the microbiological assessment of periprosthetic tissue is an important criterion for diagnosis of PJI [4, 10, 11]. Most clinical microbiology laboratory diagnostic methods for PJI are based on culturing bacteria on agar plate and in enrichment broth. These methods are labor intensive, involve subculturing and require daily inspection of enrichment broths. Furthermore, low sensitivity and lack of specificity leads to 10 to 30% false-negative results [3, 12]. These numbers are not surprising. The accurate diagnosis of PJI is

* Correspondence: adriana.m.moreno@uit.no
[1]Research Group for Host-Microbe Interactions, Department of Medical Biology, Faculty of Health Sciences, UiT, The Arctic University of Norway, Tromsø, Norway
Full list of author information is available at the end of the article

challenging due to various factors that affect culture results such as: low bacterial inoculum concentration, prior antibiotic therapy, formation of biofilms, and presence of fastidious slow-growing microorganisms [4, 13, 14]. Hence, improved methods for culturing periprosthetic tissue for diagnosis of PJI are urgently needed.

Inoculation of blood culture bottles (BCBs) is often used for detection of microorganisms. Clinical microbiology laboratories use this semiautomated system with continuous monitoring for microbiological diagnosis, mainly from blood and other body fluids [15]. Previous studies have demonstrated the potential of microbiological detection using BCBs for culturing specimens related to PJI, such as synovial fluid [2, 16], sonication fluid [5, 17–19], and prosthetic joint tissue (PJT) [5, 20–22]. However, there is little data on the culture of PJT samples on BCBs and more evidence is needed for safe patient management before method implementation into the clinical setting [12].

The most common BCB systems used worldwide are BD BACTEC (Becton Dickinson Instrument Systems, Sparks, MD) and BacT/Alert® (bioMérieux, Marcy l'Etoile, France) [15, 23, 24]. The systems are similar as they detect CO_2 production and change in pH but differ in culture media composition and additives. There are many technical factors that could affect the sensitivity of blood culture systems, such as the clinical sample type, volume of blood required for culture, and timing of blood culture media, among others. In order to ensure laboratory quality practice, it is important to verify, evaluate and optimize the use of each BCB system when users practice a subprocess that differs from the manufacturer's guidelines, such as the implementation of BCB for clinical samples other than blood [20].

To the best of our knowledge, most studies on PJI to date have used the BD BACTEC BCB system [5, 17, 19, 22]. Here, we evaluate use of the BacT/Alert® Virtuo BCB system for culturing periprosthetic tissue specimens by the analysis of spiked samples (i.e. simulated PJT specimens with known bacteria) and clinical tissue samples. The use of BCBs emerges as an attractive tool for accurate and timely diagnosis of PJI, leading to improvement in outcome for this challenging type of infection.

Methods

Study design and samples

A prospective laboratory study was conducted over a 11-month period (August 2017–June 2018). Periprosthetic tissue specimens from hip, knee, elbow, ankle and shoulder, belonging to patients with suspicion of PJI were routinely submitted to the Department of Microbiology and Infection Control at the University Hospital of North Norway (UNN), Tromsø, Norway.

Samples were processed using routine standard microbiological procedures. Excess tissue samples (n = 158) from 62 patients were evaluated through the BacT/Alert® Virtuo BCB system. The mean number of specimens received per patient was 2.5 (mode 2, range 1–5).

The BCB method was based on, and modified from, similar methods used in previous studies [12, 20, 21] while the conventional method was an already validated in-house method.

Ethics statement

The work was performed in compliance with the ethical guidelines established by UiT- The Arctic University of Norway. The project was evaluated by the Regional Committee for Medical and Health Research Ethics, Norway (document no. 2016/1247/REK nord), concluding that ethical approval was not required. There were no ethical issues to consider due to use of anonymous clinical samples and development of methodological procedures.

Spiking experiments

Six bacterial species reported as common microbiological causes of PJI were used for spiking experiments. *Escherichia coli* ATCC 25922, *Staphylococcus aureus* ATCC 25923, *Enterococcus faecalis* ATCC 29212, *Staphylococcus epidermidis* ATCC 12228, *Bacteroides fragilis* ATCC 25285 and *Cutibacterium* (formerly *Propionibacterium*) *acnes* (clinical sample).

Excess material of a native femoral head and surrounding tissue from an anonymous donor was crushed, sterilized (irradiated to 25Gy) and tested for contamination by culturing on agar plates. Fresh bacterial cultures were suspended in NaCl 0.85% or in tryptic soy broth (TSB) for aerobic and anaerobic strains, respectively, to a 0.5 McFarland density. These were further diluted to bacterial suspensions of approximately 10^3 colony forming units (CFU)/mL.

A piece of sterilized crushed native femoral head and tissue (≈1 cm^3), i.e. a simulated PJT sample, was transferred to a 15 mL tube containing four mL of glucose broth and five ceramic beads. Five hundred microliters of the bacterial suspension (≈500 CFU) was added to the glucose broth containing the tissue. After two minutes, the mixture was homogenized using a FastPrep-24 instrument (MP Biomedicals, France) to 6.0 M/sec (meters per second) for 40 s.

Subsequently, a Bact/Alert® BCB (FA Plus for aerobic or FN Plus for anaerobic bacteria, bioMérieux, Marcy l'Etoile, France) was inoculated with one mL of the homogenized suspension containing approximately 90–150 CFU/mL and four mL of horse blood (TCS Biosciences Ltd), for a final bacterial concentration of approximately 100 CFU/bottle. BCBs were incubated in the BacT/Alert® Virtuo Microbial Detection System until signaling for positivity, or for a maximum of 12 days. Time to positivity was recorded. When growth was detected,

one drop (≈50 μl) of the BCB medium was subcultured on suitable media to confirm pure culture. Spiking experiments were performed in triplicates and two repeats were done within a time period of one month.

In total, 36 BCBs were spiked, corresponding to 24 BacT/Alert FA plus and 12 BacT/Alert FN plus bottles. Inoculum densities, viability, and purity were checked at different time points during the process by standard microbiological methods. Tissue sterility control was performed for each bottle type by adding sterilized tissue and horse blood and incubating the bottles for 12 days as mentioned above. An overview of the process is provided in Fig. 1.

Analysis of clinical specimens by the conventional and the BCB method

The BCB based method was run in parallel with the conventional diagnostic method. For the conventional method, several operators in the clinical microbiological laboratory were involved in reading of broth and agar plates as part of the routine procedures, while in the BCB method, only one person was involved in the experiments. For both methods, each tissue specimen (≈1 cm^3) inoculated in four mL of glucose broth containing five ceramic beads, was homogenized in a FastPrep-24 instrument at 6.0 M/sec (meters per second) for 40 s (Fig. 1).

In the conventional method, 0.1 mL of the previously homogenized solution was inoculated onto a set of agar plates: blood, lactose, chocolate, Sabouraud and anaerobic blood agar (Fig. 1). All agar plates (aerobic and anaerobic) and the remaining homogenized solution were incubated for five days at 37 °C under aerobic conditions, with the exception of the anaerobic blood agar plates, which were incubated in an anaerobic jar system (Anoxomat® Mark II). Aerobic agar plates and the remaining broth were visually inspected daily, while the anaerobic plates where inspected after five days of incubation. The broth was inspected for opacity (almost always cloudy) and subcultured on aerobic agar for three days at 37 °C. Time to detection (TTD) for the conventional method was recorded (defined as the time until growth is detected on the subsequent agar plate cultures). Conventional microbiological methods for identification of bacteria and fungi were performed on positive subcultures.

In the BCB method, a blood culture set, corresponding to an aerobic bottle (BacT/Alert FA Plus) and an anaerobic bottle (BacT/Alert FN Plus) was inoculated with one mL of the homogenized tissue specimen using a sterile syringe. Inoculated BCBs were enriched with four mL of horse blood and incubated in the Virtuo system up to 12 days. TTD, defined as the time when the BacT/Alert® Virtuo blood culture system signaled positive, was recorded, and one drop from the bottle was subcultured on the agar plate set mentioned above.

For both methods, the identification of bacteria was done using matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF® MS Bruker Daltonics - microflex™) and standard microbiological procedures (Fig. 1).

Statistical analysis

Descriptive statistics for categorical variables were based on percentages and frequencies, while continuous variables were based on means, standard deviations (SDs), medians and interquartile ranges (IQRs). In addition, the McNemar's test was used to evaluate if the differences between the methods were statistically significant. Data were analyzed utilizing GraphPad software 7.0e (GraphPad Software Inc., CA, US).

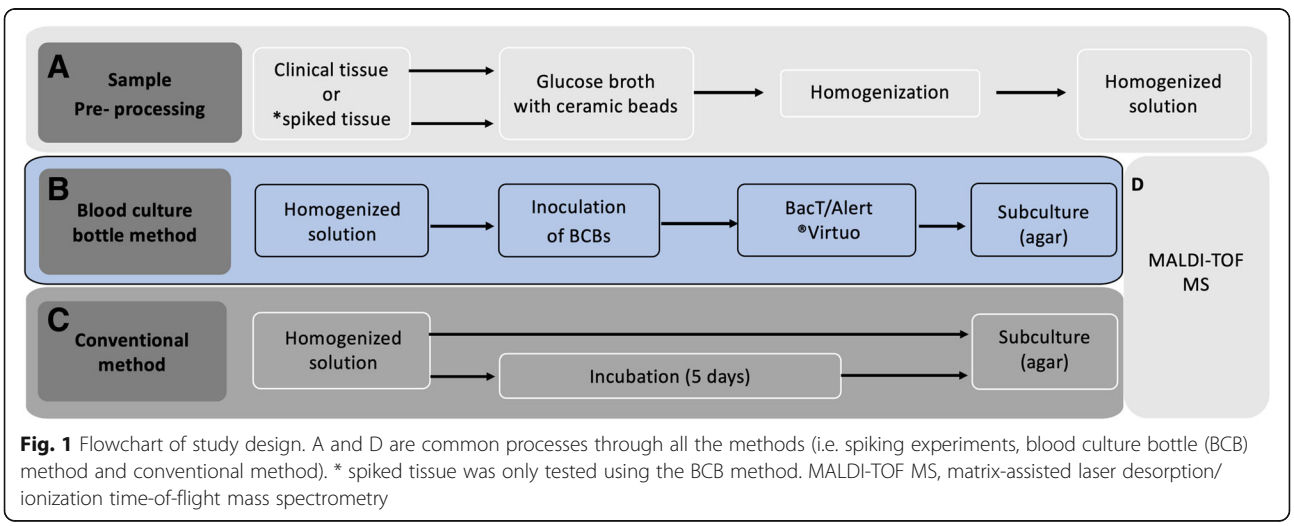


Fig. 1 Flowchart of study design. A and D are common processes through all the methods (i.e. spiking experiments, blood culture bottle (BCB) method and conventional method). * spiked tissue was only tested using the BCB method. MALDI-TOF MS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry

Results

BacT/Alert BCB system for detecting PJI pathogens - spiking

The detection rate in spiked blood culture bottles was 100% ($n = 36$) for both blood culture bottle types. BacT/Alert FA Plus bottles inoculated with aerobic bacteria flagged positive for growth before 15 h (h). The mean time to detection was 10.1 ± 2.2 h with a minimum of 7.9 h and a maximum of 14.1 h. For the BacT/Alert FN Plus bottles inoculated with anaerobic bacteria, there was a remarkable difference in TTD between the two strains tested. *C. acnes* was detected approximately nine hours later than *B. fragilis*. The mean TTD for bottles inoculated with tissue spiked with *B. fragilis* was 25.2 ± 1.1 h with a minimum TTD of 24.4 h and a maximum of 26.7 h, while for *C. acnes*, the mean TTD was 209.2 ± 18.1 h with a minimum and maximum TTD of 175.2 h and 223.2 h, respectively. No difference was observed between the two repeats performed (SD 2.3–2.2), suggesting that the method is reproducible (Additional file 1).

BacT/Alert BCB system effectively detected pathogens from PJT specimens

After having confirmed that the BCB system can effectively detect bacteria commonly found in PJI, we next wanted to test the method on clinical specimens. Each clinical specimen ($n = 158$) was inoculated into two bottles. Therefore, 316 BCBs were included in total for the study, comprising 158 bottles per blood culture type.

Eighty (25.3%) BCBs signaled positive for growth, comprising 44 aerobic and 36 anaerobic bottles. Positive BCBs belonged to 46 (29.1%) clinical tissue specimens from 24 patients (38.70%). Two subcultures of positive BCBs were negative after culturing on agar plates and therefore classified as false positives. These were included as negative samples during all analyses.

Organisms were identified from both bottles in 73.9% of the cases, from the aerobic bottle only in 21.8%, and from the anaerobic bottle only in 4.3%. For aerobic cultures, 89% of the microorganisms were detected within 24 h and 100% within 40 h. For the anaerobic cultures, 97% of the microorganisms were detected within 24 h, and 100% within 31 h.

The mean TTD for aerobic bottles was 13.9 ± 7.8 h within a range of 3.8–39.3 h. For the anaerobic bottles, the mean, minimum and maximum TTD was 11.3 ± 5.4 h, 4 h, and 30.9 h, respectively (Additional file 1).

BCB method yielded faster results compared with conventional method

Use of BCB for analyzing clinical tissue samples from patients with suspicion of PJI was compared with the conventional method according to method sensitivity, TTD and the bacterium (monomicrobial) or bacteria (polymicrobial) identified. For these analyses, a sample was considered positive by the BCB method when one or two of the bottle types flagged positive.

In total, 158 periprosthetic tissue samples from 62 patients were analyzed (Table 1). By using BCBs, 112 samples were negative and 46 positive, belonging to 24 patients. By the conventional method, 114 samples were negative and 44 positive corresponding to 23 patients (Table 1).

Sensitivity was calculated without considering patient clinical data (Additional file 2). The BCB method appeared slightly more sensitive (79%) than the conventional method (76%). A two-tailed P value of 0.844 was obtained, which means that this difference is not statistically significant. There was an 84% agreement rate between the two methods (Table 2).

TTD was recorded for the conventional method through daily visual inspection of agar plates and the glucose broth for bacterial growth. Fifty-nine percent of the bacteria were detected within the first 48 h, consisting of 27 and 32% of bacteria detected on the first and second day of incubation, respectively. The TTD was significantly reduced using the BacT/Alert FA plus and BacT/Alert FN Plus bottles compared to the conventional method. More than 80% of the bacteria were detected in less than 20 h using BCB, compared to the 5 days needed to obtain a similar percentage by the conventional method (Fig. 2).

Microbiology

In total, 46 out of 158 prosthetic joint tissue (PJT) specimens belonging to 24/62 (38.7%) patients had microorganisms detected by BCBs (Table 1). Positive samples were mostly monomicrobial (88.7%), with a low rate of polymicrobial (11.3%). The two methods gave similar bacterial species in 89% of the positive samples, while 11% of the samples gave different bacterial species. The most prevalent microorganisms found using both methods, were *Staphylococcus aureus*, *Enterococcus faecalis*, *Streptococcus mitis/oralis*, *Streptococcus dysgalactiae* and *Staphylococcus epidermidis* (Fig. 2, Additional file 3).

Discussion

Conventional microbiological culture remains the benchmark in PJI diagnosis, despite the longer time period associated with culture-based methods. Alternative approaches that are both effective and practical should be considered for use in the clinical routine setting [25].

Here, we present an evaluation of the use of the BacT/Alert® Virtuo blood culture system for microbiological analyses of periprosthetic tissue specimens as a tool that can be beneficially used for improving and accelerating the diagnosis of PJI. Our analyses included spiked and clinical samples, as well as a comparison of the BCB

Table 1 Results from culturing periprosthetic tissue samples using the blood culture bottle (BCB) method and the conventional method

Tissue samples	BCB method		Conventional method	
	No. of samples (%)	No. of patients (%)	No. of samples (%)	No. of patients (%)
Positive	46 (29.1)	24 (38.7)	44 (27.8)	23 (37.1)
Negative	112 (70.9)	38 (61.3)	114 (72.2)	39 (62.9)
Total	158 (100)	62 (100)	158 (100)	62 (100)

method with the local routine diagnostic method, for TTD and organism(s) identification. The translation of this test system into clinical application at the local level was the main goal of the project.

Currently, a microbiological definition of PJI is the isolation of two or more identical isolates from multiple specimens, or the isolation of one highly virulent organism from a single sample [9]. During this study, we worked with excess material from the clinical microbiology laboratory and the number of specimens accessible per patient varied from 1 to 5 (mean of 2.5). Thus, we have sampled a sub-optimal number of tissue specimens which may have led to diminished sensitivity. According to Peel et al. [26] three PJT specimens obtained and inoculated into BCBs will give the greatest accuracy of PJI diagnosis. Alternatively, four PJT specimens should be obtained and cultured using standard plate and broth cultures. We only evaluated the microbiological status of each sample regardless of the clinical and histopathological requirements needed to be catalogued as a true PJI.

In microbiological diagnostics, culture media and incubation time might have a high influence on test sensitivity. BCBs from different manufacturers have different compositions which could influence their performance [27]. Most of the studies using BCBs for diagnosis of PJI have used the BD BACTEC BCB system [2, 16, 22, 25, 28]. For the BacT/Alert system, only evaluation of synovial fluid has been reported [17, 28]. To the best of our knowledge, this is the first study evaluating the BacT/Alert system for culturing PJT specimens. This is relevant since there are many routine laboratories working with the BacT/Alert system from bioMérieux that could favorably implement this method into their routine procedures.

Table 2 Concordance between culturing clinical tissue samples using the blood culture bottle (BCB) method and the conventional method

Method		No. of tissue samples (%)
Conventional	BCB	
+	+	32 (20.2)
–	–	100 (63.3)
+	–	12 (7.6)
–	+	14 (8.9)
Total		158 (100)

The optimal BCB incubation time is a matter of debate [29]. Commonly, most BCBs with samples obtained from sterile sites are incubated for five days [30]. Some studies suggest prolonging the incubation time up to 14 days, to increase the method sensitivity for anaerobes and slow-growing bacteria [30, 31], while others report that longer incubation time does not increase the method sensitivity [32]. In our study, BCBs were incubated for a period of up to 12 days. During our experiments, most of the bacteria could be detected within the first two days of incubation (Fig. 2), except for the spiked *C. acnes,* which needed a mean time of 209 h (8.7 days) to be detected. In the analysis of clinical specimens, just one clinical sample contained *C. acnes* by the conventional method (possible contamination) while there was not a single sample positive for *C. acnes* or other anaerobic bacteria, using the BCB method.

Recent studies have reported that the BacT/Alert FN Plus performed poorly with regard to TTD for anaerobic bacteria including *C. acnes* [24, 33]. In our case, since the spiked anaerobic bacteria *C. acnes* (clinical strain) and *B. fragilis* (ATCC 25285) could be detected by the BacT/Alert® Virtuo system, the possible lack of sensitivity for the BCB method was discarded. Instead, lack of anaerobic bacteria in the clinical tissue samples may be due to a low bacterial load, to the absence of bacteria in the sample, or to inadequate sample storage and transport (anaerobic device was not used). Overall, our results suggest that BacT/Alert® Virtuo system is able to detect relevant bacteria for PJI, including anaerobic bacteria, and that longer incubation times beyond eight days may increase the detection rate of *C. acnes.*

Despite the results mentioned above, at the general clinical level, long incubation times are not convenient in the routine clinical setting. Early results are significant for the patient as well as for the clinician. *C. acnes* is also a common contaminant of bacterial cultures and its role in PJI is not well defined [30]. We conclude that further research is needed before implementation of long incubation times to increase the detection rate of low prevalent slow-growing bacteria such as *C. acnes*.

In this study, we also compared the BCB method and the local conventional method. Use of BCBs yielded faster results than the conventional method. Shorter time to detection, when using a BCB system, has been reported earlier for different sample types (body fluids and

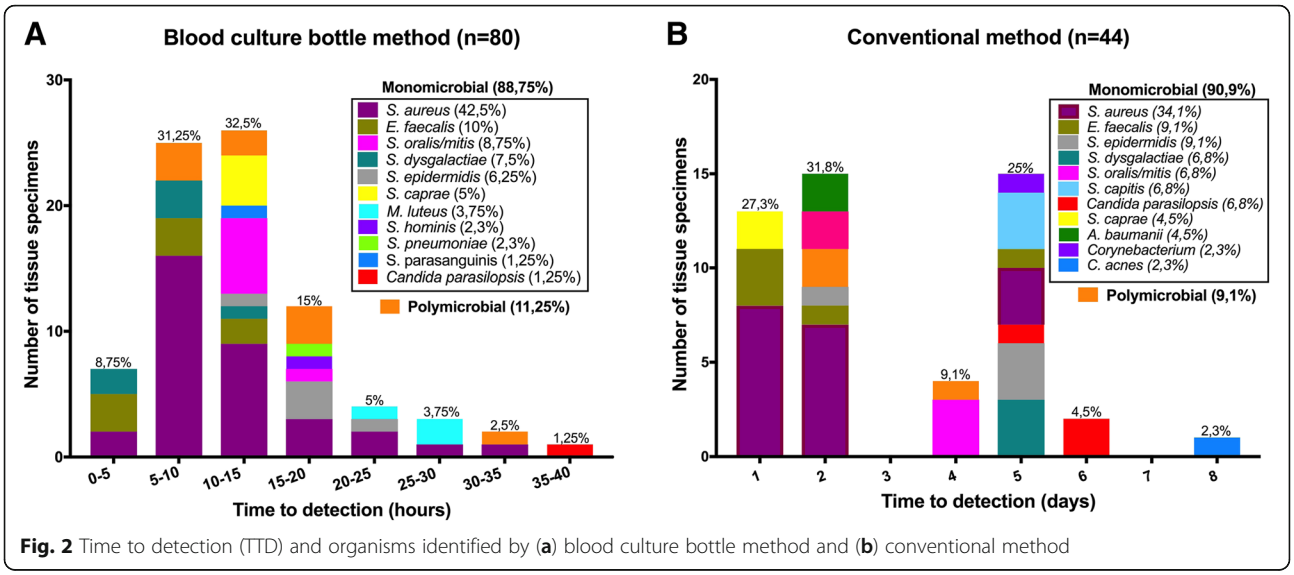


Fig. 2 Time to detection (TTD) and organisms identified by (**a**) blood culture bottle method and (**b**) conventional method

tissue) [20, 26, 34]. Peel et al. [12] compared a BCB system with standard agar and thioglycolate broth culture, yielding faster results using BCBs and showing a 47% improvement of sensitivity.

Additionally, we have used horse blood as enrichment supplement for the BCBs, which has previously been shown to produce high positivity rates and shortening of time to detection [35]. The use of horse blood does not significantly influence the performance of blood culture systems [33, 36].

Also, the BCB method was as specific as the use of agar and broth. These findings are very interesting for cases in which rapid diagnostic methods are applied directly from positive BCBs, shortening the time for pathogen identification and to determine antimicrobial susceptibility. The BCB method was shown to be reproducible during spiking experiments, documented by the fact that there were no differences observed in the TTD from the repeats performed at different time points (Additional file 1).

