Journal of Microbiological Methods 230-231 (2025) 107103

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Methodological aspects of investigating the resistome in pig farm environments

Valeriia Ladyhina [a,c,*], Elisabeth Rajala [a], Susanna Sternberg-Lewerin [a], Leila Nasirzadeh [a,d,e], Erik Bongcam-Rudloff [a], Johan Dicksved [b]

[a] Department of Animal Biosciences, Swedish University of Agricultural Sciences, Uppsala, Sweden
[b] Department of Applied Animal Science and Welfare, Swedish University of Agricultural Sciences, Uppsala, Sweden
[c] Uppsala Antibiotic Center, Uppsala University, Uppsala, Sweden
[d] Bioinformatics Unit, Core Facility (KEF), Faculty of Medical and Health Sciences (BKV), Linköping University, Linköping, Sweden
[e] Clinical Genomics Linköping, SciLife Laboratory, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

ARTICLE INFO

Keywords:
Antimicrobial resistance genes
Metagenomics
Microbiome
Environmental samples
Nanopore sequencing
Illumina sequencing

ABSTRACT

A typical One Health issue, antimicrobial resistance (AMR) development and its spread among people, animals, and the environment attracts significant research attention. The animal sector is one of the major contributors to the development and dissemination of AMR and accounts for more than 50 % of global antibiotics usage. The use of antibiotics exerts a selective pressure for resistant bacteria in the exposed microbiome, but many questions about the epidemiology of AMR in farm environments remain unanswered. This is connected to several methodological challenges and limitations, such as inconsistent sampling methods, complexity of farm environment samples and the lack of standardized protocols for sample collection, processing and bioinformatical analysis. In this project, we combined metagenomics and bioinformatics to optimise the methodology for reproducible research on the resistome in complex samples from the indoor farm environment. The work included optimizing sample collection, transportation, and storage, as well as DNA extraction, sequencing, and bioinformatic analysis, such as metagenome assembly and antibiotic resistance gene (ARG) detection. Our studies suggest that the current most optimal and cost-effective pipeline for ARG search should be based on Illumina sequencing of sock sample material at high depth (at least 25 M 250 bp PE for AMR gene families and 43 M for gene variants). We present a computational analysis utilizing MEGAHIT assembly to balance the identification of bacteria carrying ARGs with the potential loss of diversity and abundance of resistance genes. Our findings indicate that searching against multiple ARG databases is essential for detecting the highest diversity of ARGs.

1. Introduction

Antimicrobial resistance (AMR) is a typical One Health problem that imposes a considerable burden on global health and economy (Murray et al., 2022). AMR in pathogenic bacteria poses substantial challenges to effective treatment, leading to increased healthcare costs both in human and animal medicine. Globally, intensive livestock production is among the major consumers of antimicrobials (Van Boeckel et al., 2015). The increasing prevalence of resistant bacteria in livestock raises concerns for animal health. In addition, livestock may serve as a reservoir for resistant zoonotic bacteria and resistance genes that can transfer to human pathogens (Tang et al., 2017). Antimicrobial usage (AMU) is recognized to be one of the main causes for the emergence of AMR in bacteria found in humans, animals, and the environment. Recent studies on livestock have shown that there are additional risk factors such as different biosecurity measures, and co-selection of antibiotic resistance genes (ARGs) in the presence of biocide and metal resistance genes within the farm environment (Horie et al., 2021; Li et al., 2022).

Most surveillance of AMR is based on culturing of indicator bacteria. This approach provides an insight into the phenotypic resistance but may not be representative for the overall occurrence of ARGs in the commensal microbiota as it exludes non-culturable bacteria (Andersen et al., 2017). Another popular screening method is quantitative real-time polymerase chain reaction (qPCR) focused on a limited selection

* Corresponding author at: Swedish University of Agricultural Sciences, SE-750 07 Uppsala, Sweden.
E-mail addresses: valeriia.ladyhina@slu.se (V. Ladyhina), elisabeth.rajala@slu.se (E. Rajala), susanna.sternberg-lewerin@slu.se (S. Sternberg-Lewerin), leila.nasirzadeh@liu.se (L. Nasirzadeh), erik.bongcam@slu.se (E. Bongcam-Rudloff), johan.dicksved@slu.se (J. Dicksved).

https://doi.org/10.1016/j.mimet.2025.107103
Received 17 December 2024; Received in revised form 12 February 2025; Accepted 12 February 2025
Available online 13 February 2025

