



Article

Culture-Free Detection of Antibiotic Resistance Markers from Native Patient Samples by Hybridization Capture Sequencing

Ines Ferreira ^{1,2}, Sarah Lepuschitz ¹, Stephan Beisken ¹, Giuseppe Fiume ¹, Katharina Mrazek ¹, Bernhard J. H. Frank ³ , Silke Huber ⁴ , Miriam A. Knoll ⁴, Arndt von Haeseler ^{2,5}, Arne Materna ¹, Jochen G. Hofstaetter ³, Andreas E. Posch ¹ and Johannes Weinberger ^{1,*}

¹ Ares Genetics GmbH, 1030 Vienna, Austria; ines.ferreira@ares-genetics.com (I.F.); sarah.lepuschitz@ares-genetics.com (S.L.); stephan.beisken@ares-genetics.com (S.B.); fiume.gius@gmail.com (G.F.); kathi.mrazek@live.at (K.M.); arne.materna@ares-genetics.com (A.M.); andreas.posch@ares-genetics.com (A.E.P.)

² Center for Integrative Bioinformatics Vienna, Max Perutz Laboratories, University of Vienna, 1030 Vienna, Austria; arndt.von.haeseler@univie.ac.at

³ Michael Ogon Laboratory for Orthopaedic Research, Orthopaedic Hospital Vienna-Speising, 1130 Vienna, Austria; Bernhard.Frank@oss.at (B.J.H.F.); Jochen.Hofstaetter@oss.at (J.G.H.)

⁴ Institute of Hygiene and Medical Microbiology, Medical University of Innsbruck, 6020 Innsbruck, Austria; silke.huber@i-med.ac.at (S.H.); miriam.knoll@i-med.ac.at (M.A.K.)

⁵ Bioinformatics and Computational Biology, Faculty of Computer Science, University of Vienna, 1090 Vienna, Austria

* Correspondence: johannes.weinberger@ares-genetics.com; Tel.: +43-664-8810-3506



Citation: Ferreira, I.; Lepuschitz, S.; Beisken, S.; Fiume, G.; Mrazek, K.; Frank, B.J.H.; Huber, S.; Knoll, M.A.; von Haeseler, A.; Materna, A.; et al. Culture-Free Detection of Antibiotic Resistance Markers from Native Patient Samples by Hybridization Capture Sequencing. *Microorganisms* **2021**, *9*, 1672. <https://doi.org/10.3390/microorganisms9081672>

Academic Editor: Ralf Matthias Hagen

Received: 2 July 2021
Accepted: 3 August 2021
Published: 5 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The increasing incidence of antimicrobial resistance (AMR) is a major global challenge. Routine techniques for molecular AMR marker detection are largely based on low-plex PCR and detect dozens to hundreds of AMR markers. To allow for comprehensive and sensitive profiling of AMR markers, we developed a capture-based next generation sequencing (NGS) workflow featuring a novel AMR marker panel based on the curated AMR database ARESdb. Our primary objective was to compare the sensitivity of target enrichment-based AMR marker detection to metagenomics sequencing. Therefore, we determined the limit of detection (LOD) in synovial fluid and urine samples across four key pathogens. We further demonstrated proof-of-concept for AMR marker profiling from septic samples using a selection of urine samples with confirmed monoinfection. The results showed that the capture-based workflow is more sensitive and requires lower sequencing depth compared with metagenomics sequencing, allowing for comprehensive AMR marker detection with an LOD of 1000 CFU/mL. Combining the ARESdb AMR panel with 16S rRNA gene sequencing allowed for the culture-free detection of bacterial taxa and AMR markers directly from septic patient samples at an average sensitivity of 99%. Summarizing, the newly developed ARESdb AMR panel may serve as a valuable tool for comprehensive and sensitive AMR marker detection.

Keywords: target enrichment; antimicrobial resistance; next generation sequencing; human pathogens; urinary tract infection; synovial fluid; infectious disease surveillance; molecular detection

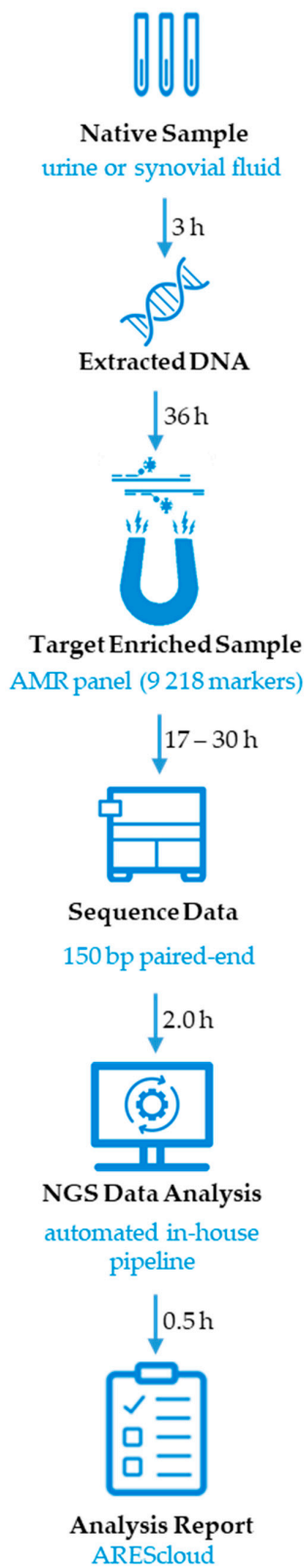
1. Introduction

The burden of antimicrobial resistance (AMR) and its accelerating progression has been acknowledged worldwide by leading health institutes such as the WHO and the CDC [1–3]. Besides the need for new antibiotics, efforts in the field of AMR surveillance and diagnostics are essential to address the rise of AMR [4].