S. aureus, *E. faecalis* and coagulase-negative staphylococci are among the most common causes of PJI [3, 12], findings that agree with our results. Less commonly recognized pathogens, such as *S. pneumoniae* and *Candida parapsilosis* have also been confirmed to be associated with PJI [37–39].

The BCB method and the conventional method were concordant in most cases. However, in some cases different microbiological agents were identified comparing the two methods. This finding may indicate a higher sensitivity of the BCB method for detection of polymicrobial samples, since a higher number of species was found in samples sub-cultured from positive BCBs (Additional file 3). This result is similar to the report by Velay et al. [22] using the BACTEC BCB system. In cases where the methods yielded different microorganism(s), discrepancies were analyzed (when possible) by comparing the results obtained from other samples belonging to the same patient and/or by the presence of a virulent bacterium (e.g. *S. aureus*). We conclude that in most of the cases, use of the BCB method was more accurate.

There are certain disadvantages using BCBs, e.g. cost of BCBs, cost of waste disposal and capacity problems [6, 26]. There is also a contamination risk when inoculating the homogenized material and horse blood into the bottles, and careful manipulation should be applied.

Our study had some technical limitations. (1) Low total number of clinical samples collected ($n = 158$) and limited number of samples per patient (mean 2.5); (2) Limited clinical data about the patients making it difficult to define a sample selection criterion to distinguish between true PJI and contaminations; (3) the inoculum volume for the BCB (1 mL) was ten times higher than that inoculated onto the agar plates (0,1 mL) which could partially explain the faster detection rate observed with the BCB method, and (4) As for the time of detection, comparative studies of this type contain a bias related to the reading frequency which differed between the two methods. In the BCB method, reading was done automatically every 10 min, whereas in the conventional method, reading was performed manually only once daily to comply with local laboratory practice. This confirms the utility and advantages of the BCB Virtuo system, including automatization, mechanical loading and unloading of bottles, faster detection of microorganisms present in the samples, minimizing the risk of contamination, in addition to reduce labor requirements in the clinical laboratory [23].

Conclusion

In summary, we present a laboratory procedure that may be an important tool in the diagnosis of PJI. We recommend using the BacT/ALERT® BCB system for culturing periprosthetic tissue as a laboratory procedure that can reliably and rapidly detect bacteria commonly found in PJI, thus facilitating early clinical decision making.

Additional files

Additional file 1: Table S1. Analysis of Time to detection (TTD) obtained from the spiking experiments. **Figure S1.** Boxplots from analysis of Time to detection (TTD) obtained from the spiking experiments. (DOCX 354 kb)

Additional file 2: Table S2. Conventional and BCB method sensitivity. (DOCX 13 kb)

Additional file 3: Table S3. Organism(s) identified by the blood culture bottle (BCB) method and the conventional method. (DOCX 96 kb)

Abbreviations

ATCC: American type culture collection; BCB: Blood culture bottle; CFU: Colony forming units; IQRS: Interquartile ranges; MALDI-TOF MS: Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; PJI: Prosthetic joint infection; PJT: Prosthetic joint tissue; SD: Standard deviation; TTD: Time to detection

Acknowledgements

We thank the Department of Microbiology and Infection Control, University Hospital of North Norway, Tromsø, especially Bodil Pedersen for assistance in the use of the BCB system, and Silje Nysted Hagen and Maria Frost for coordinating the samples collection at the preanalytical laboratory. We also thank Dr. Bjarne A. Olsen, Department of Orthopedics, University Hospital of North Norway, Tromsø, for providing the tissue material, and Dr. Turid Hellevik from the Department of Radiation Oncology, University Hospital of North Norway, Tromsø, for performing the tissue sterilization for the spiking experiments.

Declarations

A partial content of this article was presented as a poster at the 28th European Congress of Clinical Microbiology and Infectious Diseases ECCMID, which took place in Madrid, Spain, 21–24 April 2018. The poster was entitled *Evaluation of BacT/Alert Virtuo blood culture system for culturing of periprosthetic tissue specimens*.

Author's contributions

AS performed the experiments and prepared the drafted manuscript. AS, AMH, GSS and JES performed the study design and AS, MEOR, MJ, JES, GSS and AMH assisted in scientific and technical design of experiments. AS, AMH and GSS analysed and interpreted the data results. All authors reviewed, gave inputs and approved the final manuscript.

Funding

This work was supported by grants from joint "Miljøstøtte" financed by Strategisk-HN05–14 (Helse Nord RFH) and Faculty of Health Sciences A20389 (2014–2017), and the National Graduate School in Infection Biology and Antimicrobials (grant number 249062). The publication charges for this article have been funded by a grant from the publication fund of UiT The Arctic University of Norway. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

The work was performed in compliance with the ethical guidelines established by UiT- The Arctic University of Norway. The project was evaluated by the Regional Committee for Medical and Health Research Ethics, Norway (document no. 2016/1247/REK nord), concluding that ethical approval was not required. There were no ethical issues to consider due to use of anonymous clinical samples and development of methodological procedures.

Consent for publication

No applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Research Group for Host-Microbe Interactions, Department of Medical Biology, Faculty of Health Sciences, UiT, The Arctic University of Norway, Tromsø, Norway. [2]Department of Microbiology and Infection Control, University Hospital of North Norway, Tromsø, Norway.

Received: 23 January 2019 Accepted: 19 June 2019
Published online: 10 July 2019

References

1. Moore AJ, Blom AW, Whitehouse MR, Gooberman-Hill R. Deep prosthetic joint infection: a qualitative study of the impact on patients and their experiences of revision surgery. BMJ Open. 2015;5:e009495.
2. Hughes JG, Vetter EA, Patel R, Schleck CD, Harmsen S, Turgeant LT, Cockerill FR 3rd. Culture with BACTEC Peds plus/F bottle compared with conventional methods for detection of bacteria in synovial fluid. J Clin Microbiol. 2001;39:4468–71.
3. Tande AJ, Patel R. Prosthetic joint infection. Clin Microbiol Rev. 2014;27:302–45.
4. Yan Q, Karau MJ, Greenwood-Quaintance KE, Mandrekar JN, Osmon DR, Abdel MP, Patel R. Comparison of diagnostic accuracy of Periprosthetic tissue culture in blood culture bottles to that of prosthesis sonication fluid culture for diagnosis of prosthetic joint infection (PJI) by use of Bayesian latent class modeling and IDSA PJI criteria for classification. J Clin Microbiol. 2018;56.
5. Dudareva M, Barrett L, Figtree M, Scarborough M, Watanabe M, Newnham R, Wallis R, Oakley S, Kendrick B, Stubbs D, et al. Sonication versus tissue sampling for diagnosis of prosthetic joint and other orthopedic device-related infections. J Clin Microbiol. 2018;56.
6. Vasoo S. Improving the diagnosis of orthopedic implant-associated infections: optimizing the use of tools already in the box. J Clin Microbiol. 2018;56:e01379–18.
7. Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. J Bone Joint Surg Am. 2007;89(4):780–5.
8. Dale H, Fenstad AM, Hallan G, Havelin LI, Furnes O, Overgaard S, Pedersen AB, Karrholm J, Garellick G, Pulkkinen P, et al. Increasing risk of prosthetic joint infection after total hip arthroplasty. Acta Orthop. 2012;83:449–58.
9. Parvizi J, Zmistowski B, Berbari EF, Bauer TW, Springer BD, Della Valle CJ, Garvin KL, Mont MA, Wongworawat MD, Zalavras CG. New definition for periprosthetic joint infection: from the workgroup of the musculoskeletal infection society. Clin Orthop Relat Res. 2011;469:2992–4.
10. Parvizi J, Tan TL, Goswami K, Higuera C, Della Valle C, Chen AF, Shohat N. The 2018 definition of Periprosthetic hip and knee infection: an evidence-based and validated criteria. J Arthroplast. 2018;33:1309–1314 e1302.
11. Osmon DR, Berbari EF, Berendt AR, Lew D, Zimmerli W, Steckelberg JM, Rao N, Hanssen A, Wilson WR. Diagnosis and management of prosthetic joint infection: clinical practice guidelines by the Infectious Diseases Society of America. Clin Infect Dis. 2013;56:e1–e25.
12. Peel TN, Dylla BL, Hughes JG, Lynch DT, Greenwood-Quaintance KE, Cheng AC, Mandrekar JN, Patel R. Improved diagnosis of prosthetic joint infection by culturing Periprosthetic tissue specimens in blood culture bottles. MBio. 2016;7:e01776–15.
13. Esteban J. P-JC, Pérez-Tanoira R., Gómez-Barrena E. London: Microbiological Diagnosis of Prosthetic Joint Infection. In *Infected Total Joint Arthroplasty* (Springer BD ed.: In: Trebše R. (eds); (2012).
14. Berbari EF, Marculescu C, Sia I, Lahr BD, Hanssen AD, Steckelberg JM, Gullerud R, Osmon DR. Culture-negative prosthetic joint infection. Clin Infect Dis. 2007;45:1113–9.

15. Weinstein MP. Current blood culture methods and systems: clinical concepts, technology, and interpretation of results. Clin Infect Dis. 1996;23: 40–6.
16. Font-Vizcarra L, García S, Martínez-Pastor JC, Sierra JM, Soriano A. Blood culture flasks for culturing synovial fluid in prosthetic joint infections. Clin Orthop Relat Res. 2010;468:2238–43.
17. Portillo ME, Salvado M, Trampuz A, Siverio A, Alier A, Sorli L, Martinez S, Perez-Prieto D, Horcajada JP, Puig-Verdie L. Improved diagnosis of orthopedic implant-associated infection by inoculation of sonication fluid into blood culture bottles. J Clin Microbiol. 2015;53:1622–7.
18. Janz V, Wassilew GI, Carsten P. Incubation of sonicate fluid in blood culture bottles leads to an improved and quicker rate of bacterial isolation. Orthop Proc. 2016;98-B:71.
19. Shen H, Tang J, Wang Q, Jiang Y, Zhang X. Sonication of explanted prosthesis combined with incubation in BD bactec bottles for pathogen-based diagnosis of prosthetic joint infection. J Clin Microbiol. 2015;53:777–81.
20. Minassian AM, Newnham R, Kalimeris E, Bejon P, Atkins BL, Bowler IC. Use of an automated blood culture system (BD BACTEC™) for diagnosis of prosthetic joint infections: easy and fast. BMC Infect Dis. 2014;14:233.
21. Hughes HC, Newnham R, Athanasou N, Atkins BL, Bejon P, Bowler IC. Microbiological diagnosis of prosthetic joint infections: a prospective evaluation of four bacterial culture media in the routine laboratory. Clin Microbiol Infect. 2011;17:1528–30.
22. Velay A, Schramm F, Gaudias J, Jaulhac B, Riegel P. Culture with BACTEC Peds plus bottle compared with conventional media for the detection of bacteria in tissue samples from orthopedic surgery. Diagn Microbiol Infect Dis. 2010;68:83–5.
23. Jacobs MR, Mazzulli T, Hazen KC, Good CE, Abdelhamed AM, Lo P, Shum B, Roman KP, Robinson DC. Multicenter clinical evaluation of BacT/Alert Virtuo blood culture system. J Clin Microbiol. 2017;55:2413–21.
24. Mueller-Premru M, Jeverica S, Papst L, Nagy E. Performance of two blood culture systems to detect anaerobic bacteria. Is there any difference? Anaerobe. 2017;45:59–64.
25. Geller JA, MacCallum KP, Murtaugh TS, Patrick DA Jr, Liabaud B, Jonna VK. Prospective comparison of blood culture bottles and conventional swabs for microbial identification of suspected Periprosthetic joint infection. J Arthroplast. 2016;31:1779–83.
26. Peel TN, Spelman T, Dylla BL, Hughes JG, Greenwood-Quaintance KE, Cheng AC, Mandrekar JN, Patel R. Optimal Periprosthetic tissue specimen number for diagnosis of prosthetic joint infection. J Clin Microbiol. 2017;55:234–43.
27. Fiori B, D'Inzeo T, Di Florio V, De Maio F, De Angelis G, Giaquinto A, Campana L, Tanzarella E, Tumbarello M, Antonelli M, et al. Performance of two resin-containing blood culture media in detection of bloodstream infections and in direct matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) broth assays for isolate identification: clinical comparison of the BacT/Alert plus and Bactec plus systems. J Clin Microbiol. 2014;52:3558–67.
28. Bourbeau P, Riley J, Heiter BJ, Master R, Young C, Pierson C. Use of the BacT/Alert blood culture system for culture of sterile body fluids other than blood. J Clin Microbiol. 1998;36:3273–7.
29. Rieber H, Frontzek A, Jerosch J, Alefeld M, Strohecker T, Ulatowski M, Morawietz T, Hinsenkamp S, Bell A, Kucukkoylu D, Frommelt L. Periprosthetic joint infection caused by anaerobes. Retrospective analysis reveals no need for prolonged cultivation time if sensitive supplemented growth media are used. Anaerobe. 2018;50:12–8.
30. Butler-Wu SM, Burns EM, Pottinger PS, Magaret AS, Rakeman JL, Matsen FA 3rd, Cookson BT. Optimization of periprosthetic culture for diagnosis of Propionibacterium acnes prosthetic joint infection. J Clin Microbiol. 2011;49:2490–5.
31. Schafer P, Fink B, Sandow D, Margull A, Berger I, Frommelt L. Prolonged bacterial culture to identify late periprosthetic joint infection: a promising strategy. Clin Infect Dis. 2008;47:1403–9.
32. Bernaldo de Quirós JCL, Sánchez-Carrillo C, Cercenado E, Fron C, Catalán P, Bouza E. Clinical impact of a 5-day versus 7-day blood culture incubation period. Clin Microbiol Infect. 1996;2:143–5.
33. Almuhayawi M, Altun O, Abdulmajeed AD, Ullberg M, Ozenci V. The performance of the four anaerobic blood culture bottles BacT/ALERT-FN, –FN plus, BACTEC-plus and -lytic in detection of anaerobic Bacteria and identification by direct MALDI-TOF MS. PLoS One. 2015;10:e0142398.
34. Alfa MA, Degagne P, Olson N, Harding GK. Improved detection of bacterial growth in continuous ambulatory peritoneal dialysis effluent by use of BacT/Alert FAN bottles. J Clin Microbiol. 1997;35:862–6.
35. Nylen T, Saeedi B, Borg C, Ullberg M, Ozenci V. The performance of 4 different supplements and 5 blood culture bottles types in detection of bacteria and Candida spp. in simulated sterile body fluid cultures. Diagn Microbiol Infect Dis. 2013;77:1–4.
36. Altun O, Almuhayawi M, Luthje P, Taha R, Ullberg M, Ozenci V. Controlled evaluation of the new BacT/Alert Virtuo blood culture system for detection and time to detection of Bacteria and yeasts. J Clin Microbiol. 2016;54:1148–51.
37. Raad J, Peacock JE, Jr.: Septic arthritis in the adult caused by Streptococcus pneumoniae: a report of 4 cases and review of the literature. Semin Arthritis Rheum 2004, 34:559–569.
38. Bertani A, Drouin C, Demortiere E, Gonzalez JF, Candoni P, Di Schino M. A prosthetic joint infection caused by Streptococcus pneumoniae: a case report and review of the literature. Rev Chir Orthop Reparatrice Appar Mot. 2006;92:610–4.
39. Cobo F, Rodríguez-Granger J, Sampedro A, Aliaga-Martínez L, Navarro-Marí JM. Candida prosthetic joint infection. A review of treatment methods. J Bone Jt Infect. 2017;2:114–21.

Publisher's Note

Additional file 1: Table S1. Analysis of Time to detection (TTD) obtained from the spiking experiments.

Statistical parameter	Aerobic strains in BacT/ALERT FA Plus[a]			Anaerobic strains in BacT/ALERT FA Plus[b]					
				Bacteroides fragilis			*P. acnes*		
	Setup 1 (hrs)	Setup 2 (hrs)	Setup 1 & 2 (hrs)	Setup 1 (hrs)	Setup 2 (hrs)	Setup 1 & 2 (hrs)	Setup 1 (hrs)	Setup 2 (hrs)	Setup 1 & 2 (hrs)
Min	8	7,9	7,9	24,5	24,4	24,4	213,6	175,2	175,2
1st Quartile	8,6	8,42	8,42	24,6	24,45	24,5	218,4	186,9	206,4
Median	9,65	9,2	9,35	24,7	24,5	24,6	223,2	204	214,8
Mean	10,22	9,93	10,07	25,3	25,13	25,22	220	198,4	209,2
3rd Quartile	10,97	10,5	10,95	25,7	25,5	26,05	223,2	210	221,4
Max	14,1	14	14,1	26,7	26,5	26,7	223,2	216	223,2
SD	2,31	2,21	2,22	1,21	1,18	1,07	5,54	20,96	18,11

Additional file 1: Figure S1. Boxplots from the analysis of Time to detection (TTD) obtained from the spiking experiments.

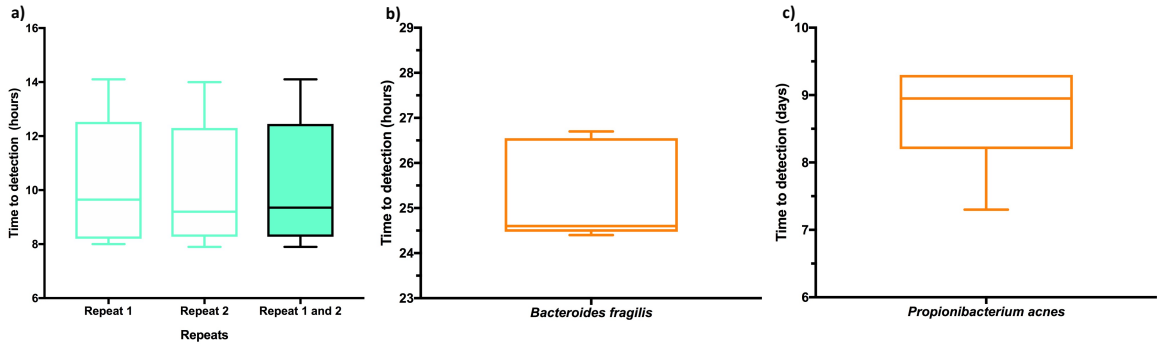


Figure S1: Boxplots from the analysis of Time to detection (TTD) obtained from the spiking experiments. **a)** Time to detection for the BacT/Alert FA Plus bottles inoculated with the aerobic bacteria; **b)** Time to detection for the BacT/Alert FN Plus bottles inoculated with *B. fragilis*; **c)** Time to detection for the BacT/Alert FN Plus bottles inoculated with *P. acnes*.

Additional file 2: Table S2. Conventional and BacT/Alert blood culture system sensitivity

Method	True positives	Sensitivity
Conventional	(32+12) / (44+14) x100	75,86%
BacT/Alert blood culture system	(32+14) / (46+12) x100	79,31%

Additional file 3: Table S3. Organism(s) identified by the blood culture bottle (BCB) method and the conventional method.

Organism (s)	Methods								
	Blood culture bottle method						Conventional Method		
	No of patients	Total No. of positive samples	Total No. of positive BCBs	%	Blood culture bottle type		Total No. of positive samples	No of patients	%
					FA	FN			
Monomicrobial		**43**	**74**	**90,24**			**40**	**22**	**95,00**
Staphylococcus aureus	9	18	36	43,90	18	18	15	7	35,71
Enterococcus faecalis	2	5	9	10,98	4	5	4	2	9,52
Stretococcus oralis/mitis	2	4	7	8,54	4	3	3	1	7,14
Streptococcus dysgalactiae	1	3	6	7,32	3	3	3	1	7,14
Staphylococcus epidermidis	4	4	5	6,10	2	3	4	4	9,52
Staphylococcus caprae	1	2	4	4,88	2	2	2	1	4,76
Micrococcus luteus	3	3	3	3,66	3	0	0	0	0,00
Staphylococcus hominis	1	1	1	1,22	1	0	0	0	0,00
Streptococcus parasanguinis	1	1	1	1,22	0	1	0	0	0,00
Stretococcus pneumoniae	1	1	1	1,22	1	0	0	0	0,00
Candida parapsilosis	2	1	1	1,22	1	0	3	1	7,14
Staphylococcus capitis	0	0	0	0,00	0	0	3	2	7,14
Corynebacterium	0	0	0	0,00	0	0	1	1	2,38
Acinetobacter baumannii	0	0	0	0,00	0	0	2	1	4,76
Propionibacterium acnes	0	0	0	0,00	0	0	1	1	2,38
Polymicrobial		**8**	**8**	**9,76**			**2**	**2**	**5,00**
Enterococcus faecalis + Pseudomonas aeruginosa + Dichelobacter nodosus + Campylobacter jejuni	1	1	1	1,22	1	0	0	0	0,00
Enterococcus faecalis + Serratia licuefaciens	1	1	1	1,22	0	1	0	0	0,00
Acinetobacter pittii + S. epidermidis + Dichelobacter nodosus	1	1	1	1,22	1	0	0	0	0,00
Staphylococcus condimenti + Burkholderia phymatum	1	1	1	1,22	0	1	0	0	0,00
Enterococcus faecalis + Acinetobacter pittii	1	1	1	1,22	1	0	0	0	0,00
Staphylococcus epidermidis + Staphylococcus hominis	1	1	1	1,22	1	0	0	0	0,00
Candida parapsilosis + Micrococcus luteus	1	1	1	1,22	1	0	0	0	0,00
Staphylococcus epidermidis + Staphylococcus dysgalactiae	1	1	1	1,22	1	0	0	0	0,00
Enterococcus faecalis + Acinetobacter baumannii	0	0	0	0,00	0	0	1	1	2,38
Coagulase negative Staphylococcus + *Streptococcus viridans*	0	0	0	0,00	0	0	1	1	2,38
Indeterminate	**2**	**0**	**0**	0,00	1	1	0	0	0,00
Total			**82**	**100,0**	**46**	**38**	**42**		**100,00**

*FA: Aerobic bottle BacT/Alert FA Plus, FN: Anaerobic bottle BacT/Alert FN Plus

Paper II

frontiers in Microbiology

METHODS
published: 17 July 2020
doi: 10.3389/fmicb.2020.01687

Shotgun-Metagenomics on Positive Blood Culture Bottles Inoculated With Prosthetic Joint Tissue: A Proof of Concept Study

Adriana Sanabria[1], Erik Hjerde[2], Mona Johannessen[1], Johanna Ericson Sollid[1], Gunnar Skov Simonsen[1,3] and Anne-Merethe Hanssen[1]*

[1] Research Group for Host-Microbe Interactions, Department of Medical Biology, Faculty of Health Sciences, UiT – The Arctic University of Norway, Tromsø, Norway, [2] Department of Chemistry, Centre for Bioinformatics, UiT – The Arctic University of Norway, Tromsø, Norway, [3] Department of Microbiology and Infection Control, University Hospital of North Norway, Tromsø, Norway

Clinical metagenomics is actively moving from research to clinical laboratories. It has the potential to change the microbial diagnosis of infectious diseases, especially when detection and identification of pathogens can be challenging, such as in prosthetic joint infection (PJI). The application of metagenomic sequencing to periprosthetic joint tissue (PJT) specimens is often challenged by low bacterial load in addition to high level of inhibitor and contaminant host DNA, limiting pathogen recovery. Shotgun-metagenomics (SMg) performed directly on positive blood culture bottles (BCBs) inoculated with PJT may be a convenient approach to overcome these obstacles. The aim was to test if it is possible to perform SMg on PJT inoculated into BCBs for pathogen identification in PJI diagnosis. Our study was conducted as a laboratory method development. For this purpose, spiked samples (positive controls), negative control and clinical tissue samples (positive BCBs) were included to get a comprehensive overview. We developed a method for preparation of bacterial DNA directly from PJT inoculated in BCBs. Samples were processed using MolYsis5 kit for removal of human DNA and DNA extracted with BiOstic kit. High DNA quantity/quality was obtained, and no inhibition was observed during the library preparation, allowing further sequencing process. DNA sequencing reads obtained from the BCBs, presented a low proportion of human reads (<1%) improving the sensitivity of bacterial detection. We detected a 19-fold increase in the number of reads mapping to human in a sample untreated with MolYsis5. Taxonomic classification of clinical samples identified a median of 96.08% (IQR, 93.85–97.07%; range 85.7–98.6%) bacterial reads. Shotgun-metagenomics results were consistent with the results from a conventional BCB culture method, validating our approach. Overall, we demonstrated a proof of concept that it is possible to perform SMg directly on BCBs inoculated with PJT, with potential of pathogen identification in PJI diagnosis. We consider this a first step in research efforts needed to face the challenges presented in PJI diagnoses.