of ARGs (Graesboll et al., 2019; Luiken et al., 2022; Yang et al., 2022). However, this method is insufficient to obtain a comprehensive picture of the bacterial community and the resistome. Recent technological developments in molecular biology have increased our ability to investigate and quantify the gene content of microbial communities by whole metagenome sequencing. The main advantage of using metagenomics is the ability to look at the whole microbiome community including bacterial species that cannot be cultured in vitro and the resistome, including both known and novel ARGs in samples, improving our understanding of microbial communities (Noyes et al., 2016; Pillay et al., 2022). Although metagenomics is a very promising technique for resistance analysis, there is still only a limited number of studies that use this technique, escpecially in the area of non-clinical longitudinal research. This can be explained by the complexity of samples as well as a lack of standardized protocols for sample collection and further sample manipulation, including steps of sequencing and bioinformatic analysis. Previous methodology-based studies usually focused on comparison of one specific step of the resistome analysis pipeline. Many studies address the problem of randomisation and pooling of individual animal samples or fresh droppings (Yamamoto et al., 2014; Munk et al., 2017; Andersen et al., 2021). Another question for debate is the selection of sequencing technique: short reads, long read based sequencing or a combination of both methods. Most of the currently published studies on the resistome in farm environments are based on shotgun metagenome sequencing (Illumina) of individual fecal samples, which is still seen as a golden standard in metagenome analysis of resistance due to the possibility to obtain big amounts of good quality sequence data at a low cost (Luiken et al., 2019; Van Gompel et al., 2019; Mencia-Ares et al., 2020; Stevens et al., 2023). However, the length of the short reads is not enough to cover the length of most ARGs and, therefore, a search on raw Illumina reads results in missing genomic context and can lead to appearance of false positives due to spurious mapping (Boolchandani et al., 2019). This problem can be solved in two ways: either perform assembly of short reads or use a sequencing technique based on long reads. There is a list of studies that suggest that long read sequencing (Oxford Nanopore Technologies) is the future for resistome research as it helps to avoid the computationally heavy step of assembly while providing enough of genome context information to not only recover ARGs but also their connection to the microbiome (Weinmaier et al., 2023; Sierra et al., 2024; Slizovskiy et al., 2024). However, long-read sequencing suffers from low-accuracy base calling and the excessive costs of high-depth sequencing needed for high complexity samples (Brown et al., 2021; Zhao et al., 2023; Abramova et al., 2024). Therefore, short-read sequencing is often required to improve genome accuracy, which increases costs and turnaround time (Brown et al., 2021; Zhao et al., 2023; Abramova et al., 2024). As stated earlier, another way to tackle the problem of detection of false positives is to perform assembly, however this is hampered by the tendency of metagenomic assemblies to break around ARGs (Abramova et al., 2024). Another issue is that each additional step of analysis, such as assembly or binning, results in some loss of data. Furthermore, different assemblers employ varying algorithms for constructing contigs, making the choice of assembler a trade-off between the quality of the resulting assemblies, computational requirements, the time required for assembly and the sequencing depth that the assembler can handle. Previous research also showed that sample complexity may affect the choice of the most suitable assembler (Abramova et al., 2024). Therefore, constant development of the technology of long read sequencing as well as new and improved assembly tools requires the re-evaluation of existing pipelines for ARG search for each specific sample type. Another important question that is often addressed in many studies but not thoroughly evaluated is the sequencing depth required to obtain an adequate representation of the resistome. To our knowledge, only one study has specifically addressed this question (Gweon et al., 2019). However, changes in sample material and the anticipated baseline levels of resistance may necessitate recalculating the appropriate sequencing depth. The last but not least step in resistome analysis is the search for ARGs and choice of a suitable database. There is a list of articles that compare advantages and disadvantages of different ARG-databases based on sample and goal for research (de Abreu et al., 2020; Papp and Solymosi, 2022). While metagenomic sequencing has advanced resistome analysis, most previous studies have focused on individual steps of the pipeline, such as sequencing techniques or assembly methods, rather than addressing the pipeline as a whole. We propose that a comprehensive approach, tailored to sample type, complexity, and research goals, is crucial for accurate and reliable ARG detection.

To our knowledge, this study is the first to focus on a comprehensive evaluation of different steps in the environmental metagenomics research pipeline with the aim of advancing farm resistome analysis. This included evaluating various sample transportation temperatures, comparing usage of short and long-read sequencing techniques for microbiome analysis as well as evaluation of tools for assembly and ARG search in the bioinformatical pipeline for investigation of the environmental resistome.

2. Materials and methods

2.1. Sampling and experimental design

The material used in this study originated from environmental sock samples (Fig. 1) collected from the Swedish University of Agricultural Sciences (SLU) pig farm at the Swedish Livestock Research Centre in Lövsta, outside Uppsala, Sweden. Samples were collected in March 2022 from four pens, each housing 10 to 13 three-month-old pigs from the same batch. The pens were located in different parts of the same room. All pigs were managed the same way, except that the pigs in pen number one had recently been treated with penicillin. Sterile gauze socks were put on clean boots covered with clean disposable foot cover. One pair of socks was used in each pen and in every pen, 100 steps were taken, turning the socks 90° after every 25 steps. The study was divided into two parts.

2.1.1. Part 1

Technical aliquots of one sample (Sample 0) were used for comparison of three commercial DNA extraction kits in terms of their influence on the taxonomic composition of the sample as well as compatibility with long-read sequencing (Oxford Nanopore Technologies plc., Oxford, UK (ONT)), which relies on high-quality DNA with long, intact fragments for optimal performance. This sample was handled on the day of collection and kept at −80 °C for long-term storage.

2.1.2. Part 2

Seven samples were used for investigation of the effect of

Fig. 1. Sock sampling: 20 cm long elastic tubular retention bandage over single use plastic boot coverage.

temperature during transportation and sequencing technique on the taxonomical profile and abundance of ARGs, as well as for the development of a bioinformatics pipeline, addressing crucial steps of assembly and ARG search. To simulate different transportation methods, we compared three schemes of storage before extraction of sock samples in the lab. For simulation of transportation of material in the best conditions (on dry ice), sock samples were handled on the day of sampling. To mock sample transportation via ordinary mail, a two-day storage was done at two different temperatures: room temperature (RT, +20 °C) for samples that would be transported without cooling, and + 4 °C to mock sample transportation on ice pack. For handling of the samples, the socks were soaked in buffered peptone water (50 ml per sock) and put in a stomacher to extract sample material. The resulting suspension was concentrated by centrifuging (3000 g for 10 min) and the pellet was resuspended in 15 ml of supernatant. After this, samples were transferred to 1.8 ml cryotubes for storage at −80 °C. A more detailed description of the study procedure can be found in Table 1.

2.2. DNA extraction

DNA was extracted from three technical replicates of sample 0 in order to compare three commercially available kits: (1) ZM = ZYMOBiomics MagBead DNA/RNA (Zymo Research, Irvine, CA, USA), (2) EZ = E.Z.N.A Universal Pathogen Kit (Omega Bio-tek, Norcross, GA, USA) and (3) MP = MagPure Stool DNA LQ Kit (Magen Biotechnology, Guangzhou, Guangdong, China). All three kits employ a bead beating step in order to achieve mechanical lysis which was performed according to the protocol settings either using Precellys Evolution homogenizer (10,000 RPM, 4 cycles × 60 s with a 60 s pause in between each cycle for ZM and MP) or using a Vortex-Genie 2 (Scientific Industries, Inc., Bohemia, NY, USA) at the highest speed for 5 min (EZ). Two out of three kits are based on magnetic bead purification (ZM and MP), while EZ includes a step of DNA purification with a MicroElute Column. The extraction protocols were carried out following the manufacturers' guidelines.

