



Review

Combination of Whole Genome Sequencing and Metagenomics for Microbiological Diagnostics

Srinithi Purushothaman ^{1,2} , Marco Meola ^{1,2,3} and Adrian Egli ^{1,2,4,*}

¹ Applied Microbiology Research, Department of Biomedicine, University of Basel, 4031 Basel, Switzerland

² Institute of Medical Microbiology, University of Zurich, 8006 Zurich, Switzerland

³ Swiss Institute of Bioinformatics, University of Basel, 4031 Basel, Switzerland

⁴ Clinical Bacteriology and Mycology, University Hospital Basel, 4031 Basel, Switzerland

* Correspondence: aegli@imm.uzh.ch; Tel.: +41-44-634-27-00

Abstract: Whole genome sequencing (WGS) provides the highest resolution for genome-based species identification and can provide insight into the antimicrobial resistance and virulence potential of a single microbiological isolate during the diagnostic process. In contrast, metagenomic sequencing allows the analysis of DNA segments from multiple microorganisms within a community, either using an amplicon- or shotgun-based approach. However, WGS and shotgun metagenomic data are rarely combined, although such an approach may generate additive or synergistic information, critical for, e.g., patient management, infection control, and pathogen surveillance. To produce a combined workflow with actionable outputs, we need to understand the pre-to-post analytical process of both technologies. This will require specific databases storing interlinked sequencing and metadata, and also involves customized bioinformatic analytical pipelines. This review article will provide an overview of the critical steps and potential clinical application of combining WGS and metagenomics together for microbiological diagnosis.



Citation: Purushothaman, S.; Meola, M.; Egli, A. Combination of Whole Genome Sequencing and Metagenomics for Microbiological Diagnostics. *Int. J. Mol. Sci.* **2022**, *23*, 9834. <https://doi.org/10.3390/ijms23179834>

Academic Editors: Gary A. Toranzos and Tasha M. Santiago-Rodríguez

Received: 4 July 2022

Accepted: 26 August 2022

Published: 30 August 2022

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



Copyright: © 2022 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/>).

Keywords: whole genome sequencing; metagenomics; epidemiology; surveillance; antimicrobial resistance; transmission; database; bioinformatics; combination

1. Introduction

Three questions guide diagnostics in clinical microbiology: (i) Species identification: “who is there?”, (ii) biological functionality: “what are microorganisms doing?”, and (iii) interaction between microorganisms and the host: “are they linked?” [1]. These questions allow the assessment of different clinical scenarios using specific culture- or genome-based technologies. Culture-based microbiological diagnostic assays usually provide reliable species identification and allow the detection of mixed infections on agar plates. Culture also delivers established phenotypic readouts such as antibiotic susceptibility [2]. However, important shortcomings include the following: (i) Most phenotypic methods are time-consuming, requiring up to 72 h to obtain results [3,4]; (ii) many microbes are fastidious or non-culturable, thereby remain undetected, introducing a diagnostic bias [5]; (iii) information on virulence is often scarce due to a lack of standardized phenotypic readouts; and (iv) phenotypes for antimicrobial resistance profiling can be influenced by numerous parameters, e.g., type of agar plates, age of colonies, and environmental factors such as temperature and pH [6]. Genome-based diagnostics may help to overcome some of these limitations.

Since the first fully sequenced bacterial genome of *Haemophilus influenzae* type B became available in 1995 [7], sequencing technologies have rapidly evolved and are now used in patient care and infection control management [8]. Whole genome sequencing (WGS) and metagenomics use different approaches to determine the genomic content in a sample. WGS aims to analyze the whole genome of a single bacterial colony, while amplicon-based marker gene sequencing (e.g., 16S/ITS) or shotgun metagenomics focuses

on microbial communities within a sample, usually without culture [9,10]. The costs per sequenced nucleotide have substantially decreased during the past decade [11,12] due to the expansion of sequencing capacities, the development of cost-effective technologies, advances in laboratory automation, and the progression of standardized workflows [13,14]. Today, knowledge and utilization have increased to a point where WGS can be applied in clinical microbiological diagnostics and surveillance not only in high-resource laboratories, but also in limited-resource environments [15–17].

Successfully linking the genotype to the phenotype for clinical applications requires a profound understanding of the diagnostic process, for example, when and how linking both types of information is appropriate. Sequencing capacity was further boosted during the COVID-19 pandemic [10,18], with more than 13.9 million SARS-CoV-2 genomes sequenced and made publicly available (www.covid19dataportal.org; accessed on 9 August 2022). This massive sequencing effort also resulted in a high degree of standardized analytical protocols and increased awareness of quality control of the sequencing data [19].

The choice of the sequencing approach (WGS vs. metagenomics) is dependent on the clinical question and demand, e.g., rapid result, need for acute management vs. high-resolution result for typing. In addition, the available sample type influences the choice of technology. Figure 1 provides a decision flow chart for selecting the suitable sequencing technology.

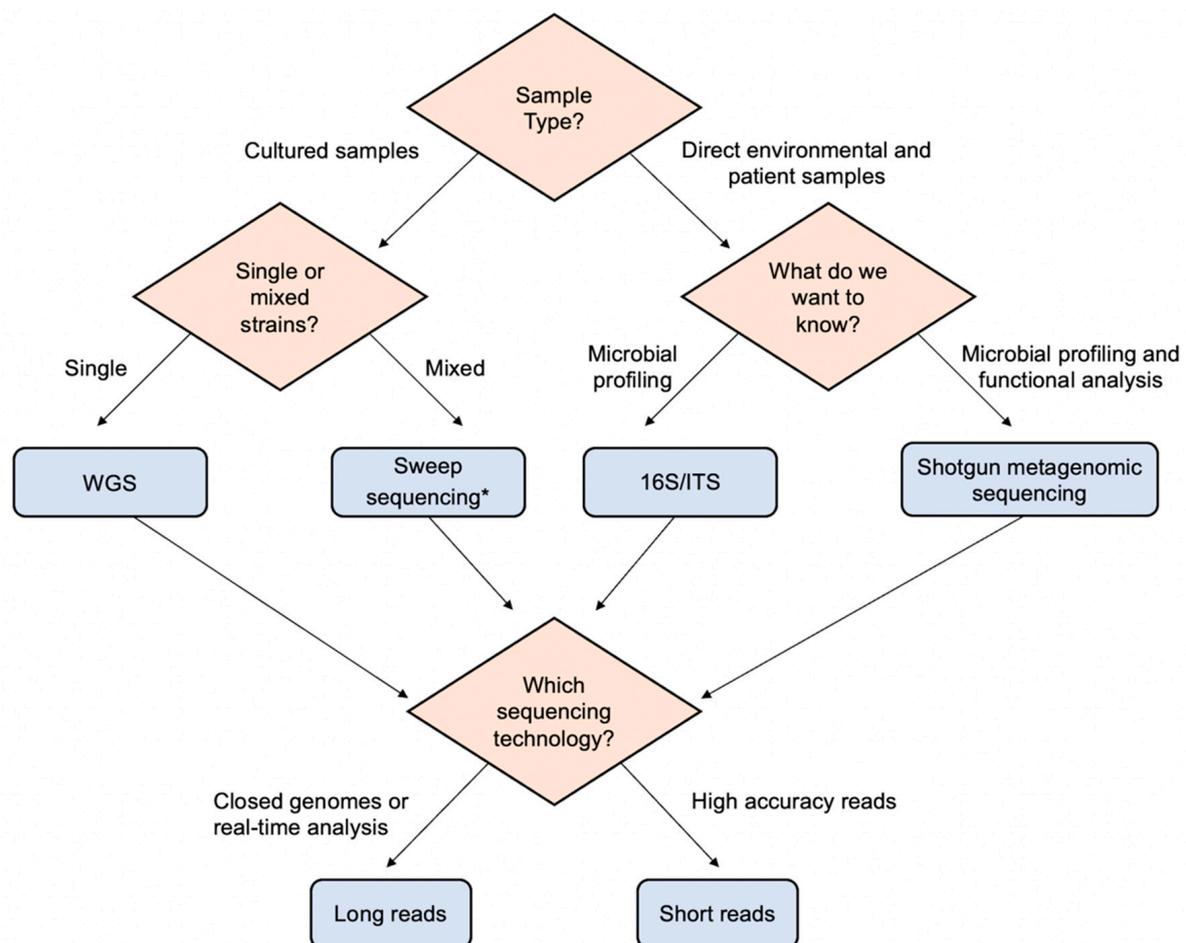


Figure 1. Decision tree for the selection of sequencing technology. The type of samples and research questions guide the selection of suitable sequencing strategies and sequencing platforms. The diamonds represent the checkpoints/questions, and the blue rectangles represent the sequencing strategies. * Sweep sequencing is a technique in which concurrent sequencing of multiple colonies at the same time is possible—this could also be considered a metagenomic approach.

With increasing numbers of available WGS and metagenomic datasets, the question arises as to whether a combination of both potentiates the technical resolution and the information produced for disentangling the microbial communities present in various environments and their transmission between them. While WGS still largely generates data from isolates grown on plates, culture-independent 16S metagenomics drastically reduces the time required for species identification, thanks to the more rapid and direct marker-gene sequencing and available high-quality databases. In the case of pathogens, WGS can elucidate risk potentials such as strain relatedness for outbreak detection, as well as the presence of genes potentially encoding antimicrobial resistance (AMR). Shotgun metagenomics can combine the advantages of both these methods and also shed light on the pathogens circulating in the environment. However, this comes with high costs, low standardization, and limited sequencing depth, often resulting in fragmented information from single bacterial genomes [20], and challenging bioinformatics hampers its application in clinics so far. A successful combination may allow the application of new statistical approaches such as data mining and enable the exploration of individual pathogens and host–pathogen interactions, as well as the complex interplay of microbial communities. This could pave the way for a more personalized risk assessment of colonizing or virulent pathogens and the transmission dynamics of microbes between different environments. How could a combined approach be applied to healthcare? What are the technical and analytical requirements? What are the potential use cases?

In this review, we will try to answer and discuss these critical questions.

2. Focus on Individual Microorganisms

In bacteriology and mycology routine diagnostics, WGS is most commonly applied to single bacterial isolates [21]. Nowadays, with the help of advanced molecular biology techniques, difficult-to-culture microorganisms can also be sequenced using WGS. For example, the sequencing of *Mycobacterium tuberculosis* from liquid media using the Mycobacteria Growth Indicator Tube (MGIT) [22] has been established, although with a slower turnaround time compared to targeted DNA enrichment directly from sputum [23]. Viruses or intracellular bacteria can be amplified or enriched using (i) pathogen-specific PCRs (e.g., for SARS-CoV-2 or Influenza) [24,25] or (ii) bead-based DNA pull-down assays, e.g., for *Chlamydia trachomatis* [26], human papillomaviruses [27], or Noroviruses [28]. Sweeps from solid agar plates with multiple mixed species can also be used, as shown for the mGEMS and mSweep bioinformatic pipelines, which were validated with *Escherichia coli*, *Enterococcus faecalis*, and *Staphylococcus aureus* [29].

Quality control. After sequencing, the typical analysis workflow includes quality assessment of the sequenced data (raw fastq files), followed by preprocessing of the sequenced reads (adapter trimming and filtering low-quality reads) and assembly [30–32]. The quality control (QC) monitors the following critical parameters: Read accuracy (Q30 score to measure the probability of incorrect base calling), genome coverage, genome completeness, and the number of contigs [33]. QC is followed by the assembly of the reads to generate contigs and possibly circularize the genome. Assembly is either performed de novo, against a reference genome map, or as a split k-mer analysis [34], which all may influence the subsequent typing resolution. Short-read and long-read sequencing can be used together to form hybrid assemblies, which allow the generation of high-quality genomes of single pathogens [35,36]. GC biases (GC-poor and GC-rich regions), which arise due to the genomic composition of the microorganisms, lead to uneven coverage during sequencing and might affect the resulting assemblies [37,38]. The library preparation techniques used for sequencing can also impact the assemblies [39,40]. In addition, contamination from different or even the same species needs to be excluded to ensure the reliability of the downstream analyses [41]. After assembly of the raw data, it is difficult to attribute the cause of errors. Specific QC pipelines such as AQUAMIS (Assembly-based Quality Assessment for Microbial Isolate Sequencing) allow for automating this process [42].