Keywords: shotgun-metagenomics, clinical metagenomics, prosthetic joint infection, blood culture bottles, prosthetic joint tissue

OPEN ACCESS

Edited by:
John W. A. Rossen, University Medical Center Groningen, Netherlands

Reviewed by:
Nur A. Hasan, University of Maryland, United States
Nicholas D. Sanderson, University of Oxford, United Kingdom

****Correspondence:***
Adriana Sanabria
adriana.m.moreno@uit.no

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *20 February 2020*
Accepted: *29 June 2020*
Published: *17 July 2020*

Citation:
Sanabria A, Hjerde E, Johannessen M, Sollid JE, Simonsen GS and Hanssen A-M (2020) Shotgun-Metagenomics on Positive Blood Culture Bottles Inoculated With Prosthetic Joint Tissue: A Proof of Concept Study. Front. Microbiol. 11:1687. doi: 10.3389/fmicb.2020.01687

INTRODUCTION

Metagenomic next-generation sequencing (mNGS) refers to shotgun sequencing of all available DNA and/or RNA in a sample followed by the precise taxonomic identification and classification of each sequence (Couto et al., 2018; Rutanga et al., 2018). The application of shotgun metagenomics (SMg) to clinical samples to recover information of clinical relevance is emerging and it is known as clinical metagenomics (Forbes et al., 2018).

Shotgun-metagenomics has a huge potential, particularly in areas where conventional diagnostic methods have limitations such as in prosthetic joint infection (PJI) (Thoendel et al., 2018). It is a promising approach opening huge opportunities for detecting, identifying and characterizing all potential pathogens, providing at the same time additional inputs on important characteristics for clinical management, such as antibiotic resistance determinants, virulence factors, and bacterial evolution (Wilson et al., 2019).

To date, several studies have provided evidence of the potential and successful applications of clinical metagenomics in a variety of clinical specimens including urine (Hasman et al., 2014; Schmidt et al., 2017), respiratory secretions (Nakamura et al., 2009; Bogaert et al., 2011; Schlaberg et al., 2017; Langelier et al., 2018), cerebrospinal fluid (Schlaberg et al., 2017; Miller et al., 2019), stool (Zhou et al., 2016), blood (whole blood, serum, and plasma), and tissue (Ruppé et al., 2017). Several reviews have summarized the advances, limitations and challenges in the field (Padmanabhan et al., 2013; Simner et al., 2018; Chiu and Miller, 2019).

There are many issues that need to be addressed before performing clinical metagenomics in real time directly from clinical samples as an integral part of routine diagnostic testing. Among them are the technical challenges related to sample preparation protocols, to process and analyze clinical specimens. In this regard, it is important to consider that different specimen types present their unique and specific challenges reflecting their matrix and concentrations of the target pathogen(s) and normal (resident) flora (Bachmann et al., 2018). Therefore, when performing sequencing directly from a clinical sample, special attention has to be paid to (1) contamination of host DNA and other microorganisms, (2) low abundance of the target organism present in the sample and (3) the presence of DNA amplification inhibitors and other sample variables (i.e., sample type, matrix) (Mulcahy-O'Grady and Workentine, 2016; Bachmann et al., 2018).

Prosthetic joint infection (PJI) is a serious complication after joint implantation, occurring in 1–2% of primary arthroplasties and 4% in revision arthroplasties (Ong et al., 2009; Corvec et al., 2012; Izakovicova et al., 2019). The infection is associated with high morbidity rates and diagnosis is challenging due to imperfect definition and imperfect diagnostic techniques (Tande and Patel, 2014; Park and Patel, 2018). Undiagnosed PJI cases due to culture-negatives have been estimated to a proportion of 10–30%, for example due to ongoing empirical antibiotic treatment (Tande and Patel, 2014; Peel et al., 2016). Culturing the causative pathogen takes longer time and is problematic in less virulent, fastidious and slow growing organisms (Schafer et al., 2008).

The diagnosis of PJI is not standardized (Parvizi et al., 2011). However, the microbiological assessment of periprosthetic tissue is the most important method and to date the gold standard for diagnosing PJI (Parvizi et al., 2011; Yan et al., 2018). Due to low sensitivity and specificity, rapid and accurate diagnosis is still a challenge. Additionally, methods for rapid pathogen identification directly on clinical samples such as multiplex PCR, MALDI-TOF and whole genome sequencing (WGS) have been developed (Greenwood-Quaintance et al., 2014; Patel, 2015; Tagini and Greub, 2017). However, they are still dependent on pure microbial culture in addition to the fact that some of them do not give information beyond species identification (Török et al., 2013). Research and development of new diagnostic methods that overcome these limitations are required.

In a previous study, we assessed the use of a BacT/Alert® Virtuo blood culture system for culturing periprosthetic tissue specimens (Sanabria et al., 2019). The blood culture bottle (BCB) method was found to detect a wider range of bacteria more rapidly than the conventional microbiological method. Furthermore, previous studies have demonstrated the potential of microbiological detection using BCBs for culturing specimens related to PJI, such as synovial fluid (Hughes et al., 2001; Font-Vizcarra et al., 2010), sonication fluid (Portillo et al., 2015; Shen et al., 2015; Janz et al., 2016; Dudareva et al., 2018), and prosthetic joint tissue (PJT) (Velay et al., 2010; Minassian et al., 2014; Peel et al., 2016, 2017; Dudareva et al., 2018). The evaluation of SMg in the diagnosis of PJI remains limited (Zhang et al., 2019). To date, studies have investigated the application of SMg on tissue specimens (Ruppé et al., 2017), synovial fluid (Ivy et al., 2018) and sonication fluid (Thoendel et al., 2016, 2017, 2018; Street et al., 2017; Sanderson et al., 2018; Yan et al., 2019), where the main obstacle has been a high background of genetic material mainly derived from the host, hampering the detection of pathogens (Zhang et al., 2019).

So far, there are no studies about SMg on BCBs inoculated with any specimens related to PJI. We believe that the combination of inoculation of BCBs with PJT specimens followed by direct DNA sequencing may be a beneficial strategy. Here, we developed a proof of concept study with the aim of evaluating the use of SMg on BCBs inoculated with PJT for pathogen identification in diagnosis of PJI. Our study was conducted as a laboratory method development, including PJT specimens and appropriate controls. The aim was to test if SMg technology works on this specific type of specimens, and for this purpose develop a method for preparation of high-quality bacterial DNA from PJT for downstream SMg, establish a bioinformatics pipeline, and compare SMg results with conventional culture method results. SMg was performed to investigate if it was possible to obtain an acceptable high number of bacterial reads, genome coverage and genome sequencing depth for identification of PJI pathogens.

Abbreviations: ATCC, American type culture collection; BCBs, Blood culture bottles; ENA, European Nucleotide Archive; IQR, Interquartile ranges; MALDI-TOF, Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; mNGS, metagenomic next-generation sequencing; PJI, Prosthetic joint infection; PJT, Periprosthetic joint tissue; SMg, Shotgun-metagenomics; WGS, Whole genome sequencing.

MATERIALS AND METHODS

Ethics Statement

This study was performed in accordance with the ethical guidelines established by The Arctic University of Norway (UiT). The project has been evaluated by the Regional Committee for Medical and Health Research Ethics, Rec North, Norway (document no. 2016/1247/REK nord), concluding that ethical approval was not required. There were no ethical issues to consider due to use of anonymous clinical samples and development of methodological procedures.

Sample Collection

A sample collection of positive BCBs was obtained from a previous study, where we evaluated the use of Bact/Alert® BCBs (bioMérieux, Marcy l'Etoile, France) for culturing PJT (Sanabria et al., 2019). In brief, three different types of BCB samples were obtained during that study: BCBs inoculated with PJT clinical samples from patients with suspicion of PJI, BCBs inoculated with tissue spiked with bacterial species reported as common microbiological causes of PJI and a negative control which was prepared by inoculating sterilized tissue (irradiated to 25 Gy) from a crushed native femoral head. For further details on the BCB sample preparation method, see Sanabria et al. (2019).

Bact/Alert® BCBs were enriched with 4 mL of horse blood, which has previously been shown to produce high positivity rates and shortening of time to detection (Nylén et al., 2013) and the PJT was analyzed in parallel with the conventional diagnostic method. Bacterial identification was performed using matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF® MS Bruker Daltonics – microflex™).

In this study, 25 positive aerobic BCBs (BacT/Alert FA Plus) inoculated with PJT were used initially to test two different DNA sample preparation methods. Subsequently, DNA from nine BCBs were selected for SMg. Selection criteria included: samples with presence of a single species of microorganism (monomicrobial) reported as common microbiological cause of PJI and high quality/quantity of total DNA concentration of at least 1 ng/μL bacterial DNA. Samples sequenced included: samples with *Staphylococcus aureus* (n = 6), *Staphylococcus epiderimidis* (n = 2), *Enterococcus faecalis* (n = 1), negative control (NC, n = 1), and spiked samples (positive controls, PC1-3) (n = 3). The negative control contained DNA extracted from a BCB enriched with horse blood and inoculated with sterilized tissue (irradiated to 25 Gy) from a donor with no suspicion of infection. The three spiked samples were BCBs inoculated with tissue spiked with a suspension of approximately 90–150 CFU/mL of *Escherichia coli* ATCC 25922, *S. aureus* ATCC 25923, or *S. aureus* ATCC 25923, and *Cutibacterium acnes* (clinical sample), respectively.

In addition, DNA from one of the nine clinical samples was sequenced three times (sample 1: S1a, S1b, S1c) in order to evaluate the impact of the sample preparation method on the metagenomic results, and to determine a suitable number of samples that could be multiplexed in one lane on the flow cell during the sequencing process.

Figure 1 gives an overview of the number of samples and the sample types included in each step of the process. An overview of all the samples sequenced through SMg, including the controls and their features are listed in **Table 1**.

DNA Preparation

Total gDNA was extracted from the blood culture bottles using the QIAamp BiOstic Bacteremia DNA Kit (Qiagen, Hilden, Germany). Samples were pre-treated using MolYsis™ Basic5 kit (Molzym, Bremen, Germany) to deplete human DNA from the samples before DNA extraction. In order to find the most suitable procedure for extracting DNA from BCBs, DNA was extracted with or without pretreatment with the MolYsis5 kit. Sample pre-treatment and DNA extraction methods were according to the manufacturer's instructions and the procedures were evaluated on the basis of DNA quantity and quality. Total DNA concentration was measured using a Qubit dsDNA HS Assay Kit (Life Technologies, Carlsbad, CA, United States) and DNA quality (OD_{260}/OD_{280} and OD_{260}/OD_{230}) determined by Nanodrop.

qPCR

Each DNA sample was fold diluted 1:100 and 1:1000 after determining these as suitable dilution factors to avoid inhibition during amplification. Bacterial DNA concentration was calculated by qPCR targeting the 16S rDNA gene. The standard curve consisted of a 10-fold dilution series of a mix of gDNA from bacterial species common to PJI: *S. aureus* ATCC25923, *E. coli* ATCC25922 and *S. epidermidis* ATCC 12228. The primers and the probe used were as follows: forward primer 5′-CGA AAG CGT GGG GAG CAA A-3′; reverse primer 5′-GTT CGT ACT CCC CAG GCG G-3′; probe 5′-(FAM)-ATT AGA TAC CCT GGT AGT CCA-(MGB)-3′ (Bogaert et al., 2011). The 20-μL PCR master mix consisted of 2X TaqMan Universal master mix II with UNG, 10 μM of each primer (0.8 μL), 5 μM MGB probe (0.8 μL), 2.6 μL DNA free water, and 5 μL of template DNA (Cremers et al., 2014).

Before adding the DNA template, 1M DTT and dsDNase (0.5 μL) from PCR decontamination kit (ArcticZymes, Tromsø, Norway) were added. Samples were then decontaminated by incubation at 37°C for 20 min followed by 60°C for 20 min to inactivate the dsDNase. Amplification was preformed using a 7300 Real-time PCR system (Applied Biosystems, Foster City, CA 94404 USA) under the following conditions: 50°C for 2 min, 95°C for 10 min, 50 cycles of 95°C for 15 s and 60°C for 1 min. The gDNA samples were stored at −20°C until further use.

Metagenomic Sequencing

Sequencing libraries were prepared using the ThruPLEX® DNA-seq Kit (*Rubicon* Genomics, United States) following the manufacturer's instructions. Approximately 100 ng of DNA was used as input for library preparation from the clinical and spiked samples, while for the negative control, a little more than 1 ng was used. The sequencing process was performed at the Norwegian Sequencing Centre, Oslo, using a MiSeq sequencer (Illumina Inc., San Diego, CA, United States) with v2 chemistry and 500 cycles for 250 bp paired-end sequencing. Samples were multiplexed

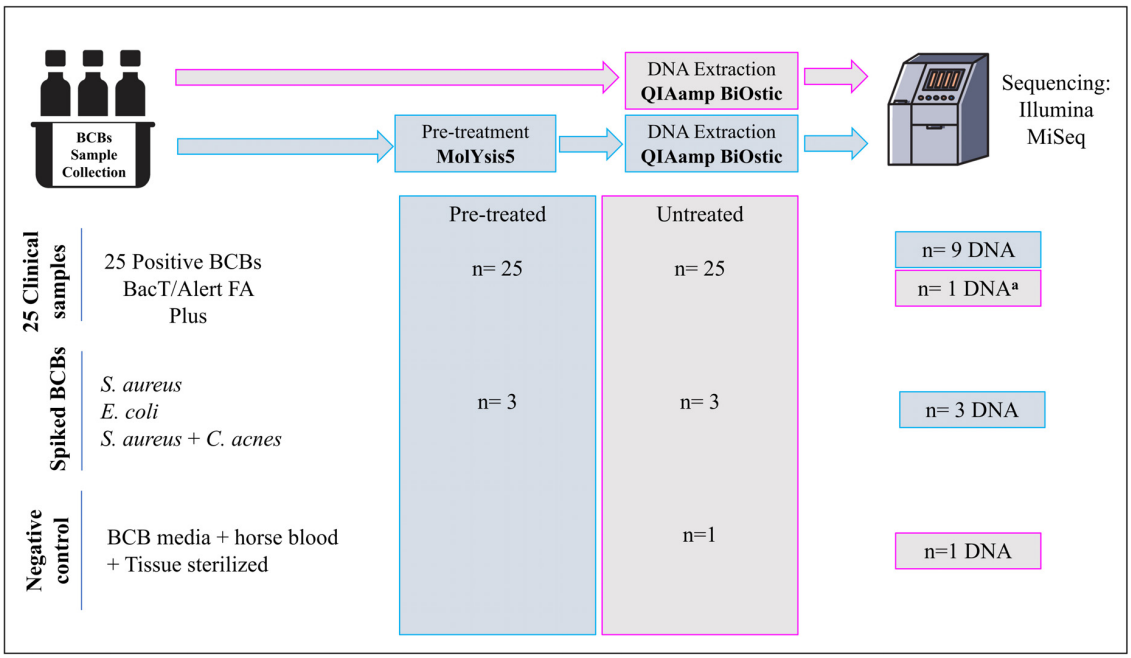


FIGURE 1 | Overview of the number of samples included in each step of the process. [a]DNA from a BCB sample untreated with molYsis5. Treated: BCB samples pre-treated with MolYsis5 before DNA extraction. Untreated: BCB samples with no pretreatment with MolYsis5 before DNA extraction.

TABLE 1 | Description of samples sequenced through SMg.

Sample		Sample type inoculated in BCB	DNA preparation method	Microorganism identified Laboratory[a] (MALDI-TOF)	
No	Code			BCB	Conventional
1	S1a	PJT (clinical sample)	MolYsis5 + BiOstic	*S. aureus*	*S. aureus*
	S1b	PJT (clinical sample)	MolYsis5 + BiOstic	*S. aureus*	*S. aureus*
	S1c	PJT (clinical sample)	BiOstic	*S. aureus*	*S. aureus*
2	S2	PJT (clinical sample)	MolYsis5 + BiOstic	*S. aureus*	*S. aureus*
3	S3	PJT (clinical sample)	MolYsis5 + BiOstic	*S. aureus*	*S. aureus*
4	S4	PJT (clinical sample)	MolYsis5 + BiOstic	*S. aureus*	*S. aureus*
5	S5	PJT (clinical sample)	MolYsis5 + BiOstic	*S. aureus*	*S. aureus*
6	S6	PJT (clinical sample)	MolYsis5 + BiOstic	*E. faecalis*	*E. faecalis*
7	S7	PJT (clinical sample)	MolYsis5 + BiOstic	*S. aureus*	No growth
8	S8	PJT (clinical sample)	MolYsis5 + BiOstic	*S. epidermidis*	No growth
9	S9	PJT (clinical sample)	MolYsis5 + BiOstic	*S. epidermidis*	No growth
10	PC1	Tissue spiked with *S. aureus*	MolYsis5 + BiOstic	*S. aureus*	NA
11	PC2	Tissue spiked with *E. coli*	MolYsis5 + BiOstic	*E. coli*	NA
12	PC3	Tissue spiked with *S. aureus* and *C. acnes*	MolYsis5 + BiOstic	*S. aureus*	NA
13	NC	Tissue sterile	BiOstic	NA	NA

[a]*Results from our previous study (Sanabria et al., 2019) including BCB method and conventional method. S1–S9, clinical samples; PC1-3, spiked samples (positive controls); NC, negative control; NA, not applicable.*

with 3 or 4 samples per lane. To estimate how many samples could be multiplexed in one MiSeq lane, one of the samples was run on a single lane in a pilot assay.

Bioinformatic Data Analysis

The bioinformatic analysis can be summarized in two main steps: reads preprocessing (**Figure 2A**) and taxonomic analyses (**Figure 2B**).

Obtained raw reads were checked for quality using FastQC software v0.11.8[1]. Optical duplicates in fastq files were removed using the program Clumpify v38.36 from BBTools suite[2] with default parameters. Adapter sequences were trimmed off and the poor-quality reads were removed using BBDuk of BBTools

[1]http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
[2]https://jgi.doe.gov/data-and-tools/bbtools/

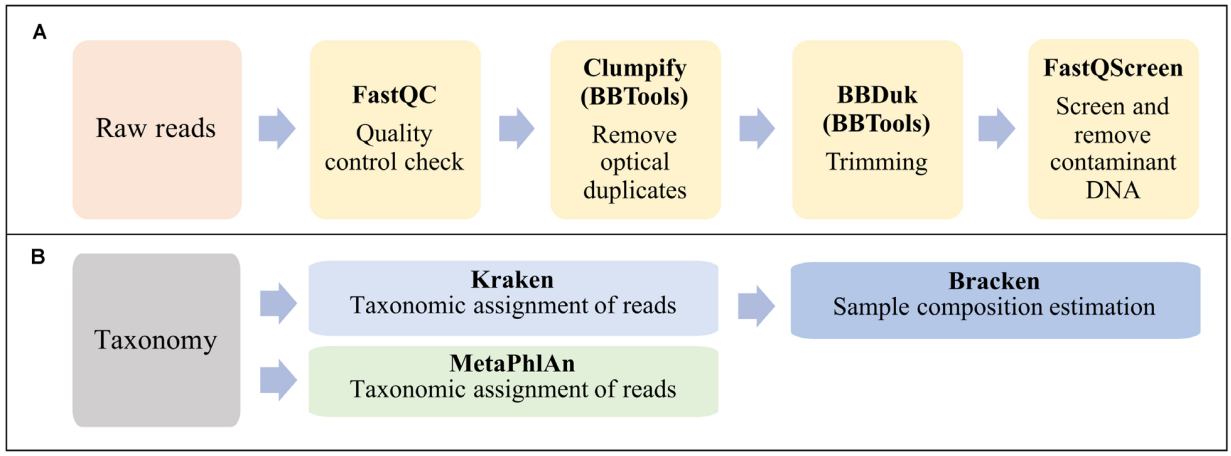


FIGURE 2 | Workflow summarizing the bioinformatic analysis in this study. **(A)** Reads preprocessing. **(B)** Taxonomic classification and sample composition estimation.

suite. The minimal length and Phred score were set to 50 and 20 nucleotides, respectively. Contaminant DNA was identified and removed by mapping all reads against the reference genomes of human GRCh38.p13 (GCF_000001405.39), horse (GCF_002863925.1) and the PhiX phage (*Escherichia* virus phiX174, GCF_000819615.1) using Bowtie2 aligner in FastQ Screen v0.13.0 (Wingett and Andrews, 2018).

Preprocessed PE reads were classified using two established tools for taxonomic profiling of metagenomic samples: Kraken (Salzberg and Wood, 2014), which is based on exact alignments of genomic k-mers, and MetaPhlAn2 which is based on clade-specific marker genes (Segata et al., 2012). However, in this work we have chosen to focus more on the Kraken analyses.

Taxonomic classification with MetaPhlAn2 v2.7.7 was executed with default parameters and using the database provided by the tool, as of July 2019. Taxonomic classification with Kraken v1.1.1 was performed using default parameters against the 8 GB dustmasked miniKraken database constructed from bacterial, archaeal, and viral genomes in Refseq, as of Oct. 18, 2017. Re-estimation of abundance was done using Bracken (Lu et al., 2017). The same pipeline for taxonomical assignment and species abundance estimation was applied to all the samples after downsizing the number of reads by randomly sampling (using SeqKit v0.11.0.) a proportion of reads from the full dataset. Subsamples from 10 to 100% of the reads were extracted and analyzed. In addition, the effect of setting a threshold for species detection level when using Kraken classifier, was evaluated by setting different threshold values (0, 10, 100, 500, 1000, 1500, 2500, 3000, 3500, and 4000 reads) on the spiked and clinical samples.

Finally, as a control to test false taxonomic assignments by Kraken, sequence reads belonging to the spiked samples and the negative control were aligned against reference genomes belonging to the top species hits from Kraken classification, using Burrows–Wheeler aligner-maximum exact matches (BWE-MEM) v0.7.17 (Li and Durbin, 2009).