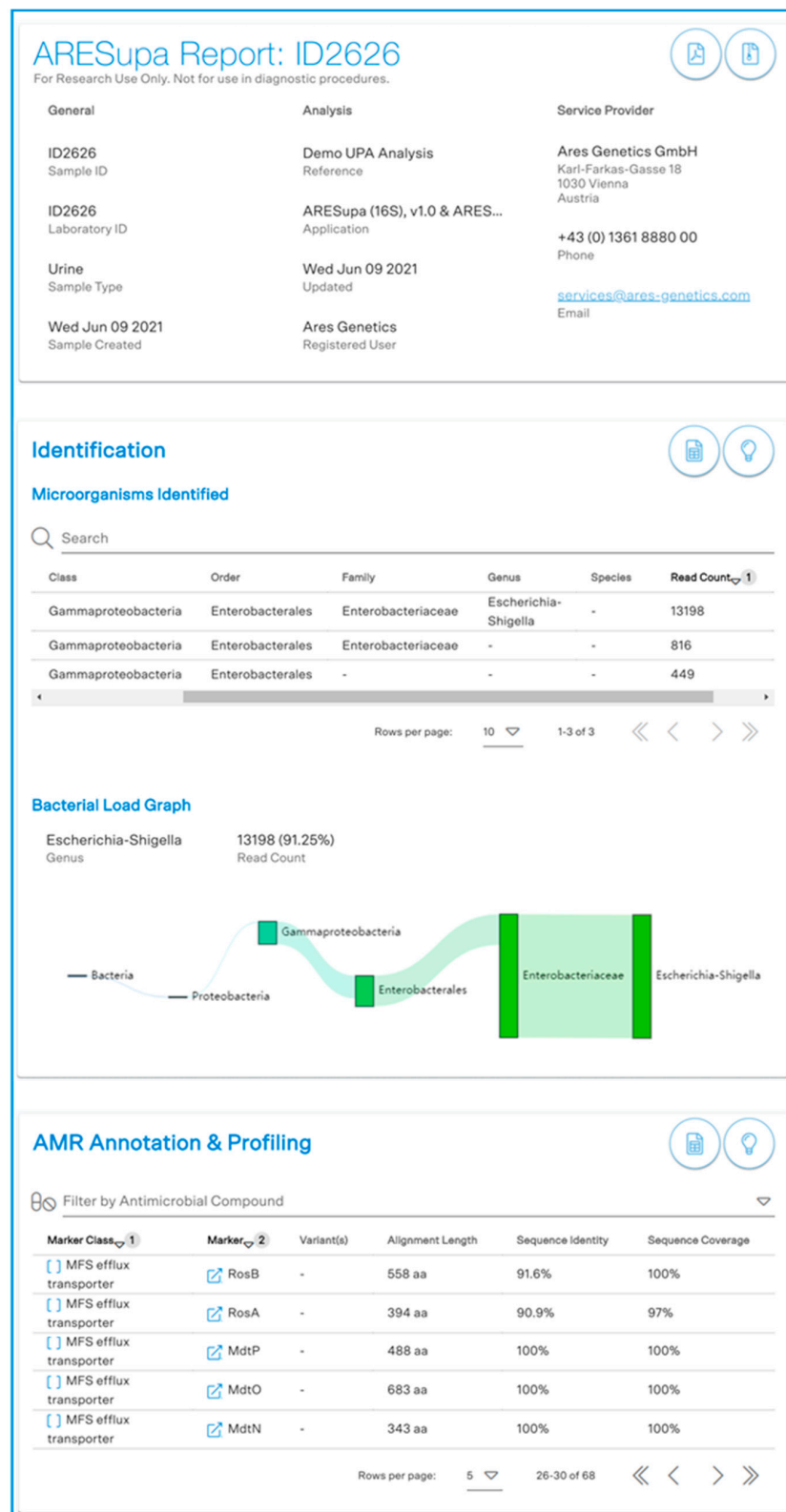
One unmet need to inform AMR surveillance is to comprehensively define resistomes in emerging multidrug-resistant human pathogens [5]. The WHO report on AMR surveillance [1] states that next generation sequencing (NGS) of bacterial isolates serves as a valuable tool for molecular AMR profiling. In order to obtain a more comprehensive picture of the prevalence, transmission, and composition of AMR, it would be beneficial to

extend AMR profiling beyond bacterial culture and enable AMR profiling directly from native samples. Thereby, monitoring gaps caused by non-culturable bacteria or false negative culture results could be reduced. However, culture-free AMR profiling methods need to overcome certain challenges [6]. For example, the fraction of host DNA in clinical specimens outweighs microbial DNA by nine orders of magnitude. Ultra-deep sequencing has been proposed as a solution to address this issue, but excessive sequencing costs and an uncertain performance when screening for individual AMR markers in low abundance microbes hamper routine use [7]. Other culture-free technologies such as PCR-based or loop mediated isothermal amplification (LAMP)-based assays have been demonstrated to sensitively detect AMR markers; however, their target space usually comprises only a limited set of markers for only a few dozen AMR markers [8].

To address these constraints for culture-free AMR marker detection, we established a capture-based next generation sequencing (NGS) workflow selectively enriching 9218 AMR markers catalogued in ARESdb for samples with low bacterial biomass or a high background of host DNA. In order to assess the sensitivity of the ARESdb AMR panel, we determined the limit of detection (LOD) by spike-in experiments in different native specimens, and subsequently performed an initial proof-of-concept in a set of septic patient samples. In addition, the workflow was complemented with 16S rRNA gene sequencing to enable taxonomic profiling (Figure 1).



(a)



(b)

Figure 1. Workflow from sample to result. (a) Flow chart of the target enrichment workflow established and tested for synovial fluid and urine samples. (b) Example of the standardized sample report implemented on AREScloud to summarize AMR marker profiling results and taxonomic identification via 16S rRNA sequencing.

2. Materials and Methods

2.1. Clinical Samples and Bacterial Isolates

2.1.1. Aseptic Urine and Synovial Fluid Samples

Aseptic urine samples were collected from three healthy male and three healthy female probands. To investigate the differences among female and male samples, one urine pool sample was created from male probands and one urine pool sample from female probands. Further, aseptic synovial fluids from six patients who underwent revision arthroplasty after total knee joint replacement were included. Synovial samples were retrieved under sterile conditions by aspiration of the affected joint in the operating theater. Samples were collected between 09/2019 and 07/2020 at the Orthopaedic Hospital Vienna-Speising. Synovial fluids were considered as aseptic according to the following criteria: aseptic indication for revision arthroplasty, no clinical signs of infection, negative laboratory infection parameters, negative aerobic and anaerobic microbiological cultures and fungi cultures, and aseptic histological results. After joint aspiration, six aseptic synovial fluid samples were stored at $-80\text{ }^{\circ}\text{C}$ and shipped frozen on dry ice to the testing laboratory for further analysis. As volumes of synovial fluid were limited (five to eight milliliters per patient) and in order to obtain sufficient sample material for all planned spike-in experiments, synovial fluids were pooled from all available patient samples.

To determine the LOD, reference strains of two Gram-negative (*E. coli* ATCC 35218, *K. quasipneumoniae* ATCC 700603) and two Gram-positive pathogens (*S. aureus* ATCC BAA-2312, *E. faecium* ATCC 700221) were used for spike-in experiments. All samples were processed as duplicates and native blank control samples with no bacterial spike-in were included. Bacteria present in unspiked aseptic sample material, i.e., controls, were considered background and removed for the analysis of the LOD. Values for performance metrics were averaged across replicates.