For further comparison of effects of storage temperature and sequencing technique, the EZ kit was used for DNA isolation but with minor adjustments compared with the manufacturers protocol: 1) After bead beating disruption, the entire volume of supernatant in the disruptor tube was transferred to a centrifuge tube and diluted in RBB buffer (in volume equal to double volume of supernatant); 2) the whole volume of the obtained sample was transferred and centrifuged in a MicroElute column. These protocol modifications markedly increased the yield of extracted DNA (2–3 fold) without compromising its quality. This is particularly advantageous for nanopore sequencing, which often requires a DNA cleanup step that can result in the loss of at least half of the initial DNA quantity. We also compared two bead-beating options: via vortexing at the maximum speed for 5 min and using Precellys Evolution homogenizer (10,000 RPM, 4 cycles × 60 s with a 60 s pause) and its compatibility for short and long reads sequencing. A 1 % agarose gel electrophoresis was used for estimation of the length of DNA fragments obtained by using these two bead-beating schemes. Based on the results, aliquots of samples obtained via Precellys Evolution homogenizer were used for further Illumina sequencing, while vortexing was used for DNA extraction for nanopore sequencing. The DNA quality was assessed using a NanoDrop spectrophotometer, and Qubit dsDNA HS and BR Assay Kits were used for quantification of the concentration of extracted DNA.

Table 1
Overview of the procedure comparisons done in the study.

Study part	Sample	Pig pen	Storage prior to handling	DNA extraction kit	Sequencing technique
1	Sample 0	0	0d	EZ	ONT
1	Sample 0	0	0d	ZM	ONT
1	Sample 0	0	0d	MP	ONT
2	Sample 1	1	0d	EZ	ONT, Illumina
2	Sample 2	1	2d at +4 °C	EZ	ONT, Illumina
2	Sample 3	1	2d at RT	EZ	ONT, Illumina
2	Sample 4	2	0d	EZ	ONT, Illumina
2	Sample 5	2	2d at RT	EZ	ONT, Illumina
2	Sample 6	3	2d at +4 °C	EZ	ONT, Illumina
2	Sample 7	3	2d at RT	EZ	ONT, Illumina

RT – room temperature, EZ - E.Z.N.A Universal Pathogen Kit, ZM - ZYMOBiomics MagBead DNA/RNA, MP - MagPure Stool DNA LQ Kit, ONT – Oxford Nanopore Technology.

2.3. DNA sequencing

2.3.1. Nanopore sequencing

Nanopore sequencing was performed on all samples of the study. One nanopore library preparation per sample was constructed using the ligation sequencing kit SQK-LSK109 (Oxford Nanopore Technologies (ONT), Oxford, UK). Sequencing of samples for kit comparison (Part 1) was performed on MinION Flow Cells R9.4.1, while Flow Cells with chemistry type R10.4.1 were used for experiments on sequencing technique comparison (Part 2). Sequencing proceeded for 72 h using MinKNOW software to collect raw sequencing data. Fast5 files were basecalled using Guppy (v.6.1.7) and output DNA sequence reads were saved to fastq files.

2.3.2. Illumina sequencing

Illumina sequencing was used on samples from Part 2 of the study. Following the manufacturer's instruction, we constructed one DNA paired-end (PE) library with an insert size of 550 base pairs per sample with TruSeq DNA PCR-free Library Prep (Illumina Inc., San Diego, CA, USA). Shotgun metagenomic sequencing was carried out on an Illumina NovaSeq 6000 instrument (Illumina Inc.) using an SP flowcell (250 bp PE) with v1.5 sequencing chemistry (Illumina Inc.) at the National Genomic Infrastructure (NGI)/ the SNP&SEQ Technology platform, Uppsala, Sweden.

2.4. Data analysis

2.4.1. Illumina reads quality filtering

Raw reads were filtered to remove adaptor contamination, low-quality reads (<20, base call accuracy ~99 %), reads shorter than 50 bp and host genomic DNA. Pre-processing of raw reads by sequence quality was performed with fastp v0.19.5 (Chen et al., 2018; Danecek et al., 2021; Chen, 2023). The clean Illumina sequences were screened with bowtie2 v.2.5.2 (Langmead and Salzberg, 2012) and Samtools v.1.3.1 (Danecek et al., 2021) against the pig reference genome (*Sus scrofa* 11.1, NCBI) downloaded with pre-built indexes for bowtie2 alignment (https://genome-idx.s3.amazonaws.com/bt/Sscrofa11.1.zip) to remove contamination with host genome sequences.

2.4.2. ONT reads' quality filtering

ONT raw reads were filtered from low quality reads (<15, base call accuracy ~96.8 %). For cleaning of host contamination from long reads we performed alignment against the pig reference genome with minimap2 v2.1-r311 (Li, 2018; Li, 2021) followed by Samtools v.1.3.1.

2.4.3. Taxonomic annotation

Taxonomic annotation was performed on trimmed reads obtained

with Illumina and ONT in Kaiju v1.10.1 (Menzel et al., 2016) using Kaiju databases Refseq_nr (2023_06_17).

2.4.4. Assembly of bacterial genomes

Three types of assemblies were performed: assembly of short reads (Illumina), assembly of long reads (ONT) and hybrid assembly (combination of short and long reads). For each type of assembly, two tools were compared. The names of tools and assembly algorithms can be found in Table 2. We also used metaplasmidSPAdes as a tool for plasmid search.

2.4.5. Quality assessment of obtained assemblies

To compare the performance of different assemblers, QUAST v5.2.0. was used to assess contig statistics (Mikheenko et al., 2018). We enabled the MetaQUAST mode to obtain Genome fraction statistics for the 20 most abundant bacterial species presented in pig farm environment (Mikheenko et al., 2016; Chen et al., 2021).

2.4.6. Assembly of metagenome-assembled genomes (MAG)s and quality assessment

MAGs assembly was performed on contigs obtained from MEGAHIT. MAGs were assembled by back-mapping of trimmed reads to obtained contigs using Bowtie2. Binning of contigs was performed using MetaBat v2.12.1 on mapped reads with a minimal length of 1500 bp (Kang et al., 2015). Quality assessment of obtained bins was performed in CheckM v1.0.7 and bins with completeness <50 % and contamination level > 10 % were classified as bad quality bins and excluded from further analysis.

2.4.7. ARG search

The search for ARGs was performed both on contigs obtained with five assembly tools (hybridSPAdes, metaSPAdes, MEGAHIT, metaFlye, metaplasmidSPAdes) and bins obtained from MEGAHIT contigs using ABRicate v1.0.1 (Seemann). The analysis included a search against five ARG databases: ARG-ANNOT (Gupta et al., 2014), CARD (Jia et al., 2017), MEGARes (Doster et al., 2020), NCBI AMRFinderPlus (Feldgarden et al., 2019) and ResFinder (Zankari et al., 2012). Database versions, number of ARGs and date of data retrieval are presented in Table 3. Minimum DNA identity and minimum DNA coverage were set at 80 %. A dictionary of ARGs that includes standard gene name, description of mechanism of work, the resistance pattern (different antibiotics classes, metals and biocides and antibiotic gene family) was manually created to unify the output from 5 different databases as well as to avoid the appearance of duplicates during ARG search.