Statistical Analysis

Descriptive statistics for categorical variables were based on percentages and frequencies, while continuous variables were based on means, standard deviations (SDs), medians and interquartile ranges (IQRs). In addition, Wilcoxon rank sum test was used to evaluate if the differences between the DNA sample preparation methods were statistically significant. The differences were considered statistically significant with *P*-values lower than 0.05. Data were analyzed utilizing GraphPad Prism software, version 8.3.0 (GraphPad Software Inc., CA, United States).

Classification results from the metagenomics experiments were explored using the Pavian R package version 0.8.4 (Breitwieser and Salzberg, 2019) by using their data tables, heatmaps, and Sankey flow diagrams.

RESULTS

Establishment of Method for Bacterial DNA Isolation From BCBs Inoculated With PJT

Since there is no standard procedure for the isolation of DNA from BCBs inoculated with PJT, we initially examined the performance of QIAamp BiOstic Bacteremia DNA Kit pre-treated and untreated with the MolYsis basic5 kit.

DNA from 25 positive BCBs (BacT/Alert® FA Plus, aerobic bottles) were isolated using the BiOstic kit with or without pretreatment with MolYsis5. Comparison between the two methods was based on the total DNA yield and the DNA quality (**Supplementary Table S1**). DNA concentration measurements showed that samples treated with MolYsis5 yielded a higher mean DNA concentration (DNA 84.23 ng/μL, $p = 0.0069$), than untreated samples (65.28 ng/μL) (**Supplementary Tables S2, S3**). Both procedures yielded

relatively pure DNA with median absorbance ratios of 1.81 (IQR = 1.67–1.90) and 1.88 (IQR = 1.8–2.02), respectively (**Supplementary Figures S1A,B**).

In order to evaluate the level of bacterial DNA within total DNA, qPCR was performed. Amplification on undiluted DNA extracts failed to amplify in all the cases, while amplification of DNA diluted 1:100 and 1:1000, revealed that the average bacterial DNA concentration was 21.35 ng/μL in untreated samples and 28.53 ng/μL in the MolYsis treated samples (**Supplementary Figures S1C,D**). All DNA samples contained at least 1 pg bacterial DNA/μL meeting the requirements needed to be considered eligible for SMg. DNA from negative controls contained 0.79% (3.6 pg/μL) and 0.04% (0.2 pg/μL) bacterial DNA without and with MolYsis treatment, respectively. The percentage of bacterial DNA was significantly different between the pre-treated and the untreated samples ($p = 0.0207$) (**Supplementary Table S4**). In conclusion, these studies reveal that high quality microbial gDNA was obtained from PJT on BCB samples pre-treated and untreated with MolYsis5.

SMg Pilot Study – DNA Sample Preparation Method for Further Assays

Two DNA samples (from the same BCB inoculated with a PJT clinical sample) extracted using each of the two DNA sample preparation methods [pre-treated (S1b) and untreated (S1c)], were sequenced using illumina MiSeq sequencing technology. In total, 14,785,194 and 14,078,494 raw reads were obtained from the DNA pre-treated and untreated with MolYsis5, respectively. The two samples were analyzed for the presence of contaminant DNA from human and horse sources in order to determine the proportion of contaminant reads. After preprocessing raw reads, we mapped the remaining reads against a set of human and horse reference genomes. We detected a 16-fold increase in the number of reads mapping to horse and a 19-fold increase in the number of reads mapping to human in the sample untreated with MolYsis5 (**Table 2** and **Figure 3**).

One of the key issues in clinical metagenomics is to remove the host DNA without a substantial loss of bacterial DNA. Samples treated with MolYsis5 before DNA extraction with QIAamp BiOstic seemed to be a suitable DNA preparation method to continue with SMg from BCB.

Shotgun-Metagenomics

Shotgun-metagenomics sequencing of DNA from clinical samples resulted in a mean number of 10,277,311 reads per sample (range 7,236,776–16,172,074). Sequencing of DNA from spiked samples produced a mean number of 9,831,669 (range 6,857,300–11,884,076) reads, and 11,192,852 total reads were obtained from sequencing the negative control (DNA without MolYsis5 treatment) (**Supplementary Table S5**). An additional negative control pre-treated with MolYsis5 was sent for sequencing, but no good library preparation could be obtained due to low DNA yield (results not shown). Thus, only the untreated negative control was used in the analyses.

Extracting Sufficient Bacterial DNA While Removing DNA From Other Sources

In this study, raw reads were screened for the presence of reads from the host, from the horse blood (added to the bottles to enrich the BCB media), and from PhiX phage used as control during the sequencing process.

Overall, DNA from clinical samples presented a mean proportion of reads classified as human (0.07%), horse (0.61%), and PhiX (0.08%) lower than one percent. Similar results were observed when analyzing DNA from the spiked samples, where the mean proportion of reads mapping to the reference genomes of human were 0.26%, horse 0.56%, and to the PhiX phage 0.10%. For both cases, more than 99% of the reads did not map to any of the reference genomes used during the alignment (**Figure 3**). A different pattern was observed when analyzing the negative control, the mean proportion of reads mapped to human (1.4%) and horse (87.6%), while for the PhiX phage, no difference was observed (0.01%). Two percent of the reads did not map to any of the reference genomes (horse, human and PhiX) and were then used for taxonomical analyses (**Figure 3** and **Supplementary Table S5**). Overall, samples untreated with MolYsis5 (NC and S1C) presented more reads mapping to the horse and human genomes. In addition, the proportion of reads mapping to the horse genome is higher than the human genome in all the samples (**Figure 3**).

After preprocessing the raw data, we obtained a mean number of 9,510,050 (92.86%) reads from the clinical samples, and 7,936,436 (86.80%) from the spiked samples, while from the negative control only 187,094 reads (1.67%) remained for taxonomical analyses (**Supplementary Table S5**).

Kraken was used to assign a taxonomic label to each read for estimating the relative proportion of species in the samples and to determine the relative amount of sequences that came from the known bacteria identified in the sample. Taxonomic classification by Kraken when analyzing data from clinical samples identified a median of 96.08% (IQR, 93.85–97.07%; range 85.7–98.6%) bacterial reads. Spiked samples presented a median of 98.7% (IQR, 96.90–98.70) bacterial reads, while for the negative control, the proportion of bacterial reads was 30% (28,058 reads) (**Supplementary Table S6**). In addition, when comparing the pre-treated (sample S1b) and the untreated sample (sample S1c) more reads were classified as bacterial reads (95.74%) for the pre-treated sample than for the untreated sample (86.89%). Overall, two percent more bacterial reads were obtained from the spiked samples than in the clinical samples. Additionally, treatment of the sample with MolYsis5 prior to DNA extraction resulted in a higher proportion of bacterial reads when compared with an untreated sample.

When estimating the number of reads classified at the genus and species level (inferring the abundance of the number of individuals from each taxon by correction for genome length into abundance estimates by Bracken), it was on average higher than 97.7% for both taxonomic levels (genus and species) and for both sample types (DNA from spiked and clinical specimens). In the negative control, 30.1% of the reads were classified to

TABLE 2 | Reads affiliated to human and horse genomes when sample was treated (1b) and untreated with MolYsis5 (1c).

Sample code	DNA preparation method	Reads analyzed	Human reads (%)	Horse reads (%)
Treated (1b)	MolYsis5 + BiOstic	13,417,582	14,521 (0.01)	135,751 (0.9)
Untreated (1c)	BiOstic	12,698,920	249,520 (0.1)	2,365,316 (16.7)

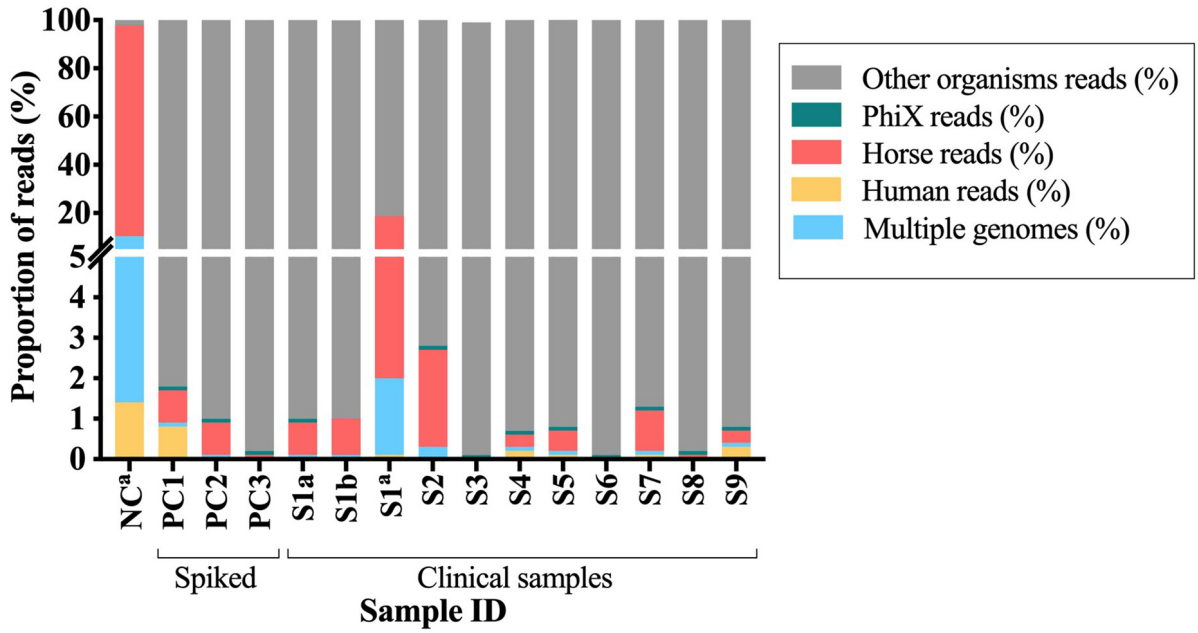


FIGURE 3 | Proportion of the contaminant reads that mapped to PhiX, horse and human reference genomes or to multiple genomes. S1–S9, clinical samples; PC1-3, spiked samples (positive controls); NC, negative control. [a]Samples untreated with Molysis5.

the genus and species level while the rest remained unclassified (**Supplementary Table S7**).

Identification of Bacterial Species Reported as Common Cause of PJI by SMg

Nine of the PJT clinical samples in this study had previously been analyzed using a conventional and a BCB method (Sanabria et al., 2019). Bacterial identification for both methods was done using MALDI-TOF. According to the results obtained from the BCB method, *S. aureus* was identified in sample 1 (S1b), 2 (S2), 3 (S3), 4 (S4), and 7 (S7). *Enterococcus faecalis* was identified in sample 6 (S6), while in sample 8 (S8) and 9 (S9) *Staphylococcus epidermidis* was found (**Table 3**).

All bacterial species identified from the BCB culture method were detected by Kraken and MetaPhlAn, and they were found to be the most abundant species present in the sample (**Supplementary Figure S2**). The mean rate of the most abundant bacteria in the clinical samples were 97.9 and 99.2% for the spiked samples when using Kraken/Bracken. Similar results were found using MetaPhlAn2, with 99.2 and 96.9% mean rates for clinical and spiked samples, respectively.

The conventional culture results were all negative in samples S7–S9, while the BCB method and the SMg were consistent concerning bacterial species found (**Table 3**). Bacteria detected by the conventional methods and by SMg (Kraken/Bracken and MetaPhlan2) are listed in **Table 3**. The taxonomic classification results for each sample type excluding the negative control as estimated by Kraken/Braken are presented in **Figure 4**. Similarly, a heat map representing the relative abundance at the species level as estimated by MetaPhlAn2 can be found in **Supplementary Figure S3**.

Additionally, we tested the influence of downsizing the number of reads (by subsampling reads) in the estimation of abundance of species present in both the clinical and spiked samples. We observed that the number of species detected is higher at increasing sequencing depths and that the taxonomical distribution did not change. The potential pathogen species found in culture could still be detected even if a smaller proportion of reads were subsampled (10%). A rarefaction curve was used as a qualitative method to estimate the species abundance as a function of sequencing depth (**Supplementary Figure S4**). Rarefaction curve reaches plateau just for a few samples suggesting that saturation in sequencing was not achieved and deeper sequencing is required to detect all the species present. However, estimating the diversity is not the end point in PJI diagnoses.

TABLE 3 | Microorganisms identified in the samples using laboratory and SMg methods.

Sample		Microorganism identified			
		Laboratory[a] (MALDI-TOF)		SMg	
No	Code	BCB	Conventional	Kraken/Bracken	MetaPhlAn2
1	S1b	*S. aureus*	*S. aureus*	*S. aureus* (99.6%)	*S. aureus* (100%)
2	S2	*S. aureus*	*S. aureus*	*S. aureus* (99.5%)	*S. aureus* (100%)
3	S3	*S. aureus*	*S. aureus*	*S. aureus* (99.9%)	*S. aureus* (100%)
4	S4	*S. aureus*	*S. aureus*	*S. aureus* (89.8%)	*S. aureus* (100%)
5	S5	*S. aureus*	*S. aureus*	*S. aureus* (99.8%)	*S. aureus* (100%)
6	S6	*E. faecalis*	*E. faecalis*	*E. faecalis* (99.9%)	*E. faecalis* (99.9%)
7	S7	*S. aureus*	No growth	*S. aureus* (99.9%)	*S. aureus* (100%)
8	S8	*S. epidermidis*	No growth	*S. epidermidis* (95.8%)	*S. epidermidis* (99.9%)
9	S9	*S. epidermidis*	No growth	*S. epidermidis* (97.7%)	*S. epidermidis* (99.9%)
10	PC1	*S. aureus*	NA	*S. aureus* (99.9%)	*S. aureus* (100%)
11	PC2	*E. coli*	NA	*E. coli* (99.7%)	*E. coli* (90.7%)
12	PC3	*S. aureus*	NA	*S. aureus* (97.3%)	*S. aureus* (100%)

[a]*Results from our previous study (Sanabria et al., 2019), BCB method and conventional method. PJT, Periprosthetic joint tissue; BCB, Blood culture bottle; SMg, shotgun-metagenomics; S1–S9, clinical samples; PC1-3, spiked samples (positive controls); NC, negative control; NA, not applicable.*

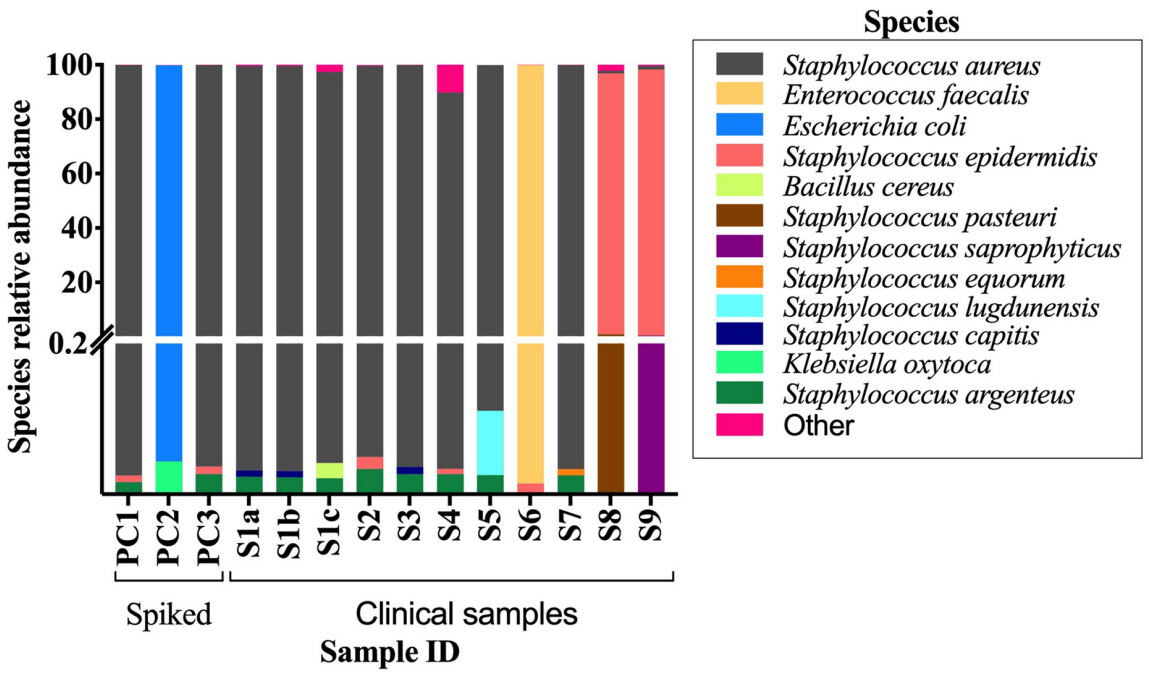


FIGURE 4 | Stacked bar plot displaying relative abundance of bacteria at the species level as estimated by Kraken/Bracken in positive BCBs with PJT. Y-axis is split in two parts, and each part with their own linear scale. From 0 to 0.2 the scale highlights the small values of species relative abundance. S1-S9, clinical samples; PC1-3, spiked samples (positive controls).

Reads belonging to spiked samples were mapped against the reference genomes of the respective strains used for spiking. For *S. aureus* and *E. coli*, more than 98% of the genome was covered with at least 4× depth. We obtained a genome coverage depth of 775X for the sample spiked with *S. aureus* (PC1) and 209X for the sample spiked with *E. coli* (PC2). In the sample spiked with *S. aureus* and *C. acnes*, no reads were observed when mapping against *C. acnes* reference genome; instead 99.98% of the reads mapped to the *S. aureus* reference genome, covering 99.5% of its genome (**Supplementary Table S11**).

Other Bacteria and Possible Contaminants Detected by SMg

In addition to the metagenomic reads belonging to the same bacteria found by the conventional methods, reads belonging

to other bacterial species were also found by Kraken. Their respective abundance was estimated by Bracken in the clinical samples (mean: 1.98%; range 0.07–10.21%), and in the spiked samples (mean: 0.12%; range 0.06–0.23%). For each of the sample types, we provide the respective Sankey flow diagram with the classification results from Kraken (**Supplementary Figure S5**).

Reads belonging to *Staphylococcus argenteus* were found in samples, S1(a, b, c), S3 S4, S5, and S7 always at the same relative abundance (0.02%) and in samples S8 and S9 at a lower abundance (0.000422 and 6.043e-05, respectively). This bacterial species was not found in the negative control (**Figure 5**). The most abundant bacterial species found by Kraken/Bracken in the negative control were *Bacillus cereus* (81.5%), *S. aureus* (10.3%), *Bacillus weihenstephanensis* (1.2%), and *E. coli* (0.88%) (**Figure 5**). MetaPhlAn did not identify *B. cereus* and *B. weihenstephanensis,* but instead identified the yeast *Saccharomyces cerevisiae* as the most abundant microorganism found in the negative control (72.5%). *S. aureus* was also found by MetaPhlAn2 at a relative abundance of 14.7%. *E. coli* was not detected but instead we found *Escherichia* unclassified (4.1%) which means "an unknown species in the genus *Escherichia*."

When analyzing the clinical and spiked samples, we found *B. cereus* in all the samples in a very low abundance (mean, 0.0016%). It was 12-fold more abundant (0.02%) in the sample untreated with MolYsis5 (sample S1c), supporting that it is a contaminant in this study. *S. aureus* was found by Kraken/Bracken in the negative control (10.3%), as well as in samples S6, S8, S9 and in the sample spiked with *E. coli* with a very low occurrence (mean, 0.58%; range, 0.03–1.2%).

DISCUSSION

Here we demonstrate proof of concept that it is possible to perform SMg directly on PJT from BCBs with the potential of pathogen identification in PJI diagnosis. Culturing of PJT on BCB was the starting point for our study, and we validated our findings by comparison with conventional culturing methods, where we already knew the outcome, i.e., complete microbiological data from culturing were available for all samples.

To the best of our knowledge, this is the first study applying SMg to Bact/Alert® BCBs (bioMérieux) inoculated with PJT. Sample type greatly influences the composition of sequencing reads, and due to the complexity of both the BCBs and the PJT specimens, sample preparation and bacterial enrichment methods need to be specifically adapted.

It is well-known that the success of metagenomics is highly dependent on the quality and quantity of DNA extracted from a given specimen (Schlaberg et al., 2017). Therefore, sample preparation methods need to be developed and improved to effectively reduce the amount of host nucleic acids, lyse host cells to release intracellular organisms and lyse different types of microbial cells to release nucleic acids, without affecting the quantity of targeted nucleic acid in the sample (Forbes et al., 2018; Vijayvargiya et al., 2019). Here we demonstrated that the use of MolYsis5 kit together with BiOstic kit is an effective sample preparation method for performing SMg directly from BCBs inoculated with PJT. Our DNA preparation method resulted in high quality microbial DNA from all samples, both with and without human DNA depletion, allowing downstream SMg.

Thus, we conclude that preanalytical reduction of the human DNA load improved the output of targeted sequence reads. This is important since DNA samples overwhelmed with human DNA sequences affect the sensitivity for the detection of bacteria that occur in low abundance in clinical specimens. In addition, we screened the samples and removed reads belonging to the horse genome, since horse blood was used as enrichment supplement for the BCB, which may affect the sensitivity for bacterial detection as such. Moreover, this makes the method more expensive, creates the need of subtracting the human and horse sequences during the bioinformatic analysis, which delays the analyses, and it requires a significant computational power (Hasan et al., 2016). Overall, our method consistently generated high DNA yield quantity/quality, removed efficiently human/horse DNA (<1%) and there was no inhibition observed during the SMg library preparation, allowing further sequencing process.

In order to select samples suitable for sequencing, one of the main criteria that was considered, was the bacterial DNA concentration measured by qPCR. We observed that samples had to be diluted to get amplification signals. qPCR results could not be considered as selection criteria. In addition, the concentration of bacterial DNA from qPCR did not correlate with the bacterial rate obtained from the classification of the sequencing reads (median of 96.08%). This result was not surprising since there are several factors that can affect the qPCR amplification, among them: the presence of amplification inhibitors commonly found in BCBs formulations (e.g., the anticoagulant, sodium polyanetholesulfonate SPS) (Mäki, 2015), unspecific primer binding to host DNA which decreases the sensitivity and specificity, and the less likely in this case, due to an overload of eukaryotic DNA (since low horse and human DNA were found in the sequencing reads). However, the DNA preparation tool MolYsis which eliminates the human background, also removes other PCR-inhibiting substances contained in BCBs, like SPS (Gebert et al., 2008).

Methods for the preparation of bacterial DNA from BCBs (BD BACTEC, Becton Dickinson Sparks, MD) using SMg have been reported earlier but for bloodstream infection diagnosis (Anson et al., 2018). However, the methods included a pre-treatment by differential centrifugation to remove human cells and DNA purification with SPRI beads prior to sequencing. In their study, the use of BiOstic kit provided a higher DNA yield than Molysis Plus (430× greater), and the mean DNA yield obtained was lower (*E. coli*, 39.3 ng/μL and *S. aureus,* 11.5 ng/μL) than the DNA yield obtained in our study (Mean DNA concentration: 84.23 ng/μL for MolYsis5 treated samples and 65.28 ng/μL for untreated samples). These differences in result may be due to the type of clinical specimen inoculated into the BCBs, the effectiveness of the pre-processing and purification step, in addition to the fact that the BCBs belong to different brands.