2.1.2. Septic Urine Samples

Septic urine samples were residue samples collected from routinely processed urine samples derived from patients with suspected UTI (urinary tract infection) at the Institute of Hygiene and Medical Microbiology at the Medical University of Innsbruck. Isolates from patients with clinical signs of UTI were primarily selected based on macroscopic turbidity of urine samples as well as microscopic identification of bacterial structures and leukocytes in native samples. Samples were frozen immediately after plating of cultures to ensure freshness and stored before further selection. Isolates of the causative pathogen cultured on selective agar plates were frozen in skim milk. Antimicrobial inhibitor tests were performed for each sample by inoculating antibiotic-free filter plates with 10 μL of native urine and placing them on an agar plate seeded with *Bacillus subtilis* spores (Axon Lab AG, Polling, Austria). Bacterial cultures were obtained using BD CHROMagar Orientation Medium (Becton Dickinson Diagnostics, Heidelberg, Germany). Colony-forming units (CFUs) counts were read after overnight culture and were classified into categories of below 10,000 CFU/mL, approximately 10,000 CFU/mL, and more than 10,000 CFU/mL. Species identification of relevant pathogens was performed by matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS, Bruker Daltonik, Bremen, Germany) using the reference Biotyper library v4.1 (Bruker Daltonik, Bremen, Germany). Samples were included if clinically relevant monocultures were obtained. Thirteen septic urine samples as well as 13 corresponding bacterial isolates were shipped to the testing laboratory (Table 1). All samples were processed as duplicates and values for performance metrics were averaged across replicates.

Table 1. Set of septic urine samples.

| Sample ID | Gender | Sample Type | Type of UTI | Presence/Absence of Leukocytes (+/−) | Culture Pos. Detected Species | Antibiotic Treatment | CFU/mL |
|-----------|--------|-------------|-------------|--------------------------------------|-------------------------------|----------------------|---------|
| ID-1 | m | MSU | recurrent | − | <i>E. coli</i> | Fosfomycin | >10,000 |
| ID-2 | m | MSU | recurrent | + | <i>E. coli</i> ESBL | none | >10,000 |
| ID-3 | m | MSU | recurrent | + | <i>E. coli</i> | none | >10,000 |
| ID-4 | m | MSU | recurrent | + | <i>C. koseri</i> | none | >10,000 |
| ID-5 | m | CU | acute | − | <i>E. hormaechei</i> | none | >10,000 |
| ID-6 | m | MSU | recurrent | − | <i>E. coli</i> ESBL | none | >10,000 |
| ID-7 | f | MSU | recurrent | + | <i>Enterococcus</i> sp. | none | >10,000 |
| ID-8 | f | MSU | acute | − | <i>E. coli</i> | none | >10,000 |
| ID-9 | f | MSU | acute | − | <i>K. pneumoniae</i> | none | >10,000 |
| ID-10 | f | MSU | acute | + | <i>E. coli</i> | none | >10,000 |
| ID-11 | f | MSU | acute | − | <i>Enterococcus</i> sp. | none | >10,000 |
| ID-12 | m | CU | acute | + | <i>E. coli</i> ESBL | none | >10,000 |
| ID-13 | f | CU | acute | + | <i>E. coli</i> | none | >10,000 |

MSU: midstream urine; CU: catheter urine; ESBL: extended spectrum-beta-lactamase; CFU: colony forming units; f: female; m: male.

3. Wet-Lab Workflow

3.1. Cultivation, DNA Isolation, and Quantification

A stock solution for each reference strain (*E. coli* ATCC 35218, *K. quasipneumoniae* ATCC 700603, *S. aureus* ATCC BAA-2312, *E. faecium* ATCC 700221) was prepared by picking one colony from the overnight culture and resuspended in 1 mL PBS. Serial dilutions from 10^1 to 10^7 were made and an aliquot of 100 μ L was subsequently cultured overnight on plate-count agar. The calculation of colony forming units per ml (CFUs/mL) was averaged across triplicates and spike-in dilutions were prepared for 10, 100, 1000, 10,000, and 100,000 CFU/mL. Then, defined concentrations were spiked into aseptic native material (synovial fluid, male-, female-urine) and processed as duplicates.

Automated DNA extractions from native samples and bacterial isolates were performed on a QIASymphony SP instrument (QIAGEN, Hilden, Germany) using the QIASymphony DSP DNA Kit (QIAGEN) as described previously [9]. Each independent DNA extraction contained no template controls (NTC) containing molecular grade water only.

Quantitative PCR (qPCR) to assess bacterial and human DNA concentrations was performed with the CFX96 Touch Real-Time PCR System (Bio-Rad, Hercules, CA, USA). The qPCR reaction was performed in a total volume of 20 μ L using the TaqMan Universal PCR Master Mix (Applied Biosystems, Bedford, MA, USA), containing 100 nM of each of the forward and reverse primers and the fluorogenic probe. The reaction conditions for amplification of DNA were 50 °C for 2 min, 95 °C for 10 min, and 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Data analysis made use of Bio-Rad CFX Maestro Software. *S. aureus* and *H. sapiens* DNA were used as standards for determining concentrations of samples by real-time PCR. For bacterial detection [10], the reaction set included forward primer (5'-TCCTACGGGAGGCAGCAGT-3'), reverse primer (5'-GGACTACCAGGGTATCTAATCCTGTT-3'), and the probe ((6-FAM)-5'-CGTATTACCGC-GGCTGCTGGCAC-3'-(TAMRA)), and for human detection [11], the reaction set included forward primer (5'-CATGGTGAACCCCGTCTCTA-3'), reverse primer (5'-GCCTCAGCCT-CCCCGAGTAG-3'), and the probe ((6-FAM)-5'-ATTAGCCGGGCGTGGTGGCG-3'-(TAMRA)).

3.2. ARESdb AMR Panel Design

A total of 9218 markers, including 7312 AMR genes and 1906 genetic variants, were extracted from ARESdb (date: 24 June 2020) [12]. Probes (120 bp length) were proprietary designed by Roche (Pleasanton, CA, USA) to cover all genes with a tiling of one, i.e., a targeted per-base coverage of one. Because of an insufficient marker length, five AMR

markers were excluded from the panel design (minimum required marker length: 120 bp). The final panel covered a capture space of 7.74 Mbp.