2.4.8. Sequencing depth subsampling and rarefaction curves

In order to simulate the effect of sequencing depth, subsampling was performed on three samples. For this purpose, the sample function of the seqtk (v. 1.3-r106) package was used with random seed equal to 100. Subsampling was performed on reads that had already passed quality control into the following set of depth intervals: 1 M, 2 M, 4 M, 6 M, 8 M, 10 M, 20 M, 30 M, 40 M, 50 M, 60 M, 70 M, 80 M and 90 M.

Further analyses were performed in R software v4.3.1 (R Core Team, 2013), using the packages: dplyr, tidyverse and ggplot2.

Table 2
Tools and assembly algorithms assessed in the study.

Tool	Version	Assembly	Algorithm	Reference
Canu	v2.2	Long reads	Adaptive k-mer weighting and repeat separation	(Koren et al., 2017)
metaFlye	v2.9.3	Long reads	Repeat graphs	(Kolmogorov et al., 2020)
MEGAHIT	v1.2.9	Short reads	de Bruijn graph	(Li et al., 2015)
metaSPAdes	v3.15.5	Short reads	de Bruijn graph	(Nurk et al., 2017)
metaplasmidSPAdes	v3.15.5	Short reads	de Bruijn graph	(Antipov et al., 2019)
hybridSPAdes	v3.15.5	Hybrid	de Bruijn graph	(Antipov et al., 2016)
OPERA-MS	v0.8.3	Hybrid	Scaffold graph	(Bertrand et al., 2019)

Table 3
Description ARG databases used in the study.

Database	Version	Date of download	Number of ARGs
ARG-ANNOT	V5 (2019 June)	2024-Jun-18	2223
CARD	V3.2.9	2023-Jun-18	4805
MEGARes	V2.0	2023-Nov-4	6635
NCBI AMRFinderPlus	2024-05-02.2	2024-Jun-18	6863
ResFinder	V 2.3.2	2024-Jun-18	3194

3. Results

3.1. Comparison of quality and quantity using different DNA extraction kits

The Nanodrop analysis showed that the best performance was by the EZ kit, with both higher average DNA concentrations and a better A280/A260 ratio compared with the MP and ZM kits. For the other two kits, the absorbance ratio was below 1.7, which is considered as a lower DNA quality that may impact further analysis. Therefore, an additional precleaning step was performed, which reduced the DNA quantity over twofold. Agarose gel electrophoresis didn't show any visible differences among the three kits in length of obtained DNA fragments, which was approximately 20kbp for each kit. Using Precellys Evolution homogenizer for the bead-beating step of the EZ kit yielded twice as much DNA amount compared to vortexing, however, the obtained DNA fragments were smaller in size (1500 to 5000 bp) and therefore not suitable for sequencing using nanopore technology (ONT).

3.2. Effect of DNA extraction kit on detected diversity and abundance of bacteria

The performance of the three kits was further compared by ONT sequencing. Using this technique, we obtained 600 k (ZM), 3 M (MP) and 5 M (EZ) of reads per sample with similar read length distribution between the three different samples, ranging from 50 to 90,000 bases after quality trimming. The mean sequence quality of the reads after trimming remained at 20 (Phred Score/probability of incorrect call 1 in 100 bases). Taxonomic classification of the sequenced reads showed that the ZM kit extraction missed ~37 % of the total diversity on genus level, while the EZ kit resulted in the biggest variety of genera taxonomy covering ~95 % of the total taxonomic diversity (Fig. 2A). There was no clear visible effect of DNA extraction kit on the abundances of the 40 most prevalent genera in the sample (Fig. 2B).

3.3. Effect of sequencing method on bacterial taxonomy

In comparison with the ONT sequencing, Illumina sequencing yielded vastly higher numbers of reads, at least 50 M reads per sample, which were of high quality (>30 Phred score) and had a length of 250 bp. ONT yielded at least 0.8 M of reads with an average read quality (20 Phred score) and a mean length (after quality trimming) of 900-2700 bp. In total, 3813 different bacterial genera were identified across the samples, Illumina sequencing detected approximately 30 % more bacterial genera compared to ONT sequencing. For the majority of the most prevalent bacterial genera there was no difference in relative abundance obtained between the two sequencing techniques, however Illumina sequencing resulted in higher relative abundances of some genera (Fig. 3).