Identification. The assembled genomes are used to identify microbes at species or subspecies levels. Genomic-based species annotation of assembled genomes appears trivial at first but can have unexpected caveats for certain species, even for well-known pathogens [43,44]. The bacterial taxonomy is ruled by the International Code of Nomenclature of Prokaryotes (ICNP) under the supervision of the International Committee on Systematics of Prokaryotes (ICSP) based on different phylogenomic approaches [45]. The extent of sequencing and the consequent possibility for a higher-resolution distinction between the genus and species have sparked new discussions on a revision of the ICNP [46,47]. Therefore, a crucial aspect is correct annotation with well-curated and internationally accepted databases to reliably identify a bacterial species.

The Type Strain Genome Server (TYGS, (<https://www.dsmz.de/services/online-tools/tygs>, accessed on 31 May 2022) is a particularly well-curated database, and the accompanying software offers the possibility to identify potential new bacterial species [48]. The GTDB (Genome Taxonomy Database) is another highly curated, phylogenetically consistent, and genome-based taxonomy database for annotation, backed by GTDB-Tk software for genome annotations [49]. Despite these well-curated databases, identifying a bacterial species from genomic data can still be challenging. For example, currently available software tools such as PubMLST, MetaPhlan3, and Mykrobe-predictor showed variable performances in correctly identifying non-tuberculous mycobacterial species. Sensitivities ranged from 57–100% and specificities from 83–98%, which could be attributed to the different databases and algorithms used [50].

Molecular epidemiology. WGS has become the reference standard for microbial typing to address epidemiological questions. Increasingly standardized workflows and quality management have been established [51,52]. Most studies use genome comparison to a reference (mapping) or within-cluster mapping with either core genome (cg) multi-locus sequence typing (MLST), SNP-based comparison, or more recently, also split k-mer analysis (SKA) [53–56]. Pairwise genome comparisons using SKA showed a higher resolution compared to cgMLST of *Enterococcus faecium* [34]. Outbreak investigation and transmission studies benefit from large publicly available genome datasets to compare potentially related strains with non-outbreak-associated isolates. As an example, the NCBI pathogen browser covers a selection of 40 bacterial and fungal species with more than 1 million available genomes (accessed on 31 May 2022; [57]). Similarly, the Swiss Pathogen Surveillance Platform (www.SPSP.ch; accessed on 31 May 2022 [58]), the Eukaryotic Pathogen, Vector, and Host Informatics Resource (VEuPathDB) (<https://veupathdb.org/veupathdb/app/static-content/about.html>, accessed on 4 August 2022), or the European Nucleotide Archive (<https://www.ebi.ac.uk/ena>, accessed on 31 May 2022) contain genomic data on viral, fungal, and bacterial pathogens, which allow epidemiological studies with important epidemiological metadata. Outbreak analyses in hospitals often focus on antibiotic-resistant pathogens, and the value of WGS has been documented in many instances for the transmission of bacterial strains [59–63] or plasmids with multi-drug resistance genes [64].

Inference of function. Assembled genomes from WGS can be used to infer phenotypes, such as AMR and virulence [65,66]. Curated databases such as ResFinder [67] and CARD [68] are commonly used for the detection of AMR genes. The combination of highly curated databases and underlying algorithms plays a major role in prediction accuracy [69,70]. For example, the concordance of Mykrobe-based (<https://www.mykrobe.com>, accessed on 31 May 2022) AMR prediction with phenotypic testing was 94% in *Mycobacterium tuberculosis* compared to TB-profiler, MTBseq, and other benchmarked tools [71]. Single-nucleotide polymorphisms (SNPs) can be associated with a loss or gain of function. For instance, SNPs were conclusively linked to functional impairment of porins within the species *P. aeruginosa* (*oprD*; [72]) and de-regulating feedback loops affecting transcription factors were linked to beta-lactamase expression (*ampR* and *ampC*; [73,74]).

A promising new approach for identifying specific functional links is Genome-Wide Association Studies (GWAS), which have recently been translated from human genetics to microbiology [75,76]. A GWAS identifies genes, k-mers, and/or SNPs enriched in

microorganisms linked to a particular phenotype. Examples include daptomycin resistance in *Staphylococcus aureus* linked to mutations in *mprF* [77], and clinical phenotypes such as invasiveness following urinary tract infection linked to the adherence factor *papGII* of *E. coli* [78,79]. Newly developed bioinformatic GWAS pipelines such as PowerBacGWAS provide power calculations to determine statistically significant sample size association testing [80]. Clinical validation of the identified genetic markers in randomized controlled trials, similar to any other clinically used biomarker, is necessary to assess their clinical value. In the next few years, we expect that more bacterial genetic markers will lead to diagnostic applications, e.g., virulence assessment or AMR surveillance [81–83].

3. Focus on Bacterial Communities

When sequencing bacterial communities, two metagenomic sequencing approaches can be distinguished, amplicon-based and shotgun metagenomics. The former targets a marker-gene or a segment thereof, which allows for the resolution of the bacterial community structure. The latter sequences/covers representative genetic material of a specimen, usually including the DNA of the host.

Identification with amplicon-based sequencing. In recent years, amplicon-based short-read sequencing of marker genes, in particular 16S rRNA gene (16S) and its variable regions (V1–V9), became highly popular for explorative studies in ecology research. Alternative universal marker genes include the bacterial *rpoB* or the fungal internal transcribed spacer 1/2 (ITS-1/2) [84]. 16S sequencing has also been applied for clinical diagnostics, e.g., in abscesses [85], urinary tract infections [85,86], or sepsis [87], as well as for environmental studies. Oberaune et al. profiled the microbiome of intensive care units (ICUs) and found higher microbial diversity compared to culture-dependent techniques [88].

For taxonomic profiling, the 16S sequence reads are assigned to representative sequences, such as operational taxonomic units (OTUs) through clustering [89] or to amplicon sequence variances (ASVs) through a denoising algorithm [90–92]. There are various OTU clustering algorithms available [93–95]. Imprecisely clustered sequences can give rise to inaccurate OTU classifications, which has a significant impact on downstream analyses. Therefore, alternatives for OTUs have been suggested, such as pairwise alignment sequence dissimilarity (PSD), MSA-based sequence dissimilarity (MSD), and phylogenetic branch length distance (BLD) [96].

For species identification, single variable regions are usually not suitable [97], and some variable regions such as V7 are known to yield ambiguous identifications [98]. However, bioinformatics tools allow the combination of individually sequenced variable regions from one strain, thereby delivering species-level resolution within samples [99,100]. A recent promising approach uses 16S–23S de novo assembled sequencing data and a Basic Local Alignment Search Tool (BLAST) approach with a newly developed database for species identification [101]. Recently, long-read-based full rRNA operon region analysis has also been described, providing an even higher resolution [102].

Inference of function with amplicon-based sequencing. Amplicon-based sequencing lacks the possibility to directly study functional aspects of the species within a sample. However, metabolic inference approaches such as Paprica, Picrust2 [103,104], and Tax4Fun2 [105] are available. These tools use hidden-state prediction (HSP) algorithms, which allow the estimation of metabolic functions based on representative genomic content from a well-described bacterial community. A key problem in using such databases is the relatively low correlation between the relative abundance of their specific functions [106] and the high population variability, limiting these tools in applications with defined cohorts.

Identification with shotgun metagenomic sequencing. Shotgun metagenomics facilitates untargeted sequencing of all microbial genomes present in the sample [107]. The dataset generated is much more complex than amplicon-based sequencing [108]. Standardization of the methodological and analytical workflows has just started [109,110]. Proof-of-concept studies have shown the potential clinical impact in pathogen identification within culture-negative samples of, e.g., meningitis and encephalitis [111], sepsis [112–114], pneu-

monia [115,116], and prosthetic joint infections [117,118]. The detection limit of shotgun metagenomics is affected by slow-growing microorganisms or if the potential pathogen is present in low abundance. Different protocols to increase the sensitivity have been developed, e.g., a short, specific culture step for certain pathogens [117]. However, in primary non-sterile body sites, this may introduce a critical diagnostic bias. The performance in bacterial detection and identification was compared between shotgun metagenomics and 16S amplicon-based sequencing, where shotgun metagenomics showed a slightly higher sensitivity (46.3% vs. 38.8%) than 16S [119]. Similarly, Gu et al. compared shotgun metagenomic sequencing using Illumina (short-read) and Nanopore sequencing (long-read) platforms for pathogen identification and validated the results with traditional culture-based methods and also with 16S and 28S-ITS PCRs for bacterial and fungal species. They have also shown that the real-time analysis offered by Nanopore sequencing enabled a reduced turnaround time for pathogen identification [120]. The latest add-on feature with the ONT sequencing platform is “adaptive sampling”, which allows for enriching or depleting sequenced DNA from selected species selectively in a software-controlled manner during sequencing [120,121]. This is useful for clinical samples such as body fluids and swabs where human DNA largely outweighs non-human DNA, and the depletion of host DNA consequently increases the pathogen detection sensitivity [122,123]. On the bioinformatics front, new software tools such as SMAGLinker, Strainberry, and STRain Resolution ON assembly Graphs (STRONG) allow for obtaining strain-resolved genomes in microbiota samples for both short-read and long-read sequencing data. Metagenomic sequences are assigned to several bins and merged for taxonomic identification [124–126]. For RNA viruses, the RNA-dependent RNA polymerase (RdRp) can be used as a baseline core motif for species identification [127,128].

Inference of function with shotgun metagenomics. Deciphering microbiota functions is crucial for predicting clinical phenotypes. The generation of high-quality metagenomically assembled genomes (MAGs) from metagenomic datasets facilitates the understanding of microbial ecosystems by elucidating detailed metabolic pathways and horizontal gene transfer networks [124]. The prediction of function requires various bioinformatic tools and databases, of which dozens have been developed [129–131]. Since not all the genes and annotations are known yet, the databases for functional annotation (e.g., KEGG and EggNOG) are still incomplete, and thus not all metabolic functions can be inferred from the MAGs. An important aspect of functional inference is the detection of AMR genes. Knowledge about local microbiome compositions and hotspots of AMR genes present in the environment may potentially be used to trigger further investigations. Chng et al. sequenced environmental surface swabs collected from a hospital and combined short- and long-read sequencing to determine distinct ecological niches present in various regions of the hospital [132]. AMR monitoring from sewage samples is an interesting surveillance tool [133], which is not only used for pathogen surveillance, e.g., SARS-CoV-2 [134], but also for virulence monitoring [135]. Perry et al. compared sewage samples from a hospital and the surrounding community sewage plants. The authors modeled the correlation between antimicrobial usage in the hospital and resistance gene abundances and showed that sewage from the hospital has a higher percentage of Antibiotic Resistance Genes (ARGs) compared to the communal sewage samples [136]. Another interesting function to study would be the microbial association network from microbiome data in order to capture the interactions between the various identified species. The microbial association catalog (mako) is one such graph-based database compiled from 60 microbiome studies, which allows for a user-friendly network motif search [137] to infer the associations within the microbiome. The next step is to link more complex microbiological communities, on the species level or the genetic content level, to particular clinical phenotypes. A microbiome-wide association study [138–141] could help to identify crucial networks of communities and link this to, e.g., treatment response in mixed infections, the risk of colonization, or invasive infection [142]. Another interesting application of the interaction of microbiota or microbiome would be immune modulation in cancer, where specific species are enriched or show immunomodulatory effects for check-point inhibitors during cancer therapy [143–147].

4. An Integrated Approach of WGS and Metagenomic Sequencing

Comparing the currently available WGS and metagenomic technologies and approaches (refer to Table 1 for the possibilities and limitations) to study microbial features, it becomes clear that no single approach can answer all diagnostic or research questions. Both approaches show potential for complementary usage and data analysis, which may potentiate the output and provide novel insights into host–pathogen interactions, clinical outcomes, and pathogen surveillance in various environments. However, to achieve this goal, we first need to link the sequencing outputs. In a prospectively built database, e.g., during surveillance or monitoring of a patient, a unique identifier could be used to merge one or more WGS datasets with a metagenomic dataset. Ideally, additional clinical, microbiological, or epidemiological metadata would be added, such as the time and space of acquisition of each dataset and potential phenotypic readouts such as AMR or metabolic profiles. The below section discusses the data formats, clinical use cases, bioinformatics tools, and the quality control pointers for combining WGS and shotgun metagenomics sequencing data.