Our results showed that MolYsis5 kit together with BiOstic kit was an effective DNA sample preparation method for the

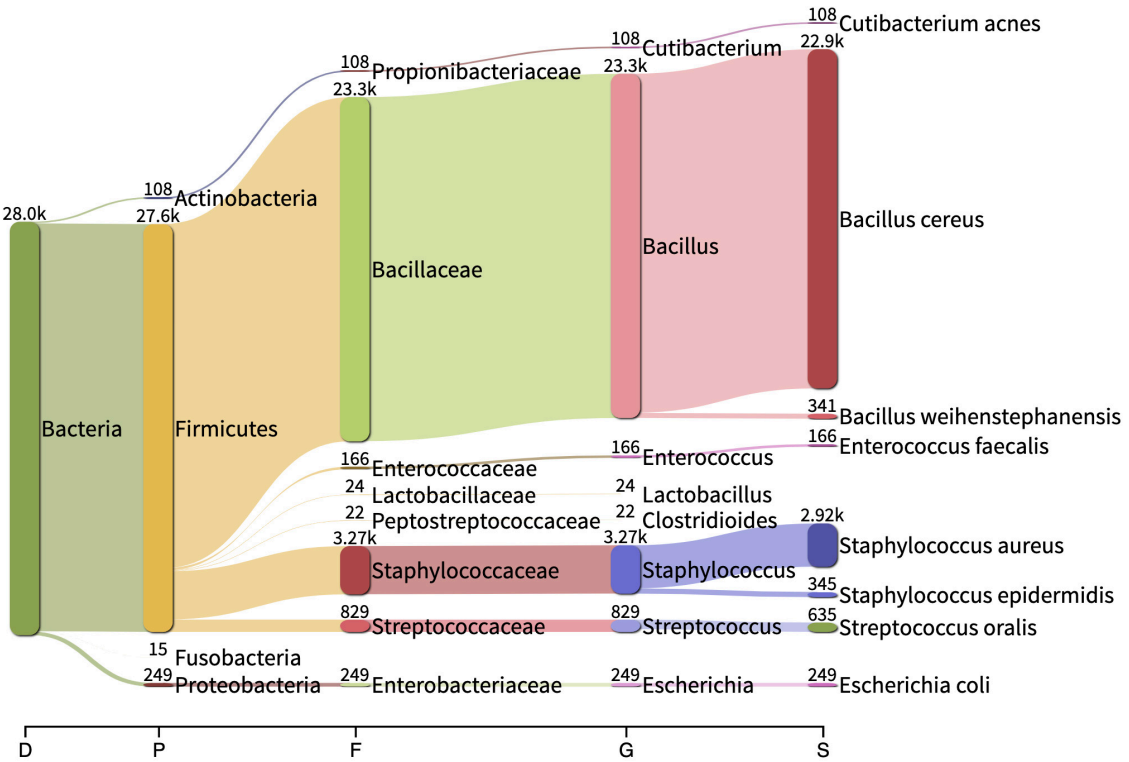


FIGURE 5 | Sankey visualization of the absolute read counts as estimated by Kraken/Bracken from the negative control sample. D, Domain; P, Phylum; F, family; G, genus; D, domain.

detection of the PJI related species tested here. Many bacterial and fungal species have been reported as being detectable after the use of MolYsis (Horz et al., 2010; Thoendel et al., 2016; Schmidt et al., 2017; Krohn et al., 2018). However, we are aware about potential limitations that should be taken into consideration when using MolYsis in the clinical diagnostics. Sample pre-treatment with MolYsis involves lysis of human cells prior to degradation of bacterial cell walls, allowing digestion of human DNA by DNAse treatment (Thoendel et al., 2016). The use of MolYsis provides a solution for the removal of host DNA and enrichment of intact microorganisms (Votintseva et al., 2017). However, free floating DNA is removed. This raises a question about which bacterial DNA is removed during the MolYsis treatment. Bacteria with weak or absent cell walls will be removed by the use of Molysis prior to DNA extraction. Among them, *Mycoplasma, Ureaplasma*, or *Chlamydia,* although unusual, are among the many organisms capable of causing PJIs (Geipel, 2009; Rieber et al., 2019). In addition, bacteria previously exposed to cell wall-targeting antibiotics will also be lysed and removed by the Molysis (Horz et al., 2010). This in turn raises some concerns about the proposed usefulness of MolYsis for the diagnosis of PJI when these bacteria are causing the infection. In these cases, sample pre-treatment with MolYsis should be avoided. DNA extraction with BiOstic without MolYsis pre-treatment can be used instead since good results were also obtained when using this approach in our study (sample S1c, untreated with MolYsis).

SMg for diagnosing PJI directly from tissue specimens has been tested (Ruppé et al., 2017). However, from a huge cohort of samples collected (n = 179), only few (n = 24) could be sequenced due to insufficient amounts of bacterial DNA recovered from the samples. SMg directly from sonication fluid (from orthopedic devices), has been tested as well, and this is the specimen related to PJI most extensively studied in metagenomics approaches (Thoendel et al., 2016, 2017, 2018; Street et al., 2017; Sanderson et al., 2018). Despite all these promising efforts on metagenomics for diagnosing PJI, direct sequencing of nucleic acids obtained from biological samples results in a high background of DNA, mainly derived from the host, hindering the detection of pathogens causing PJI. Thus, all these studies support clearly that the main challenge has been recovering enough bacterial DNA. Our approach was therefore to test the use of SMg directly on BCBs inoculated with PJT to try to solve the limitations observed when using SMg directly on PJI specimens.

In our study, predominant bacterial species in PJT from BCBs determined by SMg, were 100% concordant with the results obtained from the BCB culture method. Results were consistent with respect to both the genus and species levels. We were able to detect *S. aureus, S. epidermidis, E. faecalis, and E. coli* in the samples, indicating the potential of the method for detection of species commonly related to be the cause of PJI. The predominant species (mean rate: 97.9%), reads belonging to other bacterial

species were also found by the taxonomical classifier in a very low proportion (mean rate: 1.98%).

Apart from the development and/or improvement in the DNA sample preparation methods, one of the greatest challenges in the use of SMg for identification of pathogens is the type of controls (Couto et al., 2018). Positive controls should represent the range of organisms that can be encountered in the clinical specimen (Greninger and Naccache, 2019). In this study, three spiked samples were included as positive controls. Two of them spiked with one bacterial species (*S. aureus* or *E. coli*) and one with two bacterial species (*S. aureus* and *C. acnes*). High breadth and depth coverage were obtained when mapping the SMg reads to *S. aureus* or *E. coli* reference genomes, respectively. However, *C. acnes* was not found in the SMg taxonomical analyses of the spiked control sample. It could be several reasons for this, but we believe that it was due to mistakes in the experimental design. *S. aureus* and *C. acnes* were spiked into the BCB at the same time and we know from our previous study (Sanabria et al., 2019) that *C. acnes* grows slower (mean time to detection: 8.7 days) than *S. aureus* in the Bact/Alert® BCBs. The positive control was incubated in an aerobic BCB until positive, i.e., in this case 10.3 h after incubation. We believe that the absence of reads belonging to *C. acnes* is due to the fact that the bacterium did not have enough time to grow. In addition, *S. aureus* might be a strong competitor. Another possible explanation could be incomplete lysis of *C. acnes* during the DNA sample preparation and the lack of sensitivity of the SMg to detect the anaerobic *C. acnes*.

For the case of the negative control, we used a BCB medium enriched with horse blood and inoculated with sterilized tissue from a donor without suspicion of infection. We consider that this negative control adequately reflects the contaminant or background microorganisms originating from tissue specimens, BCB media, horse blood, reagents, the environment and other sources as from other samples and sequencing runs (Eisenhofer et al., 2019; Huang et al., 2019). Results showed that from 93,502 reads, only 28,194 could be classified (30%).

In the negative control, we found taxa reported as common contaminants (DNA extraction blank controls and no-template controls) (Eisenhofer et al., 2019), e.g., *Bacillus*, *Staphylococcus*, *Enterococcus* and *Streptococcus*. We also found some reads classified as *B. cereus* in several of the clinical samples (in a very low abundance, mean: 0.0016%) and this bacterium was found 12-fold more abundant (0.02%) in the sample not treated with Molysis5 (sample S1c). *Bacillus* spp. is often considered a contaminant when it is isolated from BCBs and in negative controls (Doern et al., 2020). However, the significance of this in SMg on samples from BCBs is unknown.

We also found some reads assigned to *S. aureus* (10.3%) and *E. coli* (0.88%) in the negative control. Since these bacteria are among the most common causes of PJI (Tande and Patel, 2014; Izakovicova et al., 2019), these reads were evaluated by aligning the reads against the reference genomes of the strain with the highest assignment number of reads. When visualizing the mapping results, reads mapped with genetic areas belonging to coding sequences annotated as RNAs with a very low coverage depth, and they were not distributed all over the genomes. These may be reads originating from laboratory, *in silico* or kit contaminants. Contamination is one of the main concerns in PJI diagnostics and even more in metagenomic sequencing.

Reports have demonstrated that even the commercial kits for DNA extraction and library preparation are potential for contamination leading to misinterpretation of sequencing data from clinical specimens (Salter et al., 2014; Eisenhofer et al., 2019). It is therefore recommended to include and sequence negative controls when performing SMg studies.

The spectrum of organisms defined as reportable by SMg assays should be defined, and organisms determined to be background contaminants or clinically insignificant should be described (Gu et al., 2019). Defining a contaminant is not clear for blood cultures in the laboratory and present a challenge for SMg (Greninger and Naccache, 2019). Many factors should be considered when interpreting the results from SMg especially because the results are highly dependent on the database used for the analysis, which could be incomplete for rare pathogens or biased toward certain organisms, in addition to the fact that contamination with normal flora and reagents are a common occurrences that can limit specificity (Gu et al., 2019). Consequently, it is very important to be careful when analyzing the clinical significance of the results.

The most common *in silico* decontamination method in practice is the removal of sequences below a determined detection threshold (Davis et al., 2018). Usually, software used for taxonomical classification such as Kraken can predict a lot of species. Although we have limited clinical data to distinguish between true PJI and contamination, we tested the effect of several thresholds on the estimation of species abundance in the samples. We observed that the number of species detected in the samples are highly dependent on the threshold value used during the analyses (**Supplementary Figure S6** and **Supplementary Table S10**). The results presented here prove that for species detection, thresholds may often lead to different inferences while interpreting the diagnostic results. Consequently, when using SMg, thresholds need to be validated in each specific case for an accurate interpretation of the results (Schlaberg et al., 2017). In our case, unfortunately we cannot allow to give an exact cut-off value, due to limited access to clinical data of our samples. However, others have set optimal thresholds for differentiating low-level contaminations from true PJI when using SMg on sonication fluid (Street et al., 2017; Ivy et al., 2018).

There is no standard method for interpreting metagenomic sequencing results. Contaminant DNA in SMg is a challenge for clinical interpretation of metagenomics data (Peel et al., 2016; Thoendel et al., 2016; Ruppé et al., 2017; Street et al., 2017; Thoendel et al., 2017; Ivy et al., 2018; Simner et al., 2018; Thoendel et al., 2018). We cannot exclude the possibility of contamination in our study. As in most studies, all our samples contained read identifications for microorganisms other than known or suspected pathogens. In order to determine if the bacterial species found are infection inducing pathogens, background contaminants or noise, we observed that there are several aspects that can help to differentiate them, among them: (i) The proportion of reads assigned to the species present. The possibility of obtaining quantified abundances of microorganisms is important for distinguishing causative

pathogens (Greninger and Naccache, 2019). In our case, we considered that a high proportion of reads belonging to the most abundant bacteria present in the samples (97.9% mean rate) could be an indicator of the bacterium causing the infection. (ii) Genome coverage and the proportion of the genome covered, higher depth and breadth coverage expected for the pathogen species with respect to other species detected by the taxonomical classifier. For pathogen detection it is even more important because sequencing depth also affects analytic sensitivity (Schlaberg et al., 2017; Couto et al., 2018). (iii) Comparison with respect to the species and to the proportions of reads found in the negative control and in the spiked samples, and (iv) To set an appropriate threshold, for pathogen detection. Regarding this, it is important to limit the number of species identified, for minimizing false-positive results, increasing the detection rate of potential true pathogens and reducing the misclassification of other species related signals as potential pathogens (Couto et al., 2018). It is important to consider that setting up cut-off values for pathogen detection may result in decreased sensitivity. Therefore, it is better to rely on the relative abundance of bacterial species in addition to the genome coverage and the proportion of the genome covered.

It is important to predict the level at which samples should be sequenced to prevent excessive sequencing and to answer our biological question (Sims et al., 2014). The relatively high cost for metagenomic sequencing is a major limitation for application in the clinical setting (Ruppé et al., 2017). Significant reduction in the cost of metagenomic sequencing is required for moving up in the diagnostic pipeline (Greninger, 2018). Multiplexing samples offers the possibility of decreasing the costs by decreasing the number of reads per sample. The question is how many reads are needed to answer a particular question (Mulcahy-O'Grady and Workentine, 2016). In this study, we wanted to assess the potential of SMg for the detection of bacterial species known as common causes of PJI. We analyzed the effect of reduction in sequencing depth (expressed as proportion of reads) for the detection of potential PJI pathogens and, we observed that even while using the minimal proportion of reads subsampled (10%) we were able to detect *S. aureus* (**Supplementary Figure S7**). This result suggests that presumably less sequencing depth is needed for detection of common PJI pathogens, more samples can be multiplexed in a sequencing lane, accomplishing a lower cost per sample. However, it is important to be aware that the impact of a lower sequencing depth to provide additional information beyond pathogen detection was not considered in this study. Nevertheless, the results obtained in this study open the possibility for studying antibiotic resistance determinants and virulence genes at further stages.

Our study has the following technical limitations: **(1)** Low total number of clinical samples analyzed trough shotgun-metagenomics (n = 9) and just one sample per patient. **(2)** Limited clinical data about the patients making it difficult to define a sample selection criterion to distinguish between true PJI and contamination; **(3)** Negative samples by conventional and BCB culture methods were not included. **(4)** Only aerobic BCBs were included. **(5)** Absence of polymicrobial samples (all samples tested were monomicrobial by culture). **(6)** In comparison with the conventional and the BCB culture methods, the application of SMg is limited by the expensive equipment and operational costs.

The use of clinical metagenomics approaches will increase during the next years in research and in medical microbiology laboratories (Deurenberg et al., 2017). The application in clinical microbiology is still in its infancy, which encourage further research on alternative and complementary tools for PJI diagnosis. There are ongoing discussions about the obstacles associated with the adoption of metagenomics in diagnostics and their clinical utility (Greninger, 2018; Chiu and Miller, 2019; Han et al., 2019). However, we do not believe that SMg can replace conventional culturing, but it can be a potential diagnostic tool to support conventional culture in cases when PJI diagnosis is challenging, e.g., with fastidious organisms, discrepancies between conventional methods, or in culture negative cases.

In conclusion, our DNA preparation method resulted in high quality microbial DNA from all PJT samples, both with and without human DNA depletion, allowing downstream SMg. By SMg we were able to identify relevant PJI pathogens, and all bacteria identified by culture were also identified through SMg. A high enough sequencing depth was obtained indicating that it is possible to multiplex samples reducing costs considerably. We achieved a high sequencing quality, low human DNA content, high number of reads and complete genome coverage of sufficient depth that technically can be used for AMR prediction, virulence gene detection and bacterial typing.

We consider this an essential step in further studies for solving the challenges presented in PJI diagnosis, e.g., when bacteria are not detected by the laboratory methods but there is still clinical signs of the presence of infection (Peel et al., 2016). It is still possible to extract DNA from a negative BCB and analyze if pathogenic bacteria are present. In fact, the results of SMg can also be valuable even when concordant with laboratory results, not only providing a guarantee that the laboratory diagnosis is correct, but also allowing extra information, e.g., detecting coinfections and/or predicting antimicrobial susceptibility. Our results can be useful for further validation and standardization for the use of SMg on BCBs inoculated with clinical samples for routine diagnostics of pathogens. It is still a long way until SMg can be used in the clinical microbiology laboratory, but this SMg approach presents an alternative tool in PJI diagnosis, complementing the currently available tools.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the European Nucleotide Archive repository (ENA) (www.ebi.ac.uk/ena), with the study accession number PRJEB36855.

ETHICS STATEMENT

This study was performed in accordance with the ethical guidelines established by the Arctic University of Norway (UiT). The project has been evaluated by the Regional Committee for Medical and Health Research Ethics, Rec

North, Norway (document no. 2016/1247/REK nord), concluding that ethical approval was not required. There were no ethical issues to consider due to use of anonymous clinical samples and development of methodological procedures. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

AS performed the experiments and the bioinformatic analysis, took part in the study design, and wrote the first version of the manuscript. AS, MJ, JS, GS, and A-MH assisted in the scientific and technical design of the experiments. EH gave advice on the bioinformatic analyzes and revised the manuscript. AS, A-MH, and GS analyzed and interpreted the data results. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from joint "Miljøstøtte" financed by Strategisk-HN05–14 (Helse Nord RFH) and Faculty of Health Sciences A20389 (2014–2017), and the National Graduate School in Infection Biology and Antimicrobials (Grant No. 249062). The publication charges for this article have been funded by a grant from the publication fund of UiT – The Arctic University of Norway. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

We thank the Department of Microbiology and Infection Control, University Hospital of North Norway, Tromsø, especially Silje Nysted Hagen and Maria Frost for coordinating the samples collection at the preanalytical laboratory. We thank Jessin Janice James Peter for her support during the bioinformatic analyses.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.01687/full#supplementary-material

REFERENCES

Anson, L. W., Chau, K., Sanderson, N., Hoosdally, S., Bradley, P., Iqbal, Z., et al. (2018). DNA extraction from primary liquid blood cultures for bloodstream infection diagnosis using whole genome sequencing. *J. Med. Microbiol.* 67, 347–357. doi: 10.1099/jmm.0.000664

Bachmann, N. L., Rockett, R. J., Timms, V. J., and Sintchenko, V. (2018). Advances in clinical sample preparation for identification and characterization of bacterial pathogens using metagenomics. *Front. Public Heal.* 6:363. doi: 10.3389/fpubh.2018.00363

Bogaert, D., Keijser, B., Huse, S., Rossen, J., Veenhoven, R., van Gils, E., et al. (2011). Variability and diversity of nasopharyngeal microbiota in children: a metagenomic analysis. *PLoS One* 6:e17035. doi: 10.1371/journal.pone.0017035

Breitwieser, F. P., and Salzberg, S. L. (2019). Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification. *Bioinformatics* 36, 1303-1304. doi: 10.1093/bioinformatics/btz715

Chiu, C. Y., and Miller, S. A. (2019). Clinical metagenomics. *Nat. Rev. Genet.* 20, 341–355. doi: 10.1038/s41576-019-0113-7

Corvec, S., Portillo, M. E., Pasticci, B. M., Borens, O., and Trampuz, A. (2012). Epidemiology and new developments in the diagnosis of prosthetic joint infection. *Int. J. Artif. Organs* 35, 923–934. doi: 10.5301/ijao.5000168

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

Cremers, A. J., Zomer, A. L., Gritzfeld, J. F., Ferwerda, G., van Hijum, S. A., Ferreira, D. M., et al. (2014). The adult nasopharyngeal microbiome as a determinant of pneumococcal acquisition. *Microbiome* 2:44. doi: 10.1186/2049-2618-2-44

Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., and Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 6:226. doi: 10.1186/s40168-018-0605-2

Deurenberg, R. H., Bathoorn, E., Chlebowicz, M. A., Couto, N., Ferdous, M., García-Cobos, S., et al. (2017). Reprint of "Application of next generation sequencing in clinical microbiology and infection prevention.". *J. Biotechnol.* 250, 2–10. doi: 10.1016/J.JBIOTEC.2017.03.035

Doern, G. V., Carroll, K. C., Diekema, D. J., Garey, K. W., Rupp, M. E., Weinstein, M. P., et al. (2020). A comprehensive update on the problem of blood culture contamination and a discussion of methods for addressing the problem. *Clin. Microbiol. Rev.* 33:e00009-19. doi: 10.1128/CMR.00009-19

Dudareva, M., Barrett, L., Figtree, M., Scarborough, M., Watanabe, M., Newnham, R., et al. (2018). Sonication versus tissue sampling for diagnosis of prosthetic joint and other orthopedic device-related infections. *J. Clin. Microbiol.* 56:e00688-18. doi: 10.1128/JCM.00688-18

Eisenhofer, R., Minich, J. J., Marotz, C., Cooper, A., Knight, R., and Weyrich, L. S. (2019). Contamination in low microbial biomass microbiome studies: issues and recommendations. *Trends Microbiol.* 27, 105–117. doi: 10.1016/j.tim.2018.11.003

Font-Vizcarra, L., García, S., Martínez-Pastor, J. C., Sierra, J. M., and Soriano, A. (2010). Blood culture flasks for culturing synovial fluid in prosthetic joint infections. *Clin. Orthop. Relat. Res.* 468, 2238–2243. doi: 10.1007/s11999-010-1254-3

Forbes, J. D., Knox, N. C., Peterson, C. L., and Reimer, A. R. (2018). Highlighting clinical metagenomics for enhanced diagnostic decision-making: a step towards wider implementation. *Comput. Struct. Biotechnol. J.* 16, 108–120. doi: 10.1016/j.csbj.2018.02.006

Gebert, S., Siegel, D., and Wellinghausen, N. (2008). Rapid detection of pathogens in blood culture bottles by real-time PCR in conjunction with the pre-analytic tool MolYsis. *J. Infect.* 57, 307–316. doi: 10.1016/J.JINF.2008.07.013

Geipel, U. (2009). Pathogenic organisms in hip joint infections. *Int. J. Med. Sci.* 6, 234–240. doi: 10.7150/ijms.6.234

Greenwood-Quaintance, K. E., Uhl, J. R., Hanssen, A. D., Sampath, R., Mandrekar, J. N., and Patel, R. (2014). Diagnosis of prosthetic joint infection by use of PCR-electrospray ionization mass spectrometry. *J. Clin. Microbiol.* 52, 642–649. doi: 10.1128/JCM.03217-13

Greninger, A. L. (2018). The challenge of diagnostic metagenomics. *Expert Rev. Mol. Diagn.* 18, 605–615. doi: 10.1080/14737159.2018.1487292

Greninger, A. L., and Naccache, S. N. (2019). Metagenomics to assist in the diagnosis of bloodstream infection. *J. Appl. Lab. Med.* 3, 643–653. doi: 10.1373/jalm.2018.026120

Gu, W., Miller, S., and Chiu, C. Y. (2019). Clinical metagenomic next-generation sequencing for pathogen detection. *Annu. Rev. Pathol. Mech. Dis.* 14, 319–338. doi: 10.1146/annurev-pathmechdis-012418-012751

Han, D., Li, Z., Li, R., Tan, P., Zhang, R., and Li, J. (2019). mNGS in clinical microbiology laboratories: on the road to maturity. *Crit. Rev. Microbiol.* 45, 668–685. doi: 10.1080/1040841X.2019.1681933