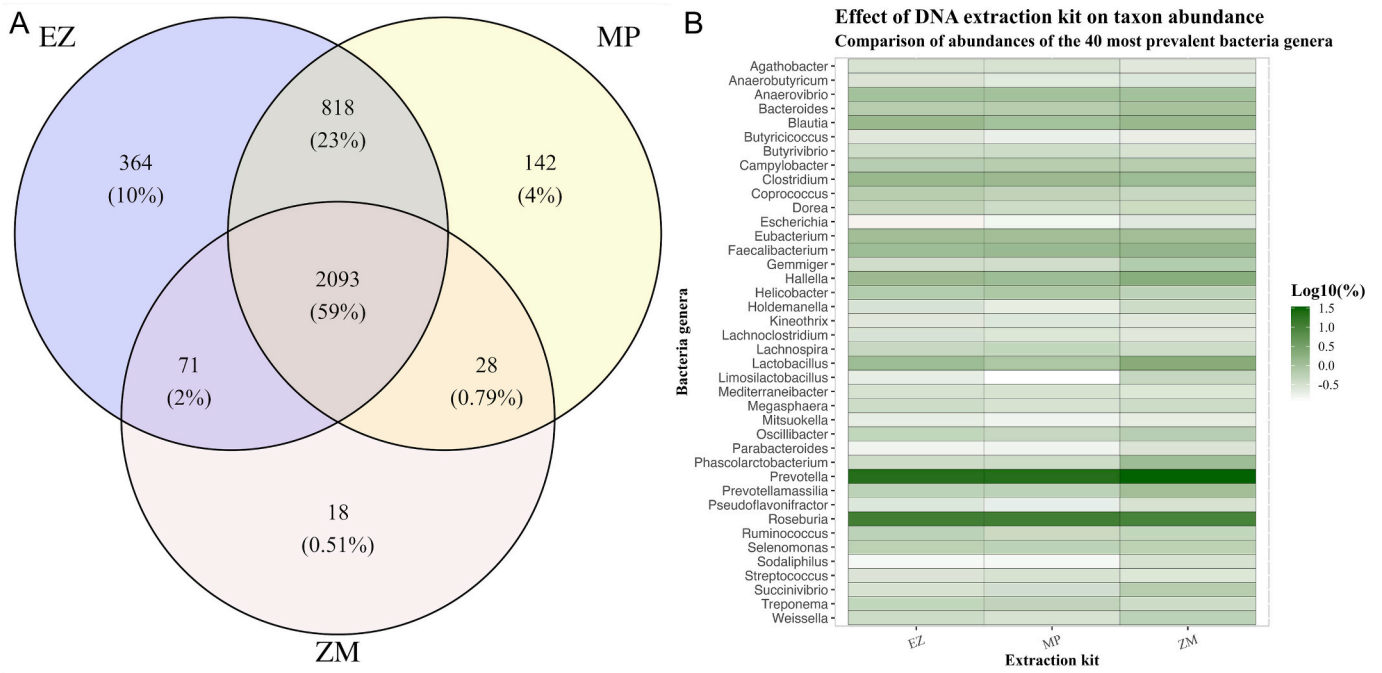

Fig. 2. Comparison of taxonomic profiles from three DNA extraction kits. **A.** Venn diagram showing the number of shared genera, and % of total genera taxonomy covered by sequencing of a sample extracted with each of the kits. EZ - E.Z.N.A Universal Pathogen Kit, ZM - ZYMOBiomics MagBead DNA/RNA, MP - MagPure Stool DNA LQ Kit. **B.** Log10-normalised abundances of the 40 most prevalent bacterial genera: for this analysis the 40 most abundant bacterial genera in the sequencing outcome from three DNA kit extraction were chosen.

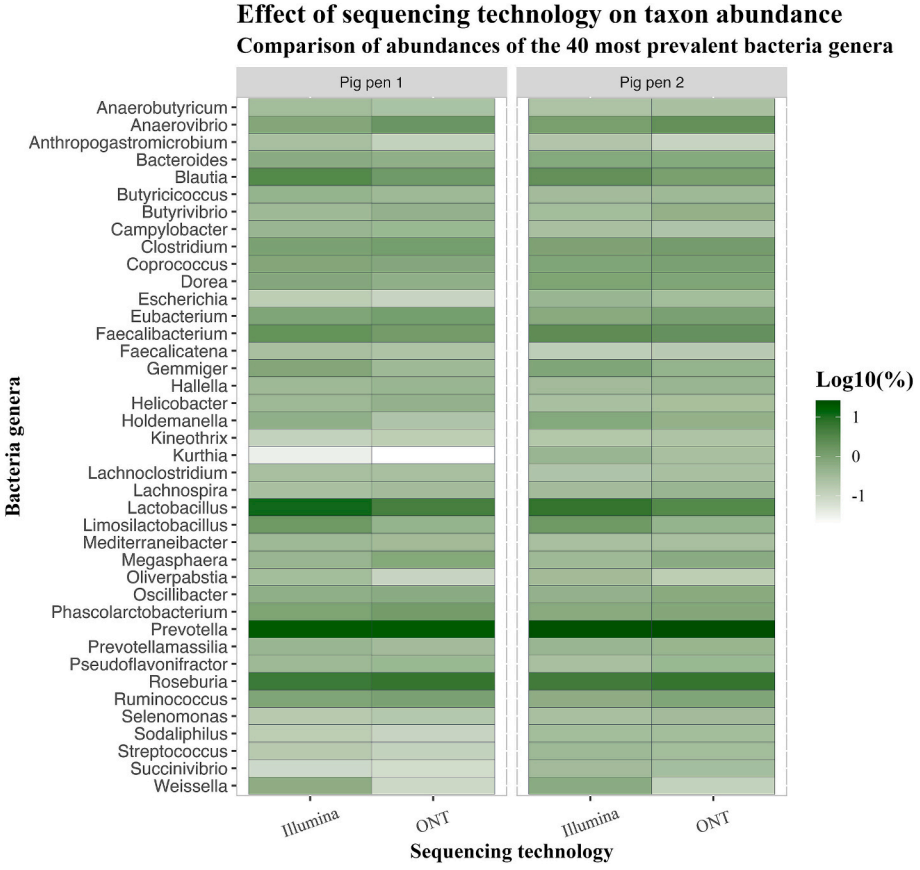

Fig. 3. Effect of sequencing method on bacterial abundance. Log10-normalised relative abundances of the 40 most prevalent bacterial genera in two different samples from pig environment.

3.4. Comparison of different assemblers

The results of assembly quality assessment are shown in Fig. 4. Comparison of assemblies was based on reference-based statistics that included genome fraction, largest alignment and total aligned length, and statistics without reference including largest contig, total length and N50. Overall, assemblers that use Illumina reads performed markedly better than assemblers based only on ONT reads, giving higher values for genome fraction and total aligned length. MEGAHIT showed the best results in genome fraction and total length of assembly.

As the main focus of the method optimisation was analysis of AMR, the assemblers were compared by ARG search. For this analysis, assemblies obtained with hybridSPAdes, metaSPAdes, MEGAHIT, metaFlye and metaplasmidSPAdes were used. There was a very limited number of ARGs (max 6 per sample) found in assemblies with metaplasmidSPAdes. Search of ARGs in metaFlye assemblies resulted in smaller numbers of ARGs within antibiotic classes as well as less varity in antibiotic classes in comparison with metaSPAdes, hybridSPAdes and MEGAHIT (Fig. 5). MetaSPAdes, MEGAHIT and hybridSPAdes resulted in a similar picture of ARGs.