Table 1. Comparison of whole-genome sequencing, marker gene-based amplicon sequencing, and shotgun metagenomic sequencing. GWAS = Genome-Wide Association Studies. SNP = Single Nucleotide Polymorphism. The symbol “\$” represents the cost of sequencing. Higher number of \$ = higher cost. The symbol “+” represents the turnaround time for the sequencing strategies. Higher number of + = longer turnaround time.

| Parameters | WGS | 16S/ITS | Shotgun Metagenomic Sequencing |
|--|--|---|--|
| Sample | Cultured or enriched microorganisms | Swabs from body sites, stool samples, body fluids or tissue samples, and sewage | Swabs from body sites, stool samples, body fluids or tissue samples fecal matter, and sewage |
| Species identification | Yes | Yes | Yes |
| Degree of resolution | Species-Strain level | Genus-Species level | Species-Strain level |
| Complete genome | Complete genome possible depending on sequencing platforms | No | Near complete to gapped genomes. |
| SNP analysis | Yes | No | Yes |
| GWAS | Yes | No | Yes |
| Identification of virulence factors and resistance genes | Yes | No | Yes |
| Microbial community profiling | No | Yes | Yes |
| Cost | \$\$ | \$ | \$\$\$ |
| Turnaround Time (TAT) | + | ++ | +++ |

Examples of combined WGS and shotgun metagenomics approaches. A combined, integrative approach allows one to look for similarities in the dataset and is the primary aspect of merging both data sources. Such a combination may be used for identification, functional readouts, and typing. The ability to quantify and detect bacterial strains within heterogeneous environments has applications in numerous fields including diagnostics [148], clinical studies for the microbiome [149], bio surveillance, One Health [150,151], outbreak investigations [150,152,153], providing insight into the spread of antibiotic resistance [152], tracking the progression of within-host bacterial evolution [153], and exploring diverse environments [154,155].

Such an integrated approach was used to investigate an outbreak of carbapenemase-producing *Enterobacter hormaechei*. The source of the strain could be allocated to the plumbing and water resources in a hospital. The combination of short- and long-read sequencing enabled the resolution of the complete plasmid of the resistance gene carrier (IMP-4) and thereby the monitoring of its transmission across the hospital environment [156]. Similarly, an outbreak of a carbapenem-resistant *Acinetobacter baumannii* was investigated and the source was again linked to the plumbing system of a hospital [157]. This outbreak investigation resulted in an internal database of circulating pathogens in the environment, and it allowed one to restrict the transmission of the resistant strain and provided information about the recurrence of the pathogens in the hospital wards.

A new addition to metagenome sequencing is the implementation of single-cell bacterial sequencing technology, which combines shotgun metagenomics and WGS for strain resolution and allows the tracking of mobile elements. For example, Zheng et al. devised Microbe-seq, single-microbe genomics to achieve sub-species resolution from the human gut microbiome. Using microfluidics, they have captured single microbes in liquid droplets, lysed the cells, and barcoded the DNA followed by whole genome amplification. Computationally, the authors have co-assembled single-amplified genomes (SAGs), recovered from the whole genome amplification of single bacterial cells, and characterized the horizontal gene transfers within the strains of the same species [158]. This approach therefore allows one to study the transfer of AMR or virulence genes. One hypothetical application for combining WGS and shotgun metagenomics data could be as follows: We can consider the identified microbiome from the shotgun metagenomic sequencing as a restricted database of microorganisms present in a clinically relevant environment. The assembled genome of a particular pathogen obtained from WGS can then be used as a query sequence, which can be searched against this database (and vice versa). During a hospital outbreak investigation, for example, this would allow the quick identification of the outbreak source, assuming environmental screening is performed at regular intervals and enriched with useful metadata. In other words, the pathogen can be traced back to an environmental origin in the hospital if a hit of high similarity is found in the database, while it is likely an introduction from outside if it cannot. This approach could also be applied for tracking the transmission of mobile genetic elements across environments. On the community front, the microbiome data obtained could also be used for identifying at-risk populations based on microbial distributions. One major drawback of such a combination is the added cost of regularly sequencing the environment of interest.

Data readouts. Typically, short-read sequencing data from Illumina machines and long-read sequencing data from Nanopore and PacBio result in fastq files after base calling from their respective sequencing data formats. Contigs, assembled genomes, or segments are usually in FASTA format. As a consequence, merging data from either platform is computationally convenient.

Possible computational methods for strain-level microbial detection from WGS and shotgun metagenomic sequencing. An inter-linked dataset allows the use of the contained data as a reference for mapping. A range of bioinformatics pipelines and methods for strain-level microbial detection in metagenome sequencing data have been developed [159]. In principle, methods are based on (i) assembly-based reconstruction and (ii) methods operating with or without a reference database. We are only at the beginning of using WGS and metagenomics in clinical settings, therefore it is crucial to have a comprehensive benchmark across different (clinical) applications to validate performances and standardize the available tools.

Assembly-based approaches identify single strains in mixed reads by whole genome assembly. Sufficient differences in the genomes are necessary to separate or cluster, e.g., bacterial variants into distinct strains [160]. This approach requires sufficient read length and sequencing depth to reach at least one variant site in most reads. Tools such as EVORhA, STRONG, StrainGE, and Strainberry deconvolute the assemblies from short- and long-reads to provide strain-level resolution [125,126,153]. Hypothetically, this offers

interesting applications in AMR surveillance, as these strains could then be compared using a curated database containing local endemic AMR-relevant strains. The assembly step in WGS can also be combined with a metagenomic dataset, which allows the identification of specific single strains within a bacterial community [161]. Zlitni et al. performed short-read and read cloud metagenomic sequencing together with metatranscriptomics to monitor the sub-strain populations within a patient's gut [162]. Meanwhile, Ivanova et al. used shotgun metagenomics with chromosome conformation capture (Hi-C) technology. This resulted in high-quality MAGs together with plasmids, as the technology also linked reads between genome and mobile genetic elements [163]. These approaches allow for identifying specific pathogens or plasmids harboring multi-drug resistance in a specific environment.

Furthermore, full genome alignment-based methods allow strain classification by aligning reads directly to a selection of reference genomes and applying stochastic models to calculate the likelihood of association between a specific read and reference [164]. Pathoscope [165] is a classification pipeline using different aligners, including GNU-MAP [166], Bowtie 2 [167], and BLAST [168], and scores for each alignment reflect the likelihood that the read source matches the reference assembly. Furthermore, a semi-quantitative assessment can be reached for strain abundance based on the number of reads mapping to each reference. Alignment-based detection works within clear and well-separated sub-lineages. However, the reference database is critical in closely related strains.

Substantial computing time can be saved by aligning a set of genetic markers, rather than the complete genome. These marker-based methods classify genetic diversity within a sample using a database of, e.g., unique genes [169], SNPs, or k-mers [170]. Pattern-based methods also require a reference database for statistical models. However, pattern-based methods initially pre-process extracted features and use these features for a new classifier algorithm, which results in substantially decreased analytical time. MIDAS is one such bioinformatic tool, applying this concept for species and strain-level taxonomic identification [171]. The k-mer-based tool GSMer identifies strains by using a strain-specific database of strain-specific k-mers, or genome-specific markers (GSM) [172]. In this tool, each strain is represented by at least 50 GSMs, and strains with less than 50 unique GSMs are excluded. Strains are only identified if there is perfect alignment for all 50 GSMs, resulting in high specificity, but potentially low sensitivity. Such an approach may be rather useful for clonal, slowly evolving strains without high rates of genetic adaptation.

Quality control. There are several quality factors to consider, before integrating the data from different omics techniques: (i) Regular update of databases: This not only includes technical and software updates but also includes the epidemiological content covered by the database, e.g., via regular shotgun metagenomic sequencing of a given environment; (ii) sequencing depth and coverage: Care should be taken to assess the quality of the data obtained from each of these omics technologies. Since the performance of these technologies is prone to bias, standardized workflows will result in reproducible read cut-off values for depth and coverage.

5. Conclusions and Future Trends

With the use of clinical WGS and metagenomics on the rise, in part due to the SARS-CoV-2 pandemic, environmental screening for microbes has become feasible and cost-effective [151]. Researchers have combined various sequencing platforms in proof-of-concept studies for pathogen identification, the characterization of virulence and resistance genes, and the typing of relatedness between bacteria, viruses, and fungi. Indeed, mapping the ecological niches of the pathogens, e.g., in the hospital [156,157] as well as the environment may potentiate the effects. It has also been shown that it is possible to integrate 16S and shotgun metagenomics for microbiome studies, where the expected readouts are taxonomic abundance, diversity, and functional annotations. While 16S can provide the identified taxonomy, shotgun metagenomic sequencing can validate it with the genomes of the identified microbes and their functional annotations [173–176]. Harmonization and standardization of the individual technologies will likely also move forward. For clini-

cal applications, clear, controlled, and reproducible protocols are necessary, reflected in the regulatory requirements such as the recently established In Vitro Diagnostic Regulation (IVDR).

So far, only a few studies have shown how to successfully integrate and use the additive and potential synergistic effects of both technologies. One limitation is the need for standardization and harmonization between protocols and workflows, and the added sequencing costs of two technologies being applied along with the costs linked to data storage and maintaining the bioinformatic pipelines and databases. Metadata with a sufficient spatio-temporal resolution (e.g., sample isolation date and geographical location) and additional epidemiological context may become important to use this potential. The access to metadata goes hand in hand with the FAIR data-sharing principles (Findable, Accessible, Interoperable, and Reusable), and the need for databases to allow for properly annotating the interlinked status between omics data types. Overall, combining WGS and shotgun metagenomics brings out complementary benefits by incorporating missing pieces of information.

Author Contributions: Conceptualization, A.E., M.M. and S.P.; writing-original draft preparation, S.P., A.E. and M.M.; writing-review and editing, A.E., S.P. and M.M.; supervision, A.E.; funding acquisition, A.E. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by Swiss National Science Foundation [Grant No. 310030_192515] to AE.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank Fanny Wegner and Helena Seth-Smith (Institute of Medical Microbiology, University of Zurich) for extensively proofreading the manuscript and sharing their valuable insights and suggestions.

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

References

1. Sharpton, T.J. An Introduction to the Analysis of Shotgun Metagenomic Data. *Front. Plant. Sci.* **2014**, *5*, 209. [[CrossRef](#)]
2. Kuper, K.M.; Boles, D.M.; Mohr, J.F.; Wanger, A. Antimicrobial Susceptibility Testing: A Primer for Clinicians. *Pharmacotherapy* **2009**, *29*, 1326–1343. [[CrossRef](#)] [[PubMed](#)]
3. Van Belkum, A.; Bachmann, T.T.; Lüdke, G.; Lisby, J.G.; Kahlmeter, G.; Mohess, A.; Becker, K.; Hays, J.P.; Woodford, N.; Mitsakakis, K.; et al. Developmental Roadmap for Antimicrobial Susceptibility Testing Systems. *Nat. Rev. Microbiol.* **2019**, *17*, 51–62. [[CrossRef](#)]
4. Smith, K.P.; Kirby, J.E. Rapid Susceptibility Testing Methods. *Clin. Lab. Med.* **2019**, *39*, 333–344. [[CrossRef](#)]
5. Lagier, J.-C.; Hugon, P.; Khelaifia, S.; Fournier, P.-E.; La Scola, B.; Raoult, D. The Rebirth of Culture in Microbiology through the Example of Culturomics to Study Human Gut Microbiota. *Clin. Microbiol. Rev.* **2015**, *28*, 237–264. [[CrossRef](#)]
6. Corona, F.; Martinez, J.L. Phenotypic Resistance to Antibiotics. *Antibiotics* **2013**, *2*, 237–255. [[CrossRef](#)]
7. Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.R.; Bult, C.J.; Tomb, J.F.; Dougherty, B.A.; Merrick, J.M. Whole-Genome Random Sequencing and Assembly of *Haemophilus influenzae* Rd. *Science* **1995**, *269*, 496–512. [[CrossRef](#)]
8. Hasman, H.; Saputra, D.; Sicheritz-Ponten, T.; Lund, O.; Svendsen, C.A.; Frimodt-Møller, 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](#)]
9. Chiu, C.Y.; Miller, S.A. Clinical Metagenomics. *Nat. Rev. Genet.* **2019**, *20*, 341–355. [[CrossRef](#)]
10. Eyre, D.W. Infection Prevention and Control Insights from a Decade of Pathogen Whole-Genome Sequencing. *J. Hosp. Infect.* **2022**, *122*, 180–186. [[CrossRef](#)]
11. Brown, B.; Allard, M.; Bazaco, M.C.; Blankenship, J.; Minor, T. An Economic Evaluation of the Whole Genome Sequencing Source Tracking Program in the U.S. *PLoS ONE* **2021**, *16*, e0258262. [[CrossRef](#)] [[PubMed](#)]
12. Dymond, A.; Davies, H.; Mealing, S.; Pollit, V.; Coll, F.; Brown, N.M.; Peacock, S.J. Genomic Surveillance of Methicillin-Resistant *Staphylococcus aureus*: A Mathematical Early Modeling Study of Cost-Effectiveness. *Clin. Infect. Dis.* **2020**, *70*, 1613–1619. [[CrossRef](#)] [[PubMed](#)]