Hasan, M. R., Rawat, A., Tang, P., Jithesh, P. V., Thomas, E., Tan, R., et al. (2016). Depletion of human DNA in spiked clinical specimens for improvement of sensitivity of pathogen detection by next-generation sequencing. *J. Clin. Microbiol.* 54, 919–927. doi: 10.1128/JCM.03050-15

Hasman, H., Saputra, D., Sicheritz-Ponten, T., Lund, O., Svendsen, C. A., Frimodt-Møller, N., et al. (2014). Rapid whole-genome sequencing for detection and characterization of microorganisms directly from clinical samples. *J. Clin. Microbiol.* 52, 139–146. doi: 10.1128/JCM.02452-13

Horz, H.-P., Scheer, S., Vianna, M. E., and Conrads, G. (2010). New methods for selective isolation of bacterial DNA from human clinical specimens. *Anaerobe* 16, 47–53. doi: 10.1016/J.ANAEROBE.2009.04.009

Huang, Z., Zhang, C., Li, W., Fang, X., Wang, Q., Xing, L., et al. (2019). Metagenomic next-generation sequencing contribution in identifying prosthetic joint infection due to Parvimonas micra: a case report. *J. Bone Jt. Infect.* 4, 50–55. doi: 10.7150/jbji.30615

Hughes, J. G., Vetter, E. A., Patel, R., Schleck, C. D., Harmsen, S., Turgeant, L. T., et al. (2001). Culture with BACTEC Peds Plus/F bottle compared with conventional methods for detection of bacteria in synovial fluid. *J. Clin. Microbiol.* 39, 4468–4471. doi: 10.1128/JCM.39.12.4468-4471.2001

Ivy, M. I., Thoendel, M. J., Jeraldo, P. R., Greenwood-Quaintance, K. E., Hanssen, A. D., Abdel, M. P., et al. (2018). Direct detection and identification of prosthetic joint infection pathogens in synovial fluid by metagenomic shotgun sequencing. *J. Clin. Microbiol.* 27:e00402-18. doi: 10.1128/JCM.00402-18

Izakovicova, P., Borens, O., and Trampuz, A. (2019). Periprosthetic joint infection: current concepts and outlook. *EFORT Open Rev.* 4, 482–494. doi: 10.1302/2058-5241.4.180092

Janz, V., Wassilew, G. I., and Perka, C. (2016). Incubation of sonicate fluid in blood culture bottles leads to an improved and quicker rate of bacterial isolation. *Orthop. Proc.* 98-B:71. doi: 10.1302/1358-992X.98BSUPP_2.ISTA2014-071

Krohn, S., Zeller, K., Böhm, S., Chatzinotas, A., Harms, H., Hartmann, J., et al. (2018). Molecular quantification and differentiation of Candida species in biological specimens of patients with liver cirrhosis. *PLoS One* 13:e0197319. doi: 10.1371/journal.pone.0197319

Langelier, C., Zinter, M. S., Kalantar, K., Yanik, G. A., Christenson, S., O'Donovan, B., et al. (2018). Metagenomic sequencing detects respiratory pathogens in hematopoietic cellular transplant patients. *Am. J. Respir. Crit. Care Med.* 197, 524–528. doi: 10.1164/rccm.201706-1097LE

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Lu, J., Breitwieser, F. P., Thielen, P., and Salzberg, S. L. (2017). Bracken: estimating species abundance in metagenomics data. *PeerJ Comput. Sci.* 3:e104. doi: 10.7717/peerj-cs.104

Mäki, M. (2015). "Bacterial and fungal DNA extraction from positive blood culture bottles: a manual and an automated protocol," in *Sepsis: Diagnostic Methods and Protocols*, ed. N. Mancini (New York, NY: Springer), 57–64. doi: 10.1007/978-1-4939-1776-1_6

Miller, S., Naccache, S. N., Samayoa, E., Messacar, K., Arevalo, S., Federman, S., et al. (2019). Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. *Genome Res.* 29, 831–842. doi: 10.1101/gr.238170.118

Minassian, A. M., Newnham, R., Kalimeris, E., Bejon, P., Atkins, B. L., and Bowler, I. C. J. W. J. W. (2014). Use of an automated blood culture system (BD BACTECTM) for diagnosis of prosthetic joint infections: easy and fast. *BMC Infect. Dis.* 14:233. doi: 10.1186/1471-2334-14-233

Mulcahy-O'Grady, H., and Workentine, M. L. (2016). The challenge and potential of metagenomics in the clinic. *Front. Immunol.* 7:29. doi: 10.3389/fimmu.2016.00029

Nakamura, S., Yang, C.-S., Sakon, N., Ueda, M., Tougan, T., Yamashita, A., et al. (2009). Direct metagenomic detection of viral pathogens in nasal and fecal specimens using an unbiased high-throughput sequencing approach. *PLoS One* 4:e4219. doi: 10.1371/journal.pone.0004219

Nylén, T., Saeedi, B., Borg, C., Ullberg, M., and Özenci, V. (2013). The performance of 4 different supplements and 5 blood culture bottles types in detection of bacteria and *Candida* spp. in simulated sterile body fluid cultures. *Diagn. Microbiol. Infect. Dis.* 77, 1–4. doi: 10.1016/j.diagmicrobio.2013.05.017

Ong, K. L., Kurtz, S. M., Lau, E., Bozic, K. J., Berry, D. J., and Parvizi, J. (2009). Prosthetic joint infection risk after total hip arthroplasty in the Medicare population. *J. Arthroplasty* 24, 105–109. doi: 10.1016/j.arth.2009.04.027

Padmanabhan, R., Mishra, A. K., Raoult, D., and Fournier, P.-E. (2013). Genomics and metagenomics in medical microbiology. *J. Microbiol. Methods* 95, 415–424. doi: 10.1016/j.mimet.2013.10.006

Park, K. H., and Patel, R. (2018). Diagnostic Methods for Prosthetic Joint Infection in Korea. *Infect. Chemother.* 50, 199–209. doi: 10.3947/ic.2018.50.3.199

Parvizi, J., Zmistowski, B., Berbari, E. F., Bauer, T. W., Springer, B. D., Della Valle, C. J., et al. (2011). New definition for periprosthetic joint infection: from the Workgroup of the Musculoskeletal Infection Society. *Clin. Orthop. Relat. Res.* 469, 2992–2994. doi: 10.1007/s11999-011-2102-9

Patel, R. (2015). MALDI-TOF MS for the diagnosis of infectious diseases. *Clin. Chem.* 61, 100–111. doi: 10.1373/clinchem.2014.221770

Peel, T. N., Dylla, B. L., Hughes, J. G., Lynch, D. T., Greenwood-Quaintance, K. E., Cheng, A. C., et al. (2016). Improved diagnosis of prosthetic joint infection by culturing periprosthetic tissue specimens in blood culture bottles. *mBio* 7:e01776. doi: 10.1128/mBio.01776-15

Peel, T. N., Sedarski, J. A., Dylla, B. L., Shannon, S. K., Amirahmadi, F., Hughes, J. G., et al. (2017). Laboratory workflow analysis of culture of periprosthetic tissues in blood culture bottles. *J. Clin. Microbiol.* 55, 2817–2826. doi: 10.1128/JCM.00652-17

Portillo, M. E., Salvadó, M., Trampuz, A., Siverio, A., Alier, A., Sorli, L., et al. (2015). Improved diagnosis of orthopedic implant-associated infection by inoculation of sonication fluid into blood culture bottles. *J. Clin. Microbiol.* 53, 1622–1627. doi: 10.1128/JCM.03683-14

Rieber, H., Frontzek, A., and Fischer, M. (2019). Periprosthetic joint infection associated with Mycoplasma hominis after transurethral instrumentation in an immunocompetent patient. Unusual or underestimated? A case report and review of the literature. *Int. J. Infect. Dis.* 82, 86–88. doi: 10.1016/J.IJID.2019.03.012

Ruppé, E., Lazarevic, V., Girard, M., Mouton, W., Ferry, T., Laurent, F., et al. (2017). Clinical metagenomics of bone and joint infections: a proof of concept study. *Sci. Rep.* 7:7718. doi: 10.1038/s41598-017-07546-5

Rutanga, J. P., Van Puyvelde, S., Heroes, A. S., Muvunyi, C. M., Jacobs, J., and Deborggraeve, S. (2018). 16S metagenomics for diagnosis of bloodstream infections: opportunities and pitfalls. *Expert Rev. Mol. Diagn.* 18, 749-759. doi: 10.1080/14737159.2018.1498786

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., et al. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* 12:87. doi: 10.1186/s12915-014-0087-z

Salzberg, S. L., and Wood, D. E. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 15:R46. doi: 10.1126/science.1093857

Sanabria, A., Røkeberg, M. E. O., Johannessen, M., Sollid, J. E., Simonsen, G. S., and Hanssen, A.-M. (2019). Culturing periprosthetic tissue in BacT/Alert® Virtuo blood culture system leads to improved and faster detection of prosthetic joint infections. *BMC Infect. Dis.* 19:607. doi: 10.1186/s12879-019-4206-x

Sanderson, N. D., Street, T. L., Foster, D., Swann, J., Atkins, B. L., Brent, A. J., et al. (2018). Real-time analysis of nanopore-based metagenomic sequencing from infected orthopaedic devices. *BMC Genomics* 19:714. doi: 10.1186/s12864-018-5094-y

Schafer, P., Fink, B., Sandow, D., Margull, A., Berger, I., and Frommelt, L. (2008). Prolonged bacterial culture to identify late periprosthetic joint infection: a promising strategy. *Clin. Infect. Dis.* 47, 1403–1409. doi: 10.1086/592973

Schlaberg, R., Chiu, C. Y., Miller, S., Procop, G. W., and Weinstock, G. (2017). Validation of metagenomic next-generation sequencing tests for universal pathogen detection. *Arch. Pathol. Lab. Med.* 141, 776–786. doi: 10.5858/arpa.2016-0539-RA

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

Segata, N., Waldron, L., Ballarini, A., Narasimhan, V., Jousson, O., and Huttenhower, C. (2012). Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* 9, 811–814. doi: 10.1038/nmeth.2066

Shen, H., Tang, J., Wang, Q., Jiang, Y., and Zhang, X. (2015). Sonication of explanted prosthesis combined with incubation in BD bactec bottles for pathogen-based diagnosis of prosthetic joint infection. *J. Clin. Microbiol.* 53, 777–781. doi: 10.1128/JCM.02863-14

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

Sims, D., Sudbery, I., Ilott, N. E., Heger, A., and Ponting, C. P. (2014). Sequencing depth and coverage: key considerations in genomic analyses. *Nat. Rev. Genet.* 15, 121–132. doi: 10.1038/nrg3642

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

Tagini, F., and Greub, G. (2017). Bacterial genome sequencing in clinical microbiology: a pathogen-oriented review. *Eur. J. Clin. Microbiol. Infect. Dis.* 36, 2007–2020. doi: 10.1007/s10096-017-3024-6

Tande, A. J., and Patel, R. (2014). *Prosthetic Joint Infection*. Washington, DC: American Society for Microbiology.

Thoendel, M., Jeraldo, P., Greenwood-Quaintance, K. E., Yao, J., Chia, N., Hanssen, A. D., et al. (2017). Impact of contaminating DNA in whole-genome amplification kits used for metagenomic shotgun sequencing for infection diagnosis. *J. Clin. Microbiol.* 55, 1789–1801. doi: 10.1128/JCM.02402-16

Thoendel, M., Jeraldo, P. R., Greenwood-Quaintance, K. E., Yao, J. Z., Chia, N., Hanssen, A. D., et al. (2016). Comparison of microbial DNA enrichment tools for metagenomic whole genome sequencing. *J. Microbiol. Methods* 127, 141–145. doi: 10.1016/j.mimet.2016.05.022

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

Török, M. E., Reuter, S., Bryant, J., Köser, C. U., Stinchcombe, S. V., Nazareth, B., et al. (2013). Rapid whole-genome sequencing for investigation of a suspected tuberculosis outbreak. *J. Clin. Microbiol.* 51, 611–614. doi: 10.1128/JCM.02279-12

Velay, A., Schramm, F., Gaudias, J., Jaulhac, B., and Riegel, P. (2010). Culture with BACTEC Peds Plus bottle compared with conventional media for the detection of bacteria in tissue samples from orthopedic surgery. *Diagn. Microbiol. Infect. Dis.* 68, 83–85. doi: 10.1016/j.diagmicrobio.2010.04.010

Vijayvargiya, P., Jeraldo, P. R., Thoendel, M. J., Greenwood-Quaintance, K. E., Esquer Garrigos, Z., Sohail, M. R., et al. (2019). Application of metagenomic shotgun sequencing to detect vector-borne pathogens in clinical blood samples. *PLoS One* 14:e0222915. doi: 10.1371/journal.pone.0222915

Votintseva, A. A., Bradley, P., Pankhurst, L., del Ojo Elias, C., Loose, M., Nilgiriwala, K., et al. (2017). Same-day diagnostic and surveillance data for tuberculosis via whole-genome sequencing of direct respiratory samples. *J. Clin. Microbiol.* 55, 1285–1298. doi: 10.1128/JCM.02483-16

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

Wingett, S. W., and Andrews, S. (2018). FastQ Screen: a tool for multi-genome mapping and quality control. *F1000Research* 7:1338. doi: 10.12688/f1000research.15931.2

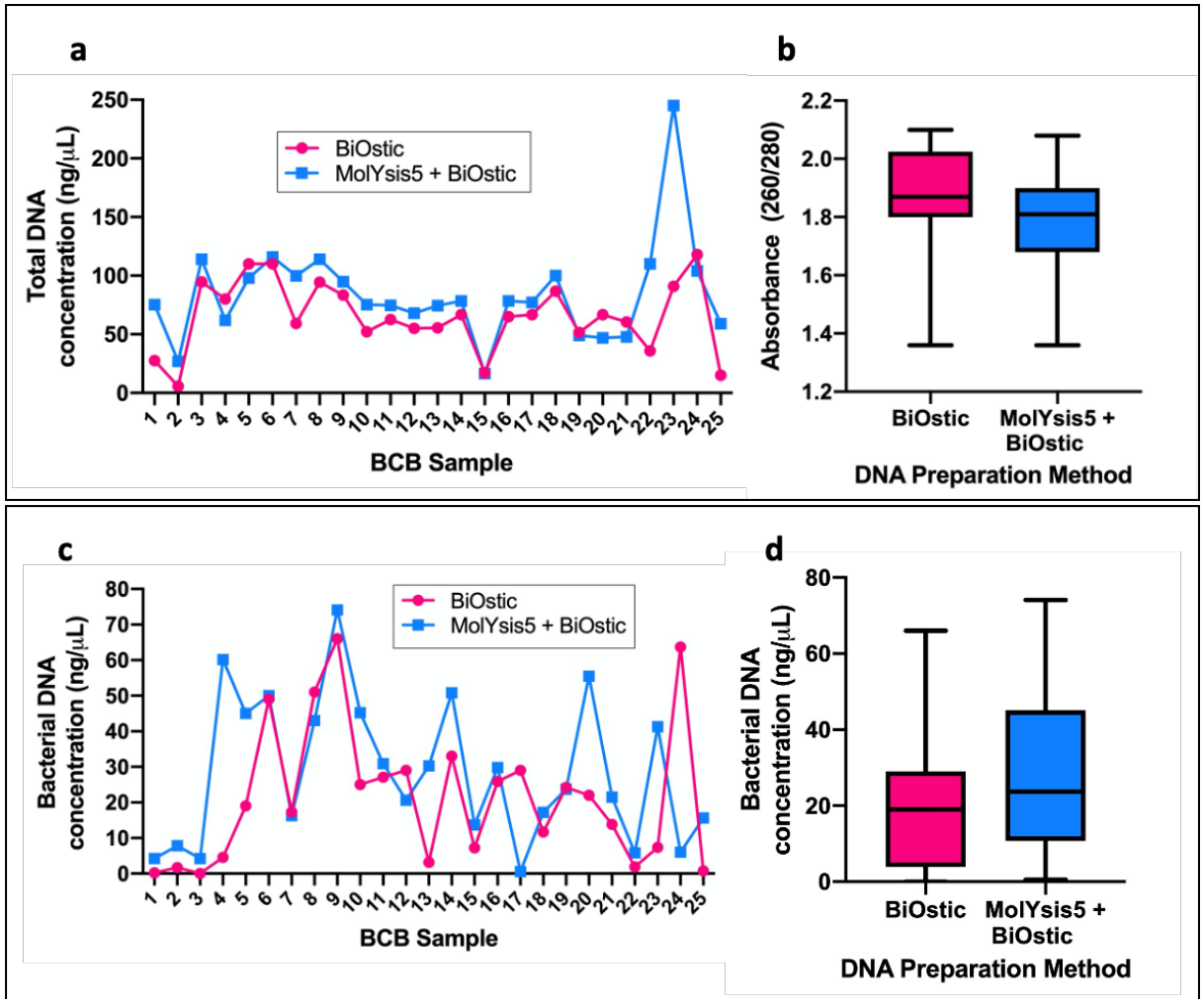
Yan, Q., Karau, M. J., Greenwood-Quaintance, K. E., Mandrekar, J. N., Osmon, D. R., Abdel, M. P., et al. (2018). Comparison of diagnostic accuracy of periprosthetic tissue culture in blood culture bottles to that of prosthesis sonication fluid culture for diagnosis of prosthetic joint infection (PJI) by use of bayesian latent class modeling and IDSA PJI criteria for. *J. Clin. Microbiol.* 56:e00319-18. doi: 10.1128/JCM.00319-18

Yan, Q., Mi, W. Y., Thoendel, M. J., Raval, Y. S., Greenwood-Quaintance, K. E., Abdel, M. P., et al. (2019). Evaluation of the cosmosid bioinformatics platform for prosthetic joint-associated sonicate fluid shotgun metagenomic data analysis. *J. Clin. Microbiol.* 57:e01182-18. doi: 10.1128/JCM.01182-18

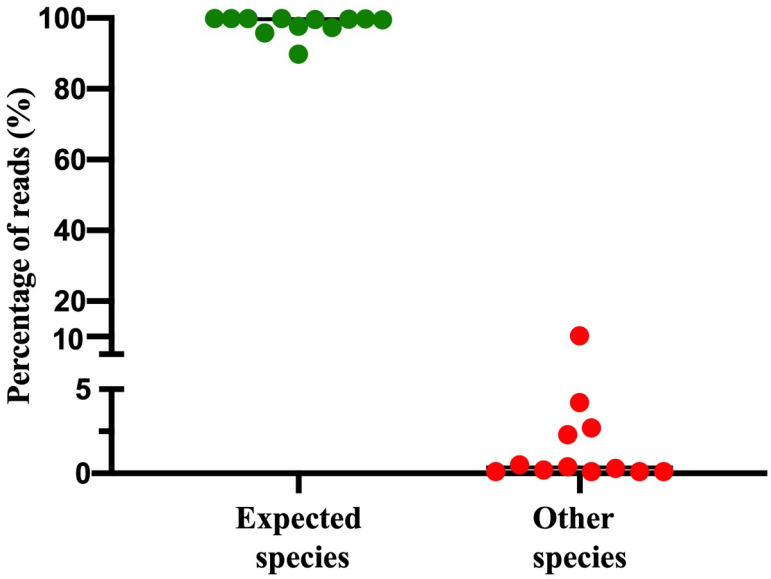
Zhang, C., Fang, X., Huang, Z., Li, W., Zhang, C., Yang, B., et al. (2019). Value of mNGS in sonication fluid for the diagnosis of periprosthetic joint infection. *Arthroplasty* 1:9. doi: 10.1186/s42836-019-0006-4

Zhou, Y., Wylie, K. M., El Feghaly, R. E., Mihindukulasuriya, K. A., Elward, A., Haslam, D. B., et al. (2016). Metagenomic approach for identification of the pathogens associated with diarrhea in stool specimens. *J. Clin. Microbiol.* 54, 368–375. doi: 10.1128/JCM.01965-15

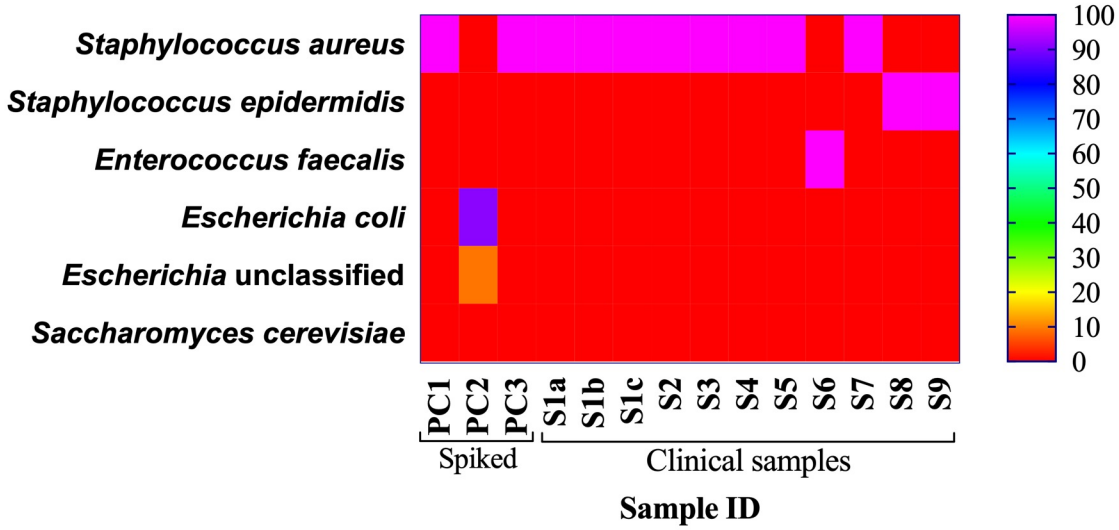
Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.



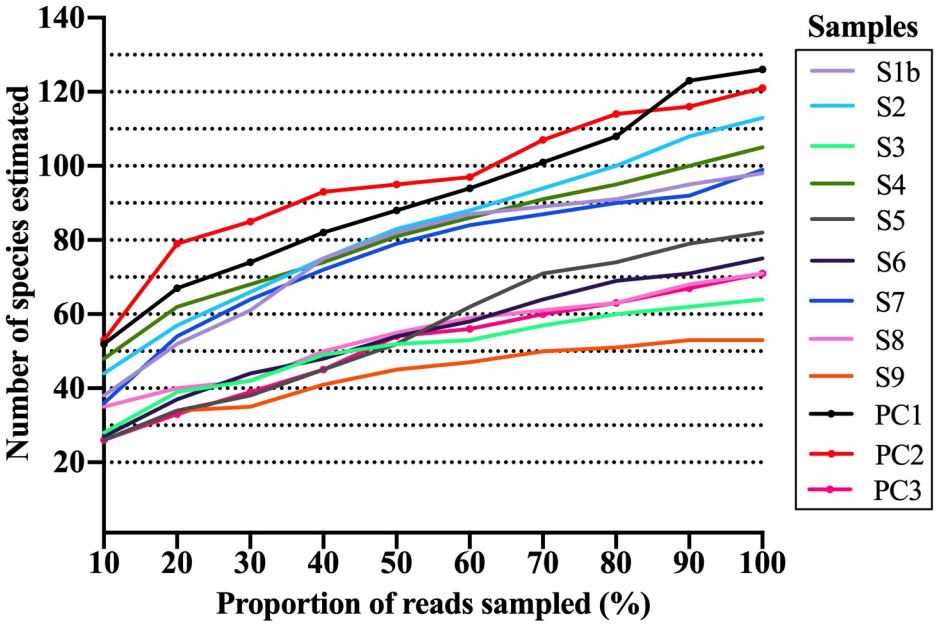
Supplementary Figure S1. Graphs comparing the two DNA preparation methods from 25 positive BCBs. **(A)** Total DNA yield in ng/µL as determined by Qubit. **(B)** Box plot comparing the 260/280 nm absorbance ratios of the resulting DNA from 25 positive BCBs as determined by spectrophotometry. An absorbance ratio of 1.8 was used as the standard for pure DNA. **(C)** Bacterial DNA yield in ng/µL as determined by qPCR. **(D)** Box plot comparing the bacterial DNA concentration ratios as determined by qPCR. Each plot reveals the median (midline) and the Inter-quartile range (IQR; box).