3.3. Library Preparation and Target Enrichment

Illumina Library preparation of DNA extracts from native samples was performed using the KAPA HyperCap Workflow v3.0 (Roche, Pleasanton, CA, USA). Briefly, 100 ng or up to 35 µL of low-concentrated DNA sample were used for library preparation. Libraries were prepared for sequencing with and without hybridization enrichment using two different sequencing lanes. The manufacturer's protocol was followed with the following modifications: (1) pre-capture PCR ranged between 13 and 17 cycles based on bacterial DNA concentrations; (2) we applied a hybridization time of 18 h; (3) pooling for hybridization was done based on CFU/mL spike in concentrations (e.g., 8-plex pool of eight 10 CFU/mL spiked in samples); and (4) post-capture PCR ranged between 17 and 20 cycles based on pre-capture PCR concentrations.

Bacterial isolate libraries were prepared for whole genome sequencing using QIAseq FX DNA Library Kit (QIAGEN, Hilden, Germany), as described previously [9].

Each independent library preparation contained a no template control (NTC) containing molecular grade water only. Paired-end sequencing was performed on an Illumina NextSeq550 instrument (Illumina, San Diego, CA, USA) using NextSeq 300-cycle Mid Output Kit v2.5 (Illumina).

3.4. 16S rRNA Amplicon Sequencing

16S amplicon library prep was performed using the QIAseq 16S/ITS Screening Panel kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol. Paired-end sequencing was performed on an Illumina MiSeq instrument (Illumina) using MiSeq 600-cycle Reagent Kit v3 (Illumina).

3.5. Dry-Lab Workflow

3.5.1. 16S rRNA Data Analysis

Primer sequences were trimmed using cutadapt v2.6 [13], followed by read QC. Reads below 200 bp with a Phred score below 10 were removed. The DADA2 (1.18.0) [14] pipeline was used to construct amplicon sequence variant (ASV) tables with "filterAndTrim" parameters set to "truncLen = c(220, 180), maxN = 20, maxEE = c(3, 7)"; DADA2 defaults were used for all other parameters. Taxonomic classification was determined using the SILVA [15] rRNA reference database v138 via "assignTaxonomy" followed by "addSpecies", assigning species information based on exact sequence matching. For reporting, ASV counts were summarized per genus and hits with less than 1% abundance were considered background and removed.

3.5.2. Bioinformatics NGS Data Pipeline

The same NGS pipeline was used for the analysis of enriched as well as metagenomics sequenced samples. Read quality was evaluated with FastQC v0.11.9 and MultiQC v1.10.1 [16,17]; reads were trimmed and quality filtered using Trimmomatic v0.39 with parameters "LEADING:10 TRAILING:10 SLIDINGWINDOW:4:15 MINLEN:36" [18], followed by removal of duplicate reads using FastUniq v1.1 [19]. Meta-assembly of reads was performed using SAUTE v 1.3.0, included in the SKESA assembler package [20]. AMR marker detection was performed on assembled contigs using protein BLAST ncbi-blast v2.9.0+ with parameter "qcov_hsp_perc 60" and subsequent filtering of the blast hits to a minimal identity of 90% [21]. If several AMR proteins aligned to the same area on a contig, the marker with the best bitscore was selected. Because the aseptic material was found to be not sterile, i.e., the aseptic material contained AMR markers also present in strains used as spike-in, markers detected on negative controls and blanks of aseptic material were excluded from the comparison to avoid inflating sensitivity. The marker sets considering

background were separately defined for aseptic synovial fluid, aseptic female urine, and aseptic male urine, thus the ground truth may vary across material types.

Whole genome sequenced (reference) isolates from the LOD and septic urine cultures were processed as described previously [9]. Detected markers were used to determine the performance of the ARESdb AMR panel.

3.5.3. Bioinformatics NGS Data Analysis

The number of trimmed and deduplicated reads, percentage of on-target reads, enrichment factor, average depth, and reads per kilobase million (RPKM) distributions of true positive (TP) and false positive (FP) markers were used for comparison. Bowtie2 v2.3.5.1 with a mapping quality of 20 (Fast gapped-read alignment with Bowtie 2) and samtools v1.7 (The Sequence Alignment/Map format and SAMtools) were used to calculate the number of reads on target. The enrichment factor was calculated as the percentage of reads on target from the enriched sample divided by the percentage of reads on target of the metagenomics sequenced sample. The average depth of each marker was calculated using bowtie2 by local alignment followed by Genomecov, bedtools v2.29.0 [22]. The RPKM was calculated by counting reads per target sequence with htseq-count v 0.13.5 [23].

$$RPKM = \frac{\#reads\ on\ gene \times 1 \times 10^{-9}}{gene\ length \times \#total\ reads}$$

For marker comparison, markers detected by whole genome sequencing (WGS) were considered the ground truth. A true positive (TP) marker was defined as a marker detected both in the reference and in the sample; a false negative (FN) marker as marker detected in the reference and not detected in the sample; and a false positive (FP) marker as marker detected in the sample being analyzed and not present in the reference.

3.5.4. Aseptic Sample Background Marker Removal

AMR markers chromosomally present in genera detected by 16S sequencing of blank aseptic samples were removed. Removed markers were not considered for comparison. Because marker removal was performed for each type of aseptic material independently, the resulting number of reference markers varied for the same species.

4. Results

We tested the performance of target-enriched sequencing with the ARESdb AMR panel versus metagenomics sequencing in different aseptic body fluid samples spiked with four different reference strains at concentrations of 10, 100, 1000, 10,000, and 100,000 CFU/mL.

4.1. ARESdb AMR Panel Sensitivity and LOD for Aseptic Synovial Fluids

Metagenomics sequencing of aseptic synovial fluid samples at an average read count per sample of $19,245,414 \pm 3,903,155$ reads proved insensitive, not detecting any AMR markers, across the range of spike-in concentrations. In contrast, AMR target-enriched samples achieved sensitivities of $91\% \pm 6\%$ across all spiked-in reference strains at a concentration of 1000 CFU/mL with $3,679,296 \pm 846,251$ reads per sample. The *E. faecium* spike-in revealed a sensitivity of 86% already at 100 CFU/mL (Table 2).

The fraction of on-target reads increased with the spiked-in amounts of DNA from 6.4% (at 10 CFU/mL) to 40.8% (at 100,000 CFU/mL) across all spiked-in species (Table S1).