13. Perez-Sepulveda, B.M.; Heavens, D.; Pulford, C.V.; Predeus, A.V.; Low, R.; Webster, H.; Dykes, G.F.; Schudoma, C.; Rowe, W.; Lipscombe, J.; et al. An Accessible, Efficient and Global Approach for the Large-Scale Sequencing of Bacterial Genomes. *Genome Biol.* **2021**, *22*, 349. [[CrossRef](#)]
14. Loman, N.J.; Constantinidou, C.; Chan, J.Z.M.; Halachev, M.; Sergeant, M.; Penn, C.W.; Robinson, E.R.; Pallen, M.J. High-Throughput Bacterial Genome Sequencing: An Embarrassment of Choice, a World of Opportunity. *Nat. Rev. Microbiol.* **2012**, *10*, 599–606. [[CrossRef](#)]
15. Smetana, J.; Brož, P. National Genome Initiatives in Europe and the United Kingdom in the Era of Whole-Genome Sequencing: A Comprehensive Review. *Genes* **2022**, *13*, 556. [[CrossRef](#)]
16. Van Belkum, A.; Rochas, O. Laboratory-Based and Point-of-Care Testing for MSSA/MRSA Detection in the Age of Whole Genome Sequencing. *Front. Microbiol.* **2018**, *9*, 1437. [[CrossRef](#)]
17. Bonsall, D.; Golubchik, T.; de Cesare, M.; Limbada, M.; Kosloff, B.; MacIntyre-Cockett, G.; Hall, M.; Wymant, C.; Ansari, M.A.; Abeler-Dörner, L.; et al. A Comprehensive Genomics Solution for HIV Surveillance and Clinical Monitoring in Low-Income Settings. *J. Clin. Microbiol.* **2020**, *58*, e00382-20. [[CrossRef](#)]
18. Aarestrup, F.M.; Bonten, M.; Koopmans, M. Pandemics- One Health Preparedness for the next. *Lancet Reg. Health Eur.* **2021**, *9*, 100210. [[CrossRef](#)]
19. Wegner, F.; Roloff, T.; Huber, M.; Cordey, S.; Ramette, A.; Gerth, Y.; Bertelli, C.; Stange, M.; Seth-Smith, H.M.B.; Mari, A.; et al. External Quality Assessment of SARS-CoV-2 Sequencing: An ESGMD-SSM Pilot Trial across 15 European Laboratories. *J. Clin. Microbiol.* **2022**, *60*, e0169821. [[CrossRef](#)]
20. Liu, J.; Wang, X.; Xie, H.; Zhong, Q.; Xia, Y. Analysis and Evaluation of Different Sequencing Depths from 5 to 20 Million Reads in Metagenome Shotgun Sequencing, with Optimal Minimum Depth Recommended. *Genome* **2022**. [[CrossRef](#)]
21. Tagini, F.; Greub, G. Bacterial Genome Sequencing in Clinical Microbiology: A Pathogen-Oriented Review. *Eur. J. Clin. Microbiol. Infect. Dis.* **2017**, *36*, 2007–2020. [[CrossRef](#)]
22. Wu, X.; Tan, G.; Sha, W.; Liu, H.; Yang, J.; Guo, Y.; Shen, X.; Wu, Z.; Shen, H.; Yu, F. Use of Whole-Genome Sequencing to Predict Complex Drug Resistance from Early Positive Liquid Cultures. *Microbiol. Spectr.* **2022**, *10*, e0251621. [[CrossRef](#)] [[PubMed](#)]
23. Doyle, R.M.; Burgess, C.; Williams, R.; Gorton, R.; Booth, H.; Brown, J.; Bryant, J.M.; Chan, J.; Creer, D.; Holdstock, J.; et al. Direct Whole-Genome Sequencing of Sputum Accurately Identifies Drug-Resistant Mycobacterium tuberculosis Faster than MGIT Culture Sequencing. *J. Clin. Microbiol.* **2018**, *56*, e00666-18. [[CrossRef](#)] [[PubMed](#)]
24. Wüthrich, D.; Lang, D.; Müller, N.F.; Neher, R.A.; Stadler, T.; Egli, A. Evaluation of Two Workflows for Whole Genome Sequencing-Based Typing of Influenza A viruses. *J. Virol. Methods* **2019**, *266*, 30–33. [[CrossRef](#)] [[PubMed](#)]
25. Stange, M.; Mari, A.; Roloff, T.; Seth-Smith, H.M.; Schweitzer, M.; Brunner, M.; Leuzinger, K.; Søgaard, K.K.; Gensch, A.; Tschudin-Sutter, S.; et al. SARS-CoV-2 Outbreak in a Tri-National Urban Area Is Dominated by a B.1 Lineage Variant Linked to a Mass Gathering Event. *PLoS Pathog.* **2021**, *17*, e1009374. [[CrossRef](#)]
26. Seth-Smith, H.M.B.; Harris, S.R.; Skilton, R.J.; Radebe, F.M.; Golparian, D.; Shipitsyna, E.; Duy, P.T.; Scott, P.; Cutcliffe, L.T.; O'Neill, C.; et al. Whole-Genome Sequences of Chlamydia trachomatis Directly from Clinical Samples without Culture. *Genome Res.* **2013**, *23*, 855–866. [[CrossRef](#)]
27. Lomsadze, A.; Li, T.; Rajeevan, M.S.; Unger, E.R.; Borodovsky, M. Bioinformatics Pipeline for Human papillomavirus Short Read Genomic Sequences Classification Using Support Vector Machine. *Viruses* **2020**, *12*, 710. [[CrossRef](#)]
28. Brown, J.R.; Roy, S.; Ruis, C.; Yara Romero, E.; Shah, D.; Williams, R.; Breuer, J. Norovirus Whole-Genome Sequencing by SureSelect Target Enrichment: A Robust and Sensitive Method. *J. Clin. Microbiol.* **2016**, *54*, 2530–2537. [[CrossRef](#)]
29. Mäklin, T.; Kallonen, T.; Alanko, J.; Samuelsen, Ø.; Hegstad, K.; Mäkinen, V.; Corander, J.; Heinz, E.; Honkela, A. Bacterial Genomic Epidemiology with Mixed Samples. *Microb. Genom.* **2021**, *7*, 691. [[CrossRef](#)]
30. Gautam, S.S.; Kc, R.; Leong, K.W.; Mac Aogáin, M.; O'Toole, R.F. A Step-by-Step Beginner's Protocol for Whole Genome Sequencing of Human Bacterial Pathogens. *J. Biol. Methods* **2019**, *6*, e110. [[CrossRef](#)]
31. Smits, T.H.M. The Importance of Genome Sequence Quality to Microbial Comparative Genomics. *BMC Genom.* **2019**, *20*, 662. [[CrossRef](#)] [[PubMed](#)]
32. Rossen, J.W.A.; Friedrich, A.W.; Moran-Gilad, J. ESCMID Study Group for Genomic and Molecular Diagnostics (ESGMD) Practical Issues in Implementing Whole-Genome-Sequencing in Routine Diagnostic Microbiology. *Clin. Microbiol. Infect.* **2018**, *24*, 355–360. [[CrossRef](#)] [[PubMed](#)]
33. Timme, R.E.; Wolfgang, W.J.; Balkey, M.; Venkata, S.L.G.; Randolph, R.; Allard, M.; Strain, E. Optimizing Open Data to Support One Health: Best Practices to Ensure Interoperability of Genomic Data from Bacterial Pathogens. *One Health Outlook* **2020**, *2*, 20. [[CrossRef](#)]
34. Higgs, C.; Sherry, N.L.; Seemann, T.; Horan, K.; Walpola, H.; Kinsella, P.; Bond, K.; Williamson, D.A.; Marshall, C.; Kwong, J.C.; et al. Optimising Genomic Approaches for Identifying Vancomycin-Resistant Enterococcus faecium Transmission in Healthcare Settings. *Nat. Commun.* **2022**, *13*, 509. [[CrossRef](#)]
35. Shropshire, W.C.; Dinh, A.Q.; Earley, M.; Komarow, L.; Panesso, D.; Rydell, K.; Gómez-Villegas, S.I.; Miao, H.; Hill, C.; Chen, L.; et al. Accessory Genomes Drive Independent Spread of Carbapenem-Resistant Klebsiella pneumoniae Clonal Groups 258 and 307 in Houston, TX. *MBio* **2022**, *13*, e0049722. [[CrossRef](#)] [[PubMed](#)]
36. Craddock, H.A.; Motro, Y.; Zilberman, B.; Khalfin, B.; Bardenstein, S.; Moran-Gilad, J. Long-Read Sequencing and Hybrid Assembly for Genomic Analysis of Clinical Isolates. *Microorganisms* **2022**, *10*, 619. [[CrossRef](#)]