3.5. Comparison of databases used for ARG search

Performing ARG search with ABRicate against 5 databases (Table 3) resulted in a total number of 188 of unique gene names, belonging to 54 families of resistance genes and conveying resistance to 15 antibiotic classes. The largest number of unique ARGs and their sequences were found with MEGARes (Table 4), followed by CARD. Databases such as MEGARes, NCBI AMRFinderPlus and ResFinder contain several sequence variants per ARG. During manual quality control of the ARG-search, it was found that the ResFinder database contains 52 instances where the same accession number is associated with at least two different ARGs, which could potentially lead to errors or misinterpretation of the data. CARD and MEGARes contain ARGs that are involved in resistance against aminocoumarins, fluoroquinolone, mupirocin and peptide antibiotics (Fig. 5). Both these databases also provide information on resistance to biocides, MEGARes additionally gives information on metal resistance genes. The list of these genes can be found in Supplementary Table 1.

3.6. Effect of assembler on ARG identification

The choice of assembler had a visible effect on ARGs detected, which was similar in all tested samples (Fig. 6). MetaplasmidSPAdes gave the worst result with a very small number of ARGs: from 0 to 6 per sample. MetaFlye produced better results than metaplasmidSPAdes (from 5 to 83 ARGs found per sample), however in comparison with search on contigs obtained with hybridSPAdes, metaSPAdes and MEGAHIT, the detected ARGs belonged to a smaller variety of different antibiotic classes as well as a lower number of genes per class. Search of ARGs in hybridSPAdes, MEGAHIT and metaSPAdes assemblies resulted in a similar picture of resistance among the samples: both in the variety of ARG classes and the number of ARGs per antibiotic class. Among the assemblers based on short Illumina reads, ARG searches in metaSPAdes yielded a smaller number of recovered genes (864 in total across 7 samples), whereas hybridSPAdes and MEGAHIT produced nearly identical results, with 883 and 882 recovered genes, respectively.

3.7. Effect of storage temperature on the diversity and abundance of ARGs in the samples

To examine whether different sample storage temperature affects the ARG patterns, we mapped MEGAHIT assemblies against five databases. The analysis did not reveal an observable effect of temperature on the diversity of detected ARGs. Fig. 7 displays the relative abundances (normalised against total length of obtained assembly) of ARGs across seven samples that were collected from three pig pens. In all samples, the largest group of ARGs was associated with the broad-spectrum antibiotic efflux pump class. Other antibiotic classes with a high proportion of associated ARGs included aminoglycosides, beta-lactams, nitroimidazoles, and tetracyclines. Overall, the use of different sample storage temperatures did not have a noticeable effect on the abundance of ARGs within any antibiotic class, with the exception of tetracycline-related genes. For these genes, a slight increase in the number of detected ARGs as the storage temperature increased was observed.

Genome statistics	Canu	metaFlye	metaSPAdes	MEGAHIT	hybridSPAdes	OPERA-MS
Genome fraction (%)	3.4	11.7	28.9	30.1	29.3	24.7
Duplication ratio	1.542	1.181	1.124	1.695	1.274	1.3
Largest alignment (kbp)	83.0	81.6	227.0	147.7	227.1	60.8
Total aligned length (Mbp)	2.4	7.0	18.2	24.4	18.6	17.3
Misassemblies						
# misassemblies	221	437	570	1572	749	692
Misassembled contigs length (Mbp)	2.5	5.3	8.2	8.7	11.3	4.4
Mismatches						
# mismatches per 100 kbp	3281.47	3982.1	2980.89	4216.28	3217.19	3209.27
# indels* per 100 kbp	401.67	694.58	124.51	216.85	139.34	284.09
Statistics without reference						
# contigs	1133	9805	785804	1133613	761156	896710
Largest contig (kbp)	220.6	655.7	377.5	430.7	456.5	146.5
Total length (Mbp)	15.1	105.3	1248.6	1466.6	1256.1	1173.6
N50	18130	18518	2206	1497	2373	1522

Fig. 4. QUAST comparison of assembly quality performed by 6 different tools. Long reads assemblers: Canu, metaFlye. Short read assemblers: MEGAHIT and metaSPAdes. Hybrid-based assemblers hybridSPAdes and OPERA-MS. Colour gradient represents the quality of obtained assemblies, where dark red indicates poorer outcome and dark blue better outcome.*indels = insertions and deletions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

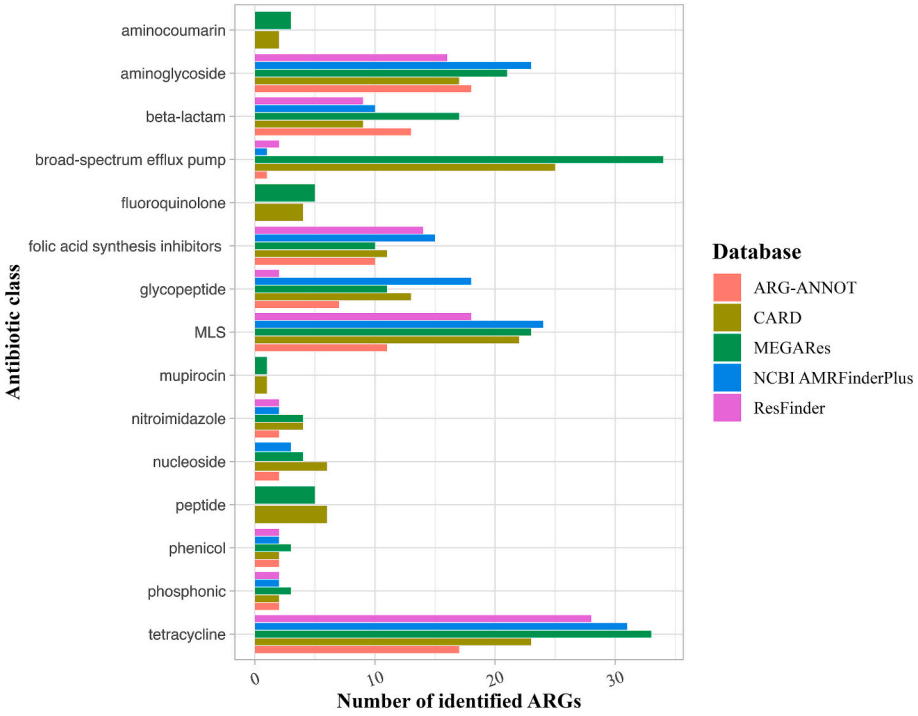

Fig. 5. Comparison of performance of search using five ARG databases. The number of identified ARGs was calculated based on the quantity of unique ARG sequences obtained from in all samples in the study.

Table 4
Identified ARGs in the study samples, based on different databases.