Table 2. Sensitivity of the ARESdb AMR panel versus metagenomics sequencing for spiked-in native materials.

| Species | [CFU/mL] | Sensitivity Synovial Fluid | | Sensitivity Male Urine | | Sensitivity Female Urine | |
|---------------------------|----------|----------------------------|------------------|------------------------|------------------|--------------------------|------------------|
| | | MG | ARESdb AMR Panel | MG | ARESdb AMR Panel | MG | ARESdb AMR Panel |
| <i>K. quasipneumoniae</i> | 10 | | 18%/5% | | 0%/0% | | 0%/4% |
| | 100 | | 53%/42% | 0%/0% | 4%/4% | | 44%/12% |
| | 1000 | 0%/0% | 97%/95% | | 40%/66% | 0%/0% | 96%/94% |
| | 10,000 | | 97%/97% | 8%/10% | 98%/98% | | 98%/98% |
| | 100,000 | | 97%/97% | 90%/98% | 98%/96% | 4%/6% | 98%/98% |
| <i>E. faecium</i> | 10 | | 33%/22% | 0%/0% | 6%/63% | | 67%/73% |
| | 100 | | 83%/89% | | 88%/94% | 0%/0% | 93%/93% |
| | 1000 | 0%/0% | 89%/89% | 25%/31% | 94%/94% | | 93%/93% |
| | 10,000 | | 94%/94% | 88%/94% | 94%/94% | 33%/53% | 93%/93% |
| | 100,000 | | 94%/94% | 94%/94% | 94%/94% | 93%/93% | 93%/93% |
| <i>S. aureus</i> | 10 | | 0%/0% | | 0%/0% | | 0%/0% |
| | 100 | | 21%/7% | | | | 8%/8% |
| | 1000 | 0%/0% | 86%/79% | 0%/0% | 100%/100% | 0%/0% | 100%/100% |
| | 10,000 | | 100%/100% | | 100%/100% | | 100%/100% |
| | 100,000 | | 100%/100% | 100%/100% | 100%/100% | | 100%/100% |
| <i>E. coli</i> | 10 | | 36%/45% | | 0%/1% | | 0%/0% |
| | 100 | | 73%/77% | 0%/0% | 0%/1% | | 36%/41% |
| | 1000 | 0%/0% | 95%/95% | | 58%/57% | 0%/0% | 97%/99% |
| | 10,000 | | 95%/95% | 1%/– | 99%/97% | | 97%/97% |
| | 100,000 | | 95%/95% | 78%/88% | 97%/97% | 1%/14% | 99%/99% |

Comparison of sensitivity for metagenomics sequencing (MG) and target-enriched sequencing (ARESdb AMR panel). Sensitivity is represented for performed duplicates (Duplicate 1/Duplicate 2).

4.2. ARESdb AMR Panel Sensitivity and LOD in Aseptic Urine

4.2.1. Male Urine

Metagenomics sequencing of aseptic male urine samples (Table 2) achieved an average sensitivity of $93\% \pm 7\%$ for AMR marker detection across all species at a concentration of 100,000 CFU/mL. AMR target enrichment sequencing outperformed metagenomics sequencing, achieving comparable or better sensitivities of on average $97\% \pm 2\%$ at already 10,000 CFU/mL for all reference strains. For two out of four reference strains, the sensitivity of target enrichment remained high down to 100 CFU/mL (91%, *E. faecium*) and 1000 CFU/mL (100%, *S. aureus*).

The percentage of reads on target for metagenomics sequencing increased across all species and concentrations from 0.004% to 0.225% compared with AMR target-enriched samples from 11.4% to 41.2% (Table S2). Of note, the percentage of reads on target obtained after enrichments depends on the enrichment and on the initial concentration of AMR marker target DNA. At 100,000 CFU/mL, true positive AMR markers were covered $99\% \pm 4\%$ by AMR target-enriched sequencing and $96\% \pm 9\%$ by metagenomics sequencing. For AMR target enrichment, the distribution of TP AMR marker depth (RPKM) was different between species, e.g., 2706 ± 1614 for *E. coli* and $26,046 \pm 18,372$ for *E. faecium* at 100,000 CFU/mL. For metagenomic sequencing, the distribution of TP AMR marker depth was on average 8 ± 4 for *E. coli* and 339 ± 419 for *E. faecium* at 100,000 CFU/mL (Figure S1).

4.2.2. Female Urine

Metagenomics sequencing of aseptic female urine samples (Table 2) revealed a sensitivity of 93% only for the *E. faecium* sample at the highest concentration of 100,000 CFU/mL. In all other samples, only single (on average, sensitivity of $7 \pm 22\%$) or no AMR markers were detected by metagenomics sequencing. In contrast, AMR target-enriched sequencing resulted in a sensitivity of on average $97 \pm 3\%$ at 1000 CFU/mL among all tested species. For *E. faecium*, the sensitivity of AMR target enrichment remained high (93%) down to 100 CFU/mL.

The percentage of reads on target for metagenomics sequencing increased across all species and concentrations from 0.015% to 0.034% compared with AMR target-enriched samples from 8.4% to 41.0% (Table S3). At 100,000 CFU/mL, true positive AMR markers were covered $99 \pm 3\%$ by AMR target-enriched sequencing and $93 \pm 11\%$ by metagenomics sequencing. For AMR target enrichment, the distribution of TP AMR marker depth (RPKM) was different between species, e.g., 2472 ± 1526 for *E. coli* and $26,694 \pm 15,585$ for *E. faecium* at 100,000 CFU/mL. For metagenomic sequencing, the distribution of TP AMR marker depth was on average 2 ± 1 for *E. coli* and 56 ± 54 for *E. faecium* at 100,000 CFU/mL samples (Figure S2).

4.3. Utility in Septic Urine Samples

For an initial proof-of-concept, sensitivity of the ARESdb AMR panel was tested against metagenomics sequencing and compared with whole genome sequenced isolates extracted from routine urine culture. The set of 13 septic urine samples spanned a total of six different species. Sample processing revealed high sensitivity of on average $99 \pm 1\%$ for both sequencing approaches. AMR target enrichment outperformed metagenomics sequencing by achieving on average 61% versus 1.5% on-target reads (Tables 1 and 3). The coverage of TP AMR targeted markers by capture-based sequencing and metagenomics sequencing was on average $99 \pm 4\%$. The distribution of target TP marker depth (RPKM) showed differences across septic samples with an average of 3162 ± 3355 for AMR target enrichment, while for metagenomics, the average depth obtained was 85 ± 35 (Figure S3).