Supplementary Figure S2. Dot plot diagram reperesenting the percentage of reads belonging to the bacterial species expected and other species (possible contaminants/miss-classified) found in the clinical samples and positive controls.

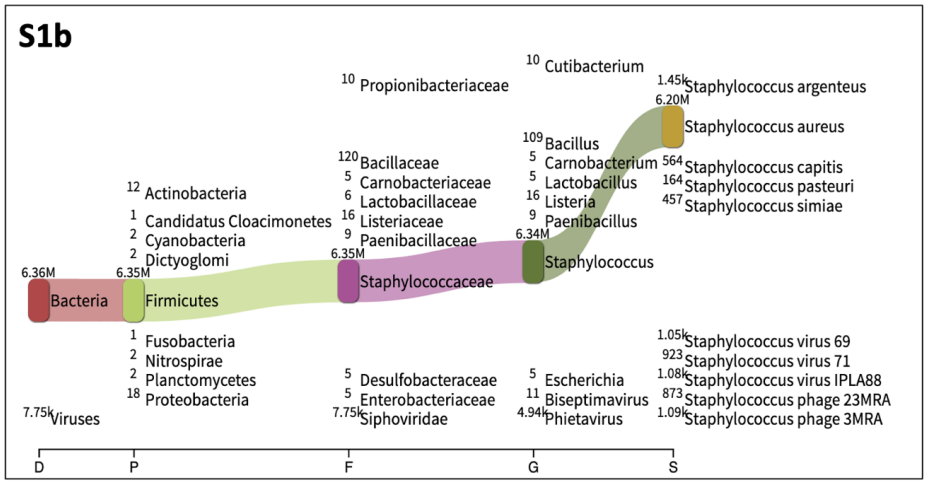


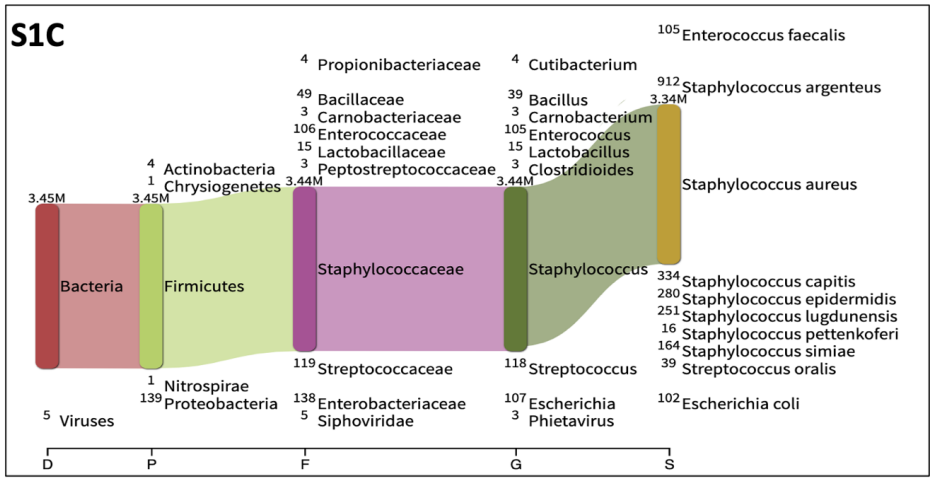
Supplementary Figure S3. Heat map representing the taxonomical composition at the species level in PJT from BCB as estimated by MetaPhlAn2. S1-S9: clinical samples, PC1-3: spiked samples (positive controls).

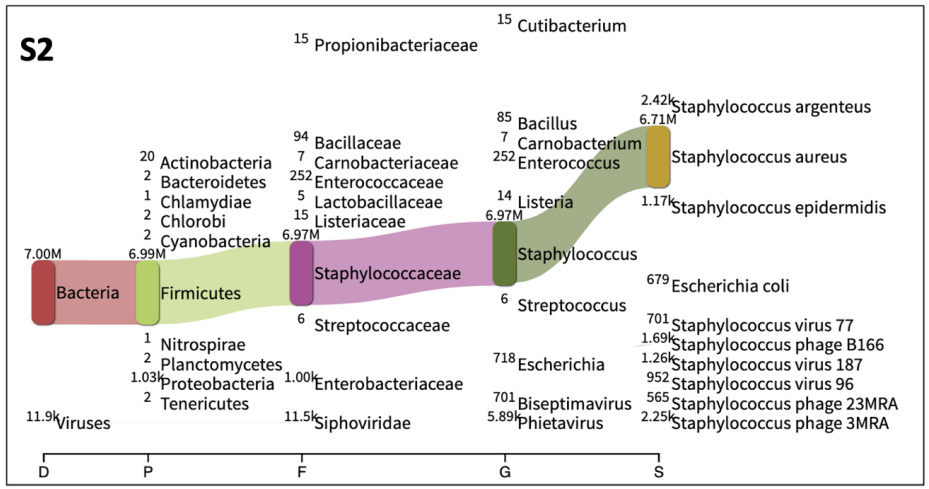


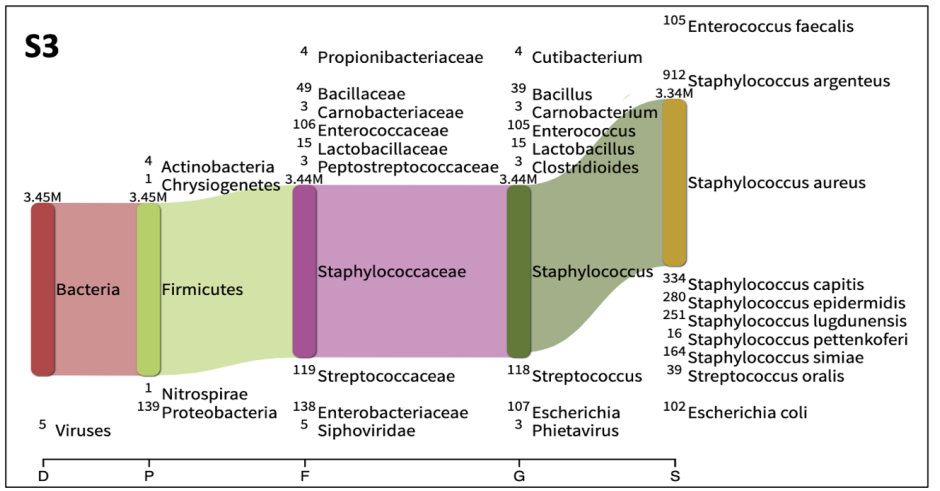
Supplementary Figure S4. Rarefaction curves of species detected in the BCBs from PJT samples. (X-axis and Y-axis are in linear scale). S1-S9: clinical samples, PC1-3: spiked samples (positive controls).

Supplementary Figure S5. Sankey flow diagram of Kraken reports from all the BCBs with PJT evaluated. S1-S9: clinical samples, PC1-3: spiked samples (positive controls), NC: negative control.

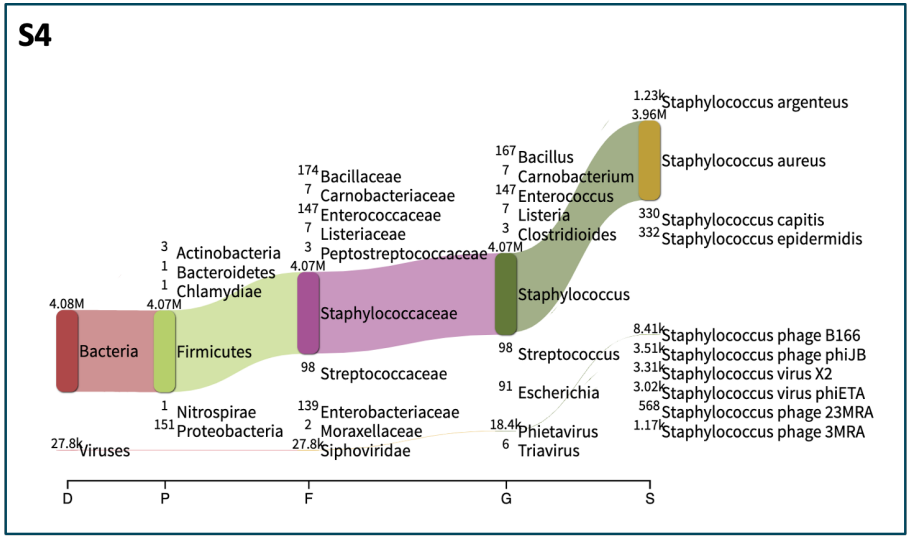


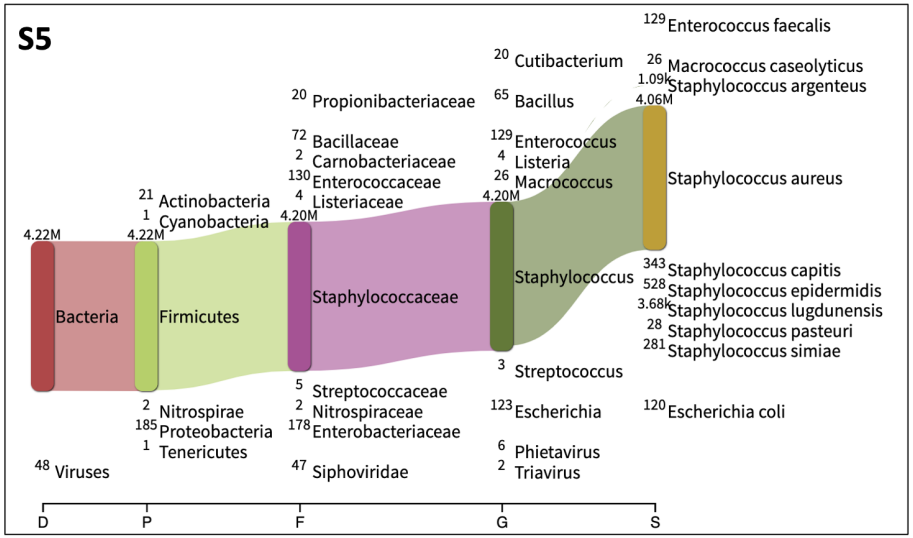


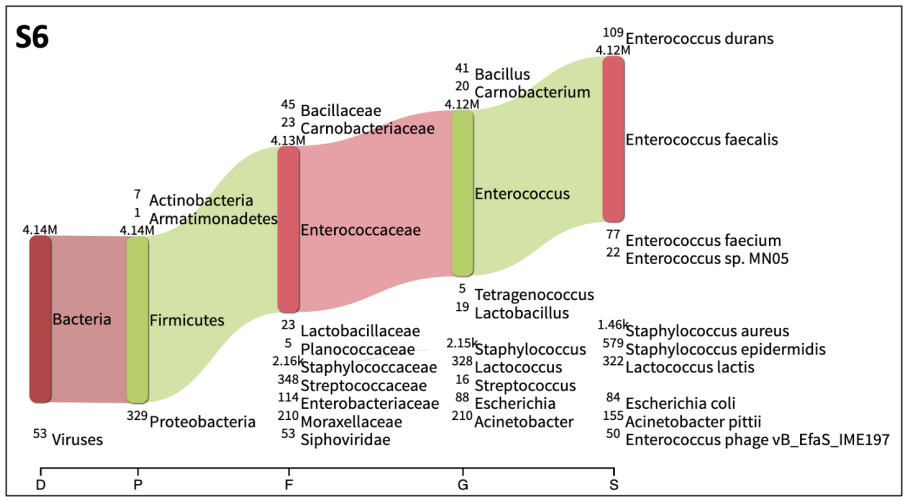


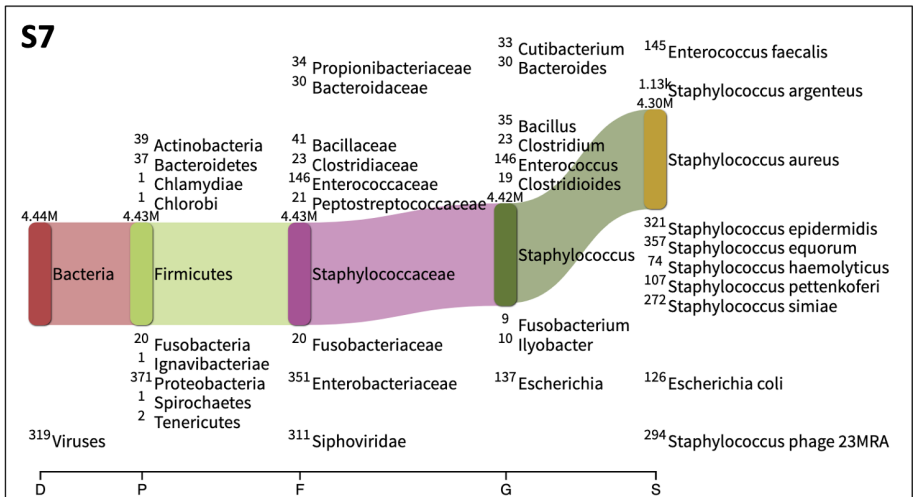


Supplementary Figure S5. Sankey flow diagram of Kraken reports from all the BCBs with PJT evaluated. S1-S9: clinical samples, PC1-3: spiked samples (positive controls), NC: negative control.

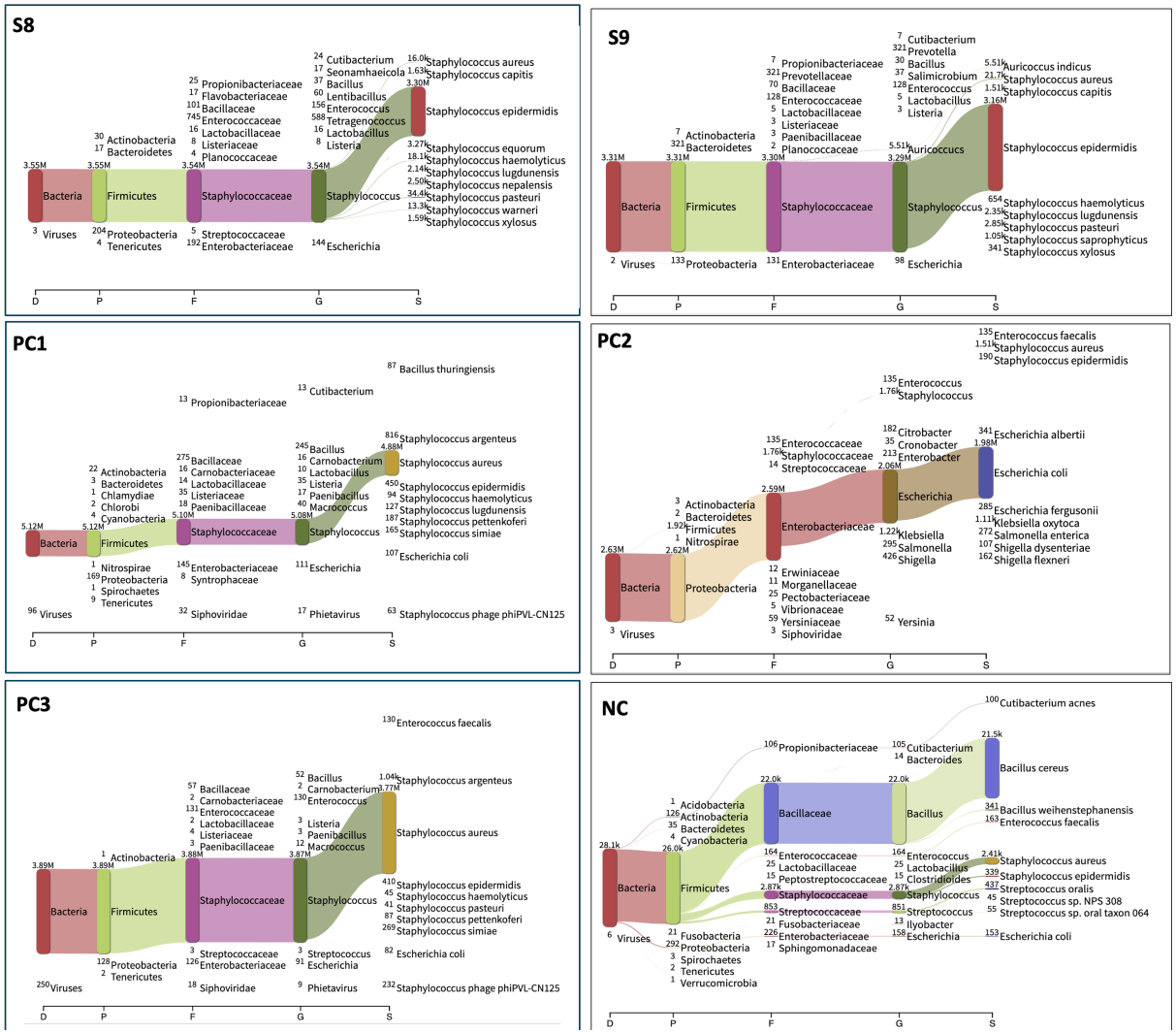


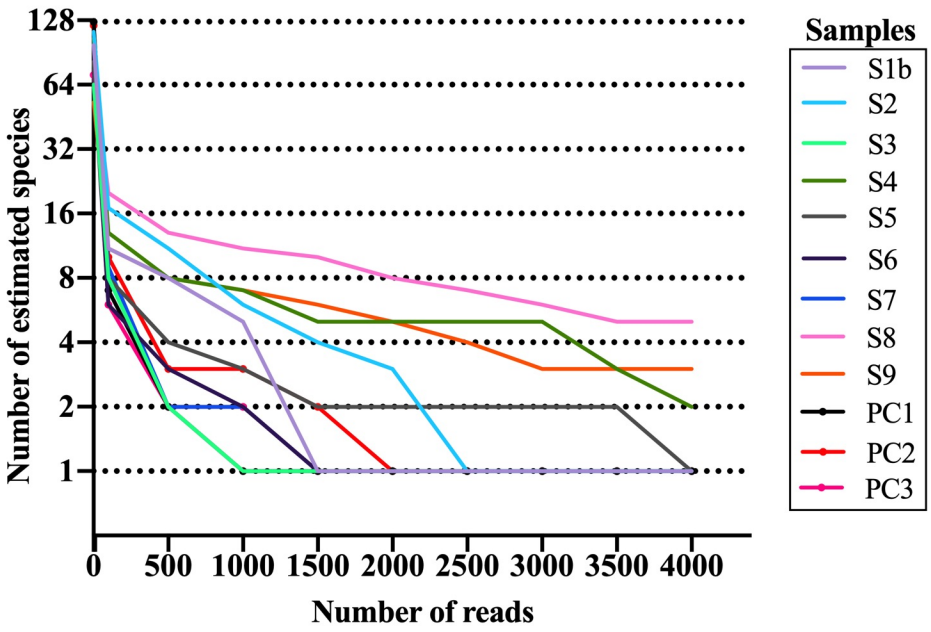




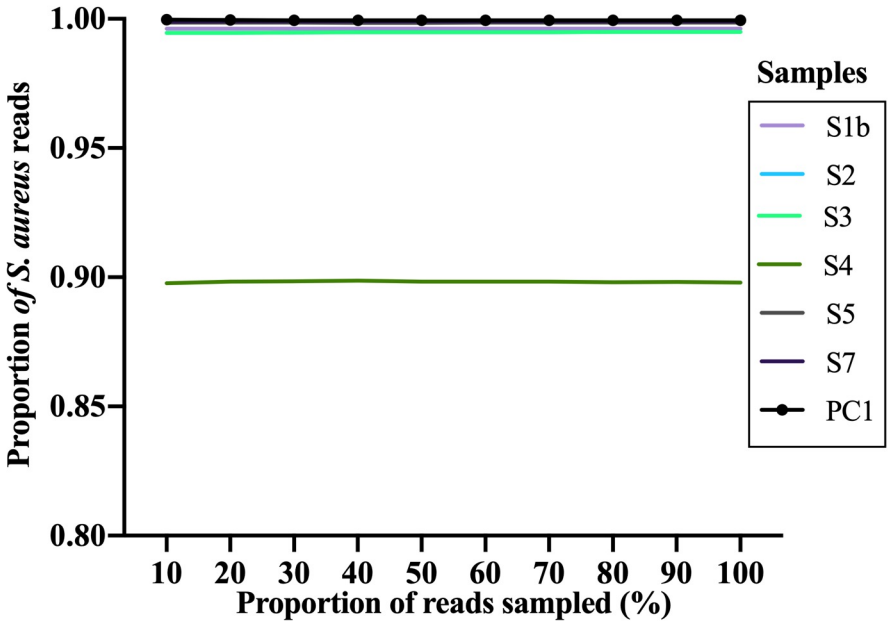


Supplementary Figure S5. Sankey flow diagram of Kraken reports from all the BCBs with PJT evaluated. S1-S9: clinical samples, PC1-3: spiked samples (positive controls), NC: negative control.





Supplementary Figure S6. The number of detected species (X-axis in linear scale) decrease as the threshold for detecting a species in number of reads increase (Y-axis in log scale). S1-S9: clinical samples, PC1-3: spiked samples (positive controls).



Supplementary Figure S7. Effect of reductionof the read dataset on the detection of S. aureusin positive BCBs with PJT samples. S1-S5, S7: clinical samples, PC1: Sample spiked with spiked *S. aureus* (positive controls).

Supplementary Table S1. DNA Extraction results obtained from the two sample preparation methods tested.