37. Uelze, L.; Borowiak, M.; Deneke, C.; Szabó, I.; Fischer, J.; Tausch, S.H.; Malorny, B. Performance and Accuracy of Four Open-Source Tools for Serotyping of Spp. Based on Whole-Genome Short-Read Sequencing Data. *Appl. Environ. Microbiol.* **2020**, *86*, e02265-19. [[CrossRef](#)]
38. Browne, P.D.; Nielsen, T.K.; Kot, W.; Aggerholm, A.; Gilbert, M.T.P.; Puetz, L.; Rasmussen, M.; Zervas, A.; Hansen, L.H. GC Bias Affects Genomic and Metagenomic Reconstructions, Underrepresenting GC-Poor Organisms. *Gigascience* **2020**, *9*, gaaa008. [[CrossRef](#)]
39. Seth-Smith, H.M.B.; Bonfiglio, F.; Cuénod, A.; Reist, J.; Egli, A.; Wüthrich, D. Evaluation of Rapid Library Preparation Protocols for Whole Genome Sequencing Based Outbreak Investigation. *Front. Public Health* **2019**, *7*, 241. [[CrossRef](#)]
40. Wick, R.R.; Judd, L.M.; Wyres, K.L.; Holt, K.E. Recovery of Small Plasmid Sequences via Oxford Nanopore Sequencing. *Microb. Genom.* **2021**, *7*, 631. [[CrossRef](#)]
41. Pightling, A.W.; Pettengill, J.B.; Wang, Y.; Rand, H.; Strain, E. Within-Species Contamination of Bacterial Whole-Genome Sequence Data Has a Greater Influence on Clustering Analyses than between-Species Contamination. *Genome Biol.* **2019**, *20*, 286. [[CrossRef](#)] [[PubMed](#)]
42. Deneke, C.; Brendebach, H.; Uelze, L.; Borowiak, M.; Malorny, B.; Tausch, S.H. Species-Specific Quality Control, Assembly and Contamination Detection in Microbial Isolate Sequences with AQUAMIS. *Genes* **2021**, *12*, 644. [[CrossRef](#)] [[PubMed](#)]
43. Riojas, M.A.; McGough, K.J.; Rider-Riojas, C.J.; Rastogi, N.; Hazbón, M.H. Phylogenomic Analysis of the Species of the Mycobacterium tuberculosis Complex Demonstrates That Mycobacterium africanum, Mycobacterium bovis, Mycobacterium caprae, Mycobacterium microti and Mycobacterium pinnipedii Are Later Heterotypic Synonyms of Mycobacterium tuberculosis. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 324–332. [[PubMed](#)]
44. Tortoli, E.; Brown-Elliott, B.A.; Chalmers, J.D.; Cirillo, D.M.; Daley, C.L.; Emler, S.; Floto, R.A.; Garcia, M.J.; Hoefsloot, W.; Koh, W.-J.; et al. Same Meat, Different Gravy: Ignore the New Names of Mycobacteria. *Eur. Respir. J.* **2019**, *54*, 1900795. [[CrossRef](#)]
45. International Code of Nomenclature of Prokaryotes. *Int. J. Syst. Evol. Microbiol.* **2019**, *69*, S1–S111. [[CrossRef](#)]
46. Oren, A.; Arahal, D.R.; Rosselló-Móra, R.; Sutcliffe, I.C.; Moore, E.R.B. Preparing a Revision of the International Code of Nomenclature of Prokaryotes. *Int. J. Syst. Evol. Microbiol.* **2021**, *71*, 4598. [[CrossRef](#)]
47. Oren, A.; Arahal, D.R.; Rosselló-Móra, R.; Sutcliffe, I.C.; Moore, E.R.B. Public Discussion on a Proposed Revision of the International Code of Nomenclature of Prokaryotes. *Int. J. Syst. Evol. Microbiol.* **2021**, *71*, 4918. [[CrossRef](#)]
48. Meier-Kolthoff, J.P.; Göker, M. TYGS Is an Automated High-Throughput Platform for State-of-the-Art Genome-Based Taxonomy. *Nat. Commun.* **2019**, *10*, 2182. [[CrossRef](#)]
49. Chaumeil, P.-A.; Mussig, A.J.; Hugenholtz, P.; Parks, D.H. GTDB-Tk: A Toolkit to Classify Genomes with the Genome Taxonomy Database. *Bioinformatics* **2019**, *36*, 1925–1927. [[CrossRef](#)]
50. Khieu, V.; Ananta, P.; Kaewprasert, O.; Laohaviroj, M.; Namwat, W.; Faksri, K. Whole-Genome Sequencing Analysis to Identify Infection with Multiple Species of Nontuberculous Mycobacteria. *Pathogens* **2021**, *10*, 879. [[CrossRef](#)]
51. Dylus, D.; Pillonel, T.; Opota, O.; Wüthrich, D.; Seth-Smith, H.M.B.; Egli, A.; Leo, S.; Lazarevic, V.; Schrenzel, J.; Laurent, S.; et al. NGS-Based Typing and Outbreak Analysis in Clinical Microbiology Laboratories: Lessons Learned From a Swiss-Wide Proficiency Test. *Front. Microbiol.* **2020**, *11*, 591093. [[CrossRef](#)] [[PubMed](#)]
52. Mellmann, A.; Andersen, P.S.; Bletz, S.; Friedrich, A.W.; Kohl, T.A.; Lilje, B.; Niemann, S.; Prior, K.; Rossen, J.W.; Harmsen, D. High Interlaboratory Reproducibility and Accuracy of Next-Generation-Sequencing-Based Bacterial Genotyping in a Ring Trial. *J. Clin. Microbiol.* **2017**, *55*, 908–913. [[CrossRef](#)] [[PubMed](#)]
53. Meinel, D.M.; Kuehl, R.; Zbinden, R.; Boskova, V.; Garzoni, C.; Fadini, D.; Dolina, M.; Blümel, B.; Weibel, T.; Tschudin-Sutter, S.; et al. Outbreak Investigation for Toxigenic Corynebacterium diphtheriae Wound Infections in Refugees from Northeast Africa and Syria in Switzerland and Germany by Whole Genome Sequencing. *Clin. Microbiol. Infect.* **2016**, *22*, e1003.e1–e1003.e8. [[CrossRef](#)] [[PubMed](#)]
54. Wüthrich, D.; Gautsch, S.; Spieler-Denz, R.; Dubuis, O.; Gaia, V.; Moran-Gilad, J.; Hinic, V.; Seth-Smith, H.M.; Nickel, C.H.; Tschudin-Sutter, S.; et al. Air-Conditioner Cooling Towers as Complex Reservoirs and Continuous Source of Infection Evidenced by a Genomic Analysis Study in 2017, Switzerland. *Euro Surveill.* **2019**, *24*, 1800192. [[CrossRef](#)] [[PubMed](#)]
55. Köser, C.U.; Holden, M.T.G.; Ellington, M.J.; Cartwright, E.J.P.; Brown, N.M.; Ogilvy-Stuart, A.L.; Hsu, L.Y.; Chewapreecha, C.; Croucher, N.J.; Harris, S.R.; et al. Rapid Whole-Genome Sequencing for Investigation of a Neonatal MRSA Outbreak. *N. Engl. J. Med.* **2012**, *366*, 2267–2275. [[CrossRef](#)]
56. Wassilew, N.; Seth-Smith, H.M.; Rolli, E.; Fietze, Y.; Casanova, C.; Führer, U.; Egli, A.; Marschall, J.; Buetti, N. Outbreak of Vancomycin-Resistant Clone ST796, Switzerland, December 2017 to April 2018. *Euro Surveill.* **2018**, *23*, 1800351. [[CrossRef](#)]
57. Timme, R.E.; Sanchez Leon, M.; Allard, M.W. Utilizing the Public GenomeTrakr Database for Foodborne Pathogen Traceback. *Methods Mol. Biol.* **2019**, *1918*, 201–212.
58. Egli, A.; Blanc, D.S.; Greub, G.; Keller, P.M.; Lazarevic, V.; Lebrand, A.; Leib, S.; Neher, R.A.; Perreten, V.; Ramette, A.; et al. Improving the Quality and Workflow of Bacterial Genome Sequencing and Analysis: Paving the Way for a Switzerland-Wide Molecular Epidemiological Surveillance Platform. *Swiss Med. Wkly.* **2018**, *148*, w14693. [[CrossRef](#)] [[PubMed](#)]
59. Xanthopoulou, K.; Peter, S.; Tobys, D.; Behnke, M.; Dinkelacker, A.G.; Eisenbeis, S.; Falgenhauer, J.; Falgenhauer, L.; Fritzenwanker, M.; Götz, H.; et al. Vancomycin-Resistant Enterococcus faecium Colonizing Patients on Hospital Admission in Germany: Prevalence and Molecular Epidemiology. *J. Antimicrob. Chemother.* **2020**, *75*, 2743–2751. [[CrossRef](#)]

60. Lee, X.J.; Elliott, T.M.; Harris, P.N.A.; Douglas, J.; Henderson, B.; Watson, C.; Paterson, D.L.; Schofield, D.S.; Graves, N.; Gordon, L.G. Clinical and Economic Outcomes of Genome Sequencing Availability on Containing a Hospital Outbreak of Resistant *Escherichia coli* in Australia. *Value Health* **2020**, *23*, 994–1002. [[CrossRef](#)]
61. Snyder, L.A.; Loman, N.J.; Faraj, L.A.; Levi, K.; Weinstock, G.; Boswell, T.C.; Pallen, M.J.; Ala'Aldeen, D.A. Epidemiological Investigation of *Pseudomonas aeruginosa* Isolates from a Six-Year-Long Hospital Outbreak Using High-Throughput Whole Genome Sequencing. *Euro Surveill.* **2013**, *18*, 20611. [[CrossRef](#)] [[PubMed](#)]
62. Lewis, T.; Loman, N.J.; Bingle, L.; Jumaa, P.; Weinstock, G.M.; Mortiboy, D.; Pallen, M.J. High-Throughput Whole-Genome Sequencing to Dissect the Epidemiology of *Acinetobacter baumannii* Isolates from a Hospital Outbreak. *J. Hosp. Infect.* **2010**, *75*, 37–41. [[CrossRef](#)] [[PubMed](#)]
63. Song, J.E.; Jeong, H.; Lim, Y.S.; Ha, E.J.; Jung, I.Y.; Jeong, W.; Choi, H.; Jeong, S.J.; Ku, N.S.; Park, E.S.; et al. An Outbreak of KPC-Producing Linked with an Index Case of Community-Acquired KPC-Producing Isolate: Epidemiological Investigation and Whole Genome Sequencing Analysis. *Microb. Drug. Resist.* **2019**, *25*, 1475–1483. [[CrossRef](#)] [[PubMed](#)]
64. Peter, S.; Bosio, M.; Gross, C.; Bezdán, D.; Gutierrez, J.; Oberhettinger, P.; Liese, J.; Vogel, W.; Dörfel, D.; Berger, L.; et al. Tracking of Antibiotic Resistance Transfer and Rapid Plasmid Evolution in a Hospital Setting by Nanopore Sequencing. *mSphere* **2020**, *5*, e00525-20. [[CrossRef](#)] [[PubMed](#)]
65. Walker, T.M.; Merker, M.; Kohl, T.A.; Crook, D.W.; Niemann, S.; Peto, T.E.A. Whole Genome Sequencing for M/XDR Tuberculosis Surveillance and for Resistance Testing. *Clin. Microbiol. Infect.* **2017**, *23*, 161–166. [[CrossRef](#)]
66. Ellington, M.J.; Ekelund, O.; Aarestrup, F.M.; Canton, R.; Doumith, M.; Giske, C.; Grundman, H.; Hasman, H.; Holden, M.T.G.; Hopkins, K.L.; et al. The Role of Whole Genome Sequencing in Antimicrobial Susceptibility Testing of Bacteria: Report from the EUCAST Subcommittee. *Clin. Microbiol. Infect.* **2017**, *23*, 2–22. [[CrossRef](#)]
67. Bortolaia, V.; Kaas, R.S.; Ruppe, E.; Roberts, M.C.; Schwarz, S.; Cattoir, V.; Philippon, A.; Allesoe, R.L.; Rebelo, A.R.; Florensa, A.F.; et al. ResFinder 4.0 for Predictions of Phenotypes from Genotypes. *J. Antimicrob. Chemother.* **2020**, *75*, 3491–3500. [[CrossRef](#)]
68. Mahfouz, N.; Ferreira, I.; Beisken, S.; von Haeseler, A.; Posch, A.E. Large-Scale Assessment of Antimicrobial Resistance Marker Databases for Genetic Phenotype Prediction: A Systematic Review. *J. Antimicrob. Chemother.* **2020**, *75*, 3099–3108. [[CrossRef](#)]
69. Papp, M.; Solymosi, N. Review and Comparison of Antimicrobial Resistance Gene Databases. *Antibiotics* **2022**, *11*, 339. [[CrossRef](#)]
70. Doyle, R.M.; O'Sullivan, D.M.; Aller, S.D.; Bruchmann, S.; Clark, T.; Coello Pelegrin, A.; Cormican, M.; Diez Benavente, E.; Ellington, M.J.; McGrath, E.; et al. Discordant Bioinformatic Predictions of Antimicrobial Resistance from Whole-Genome Sequencing Data of Bacterial Isolates: An Inter-Laboratory Study. *Microb. Genom.* **2020**, *6*, e000335. [[CrossRef](#)]
71. Hunt, M.; Bradley, P.; Lapierre, S.G.; Heys, S.; Thomsit, M.; Hall, M.B.; Malone, K.M.; Wintringer, P.; Walker, T.M.; Cirillo, D.M.; et al. Antibiotic Resistance Prediction for from Genome Sequence Data with Mykrobe. *Wellcome Open Res.* **2019**, *4*, 191. [[CrossRef](#)]
72. Lu, C.; Zhou, J.; Zhang, T.; Li, C.; Chen, J.; Fan, J.; Qu, L.; Su, X. Comparative Genomics of the Sequential *Pseudomonas aeruginosa* Isolates Obtained from the Continuous Imipenem Stress Evolution. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 10655–10667. [[CrossRef](#)]
73. Babouee Flury, B.; Ellington, M.J.; Hopkins, K.L.; Turton, J.F.; Doumith, M.; Loy, R.; Staves, P.; Hinic, V.; Frei, R.; Woodford, N. Association of Novel Nonsynonymous Single Nucleotide Polymorphisms in *ampD* with Cephalosporin Resistance and Phylogenetic Variations in *ampC*, *ampR*, *ompF*, and *ompC* in Enterobacter cloacae Isolates That Are Highly Resistant to Carbapenems. *Antimicrob. Agents Chemother.* **2016**, *60*, 2383–2390. [[CrossRef](#)]
74. Hu, D.; Fuller, N.R.; Caterson, I.D.; Holmes, A.J.; Reeves, P.R. Single-Gene Long-Read Sequencing Illuminates *Escherichia coli* Strain Dynamics in the Human Intestinal Microbiome. *Cell Rep.* **2022**, *38*, 110239. [[CrossRef](#)]
75. Yang, C.; Yang, R.-F.; Cui, Y. Bacterial Genome-Wide Association Study: Methodologies and Applications. *Yi Chuan* **2018**, *40*, 57–65.
76. Allen, J.P.; Snitkin, E.; Pincus, N.B.; Hauser, A.R. Forest and Trees: Exploring Bacterial Virulence with Genome-Wide Association Studies and Machine Learning. *Trends Microbiol.* **2021**, *29*, 621–633. [[CrossRef](#)]
77. Weber, R.E.; Fuchs, S.; Layer, F.; Sommer, A.; Bender, J.K.; Thürmer, A.; Werner, G.; Strommenger, B. Corrigendum: Genome-Wide Association Studies for the Detection of Genetic Variants Associated With Daptomycin and Ceftaroline Resistance in *Staphylococcus aureus*. *Front. Microbiol.* **2021**, *12*, 686197. [[CrossRef](#)]
78. Biggel, M.; Xavier, B.B.; Johnson, J.R.; Nielsen, K.L.; Frimodt-Møller, N.; Matheussen, V.; Goossens, H.; Moons, P.; Van Puyvelde, S. Horizontally Acquired *papGII*-Containing Pathogenicity Islands Underlie the Emergence of Invasive Uropathogenic *Escherichia coli* Lineages. *Nat. Commun.* **2020**, *11*, 5968. [[CrossRef](#)]
79. Denamur, E.; Condamine, B.; Esposito-Farèse, M.; Royer, G.; Clermont, O.; Laouenan, C.; Lefort, A.; de Lastours, V.; Galardini, M.; The Colibafii; et al. Genome Wide Association Study of *Escherichia coli* Bloodstream Infection Isolates Identifies Genetic Determinants for the Portal of Entry but Not Fatal Outcome. *PLoS Genet.* **2022**, *18*, e1010112. [[CrossRef](#)]
80. Coll, F.; Gouliouris, T.; Bruchmann, S.; Phelan, J.; Raven, K.E.; Clark, T.G.; Parkhill, J.; Peacock, S.J. PowerBacGWAS: A Computational Pipeline to Perform Power Calculations for Bacterial Genome-Wide Association Studies. *Commun. Biol.* **2022**, *5*, 266. [[CrossRef](#)]
81. Gygli, S.M.; Keller, P.M.; Ballif, M.; Blöchliger, N.; Hömke, R.; Reinhard, M.; Loiseau, C.; Ritter, C.; Sander, P.; Borrell, S.; et al. Whole-Genome Sequencing for Drug Resistance Profile Prediction in. *Antimicrob. Agents Chemother.* **2019**, *63*, e02175-18. [[CrossRef](#)]