Table 3. Performance overview for septic urine samples with confirmed mono-infection.

| Sample ID | Species | WGS Isolate | Markers | | Sensitivity | | Reads on Target | | Enrichment Factor |
|-----------|-------------------------|-------------|---------|------------------|-------------|------------------|-----------------|------------------|-------------------|
| | | | MG | ARESdb AMR Panel | MG | ARESdb AMR Panel | MG | ARESdb AMR Panel | ARESdb AMR Panel |
| ID-1 | <i>E. coli</i> | 85 | 84/84 | 84/84 | 99%/99% | 99%/99% | 1.44%/1.45% | 67%/68% | 47/47 |
| ID-2 | <i>E. coli</i> | 107 | 106/106 | 106/106 | 99%/99% | 99%/99% | 2.32%/2.67% | 60%/60% | 26/23 |
| ID-3 | <i>E. coli</i> | 104 | 102/103 | 103/102 | 98%/99% | 99%/98% | 1.50%/1.45% | 65%/65% | 43/45 |
| ID-4 | <i>C. koseri</i> | 45 | 43/43 | 43/42 | 96%/96% | 96%/93% | 0.82%/0.80% | 34%/35% | 41/44 |
| ID-5 | <i>E. hormaechei</i> | 47 | 46/46 | 46/46 | 98%/98% | 98%/98% | 1.38%/1.38% | 62%/62% | 45/45 |
| ID-6 | <i>E. coli</i> | 119 | 118/118 | 118/118 | 99%/99% | 99%/99% | 2.30%/2.29% | 60%/61% | 26/27 |
| ID-7 | <i>E. faecalis</i> | 8 | 8/8 | 8/8 | 100%/100% | 100%/100% | 0.80%/0.62% | 73%/73% | 92/118 |
| ID-8 | <i>E. coli</i> | 89 | 87/87 | 87/87 | 98%/98% | 98%/98% | 2.59%/2.64% | 69%/69% | 26/26 |
| ID-9 | <i>K. pneumoniae</i> | 51 | 50/50 | 50/50 | 98%/98% | 98%/98% | 1.14%/1.11% | 52%/52% | 45/47 |
| ID-10 | <i>E. coli</i> | 91 | 90/90 | 90/90 | 99%/99% | 99%/99% | 1.57%/1.48% | 61%/60% | 39/41 |
| ID-11 | <i>Enterococcus</i> sp. | 9 | 9/9 | 9/9 | 100%/100% | 100%/100% | 0.71%/0.72% | 69%/70% | 97/97 |
| ID-12 | <i>E. coli</i> | 109 | 108/108 | 108/107 | 99%/99% | 99%/98% | 1.48%/1.48% | 58%/58% | 39/39 |
| ID-13 | <i>E. coli</i> | 101 | 100/100 | 100/100 | 99%/99% | 99%/99% | 1.96%/1.86% | 67%/67% | 34/36 |
| AVERAGE | | 74 | 73 | 73 | 99% | 99% | 1.54% | 61% | 47 |

Comparison of performance parameters for metagenomics sequencing (MG) and target-enriched sequencing (ARESdb AMR panel). The number of ground truth markers was defined based on the whole genome sequence of the bacterial isolate (WGS isolate). Samples were processed as duplicates and represented as (Duplicate 1/Duplicate 2).

State-of-the-art 16S rRNA sequencing directly from septic urine achieved 100% sensitivity among all patient samples. For all except one processed duplicate, the top genus detected was in concordance with the initial positive culture result. Duplicate 1 of patient ID-5 revealed the correct top genus (*Enterobacter* spp.), while duplicate 2 revealed *Enterobacter* spp. as third and *Klebsiella* spp. as most abundant genus.

5. Discussion

Rapid pathogen identification combined with AMR marker profiling directly from native samples is critical to unburden infectious disease management, antibiotic stewardship, and monitor trends in antibiotic resistance [24]. Despite the fact that current de-facto standard methods like PCR are highly sensitive, their target space is limited [8]. Previous

studies have examined the utility of metagenomics sequencing in native sample types to overcome the limitations of bacterial culture [25–29]. Studies focusing on metagenomics sequencing of prosthetic joint infections have shown that the high proportion of human DNA background (>90%) represents a challenge for taxonomic identification, impeding sensitive antimicrobial resistance profiling [30–32]. Deep metagenomics sequencing has been shown to partially overcome host background and to sensitively recover microbial pathogens. However, compared with pathogen identification, considerably increased genome coverage is required to sensitively and reproducibly detect sparse genetic determinants of AMR. Therefore, routine use is hampered by sequencing depth and the associated sequencing costs [26].

Target capture-based methods have thus been proposed as a potentially viable approach for the identification of resistomes in complex native specimens [5,33]. We have selected over 9200 targets from the AMR biomarker database ARESdb and developed a target enrichment panel and NGS workflow, suitable for the processing and sensitive AMR profiling of complex sample types.

Our primary objective was to compare the sensitivity of target enrichment-based AMR marker detection to metagenomics sequencing in synovial fluid and urine samples. In the present study, the ARESdb AMR panel achieved sensitivities exceeding $91\% \pm 6\%$ at ≥ 1000 CFU/mL for AMR marker detection across all tested species for spiked-in synovial fluid samples. In addition, we tested the ARESdb AMR panel among healthy male and female urine samples separately to investigate gender-specific background differences of the urine microbiota [34]. Among both male and female spiked-in urine samples, targeted sequencing outperformed metagenomics sequencing by achieving sensitivities exceeding 90% already at 1000 CFU/mL for all species in female and for *E. faecium* and *S. aureus* in male urine spiked-in samples. The LOD of the ARESdb AMR panel is sufficient for the analysis of complicated UTI (classified at pathogen loads $>100,000$ CFU/mL) and other UTI classifications including catheter associated UTIs, which are associated with lower pathogen loads of ≥ 1000 CFU/mL [35,36].

Moreover, we demonstrated the clinical relevance of the ARESdb AMR panel by a proof-of-concept study on a set of septic urine samples. Characterized clinical samples with confirmed monoinfection and available culture results were processed to investigate the sensitivity of the AMR panel. Clinical data showed that the bacterial load in septic samples exceeded 10,000 CFU/mL. As a consequence, high sensitivity was achieved by AMR target enrichment sequencing as well as by deep metagenomics sequencing. Previous studies [27,29] suggested that metagenomics sequencing of samples originating from UTI is feasible for taxonomic identification and AMR profiling. However, according to the data presented here, the number of reads covering AMR markers was significantly higher for target enrichment sequencing compared with metagenomics sequencing (61% versus 1.5%) at comparable numbers of total reads per sample ($15,287,091 \pm 6250$ versus $18,997,857 \pm 4,996,755$). This indicates that the ARESdb AMR panel is a suitable application for lower sequencing depths. In addition, to complement the established AMR target enrichment workflow with taxonomic identification, we applied state-of-the-art 16S rRNA sequencing, which was used to confirm the initial positive culture results (100% sensitivity). Of note, the current workflow is limited by its ability to link AMR markers with individual species in a multi-infection sample. This important limitation is neither addressed by molecular methods such as metagenomics sequencing or PCR nor the capture panel presented herein and will require further technological improvements.