Sample No	DNA sample preparation methods									
	Method: BiOstic Kit					Method: MolYsis+BiOstic				
	Qubit ng/µl	Nanodrop ng/µl	Abs 260/280	Abs 260/230	Bacterial DNA ng/µl	Qubit ng/µl	Nanodrop ng/µl	Abs 260/280	Abs 260/230	Bacterial DNA ng/µl
1	27,6	23,5	1,88	1,85	0,22	75,4	81	1,8	1,97	4,2
2	5,8	6,5	2,06	1,04	1,7	27	45,8	1,94	2,22	7,8
3	94,6	70,3	1,86	2,17	0,058	114	213	1,7	1,4	4,26
4	80	103,5	1,83	2,11	4,57	62	111	1,68	1,49	60,15
5	110	104,4	1,77	1,85	19	98	103,8	1,77	1,81	45
6	110	101,6	1,8	1,95	49	116	484,6	2,04	2,29	50
7	59,2	42,5	1,86	2,21	17,3	99,8	433,6	2,08	2,3	16,3
8	94,4	81,9	1,87	2,32	51	114	109,7	1,82	1,97	43
9	83,4	540,1	2,06	2,27	66	94,8	610	1,98	2,18	74,1
10	52,2	272,7	2,09	2,2	25	75,4	259,8	1,96	2,03	45,2
11	62,6	494,6	2,07	2,26	27,1	74,6	151,3	1,85	1,99	30,9
12	55	218,6	1,99	2,27	29	68	138,4	1,88	1,96	20,61
13	55,6	383,8	2,09	2,27	3,19	74,4	134,7	1,83	2,04	30,3
14	66,8	472,9	2,1	2,33	33	78,4	162,2	1,85	2,01	50,8
15	17,1	62,7	1,91	1,77	7,2	16,6	34,8	1,66	1,23	13,7
16	65,2	178,4	1,93	2,38	25,93	78,4	295,1	1,81	1,96	29,8
17	66,8	191,2	1,95	2,4	29	77,2	298,9	1,79	1,96	0,54
18	86,8	82,2	1,7	1,21	11,7	100	149,2	1,78	1,42	17,2
19	51,6	100,2	1,36	0,85	24,2	49	143,2	1,68	1,32	23,73
20	66,8	115,6	1,79	1,8	22	47	106,9	1,56	1,07	55,47
21	60,6	65,1	1,82	2,03	13,8	47,8	65,1	1,36	0,85	21,5
22	36	16,8	1,8	1,59	1,88	110	67,4	1,92	1,92	5,82
23	91	65,4	1,85	2,04	7,39	245	193,3	1,81	1,63	41,3
24	118	216,7	1,9	2,23	63,7	104	110,5	1,49	0,98	6
25	15	32,4	1,65	0,95	0,78	59	65	1,67	1,56	15,6

Supplementary Table S2. Results from the mapping against the known sources of contamination

Sample_paper	Human counts	Human percentage	total_reads	Horse counts	Horse percentage	PhiX counts	PhiX percentage	No hits counts	No hits percentage
1	221	0.0	3246305	1554	0.04	0	0.0	3244681	99.95
1	236	0.0	3246305	1597	0.04	0	0.0	3244681	99.95
2	11918	0.42	2831673	112023	3.96	0	0.0	2718972	96.02
2	11687	0.41	2831673	111736	3.96	0	0.0	2719255	96.03
3	3976	0.13	2905432	33996	1.17	0	0.0	2871147	98.82
3	3921	0.13	2905432	34001	1.16	0	0.0	2871147	98.82
4	2404	0.09	2760218	9866	0.36	0	0.0	2749177	99.6
4	2360	0.09	2760218	9927	0.36	0	0.0	2749177	99.6
5	2690	0.09	3001487	18150	0.59	1	0.0	2982577	99.37
5	2591	0.09	3001487	18184	0.6	1	0.0	2982577	99.37
6	30097	0.98	3072977	223590	7.28	0	0.0	2841581	92.47
6	29761	0.97	3072977	223259	7.26	0	0.0	2841889	92.48
7	8482	0.29	2886161	70982	2.46	0	0.0	2814295	97.51
7	8334	0.29	2886161	70965	2.45	0	0.0	2814295	97.51
8	5430	0.2	2655317	25005	0.94	0	0.0	2627967	98.97
8	5376	0.2	2655317	25010	0.94	0	0.0	2627967	98.97
9	6977	0.13	5035684	70003	1.4	2	0.0	4965687	98.61
9	6940	0.13	5035684	70003	1.4	1	0.0	4965687	98.61
10	1150	0.04	2473824	7177	0.28	0	0.0	2466402	99.7
10	1181	0.04	2473824	7213	0.29	0	0.0	2466402	99.7
11	2215	0.07	3094383	9292	0.3	0	0.0	3083862	99.66
11	2232	0.07	3094383	9328	0.3	0	0.0	3083862	99.66
12	1171	0.05	2527129	7871	0.31	1	0.0	2518789	99.67
12	1170	0.05	2527129	7884	0.31	0	0.0	2518789	99.67
13	463	0.01	2593041	2083	0.08	0	0.0	2590707	99.91
13	466	0.01	2593041	2130	0.08	0	0.0	2590707	99.91
14	7324	0.11	6708791	67862	1.02	2627	0.04	6637677	98.94
14	7200	0.11	6708791	67889	1.02	2625	0.04	6637677	98.94
15	20725	0.27	7666695	202648	2.64	4816	0.06	7458927	97.29
15	20398	0.26	7666695	202549	2.64	4816	0.06	7458927	97.29
16	188	0.0	3502332	1071	0.03	2474	0.07	3498829	99.9
16	161	0.0	3502332	1044	0.03	2474	0.07	3498829	99.9
17	12979	0.27	4785518	18541	0.39	3213	0.07	4753455	99.33
17	12922	0.27	4785518	18557	0.39	3213	0.07	4753455	99.33
18	5511	0.12	4391574	26295	0.61	3268	0.07	4359515	99.27
18	5493	0.12	4391574	26285	0.6	3268	0.07	4359515	99.27
19	9982	0.22	4653928	53712	1.15	2995	0.06	4592496	98.68
19	10001	0.22	4653928	53691	1.15	2994	0.06	4592496	98.68
PC	3217	0.06	5276970	44054	0.84	4184	0.08	5228421	99.08
PC	3207	0.06	5276970	44081	0.84	4184	0.08	5228421	99.08

Supplementary Table S3. Results from the statistical test (Wilcoxon rank sum test) applied to the total DNA concentration results obtained, for evaluating the statistical significance amongst the two sample preparation methods tested.

Table Analyzed	DNA concentration
Column B	MolYsis5 + BiOstic
vs.	vs.
Column A	BiOstic
Wilcoxon matched-pairs signed rank test	
P value	**0,0069**
Exact or approximate P value?	Exact
P value summary	**
Significantly different (P < 0.05)?	Yes
One- or two-tailed P value?	Two-tailed
Sum of positive, negative ranks	260.5 , -64.50
Sum of signed ranks (W)	196
Number of pairs	25
Number of ties (ignored)	0
Median of differences	
Median	13
How effective was the pairing?	
rs (Spearman)	0,6837
P value (one tailed)	<0.0001
P value summary	****
Was the pairing significantly effective?	Yes

Supplementary Table S4. Results from the statistical test (Wilcoxon rank sum test) applied to the bacterial DNA concentration results obtained by qPCR, for evaluating the statistical significance amongst the two sample preparation methods tested.

Table Analyzed	qPCR_Bacterial_DNA
Column B	MolYsis5 + BiOstic
vs.	vs.
Column A	BiOstic
Wilcoxon matched-pairs signed rank test	
P value	**0,0207**
Exact or approximate P value?	Exact
P value summary	*
Significantly different (P < 0.05)?	Yes
One- or two-tailed P value?	Two-tailed
Sum of positive, negative ranks	247.5 , -77.50
Sum of signed ranks (W)	170
Number of pairs	25
Number of ties (ignored)	0
Median of differences	
Median	5,5
How effective was the pairing?	
rs (Spearman)	0,4493
P value (one tailed)	0,0121
P value summary	*
Was the pairing significantly effective?	Yes

Supplementary Table S5. Results from all the preprocessing steps. S1-S9: clinical samples, PC1-3: spiked samples (positive controls), NC: negative control.

Sample	Raw reads	after removing_optical duplicates	%	after trimming	%	Processed reads that go to filter	Human	Human %	horse	Horse %	PhiX	PhiX %
1a	34212820	34210014	99,9917984	29676882	86,74910802	29676882	1392	0	248439	0,8	1316	0,1
1b	14785194	14785194	100	13417582	90,75012475	13417582	646	0	121876	0,9	5252	0
1c	14078494	14077174	99,990624	12698920	90,20929911	12698920	6450	0,1	2122246	16,7	5982	0
2	16172074	16170958	99,9930992	15333390	94,82054186	15333390	1085	0	365165	2,4	9632	0,1
3	7404282	7403570	99,9903839	7004664	94,61197774	7004664	91	0	1855	0	4948	0,1
4	10369498	10369044	99,9956218	9571036	92,30393853	9571036	20365	0,2	31562	0,3	6426	0,1
5	9260612	9259660	99,9897199	8783148	94,85389312	8783148	4820	0,1	46398	0,5	6536	0,1
6	9426630	9425988	99,9931895	8634322	91,60124116	8634322	177	0	4018	0	4711	0,1
7	9860584	9859604	99,9900614	9307856	94,40395375	9307856	9717	0,1	97134	1	5989	0,1
8	7980152	7979396	99,9905265	7577546	94,96390454	7577546	623	0	3844	0,1	6844	0,1
9	7236776	7236012	99,9894428	6854112	94,72223097	6854112	21433	0,3	20700	0,3	5746	0,1
PC_1	11884076	11882850	99,9896837	10553940	88,81657178	10553940	340	0,8	82049	0,8	8368	0,1
PC_2	6857300	6856986	99,9954209	5374834	78,38478889	5374834	271	0	41427	0,8	6658	0,1
PC_3	8470486	8469658	99,9902249	8045360	94,99037623	8045360	1025	0	6850	0,1	5387	0,1
NC	11192852	11192462	99,9965156	10071132	89,98138211	10071132	137305	1,4	8829289	87,6	7941	0,1

Supplementary Table S5. Results from all the preprocessing steps. S1-S9: clinical samples, PC1-3: spiked samples (positive controls), NC: negative control.

Sample	Raw reads	Multiple Genomes	Multiple genomes %	Unmapped reads	Unmapped reads %	Reads after filter	%	Syncronized reads	% syncronized
1a	34212820	27817	0,1	29397918	99,1	29396700	99,05589138	29395638	85,91995047
1b	14785194	14454	0,1	13275354	98,9	13275933	98,94430308	13275394	89,78843294
1c	14078494	243195	1,9	10321047	81,3	10321172	81,27598252	10316200	73,27630356
2	16172074	39654	0,3	14917854	97,3	14917472	97,28750133	14916398	92,23552897
3	7404282	112	0	6997658	99	6997511	99,89788233	6997282	94,50318073
4	10369498	5773	0,1	9506910	99,3	9507146	99,33246516	9506784	91,68027228
5	9260612	6364	0,1	8719030	99,3	8719208	99,272015	8718812	94,14941475
6	9426630	592	0	8624824	99,9	8624813	99,88986975	8624584	91,49169958
7	9860584	10024	0,1	9184992	98,7	9184753	98,67742904	9184272	93,14125816
8	7980152	815	0	7565420	99,8	7565652	99,84303625	7565344	94,80200377
9	7236776	4213	0,1	6802020	99,2	6801830	99,23721702	6801578	93,98629998
PC_1	11884076	6341	0,1	10456842	98,2	10457098	99,08240903	10456806	87,99006334
PC_2	6857300	4318	0,1	5322160	99	5321958	99,01623008	5321658	77,60573404
PC_3	8470486	1220	0	8030878	99,8	8031112	99,82290413	8030844	94,8097193
NC	11192852	895679	8,9	200918	2	200980	1,995604863	187094	1,671548949

Supplementary Table S6. Descriptive statistics from the Taxonomic classification by Kraken when analyzing data from all type of samples. CS: clinical samples, PC: spiked samples (positive controls), all: clinical samples and spiked samples.

Parameter	Classified reads all (%)	Unclassified reads all (%)	Classified CS reads (%)	Unclassified reads CS (%)	Classified PC (%)	Unclassified PC (%)	Microbial reads all (%)	Microbial reads CS (%)	Microbial reads PC (%)
Number of values	15	15	9	9	3	3	15	9	4
Minimum	30,15	0,3671	93,97	0,5126	98,8	0,3671	30,01	86,32	30,01
25% Percentile	97,31	0,6483	96,71	1,161	98,8	0,3671	93,87	93,93	46,74
Median	98,77	1,229	98,7	1,295	99,05	0,9537	96,09	96,09	97,46
75% Percentile	99,35	2,685	98,84	3,294	99,63	1,199	97,24	97,07	98,53
Maximum	99,63	69,85	99,49	6,035	99,63	1,199	98,7	98,6	98,7
Range	69,48	69,48	5,52	5,522	0,83	0,8319	68,69	12,28	68,69
Mean	93,74	6,258	97,74	2,257	99,16	0,8399	90,82	95,05	80,91
Std. Deviation	17,66	17,66	1,749	1,751	0,4258	0,4275	17,22	3,602	33,94
Std. Error of Mean	4,559	4,559	0,5831	0,5837	0,2458	0,2468	4,446	1,201	16,97
Coefficient of variation	18.84%	282.1%	1.790%	77.57%	0.4294%	50.89%	18.96%	3.789%	41.95%
Sum	1406	93,88	879,7	20,32	297,5	2,52	1362	855,5	323,6

Parameter	Bacterial reads all (%)	Bacterial reads CS (%)	Bacterial reads PC (%)	Viral reads_all (%)	Viral reads_CS (%)	Viral reads_PC (%)
Number of values	15	9	3	15	9	3
Minimum	30,01	85,73	96,9	5,88E-05	5,88E-05	0,0001127
25% Percentile	93,82	93,85	96,9	0,0001429	0,0001111	0,0001127
Median	96,08	96,08	98	0,006226	0,001229	0,001836
75% Percentile	97,24	97,07	98,7	0,1168	0,1384	0,006226
Maximum	98,7	98,6	98,7	0,5857	0,5857	0,006226
Range	68,69	12,87	1,8	0,5856	0,5856	0,006113
Mean	90,72	94,95	97,87	0,1007	0,09688	0,002725
Std. Deviation	17,23	3,783	0,9074	0,1895	0,1929	0,003152
Std. Error of Mean	4,449	1,261	0,5239	0,04893	0,06432	0,00182
Coefficient of variation	18.99%	3.984%	0.9272%	188.1%	199.2%	115.7%
Sum	1361	854,6	293,6	1,511	0,872	0,008175

Supplementary Table S7. Kraken and Bracken general classification results (number of reads and proportion of reads classified). S1-S9: clinical samples, PC1-3: spiked samples (positive controls), NC: negative control.

Sample	Kraken general classification									
	Classified reads	%	Unclassified reads	%	Microbial reads	%	Bacterial reads	%	Viral reads	%
1a	14618131	99.46	796,88	0.5422	14082522	95.81	14065555	95.7	16967	0.1154
1b	6603673	99.49	34024	0.5126	6362912	95.86	6355162	95.74	7750	0.1168
1c	5124660	99.35	33440	0.6483	4508326	87.4	4482055	86.89	26271	0.5093
2	7286811	97.7	171388	2.298	7009491	93.98	6997563	93.82	11928	0.1599
3	3455654	98.77	42987	1.229	3449780	98.6	3449775	98.6	5	0.0001429
4	4691817	98.7	61575	1.295	4103062	86.32	4075220	85.73	27842	0.5857
5	4304158	98.73	55248	1.267	4224394	96.9	4224346	96.9	48	0.001101
6	4144003	96.1	168289	3.903	4143505	96.09	4143452	96.08	53	0.001229
7	4542010	98.91	50126	1.092	4436785	96.62	4436466	96.61	319	0.006947
8	3554393	93.97	228279	6.035	3550926	93.87	3550923	93.87	3	0.00007931
9	3309477	97.31	91312	2.685	3307061	97.24	3307059	97.24	2	0.00005881
PC_1	5209211	99.63	19192	0.3671	5123793	98	5123697	98	96	0.001836
PC_2	2635452	99.05	25377	0.9537	2626262	98.7	2626258	98.7	3	0.0001127
PC_3	3967279	98.8	48143	1.199	3891211	96.91	3890961	96.9	250	0.006226
NC	28194	30.15	65308	69.85	28064	30.01	28058	30.01	6	0.006417

Sample	Species level_braken										
	Reads_Species level	Species level (%)	Unclassified_reads	Unclassified (%)	Distributed reads	Distributed reads (%)	Reads not distributed	Reads not distributed (%)	Braken reads (distributed + species)	Specie_finally (%)	Number of species
1a	13734919	93,44868786	79688	0,542175679	883181	6,008925542	31	0,000210916	14618100	99,45761341	147
1b	6205126	93,48311621	34024	0,512587423	398527	6,00399506	20	0,000301309	6603653	99,48711127	98
1c	4384876	85,00951901	33440	0,648300731	739759	14,34169559	25	0,000484675	5124635	99,35121459	103
2	6727090	90,19724467	171388	2,297981054	559319	7,499384235	402	0,005390041	7286409	97,6966289	113
3	3340706	95,48581864	42987	1,228677078	114944	3,285389956	4	0,00011433	3455650	98,77120859	64
4	3982040	83,7725986	61575	1,295390744	707011	14,87382063	2766	0,058190025	4689051	98,64641923	105
5	4070233	93,36668803	55248	1,267328622	233893	5,365249302	32	0,000734045	4304126	98,73193733	82
6	4118881	95,51489092	168289	3,90254185	25122	0,582567229	0	0	4144003	96,09745815	75
7	4307782	93,80780534	50126	1,091561748	234214	5,100328039	14	0,000304869	4541996	98,90813338	99
8	3393299	89,7064033	228279	6,034861072	161090	4,258629879	4	0,000105745	3554389	93,96503318	71
9	3196361	93,98880672	91312	2,685023975	113113	3,326081095	3	8,82148E-05	3309474	97,31488781	53
PC_1	4883568	93,40458262	19192	0,367071934	325626	6,228020296	17	0,000325147	5209194	99,63260292	126
PC_2	1983358	74,53910041	25377	0,953725324	652087	24,50691119	7	0,000263076	2635445	99,0460116	121
PC_3	3768013	93,83853055	48143	1,198952439	199264	4,962467208	2	4,9808E-05	3967277	98,80099775	71
NC	25833	27,62828603	65308	69,8466343	2343	2,505828752	18	0,019250925	28176	30,13411478	107

Sample	Bracken_genus										
	Genus level	Genus level (%)	Unclassified	Unclassified (%)	Distributed reads	Distributed reads (%)	Reads not distributed	Reads not distributed (%)	Braken reads (distributed +genus)	Genus_finally (%)	number of genera
1a	14052828	95,61165504	79688	0,542175679	565284	3,846040015	19	0,000129271	14618112	99,45769505	67
1b	6348439	95,64219337	340,24	0,005125874	255220	3,845008291	4	6,02619E-05	6603659	99,48720166	36
1c	4492332	87,09276672	33440	0,648300731	632308	12,25854481	20	0,00038774	5124640	99,35131153	37
2	6974081	93,50891549	171388	2,297981054	312333	4,187780455	397	0,005323001	7286414	97,69669594	51
3	3441508	98,3669945	42987	1,228677078	14144	0,404271258	2	5,71651E-05	3455652	98,77126576	23
4	4084672	85,93173044	61575	1,295390744	604380	12,71470983	2765	0,058168988	4689052	98,64644027	33
5	4200265	96,34947972	55248	1,267328622	103862	2,382480549	31	0,000711106	4304127	98,73196027	32
6	4126267	95,68616875	168289	3,90254185	17736	0,411289402	0	0	4144003	96,09745815	37
7	4424055	96,33980788	50126	1,091561748	117946	2,568434384	9	0,000195987	4542001	98,90824226	55
8	3538419	93,54284485	228279	6,034861072	15972	0,422241209	2	5,28727E-05	3554391	93,96508606	34
9	3296949	96,94659092	91312	2,685023975	12526	0,368326291	2	5,88099E-05	3309475	97,31491721	34
PC_1	5082753	97,21425452	19192	0,367071934	126451	2,418539657	7	0,000133884	5209204	99,63279418	18
PC_2	2065037	77,60878283	25377	0,953725324	570410	21,43730394	5	0,000187911	2635447	99,04608676	53
PC_3	3873606	96,46821679	48143	1,198952439	93672	2,332805867	1	2,4904E-05	3967278	98,80102266	27
NC	26313	28,14164403	65308	69,8466343	1879	2,009582683	2	0,002138992	28192	30,15122671	78

Supplementary Table S8. Number of Species and genera determined by Kraken when no threshold was considered. S1-S9: clinical samples, PC1-3: spiked samples (positive controls), NC: negative control.

Sample	Number of species	Number of genera
S1b	98	36
S2	113	51
S3	64	23
S4	105	33
S5	82	32
S6	75	37
S7	99	55
S8	71	34
S9	53	18
PC1	126	62
PC2	121	53
PC3	71	27
NC	107	78

Supplementary Table S9. Descriptive statistics from the Number of Species and genus determined by Kraken when no threshold was considered. CS: clinical samples, all: clinical samples and spiked samples (positive controls).

	Number of species_all	**Number of genera_all**	**Species CS**	**Genera CS**
Number of values	13	13	9	9
Minimum	53	18	53	18
25% Percentile	71	29,5	67,5	27,5
Median	98	36	82	34
75% Percentile	110	54	102	44
Maximum	126	78	113	55
Range	73	60	60	37
Mean	91,15	41,46	84,44	35,44
Std. Deviation	23,24	17,15	20,36	11,76
Std. Error of Mean	6,447	4,758	6,787	3,92
Coefficient of variation	25.50%	41.37%	24.11%	33.18%
Sum	1185	539	760	319

Supplementary Table S10. Estimated number of species related reads when using different detection thresholds (expressed as number of reads). S1-S9: clinical samples, PC1-3: spiked samples (positive controls).

Threshold (No. of reads)	Number of species in sample											
	S1b	S2	S3	S4	S5	S6	S7	S8	S9	PC1	PC2	PC3
0	98	113	64	105	82	75	99	71	53	126	121	71
100	11	17	8	13	8	6	9	20	11	7	10	6
500	8	11	2	8	4	3	2	13	8	2	3	2
1000	5	6	1	7	3	2	2	11	7	1	3	2
1500	1	4	1	5	2	1	1	10	6	1	2	1
2000	1	3	1	5	2	1	1	8	5	1	1	1
2500	1	1	1	5	2	1	1	7	4	1	1	1
3000	1	1	1	5	2	1	1	6	3	1	1	1
3500	1	1	1	3	2	1	1	5	3	1	1	1
4000	1	1	1	2	1	1	1	5	3	1	1	1

Supplementary Table S11. Results from mapping of spiked sample reads against reference genomes. PC1-PC3: Spiked samples (positive controls).

Sample	Reference genome	Reference Genome NCBI accession number	Lenght of reference genome	Total reads	Mapped and paired reads	Unmapped reads	Total number of covered bases[a]	Coverage breadth (%)	Coverage depth
PC1	***S. aureus* ATCC 25923**	GCF_000756205.1	2 806 346	10 456 806	10 455 042	1 765	2 806 306	99,7	775X
PC2	***E. coli* ATCC 25922**	GCF_000743255.1	5 203 440	5 321 658	5 264 048	56 178	5 203 372	98,6	209X
PC3	***S. aureus* ATCC 25923**	GCF_000756205.1	5 203 440	8 030 844	8 029 324	1 440	2 806 345	99,5	583X

a Total number of covered bases with a coverage depth bigger or equal to 4X. PC1-3: spiked samples (positive controls).

Paper III