82. Juraschek, K.; Borowiak, M.; Tausch, S.H.; Malorny, B.; Käsbohrer, A.; Otani, S.; Schwarz, S.; Meemken, D.; Deneke, C.; Hammerl, J.A. Outcome of Different Sequencing and Assembly Approaches on the Detection of Plasmids and Localization of Antimicrobial Resistance Genes in Commensal. *Microorganisms* **2021**, *9*, 598. [[CrossRef](#)]
83. Wareth, G.; Linde, J.; Hammer, P.; Nguyen, N.H.; Nguyen, T.N.M.; Spletstoesser, W.D.; Makarewicz, O.; Neubauer, H.; Sprague, L.D.; Pletz, M.W. Phenotypic and WGS-Derived Antimicrobial Resistance Profiles of Clinical and Non-Clinical *Acinetobacter baumannii* Isolates from Germany and Vietnam. *Int. J. Antimicrob. Agents* **2020**, *56*, 106127. [[CrossRef](#)]
84. Case, R.J.; Boucher, Y.; Dahllöf, I.; Holmström, C.; Doolittle, W.F.; Kjelleberg, S. Use of 16S rRNA and rpoB Genes as Molecular Markers for Microbial Ecology Studies. *Appl. Environ. Microbiol.* **2007**, *73*, 278–288. [[CrossRef](#)]
85. Stebner, A.; Ensser, A.; Geißdörfer, W.; Bozhkov, Y.; Lang, R. Molecular Diagnosis of Polymicrobial Brain Abscesses with 16S-rDNA-Based next-Generation Sequencing. *Clin. Microbiol. Infect.* **2021**, *27*, 76–82. [[CrossRef](#)]
86. Ishihara, T.; Watanabe, N.; Inoue, S.; Aoki, H.; Tsuji, T.; Yamamoto, B.; Yanagi, H.; Oki, M.; Kryukov, K.; Nakagawa, S.; et al. Usefulness of next-Generation DNA Sequencing for the Diagnosis of Urinary Tract Infection. *Drug Discov. Ther.* **2020**, *14*, 42–49. [[CrossRef](#)]
87. Fida, M.; Wolf, M.J.; Hamdi, A.; Vijayvargiya, P.; Esquer Garrigos, Z.; Khalil, S.; Greenwood-Quaintance, K.E.; Thoendel, M.J.; Patel, R. Detection of Pathogenic Bacteria From Septic Patients Using 16S Ribosomal RNA Gene-Targeted Metagenomic Sequencing. *Clin. Infect. Dis.* **2021**, *73*, 1165–1172. [[CrossRef](#)]
88. Oberauner, L.; Zachow, C.; Lackner, S.; Högenauer, C.; Smolle, K.-H.; Berg, G. The Ignored Diversity: Complex Bacterial Communities in Intensive Care Units Revealed by 16S Pyrosequencing. *Sci. Rep.* **2013**, *3*, 1413. [[CrossRef](#)]
89. Edgar, R.C. Updating the 97% Identity Threshold for 16S Ribosomal RNA OTUs. *Bioinformatics* **2018**, *34*, 2371–2375. [[CrossRef](#)]
90. Mysara, M.; Leys, N.; Raes, J.; Monsieurs, P. IPED: A Highly Efficient Denoising Tool for Illumina MiSeq Paired-End 16S rRNA Gene Amplicon Sequencing Data. *BMC Bioinform.* **2016**, *17*, 192. [[CrossRef](#)]
91. 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](#)] [[PubMed](#)]
92. Amir, A.; McDonald, D.; Navas-Molina, J.A.; Kopylova, E.; Morton, J.T.; Zech Xu, Z.; Kightley, E.P.; Thompson, L.R.; Hyde, E.R.; Gonzalez, A.; et al. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems* **2017**, *2*, e00191-16. [[CrossRef](#)] [[PubMed](#)]
93. Edgar, R.C. Search and Clustering Orders of Magnitude Faster than BLAST. *Bioinformatics* **2010**, *26*, 2460–2461. [[CrossRef](#)]
94. Rognes, T.; Flouri, T.; Nichols, B.; Quince, C.; Mahé, F. VSEARCH: A Versatile Open Source Tool for Metagenomics. *PeerJ* **2016**, *4*, e2584. [[CrossRef](#)] [[PubMed](#)]
95. Kopylova, E.; Navas-Molina, J.A.; Mercier, C.; Xu, Z.Z.; Mahé, F.; He, Y.; Zhou, H.-W.; Rognes, T.; Caporaso, J.G.; Knight, R. Open-Source Sequence Clustering Methods Improve the State Of the Art. *mSystems* **2016**, *1*, e00003-15. [[CrossRef](#)]
96. Nguyen, N.-P.; Warnow, T.; Pop, M.; White, B. A Perspective on 16S rRNA Operational Taxonomic Unit Clustering Using Sequence Similarity. *NPJ Biofilms Microbiomes* **2016**, *2*, 16004. [[CrossRef](#)]
97. Johnson, J.S.; Spakowicz, D.J.; Hong, B.-Y.; Petersen, L.M.; Demkowicz, P.; Chen, L.; Leopold, S.R.; Hanson, B.M.; Agresta, H.O.; Gerstein, M.; et al. Evaluation of 16S rRNA Gene Sequencing for Species and Strain-Level Microbiome Analysis. *Nat. Commun.* **2019**, *10*, 5029. [[CrossRef](#)]
98. Cai, H.; Archambault, M.; Prescott, J.F. 16S Ribosomal RNA Sequence-Based Identification of Veterinary Clinical Bacteria. *J. Vet. Diagn. Investig.* **2003**, *15*, 465–469. [[CrossRef](#)]
99. Hoffman, C.; Siddiqui, N.Y.; Fields, I.; Gregory, W.T.; Simon, H.M.; Mooney, M.A.; Wolfe, A.J.; Karstens, L. Species-Level Resolution of Female Bladder Microbiota from 16S rRNA Amplicon Sequencing. *mSystems* **2021**, *6*, e0051821. [[CrossRef](#)]
100. Meola, M.; Rifa, E.; Shani, N.; Delbès, C.; Berthoud, H.; Chassard, C. DAIRYdb: A Manually Curated Reference Database for Improved Taxonomy Annotation of 16S rRNA Gene Sequences from Dairy Products. *BMC Genomics* **2019**, *20*, 560. [[CrossRef](#)]
101. Peker, N.; Garcia-Croes, S.; Dijkhuizen, B.; Wiersma, H.H.; van Zanten, E.; Wisselink, G.; Friedrich, A.W.; Kooistra-Smid, M.; Sinha, B.; Rossen, J.W.A.; et al. A Comparison of Three Different Bioinformatics Analyses of the 16S-23S rRNA Encoding Region for Bacterial Identification. *Front. Microbiol.* **2019**, *10*, 620. [[CrossRef](#)]
102. Seol, D.; Lim, J.S.; Sung, S.; Lee, Y.H.; Jeong, M.; Cho, S.; Kwak, W.; Kim, H. Microbial Identification Using rRNA Operon Region: Database and Tool for Metataxonomics with Long-Read Sequence. *Microbiol. Spectr.* **2022**, *10*, e0201721. [[CrossRef](#)] [[PubMed](#)]
103. Douglas, G.M.; Maffei, V.J.; Zaneveld, J.R.; Yurgel, S.N.; Brown, J.R.; Taylor, C.M.; Huttenhower, C.; Langille, M.G.I. PICRUSt2 for Prediction of Metagenome Functions. *Nat. Biotechnol.* **2020**, *38*, 685–688. [[CrossRef](#)]
104. Bowman, J.S.; Ducklow, H.W. Microbial Communities Can Be Described by Metabolic Structure: A General Framework and Application to a Seasonally Variable, Depth-Stratified Microbial Community from the Coastal West Antarctic Peninsula. *PLoS ONE* **2015**, *10*, e0135868. [[CrossRef](#)]
105. Aßhauer, K.P.; Wemheuer, B.; Daniel, R.; Meinicke, P. Tax4Fun: Predicting Functional Profiles from Metagenomic 16S rRNA Data. *Bioinformatics* **2015**, *31*, 2882–2884. [[CrossRef](#)]
106. Laroche, O.; Pochon, X.; Wood, S.A.; Keeley, N. Beyond Taxonomy: Validating Functional Inference Approaches in the Context of Fish-Farm Impact Assessments. *Mol. Ecol. Resour.* **2021**, *21*, 2264–2277. [[CrossRef](#)]
107. Quince, C.; Walker, A.W.; Simpson, J.T.; Loman, N.J.; Segata, N. Shotgun Metagenomics, from Sampling to Analysis. *Nat. Biotechnol.* **2017**, *35*, 833–844. [[CrossRef](#)]