6. Conclusions

The newly developed ARESdb AMR panel together with the established data analysis pipeline allowed to sensitively screen for >9200 AMR markers from samples with low concentrations of microbial DNA relative to human host DNA. For LOD experiments, AMR target enrichment achieved sensitivities >90% across all spiked-in reference strains at a concentration of 1000 CFU/mL. Our method was shown to be superior to metagenomics

sequencing in the LOD experiments performed, sensitively capturing the resistome at lower sequencing capacity. In addition, culture-free AMR marker detection and taxonomic identification was demonstrated on a set of septic urine samples with high sensitivity (99%). Summarizing, our findings indicate the ARESdb AMR panel can be a valuable tool for comprehensive and sensitive AMR marker detection directly from body fluids.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/microorganisms9081672/s1>, Figure S1: Marker coverage and depth of metagenomics and target enriched male urine samples; Figure S2: Marker coverage and depth of metagenomics and target enriched female urine samples; Figure S3: Marker coverage and depth of metagenomics and target enriched septic urine samples; Table S1: Performance overview for aseptic synovial fluid samples spiked with four different pathogens; Table S2. Performance overview for aseptic male urine samples spiked with four different pathogens; Table S3. Performance overview for aseptic female urine samples spiked with four different pathogens.

Author Contributions: Conceptualization, S.L., S.B., A.M., J.G.H., A.E.P. and J.W.; Methodology, I.F., S.L., S.B., G.F., K.M. and J.W.; Data Analysis, I.F., S.B., A.v.H. and J.W.; Resources, B.J.H.F., S.H., M.A.K. and J.G.H.; Data Curation, I.F.; Writing—Original Draft Preparation, I.F., S.L. and J.W.; Writing—Review & Editing, S.B., B.J.H.F., S.H., M.A.K., A.v.H., A.M., J.G.H., A.E.P. and J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Vienna Business Agency (grant number 24478239) as well as the Austrian Research Promotion Agency (FFG) (grant numbers 871243, 879570).

Institutional Review Board Statement: Septic urine samples for this study were residues of routinely processed samples and collected with approval of the ethics committee of the Innsbruck Medical University (283/4.1). For synovial fluid samples, the study was approved by the Austrian, Vienna Vinzenz group Ethical Review Board (EK10/2020).

Informed Consent Statement: As the samples are anonymized and no patient data except for sex, type of urinary tract infection and antibiotic treatment are included, the information is not identifying and no written consent from the patients was necessary.

Data Availability Statement: NGS datasets generated for this study can be found in the DDBJ/EMBL/GenBank; accession PRJNA743171.

Conflicts of Interest: I.F., S.L., S.B., G.F., K.M., A.M., A.E.P. and J.W. are employees of Ares Genetics GmbH. All other authors declare no conflict.

References

1. WHO. *GLASS Whole-Genome Sequencing for Surveillance of Antimicrobial Resistance*; WHO: Geneva, Switzerland, 2020.
2. Rochford, C.; Sridhar, D.; Woods, N.; Saleh, Z.; Hartenstein, L.; Ahlawat, H.; Whiting, E.; Dybul, M.; Cars, O.; Goosby, E.; et al. Global governance of antimicrobial resistance. *Lancet* **2018**, *391*, 1976–1978. [[CrossRef](#)]
3. WHO. *Global Tuberculosis Report 2018*; World Health Organization: Geneva, Switzerland, 2018.
4. Livermore, D.M. The need for new antibiotics. *Clin. Microbiol. Infect. Suppl.* **2004**, *10*, 1–9. [[CrossRef](#)]
5. Guitor, A.K.; Raphenya, A.R.; Klunk, J.; Kuch, M.; Alcock, B.; Surette, M.G.; McArthur, A.G.; Poinar, H.N.; Wright, G.D. Capturing the resistome: A targeted capture method to reveal antibiotic resistance determinants in metagenomes. *Antimicrob. Agents Chemother.* **2020**, *64*, e01324-19. [[CrossRef](#)]
6. Quan, J.; Langelier, C.; Kuchta, A.; Batson, J.; Teyssier, N.; Lyden, A.; Caldera, S.; McGeever, A.; Dimitrov, B.; King, R.; et al. FLASH: A next-generation CRISPR diagnostic for multiplexed detection of antimicrobial resistance sequences. *Nucleic Acids Res.* **2019**, *47*, e83. [[CrossRef](#)] [[PubMed](#)]
7. Pitt, W.G.; Alizadeh, M.; Hussein, G.A.; McClellan, D.S.; Buchanan, C.M.; Bledsoe, C.G.; Robison, R.A.; Blanco, R.; Roeder, B.L.; Melville, M.; et al. Rapid separation of bacteria from blood-review and outlook. *Biotechnol. Prog.* **2016**, *32*, 823–839. [[CrossRef](#)] [[PubMed](#)]
8. WHO. *Molecular Methods for Antimicrobial Resistance (AMR) Diagnostics to Enhance the Global Antimicrobial Resistance Surveillance System*; WHO: Geneva, Switzerland, 2019.
9. Lepuschitz, S.; Weinmaier, T.; Mrazek, K.; Beisken, S.; Weinberger, J.; Posch, A.E. Analytical Performance Validation of Next-Generation Sequencing Based Clinical Microbiology Assays Using a K-mer Analysis Workflow. *Front. Microbiol.* **2020**, *11*, 1883. [[CrossRef](#)] [[PubMed](#)]
10. Nadkarni, M.A.; Martin, F.E.; Jacques, N.A.; Hunter, N. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiology* **2002**, *148*, 257–266. [[CrossRef](#)] [[PubMed](#)]