Database	N of unique sequences	N of unique gene names	% of total genes covered
ARG-ANNOT	85	85	45
CARD	147	147	78
MEGARes	177	157	84
NCBI AMRFinderPlus	131	104	55
ResFinder	95	77	41

3.8. Comparison of ARG search results in assemblies and after binning

Fig. 8 shows the results of ARG search after assembly and after binning. It demonstrates that binning prior to ARG search decreases the variety of ARGs in comparison with search without this additional step. On average, binning reduced the total variety of ARGs by 56 %.

3.9. Sequencing depth for ARG search

The results of the simulation of sequencing with different depth followed by ARG search with ABRicate in five ARG databases are presented in Fig. 9. To achieve identification of at least 85 % of all ARG families in the samples, a sequencing depth of 25 M reads was needed (Fig. 9A). To recover 85 % of all gene variants present in the samples, the

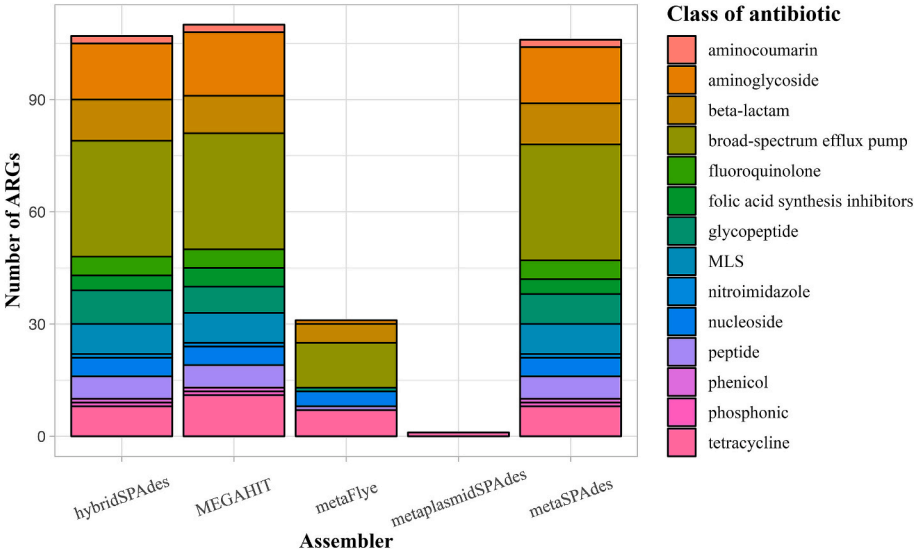

Fig. 6. Effect of assembler on the diversity and abundance of ARGs detected in the sample. The figure shows results from sample 3.

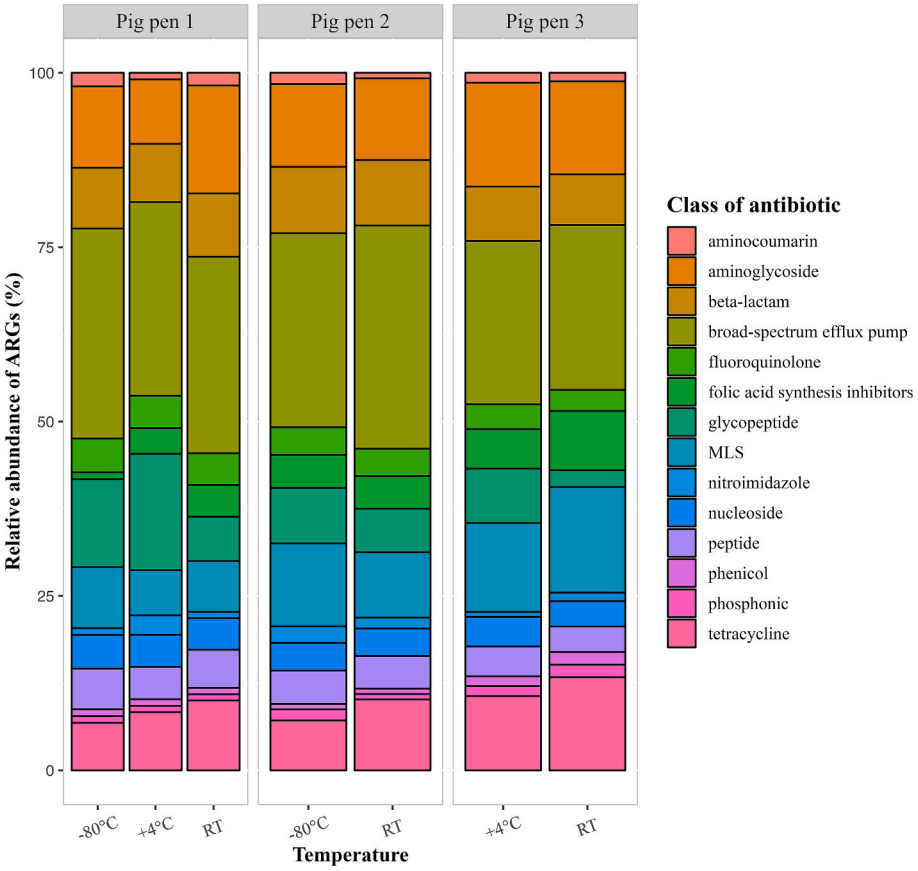

Fig. 7. Effect of sample storage temperature on the diversity of ARGs in different antibiotic classes. Samples were collected from three pig pens and stored at three different temperatures. RT -room temperature.