108. Pérez-Cobas, A.E.; Gomez-Valero, L.; Buchrieser, C. Metagenomic Approaches in Microbial Ecology: An Update on Whole-Genome and Marker Gene Sequencing Analyses. *Microb. Genom.* **2020**, *6*, mgen000409. [[CrossRef](#)]
109. Junier, T.; Huber, M.; Schmutz, S.; Kufner, V.; Zagordi, O.; Neuenschwander, S.; Ramette, A.; Kubacki, J.; Bachofen, C.; Qi, W.; et al. Viral Metagenomics in the Clinical Realm: Lessons Learned from a Swiss-Wide Ring Trial. *Genes* **2019**, *10*, 655. [[CrossRef](#)]
110. De Vries, J.J.C.; Brown, J.R.; Fischer, N.; Sidorov, I.A.; Morfopoulou, S.; Huang, J.; Munnink, B.B.O.; Sayiner, A.; Bulgurcu, A.; Rodriguez, C.; et al. Benchmark of Thirteen Bioinformatic Pipelines for Metagenomic Virus Diagnostics Using Datasets from Clinical Samples. *J. Clin. Virol.* **2021**, *141*, 104908. [[CrossRef](#)]
111. Wilson, M.R.; Sample, H.A.; Zorn, K.C.; Arevalo, S.; Yu, G.; Neuhaus, J.; Federman, S.; Stryke, D.; Briggs, B.; Langelier, C.; et al. Clinical Metagenomic Sequencing for Diagnosis of Meningitis and Encephalitis. *N. Engl. J. Med.* **2019**, *380*, 2327–2340. [[CrossRef](#)] [[PubMed](#)]
112. Shishido, A.A.; Noe, M.; Saharia, K.; Luethy, P. Clinical Impact of a Metagenomic Microbial Plasma Cell-Free DNA next-Generation Sequencing Assay on Treatment Decisions: A Single-Center Retrospective Study. *BMC Infect. Dis.* **2022**, *22*, 372. [[CrossRef](#)] [[PubMed](#)]
113. Zhang, R.; Zhuang, Y.; Xiao, Z.-H.; Li, C.-Y.; Zhang, F.; Huang, W.-Q.; Zhang, M.; Peng, X.-M.; Liu, C. Diagnosis and Surveillance of Neonatal Infections by Metagenomic Next-Generation Sequencing. *Front. Microbiol.* **2022**, *13*, 855988. [[CrossRef](#)]
114. Li, D.; Gai, W.; Zhang, J.; Cheng, W.; Cui, N.; Wang, H. Metagenomic Next-Generation Sequencing for the Microbiological Diagnosis of Abdominal Sepsis Patients. *Front. Microbiol.* **2022**, *13*, 816631. [[CrossRef](#)] [[PubMed](#)]
115. He, Y.; Fang, K.; Shi, X.; Yang, D.; Zhao, L.; Yu, W.; Zheng, Y.; Xu, Y.; Ma, X.; Chen, L.; et al. Enhanced DNA and RNA Pathogen Detection via Metagenomic Sequencing in Patients with Pneumonia. *J. Transl. Med.* **2022**, *20*, 195. [[CrossRef](#)] [[PubMed](#)]
116. Charalampous, T.; Kay, G.L.; Richardson, H.; Aydin, A.; Baldan, R.; Jeanes, C.; Rae, D.; Grundy, S.; Turner, D.J.; Wain, J.; et al. Nanopore Metagenomics Enables Rapid Clinical Diagnosis of Bacterial Lower Respiratory Infection. *Nat. Biotechnol.* **2019**, *37*, 783–792. [[CrossRef](#)]
117. Sanabria, A.; Hjerde, E.; Johannessen, M.; Sollid, J.E.; Simonsen, G.S.; Hanssen, A.-M. Shotgun-Metagenomics on Positive Blood Culture Bottles Inoculated With Prosthetic Joint Tissue: A Proof of Concept Study. *Front. Microbiol.* **2020**, *11*, 1687. [[CrossRef](#)]
118. Noone, J.C.; Helmersen, K.; Leegaard, T.M.; Skråmm, I.; Aamot, H.V. Rapid Diagnostics of Orthopaedic-Implant-Associated Infections Using Nanopore Shotgun Metagenomic Sequencing on Tissue Biopsies. *Microorganisms* **2021**, *9*, 97. [[CrossRef](#)]
119. Lamoureux, C.; Surgers, L.; Fihman, V.; Gricourt, G.; Demontant, V.; Trawinski, E.; N'Debi, M.; Gomart, C.; Royer, G.; Launay, N.; et al. Prospective Comparison Between Shotgun Metagenomics and Sanger Sequencing of the 16S rRNA Gene for the Etiological Diagnosis of Infections. *Front. Microbiol.* **2022**, *13*, 761873. [[CrossRef](#)]
120. Gu, W.; Deng, X.; Lee, M.; Sucu, Y.D.; Arevalo, S.; Stryke, D.; Federman, S.; Gopez, A.; Reyes, K.; Zorn, K.; et al. Rapid Pathogen Detection by Metagenomic next-Generation Sequencing of Infected Body Fluids. *Nat. Med.* **2021**, *27*, 115–124. [[CrossRef](#)]
121. Martin, S.; Heavens, D.; Lan, Y.; Horsfield, S.; Clark, M.D.; Leggett, R.M. Nanopore Adaptive Sampling: A Tool for Enrichment of Low Abundance Species in Metagenomic Samples. *Genome Biol.* **2022**, *23*, 11. [[CrossRef](#)] [[PubMed](#)]
122. Gan, M.; Wu, B.; Yan, G.; Li, G.; Sun, L.; Lu, G.; Zhou, W. Combined Nanopore Adaptive Sequencing and Enzyme-Based Host Depletion Efficiently Enriched Microbial Sequences and Identified Missing Respiratory Pathogens. *BMC Genom.* **2021**, *22*, 732. [[CrossRef](#)] [[PubMed](#)]
123. Marquet, M.; Zöllkau, J.; Pastuschek, J.; Viehweger, A.; Schleußner, E.; Makarewicz, O.; Pletz, M.W.; Ehrlich, R.; Brandt, C. Evaluation of Microbiome Enrichment and Host DNA Depletion in Human Vaginal Samples Using Oxford Nanopore's Adaptive Sequencing. *Sci. Rep.* **2022**, *12*, 4000. [[CrossRef](#)]
124. Arikawa, K.; Ide, K.; Kogawa, M.; Saeki, T.; Yoda, T.; Endoh, T.; Matsushashi, A.; Takeyama, H.; Hosokawa, M. Recovery of Strain-Resolved Genomes from Human Microbiome through an Integration Framework of Single-Cell Genomics and Metagenomics. *Microbiome* **2021**, *9*, 202. [[CrossRef](#)] [[PubMed](#)]
125. Vicedomini, R.; Quince, C.; Darling, A.E.; Chikhi, R. Strainberry: Automated Strain Separation in Low-Complexity Metagenomes Using Long Reads. *Nat. Commun.* **2021**, *12*, 4485. [[CrossRef](#)]
126. Quince, C.; Nurk, S.; Raguideau, S.; James, R.; Soyer, O.S.; Summers, J.K.; Limasset, A.; Eren, A.M.; Chikhi, R.; Darling, A.E. STRONG: Metagenomics Strain Resolution on Assembly Graphs. *Genome Biol.* **2021**, *22*, 214. [[CrossRef](#)] [[PubMed](#)]
127. Edgar, R.C.; Taylor, J.; Lin, V.; Altman, T.; Barbera, P.; Meleshko, D.; Lohr, D.; Novakovsky, G.; Buchfink, B.; Al-Shayeb, B.; et al. Petabase-Scale Sequence Alignment Catalyses Viral Discovery. *Nature* **2022**, *602*, 142–147. [[CrossRef](#)]
128. Zayed, A.A.; Wainaina, J.M.; Dominguez-Huerta, G.; Pelletier, E.; Guo, J.; Mohssen, M.; Tian, F.; Pratama, A.A.; Bolduc, B.; Zablocki, O.; et al. Cryptic and Abundant Marine Viruses at the Evolutionary Origins of Earth's RNA Virome. *Science* **2022**, *376*, 156–162. [[CrossRef](#)]
129. Djemiel, C.; Maron, P.-A.; Terrat, S.; Dequiedt, S.; Cottin, A.; Ranjard, L. Inferring Microbiota Functions from Taxonomic Genes: A Review. *Gigascience* **2022**, *11*, giab090. [[CrossRef](#)]
130. Belcour, A.; Frioux, C.; Aite, M.; Bretaudeau, A.; Hildebrand, F.; Siegel, A. Metage2Metabo, Microbiota-Scale Metabolic Complementarity for the Identification of Key Species. *Elife* **2020**, *9*, e61968. [[CrossRef](#)] [[PubMed](#)]
131. Huerta-Cepas, J.; Szklarczyk, D.; Heller, D.; Hernández-Plaza, A.; Forslund, S.K.; Cook, H.; Mende, D.R.; Letunic, I.; Rattei, T.; Jensen, L.J.; et al. eggNOG 5.0: A Hierarchical, Functionally and Phylogenetically Annotated Orthology Resource Based on 5090 Organisms and 2502 Viruses. *Nucleic Acids Res.* **2019**, *47*, D309–D314. [[CrossRef](#)] [[PubMed](#)]