11. Funakoshi, K.; Bagheri, M.; Zhou, M.; Suzuki, R.; Abe, H.; Akashi, H. Highly sensitive and specific Alu-based quantification of human cells among rodent cells. *Sci. Rep.* **2017**, *7*, 1–12. [[CrossRef](#)]
12. Ferreira, I.; Beisken, S.; Lueftinger, L.; Weinmaier, T.; Klein, M.; Bacher, J.; Patel, R.; von Haeseler, A.; Posch, A.E. Species identification and antibiotic resistance prediction by analysis of whole-genome sequence data by use of ARESdb: An analysis of isolates from the unyvero lower respiratory tract infection trial. *J. Clin. Microbiol.* **2020**, *58*, e00273–20. [[CrossRef](#)]
13. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMB J.* **2011**, *17*, 10. [[CrossRef](#)]
14. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13*, 581–583. [[CrossRef](#)]
15. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, D590–D596. [[CrossRef](#)]
16. Andrews, S.; Krueger, F.; Seconda-Pichon, A.; Biggins, F.; Wingett, S. Babraham Bioinformatics-FastQC a Quality Control Tool for High Throughput Sequence Data. Available online: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed on 4 August 2021).
17. Ewels, P.; Magnusson, M.; Lundin, S.; Käller, M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **2016**, *32*, 3047–3048. [[CrossRef](#)]
18. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
19. Xu, H.; Luo, X.; Qian, J.; Pang, X.; Song, J.; Qian, G.; Chen, J.; Chen, S. FastUniq: A Fast De Novo Duplicates Removal Tool for Paired Short Reads. *PLoS ONE* **2012**, *7*, 52249. [[CrossRef](#)] [[PubMed](#)]
20. Souvorov, A.; Agarwala, R.; Lipman, D.J. SKESA: Strategic k-mer extension for scrupulous assemblies. *Genome Biol.* **2018**, *19*, 153. [[CrossRef](#)]
21. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
22. Quinlan, A.R.; Hall, I.M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010**, *26*, 841–842. [[CrossRef](#)] [[PubMed](#)]
23. Anders, S.; Pyl, P.T.; Huber, W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015**, *31*, 166–169. [[CrossRef](#)]
24. Shi, X.; Kadiyala, U.; Vanepps, J.S.; Yau, S.T. Culture-free bacterial detection and identification from blood with rapid, phenotypic, antibiotic susceptibility testing. *Sci. Rep.* **2018**, *8*, 1–11. [[CrossRef](#)]
25. Greninger, A.L.; Messacar, K.; Dunnebacke, T.; Naccache, S.N.; Federman, S.; Bouquet, J.; Mirsky, D.; Nomura, Y.; Yagi, S.; Glaser, C.; et al. Clinical metagenomic identification of Balamuthia mandrillaris encephalitis and assembly of the draft genome: The continuing case for reference genome sequencing. *Genome Med.* **2015**, *7*, 1–14. [[CrossRef](#)] [[PubMed](#)]
26. Thoendel, M.; Jeraldo, P.; Greenwood-Quaintance, K.E.; Chia, N.; Abdel, M.P.; Steckelberg, J.M.; Osmon, D.R.; Patel, R. A Novel Prosthetic Joint Infection Pathogen, Mycoplasma salivarium, Identified by Metagenomic Shotgun Sequencing. *Clin. Infect. Dis.* **2017**, *65*, 332–335. [[CrossRef](#)] [[PubMed](#)]
27. Hasman, H.; Saputra, D.; Sicheritz-Ponten, T.; Lund, O.; Svendsen, C.A.; Frimodt-Moller, N.; Aarestrup, F.M. Rapid whole-genome sequencing for detection and characterization of microorganisms directly from clinical samples. *J. Clin. Microbiol.* **2014**, *52*, 139–146. [[CrossRef](#)]
28. Joensen, K.G.; Engsbjerg, A.L.; Lukjancenko, O.; Kaas, R.S.; Lund, O.; Westh, H.; Aarestrup, F.M. Evaluating next-generation sequencing for direct clinical diagnostics in diarrhoeal disease. *Eur. J. Clin. Microbiol. Infect. Dis.* **2017**, *36*, 1–14. [[CrossRef](#)]
29. Schmidt, K.; Mwaigwisya, S.; Crossman, L.C.; Doumith, M.; Munroe, D.; Pires, C.; Khan, A.M.; Woodford, N.; Saunders, N.J.; Wain, J.; et al. Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *J. Antimicrob. Chemother.* **2017**, *72*, 104–114. [[CrossRef](#)] [[PubMed](#)]
30. Thoendel, M.J.; Jeraldo, P.R.; Greenwood-Quaintance, K.E.; Yao, J.Z.; Chia, N.; Hanssen, A.D.; Abdel, M.P.; Patel, R. Identification of prosthetic joint infection pathogens using a shotgun metagenomics approach. *Clin. Infect. Dis.* **2018**, *67*, 1333–1338. [[CrossRef](#)] [[PubMed](#)]
31. Sanderson, N.D.; Street, T.L.; Foster, D.; Swann, J.; Atkins, B.L.; Brent, A.J.; McNally, M.A.; Oakley, S.; Taylor, A.; Peto, T.E.A.; et al. Real-time analysis of nanopore-based metagenomic sequencing from infected orthopaedic devices 06 Biological Sciences 0604 Genetics 06 Biological Sciences 0605 Microbiology. *BMC Genomics* **2018**, *19*, 1–11. [[CrossRef](#)]
32. Street, T.L.; Sanderson, N.D.; Atkins, B.L.; Brent, A.J.; Cole, K.; Foster, D.; McNally, M.A.; Oakley, S.; Peto, L.; Taylor, A.; et al. Molecular diagnosis of orthopedic-device-related infection directly from sonication fluid by metagenomic sequencing. *J. Clin. Microbiol.* **2017**, *55*, 2334–2347. [[CrossRef](#)]
33. Gnirke, A.; Melnikov, A.; Maguire, J.; Rogov, P.; LeProust, E.M.; Brockman, W.; Fennell, T.; Giannoukos, G.; Fisher, S.; Russ, C.; et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat. Biotechnol.* **2009**, *27*, 182–189. [[CrossRef](#)]
34. Gottschick, C.; Deng, Z.L.; Vital, M.; Masur, C.; Abels, C.; Pieper, D.H.; Wagner-Döbler, I. The urinary microbiota of men and women and its changes in women during bacterial vaginosis and antibiotic treatment. *Microbiome* **2017**, *5*, 99. [[CrossRef](#)]

-
35. Nicolle, L.E.; Evans, G.; Laverdieve, M.; Phillips, P.; Quan, C.; Rotstein, C. Complicated urinary tract infection in adults. *Can. J. Infect. Dis. Med. Microbiol.* **2005**, *16*, 349–360. [[CrossRef](#)] [[PubMed](#)]
 36. Bonkat, G.; Pickard, R.; Bartoletti, R. *EAU Guidelines on Urological Infections*; EAU Guidelines Office: Arnhem, The Netherlands, 2019.