132. Chng, K.R.; Li, C.; Bertrand, D.; Ng, A.H.Q.; Kwah, J.S.; Low, H.M.; Tong, C.; Natrajan, M.; Zhang, M.H.; Xu, L.; et al. Cartography of Opportunistic Pathogens and Antibiotic Resistance Genes in a Tertiary Hospital Environment. *Nat. Med.* **2020**, *26*, 941–951. [[CrossRef](#)] [[PubMed](#)]
133. Urban, L.; Holzer, A.; Baronas, J.J.; Hall, M.B.; Braeuninger-Weimer, P.; Scherm, M.J.; Kunz, D.J.; Perera, S.N.; Martin-Herranz, D.E.; Tipper, E.T.; et al. Freshwater Monitoring by Nanopore Sequencing. *Elife* **2021**, *10*, e61504. [[CrossRef](#)]
134. Mallapaty, S. How Sewage Could Reveal True Scale of Coronavirus Outbreak. *Nature* **2020**, *580*, 176–177. [[CrossRef](#)] [[PubMed](#)]
135. Hendriksen, R.S.; Munk, P.; Njage, P.; van Bunnik, B.; McNally, L.; Lukjancenko, O.; Röder, T.; Nieuwenhuijse, D.; Pedersen, S.K.; Kjeldgaard, J.; et al. Global Monitoring of Antimicrobial Resistance Based on Metagenomics Analyses of Urban Sewage. *Nat. Commun.* **2019**, *10*, 1124. [[CrossRef](#)]
136. Perry, M.R.; Lepper, H.C.; McNally, L.; Wee, B.A.; Munk, P.; Warr, A.; Moore, B.; Kalima, P.; Philip, C.; de Roda Husman, A.M.; et al. Secrets of the Hospital Underbelly: Patterns of Abundance of Antimicrobial Resistance Genes in Hospital Wastewater Vary by Specific Antimicrobial and Bacterial Family. *Front. Microbiol.* **2021**, *12*, 703560. [[CrossRef](#)]
137. Röttjers, L.; Faust, K. Fast and Flexible Analysis of Linked Microbiome Data with Mako. *Nat. Methods* **2022**, *19*, 51–54. [[CrossRef](#)]
138. Kishikawa, T.; Maeda, Y.; Nii, T.; Motooka, D.; Matsumoto, Y.; Matsushita, M.; Matsuoka, H.; Yoshimura, M.; Kawada, S.; Teshigawara, S.; et al. Metagenome-Wide Association Study of Gut Microbiome Revealed Novel Aetiology of Rheumatoid Arthritis in the Japanese Population. *Ann. Rheum. Dis.* **2020**, *79*, 103–111. [[CrossRef](#)]
139. Wang, D.; Doestzada, M.; Chen, L.; Andreu-Sánchez, S.; van den Munckhof, I.C.L.; Augustijn, H.E.; Koehorst, M.; Ruiz-Moreno, A.J.; Bloks, V.W.; Riksen, N.P.; et al. Characterization of Gut Microbial Structural Variations as Determinants of Human Bile Acid Metabolism. *Cell Host Microbe* **2021**, *29*, 1802–1814.e5. [[CrossRef](#)]
140. Tomofuji, Y.; Maeda, Y.; Oguro-Igashira, E.; Kishikawa, T.; Yamamoto, K.; Sonehara, K.; Motooka, D.; Matsumoto, Y.; Matsuoka, H.; Yoshimura, M.; et al. Metagenome-Wide Association Study Revealed Disease-Specific Landscape of the Gut Microbiome of Systemic Lupus Erythematosus in Japanese. *Ann. Rheum. Dis.* **2021**, *80*, 1575–1583. [[CrossRef](#)]
141. Kishikawa, T.; Tomofuji, Y.; Inohara, H.; Okada, Y. OMARU: A Robust and Multifaceted Pipeline for Metagenome-Wide Association Study. *NAR Genom Bioinform* **2022**, *4*, lqac019. [[CrossRef](#)] [[PubMed](#)]
142. Vega, L.; Herrera, G.; Muñoz, M.; Patarroyo, M.A.; Ramírez, J.D. Occurrence of in Patients with Infection. *Pathogens* **2020**, *9*, 283. [[CrossRef](#)] [[PubMed](#)]
143. Casimiro-Soriguer, C.S.; Loucera, C.; Peña-Chilet, M.; Dopazo, J. Towards a Metagenomics Machine Learning Interpretable Model for Understanding the Transition from Adenoma to Colorectal Cancer. *Sci. Rep.* **2022**, *12*, 450. [[CrossRef](#)] [[PubMed](#)]
144. Liu, N.-N.; Jiao, N.; Tan, J.-C.; Wang, Z.; Wu, D.; Wang, A.-J.; Chen, J.; Tao, L.; Zhou, C.; Fang, W.; et al. Multi-Kingdom Microbiota Analyses Identify Bacterial-Fungal Interactions and Biomarkers of Colorectal Cancer across Cohorts. *Nat. Microbiol.* **2022**, *7*, 238–250. [[CrossRef](#)]
145. Thomas, A.M.; Manghi, P.; Asnicar, F.; Pasolli, E.; Armanini, F.; Zolfo, M.; Beghini, F.; Manara, S.; Karcher, N.; Pozzi, C.; et al. Metagenomic Analysis of Colorectal Cancer Datasets Identifies Cross-Cohort Microbial Diagnostic Signatures and a Link with Choline Degradation. *Nat. Med.* **2019**, *25*, 667–678. [[CrossRef](#)]
146. Wirbel, J.; Pyl, P.T.; Kartal, E.; Zych, K.; Kashani, A.; Milanese, A.; Fleck, J.S.; Voigt, A.Y.; Palleja, A.; Ponnudurai, R.; et al. Meta-Analysis of Fecal Metagenomes Reveals Global Microbial Signatures That Are Specific for Colorectal Cancer. *Nat. Med.* **2019**, *25*, 679–689. [[CrossRef](#)]
147. Hayase, E.; Jenq, R.R. Role of the Intestinal Microbiome and Microbial-Derived Metabolites in Immune Checkpoint Blockade Immunotherapy of Cancer. *Genome Med.* **2021**, *13*, 107. [[CrossRef](#)]
148. Dekker, J.P. Metagenomics for Clinical Infectious Disease Diagnostics Steps Closer to Reality. *J. Clin. Microbiol.* **2018**, *56*, e00850-18. [[CrossRef](#)]
149. Wang, W.-L.; Xu, S.-Y.; Ren, Z.-G.; Tao, L.; Jiang, J.-W.; Zheng, S.-S. Application of Metagenomics in the Human Gut Microbiome. *World J. Gastroenterol.* **2015**, *21*, 803–814. [[CrossRef](#)]
150. Ahn, T.-H.; Chai, J.; Pan, C. Sigma: Strain-Level Inference of Genomes from Metagenomic Analysis for Biosurveillance. *Bioinformatics* **2015**, *31*, 170–177. [[CrossRef](#)]
151. Ko, K.K.K.; Chng, K.R.; Nagarajan, N. Metagenomics-Enabled Microbial Surveillance. *Nat. Microbiol.* **2022**, *7*, 486–496. [[CrossRef](#)]
152. Sukhum, K.V.; Diorio-Toth, L.; Dantas, G. Genomic and Metagenomic Approaches for Predictive Surveillance of Emerging Pathogens and Antibiotic Resistance. *Clin. Pharmacol. Ther.* **2019**, *106*, 512–524. [[CrossRef](#)] [[PubMed](#)]
153. Pulido-Tamayo, S.; Sánchez-Rodríguez, A.; Swings, T.; Van den Bergh, B.; Dubey, A.; Steenackers, H.; Michiels, J.; Fostier, J.; Marchal, K. Frequency-Based Haplotype Reconstruction from Deep Sequencing Data of Bacterial Populations. *Nucleic Acids Res.* **2015**, *43*, e105. [[CrossRef](#)] [[PubMed](#)]
154. Tringe, S.G.; Rubin, E.M. Metagenomics: DNA Sequencing of Environmental Samples. *Nat. Rev. Genet.* **2005**, *6*, 805–814. [[CrossRef](#)] [[PubMed](#)]
155. Tringe, S.G.; von Mering, C.; Kobayashi, A.; Salamov, A.A.; Chen, K.; Chang, H.W.; Podar, M.; Short, J.M.; Mathur, E.J.; Detter, J.C.; et al. Comparative Metagenomics of Microbial Communities. *Science* **2005**, *308*, 554–557. [[CrossRef](#)]
156. Roberts, L.W.; Harris, P.N.A.; Forde, B.M.; Ben Zakour, N.L.; Catchpoole, E.; Stanton-Cook, M.; Phan, M.-D.; Sidjabat, H.E.; Bergh, H.; Heney, C.; et al. Integrating Multiple Genomic Technologies to Investigate an Outbreak of Carbapenemase-Producing *Enterobacter hormaechei*. *Nat. Commun.* **2020**, *11*, 466. [[CrossRef](#)]

157. Roberts, L.W.; Forde, B.M.; Hurst, T.; Ling, W.; Nimmo, G.R.; Bergh, H.; George, N.; Hajkovicz, K.; McNamara, J.F.; Lipman, J.; et al. Genomic Surveillance, Characterization and Intervention of a Polymicrobial Multidrug-Resistant Outbreak in Critical Care. *Microb. Genom.* **2021**, *7*, mgen000530. [[CrossRef](#)]
158. Zheng, W.; Zhao, S.; Yin, Y.; Zhang, H.; Needham, D.M.; Evans, E.D.; Dai, C.L.; Lu, P.J.; Alm, E.J.; Weitz, D.A. High-Throughput, Single-Microbe Genomics with Strain Resolution, Applied to a Human Gut Microbiome. *Science* **2022**, *376*, eabm1483. [[CrossRef](#)]
159. Anyansi, C.; Straub, T.J.; Manson, A.L.; Earl, A.M.; Abeel, T. Computational Methods for Strain-Level Microbial Detection in Colony and Metagenome Sequencing Data. *Front. Microbiol.* **2020**, *11*, 1925. [[CrossRef](#)]
160. Votintseva, A.A.; Bradley, P.; Pankhurst, L.; Del Ojo Elias, C.; Loose, M.; Nilgiriwala, K.; Chatterjee, A.; Smith, E.G.; Sanderson, N.; Walker, T.M.; et al. Same-Day Diagnostic and Surveillance Data for Tuberculosis via Whole-Genome Sequencing of Direct Respiratory Samples. *J. Clin. Microbiol.* **2017**, *55*, 1285–1298. [[CrossRef](#)]
161. Marcelino, V.R.; Holmes, E.C.; Sorrell, T.C. The Use of Taxon-Specific Reference Databases Compromises Metagenomic Classification. *BMC Genom.* **2020**, *21*, 184. [[CrossRef](#)] [[PubMed](#)]
162. Zlitni, S.; Bishara, A.; Moss, E.L.; Tkachenko, E.; Kang, J.B.; Culver, R.N.; Andermann, T.M.; Weng, Z.; Wood, C.; Handy, C.; et al. Strain-Resolved Microbiome Sequencing Reveals Mobile Elements That Drive Bacterial Competition on a Clinical Timescale. *Genome Med.* **2020**, *12*, 50. [[CrossRef](#)] [[PubMed](#)]
163. Ivanova, V.; Chervenskaya, E.; Vasiluev, P.; Ivanov, A.; Tolstogonov, I.; Shafranskaya, D.; Ulyantsev, V.; Korobeynikov, A.; Razin, S.V.; Beloborodova, N.; et al. Hi-C Metagenomics in the ICU: Exploring Clinically Relevant Features of Gut Microbiome in Chronically Critically Ill Patients. *Front. Microbiol.* **2021**, *12*, 770323. [[CrossRef](#)]
164. Li, H.; Homer, N. A Survey of Sequence Alignment Algorithms for next-Generation Sequencing. *Brief. Bioinform.* **2010**, *11*, 473–483. [[CrossRef](#)]
165. Hong, C.; Manimaran, S.; Shen, Y.; Perez-Rogers, J.F.; Byrd, A.L.; Castro-Nallar, E.; Crandall, K.A.; Johnson, W.E. PathoScope 2.0: A Complete Computational Framework for Strain Identification in Environmental or Clinical Sequencing Samples. *Microbiome* **2014**, *2*, 33. [[CrossRef](#)] [[PubMed](#)]
166. Clement, N.L.; Snell, Q.; Clement, M.J.; Hollenhorst, P.C.; Purwar, J.; Graves, B.J.; Cairns, B.R.; Johnson, W.E. The GNUMAP Algorithm: Unbiased Probabilistic Mapping of Oligonucleotides from next-Generation Sequencing. *Bioinformatics* **2010**, *26*, 38–45. [[CrossRef](#)] [[PubMed](#)]
167. Langmead, B.; Salzberg, S.L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)]
168. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [[CrossRef](#)]
169. Truong, D.T.; Franzosa, E.A.; Tickle, T.L.; Scholz, M.; Weingart, G.; Pasolli, E.; Tett, A.; Huttenhower, C.; Segata, N. MetaPhlan2 for Enhanced Metagenomic Taxonomic Profiling. *Nat. Methods* **2015**, *12*, 902–903. [[CrossRef](#)]
170. Wood, D.E.; Lu, J.; Langmead, B. Improved Metagenomic Analysis with Kraken 2. *Genome Biol.* **2019**, *20*, 257. [[CrossRef](#)]
171. Nayfach, S.; Rodriguez-Mueller, B.; Garud, N.; Pollard, K.S. An Integrated Metagenomics Pipeline for Strain Profiling Reveals Novel Patterns of Bacterial Transmission and Biogeography. *Genome Res.* **2016**, *26*, 1612–1625. [[CrossRef](#)]
172. Tu, Q.; He, Z.; Zhou, J. Strain/species Identification in Metagenomes Using Genome-Specific Markers. *Nucleic Acids Res.* **2014**, *42*, e67. [[CrossRef](#)] [[PubMed](#)]
173. Laudadio, I.; Fulci, V.; Palone, F.; Stronati, L.; Cucchiara, S.; Carissimi, C. Quantitative Assessment of Shotgun Metagenomics and 16S rDNA Amplicon Sequencing in the Study of Human Gut Microbiome. *OMICS* **2018**, *22*, 248–254. [[CrossRef](#)] [[PubMed](#)]
174. Weaver, A.A.; Hasan, N.A.; Klaassen, M.; Karathia, H.; Colwell, R.R.; Shrout, J.D. Prosthetic Joint Infections Present Diverse and Unique Microbial Communities Using Combined Whole-Genome Shotgun Sequencing and Culturing Methods. *J. Med. Microbiol.* **2019**, *68*, 1507–1516. [[CrossRef](#)] [[PubMed](#)]
175. Brumfield, K.D.; Huq, A.; Colwell, R.R.; Olds, J.L.; Leddy, M.B. Microbial Resolution of Whole Genome Shotgun and 16S Amplicon Metagenomic Sequencing Using Publicly Available NEON Data. *PLoS ONE* **2020**, *15*, e0228899. [[CrossRef](#)] [[PubMed](#)]
176. Araos, R.; Battaglia, T.; Ugalde, J.A.; Rojas-Herrera, M.; Blaser, M.J.; D'Agata, E.M.C. Fecal Microbiome Characteristics and the Resistome Associated With Acquisition of Multidrug-Resistant Organisms Among Elderly Subjects. *Front. Microbiol.* **2019**, *10*, 2260. [[CrossRef](#)]