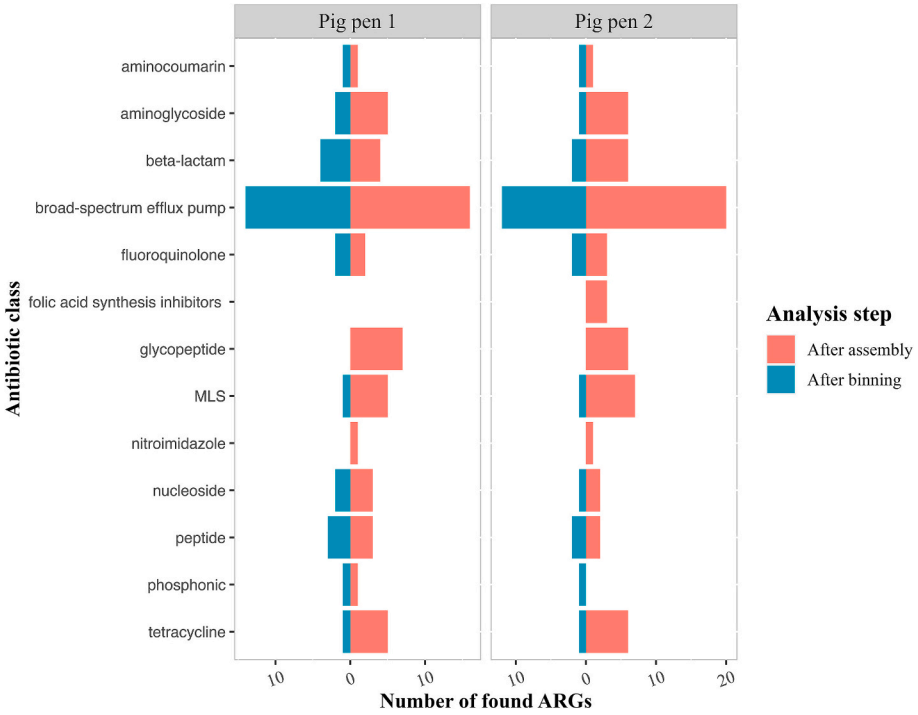

Fig. 8. Detection of ARGs after MEGAHIT based assembly and after MetaBat binning of obtained assemblies. Results presented are based on samples 1 and 4 collected from pen 1 and 2 respectively and stored at −80 °C. The number of genes identified after the assembly step and the binning step is shown on the x-axis.

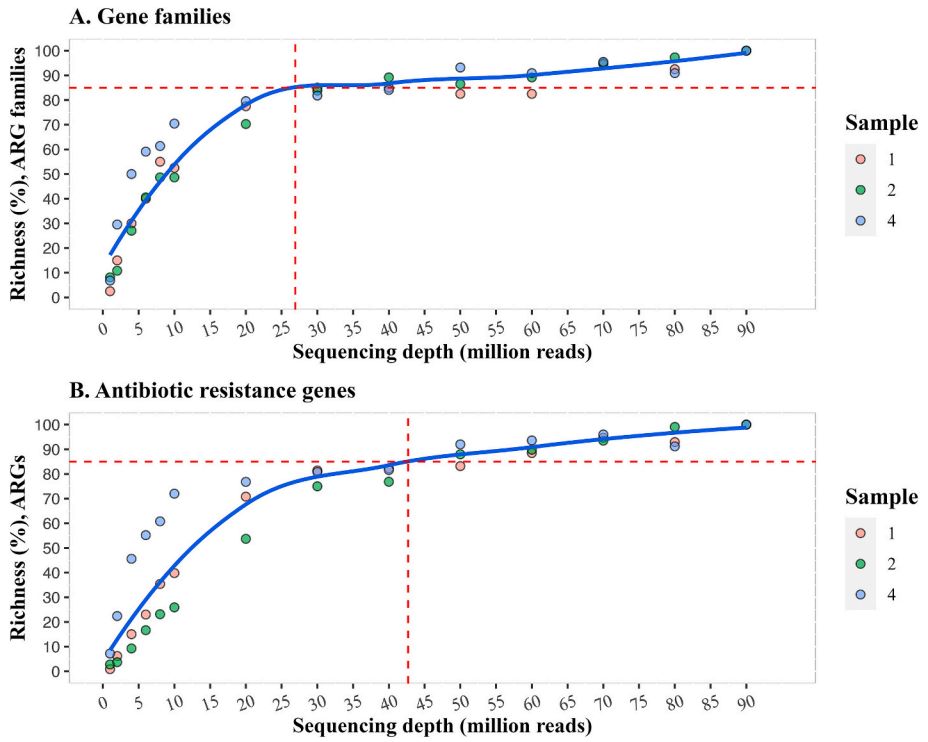

Fig. 9. Rarefaction curves for **A.** ARG gene families, and **B.** unique ARGs. The three samples with lowest concentration of ARGs were chosen for this analysis.

sequencing depth had to be increased to 43 M reads (Fig. 9B).

4. Discussion

There are several critical steps that must be optimised when investigating the farm resistome: sample collection, transportation and storage, DNA extraction and sequencing as well as bioinformatic analysis including metagenome assembly and search for ARGs. Previous studies that evaluated methodology have mainly focused on the comparison of different aspects of one step at a time from the overall pipeline for environmental resistome search. This study aimed to comprehensively evaluate methodologies across several critical steps in the environmental metagenomics research pipeline to advance farm resistome analysis. Specifically, it addressed sample transportation conditions (temperature and duration), comparing short- and long-read sequencing techniques for microbiome analysis, and assessing tools for metagenome assembly and ARG detection within the bioinformatics pipeline to ensure reproducible ARG detection from farm environments.

4.1. Sample collection and storage

We employed and evaluated the performance of sock sampling as a novel method for detecting AMR in the pig farming environment for the first time. Previously this method was concluded to be sufficient to determine the herd status regarding prevalence of bacterial pathogens (Buhr et al., 2007; Pedersen et al., 2015; Lillie-Jaschniski et al., 2023), however it has never been evaluated for suitability for ARG search. Sock sampling is an easy non-invasive method and, based on our results, it yields the needed amount of good quality bacterial material for whole-metagenome sequencing. In comparison with other sampling methods, sock sampling is a cost-effective and not very time-consuming method that helps to avoid the problem of selecting individuals to sample and provides a picture of the entire group of pigs in the pens included in the sample. There is a risk that the sock sampling method might miss low prevalent resistance genes or bacterial genera, however previous studies that compared different sampling strategies showed that a similar strategy based on scraping manure from the floor resulted in only minor differences in resistance in comparison with sampling fresh droppings or individual pig sampling (Munk et al., 2017).

The effect of sample storage temperature on the abundance and diversity of ARGs is a critical area of research, particularly for the purposes of longitudinal studies, where it is important to minimize the role of any factor that might interfere with the stability of the resistome. Previous studies reported that storage of fecal samples at ambient temperatures may exhibit notable changes in their microbiota and resistome profiles over time. For instance, (Martin de Bustamante et al., 2021) reported that fecal samples from equines stored at room temperature for extended periods showed alterations in microbial diversity, contrasting with findings in human and feline samples where no significant changes were observed within 24 h at room temperature. This suggests that the stability of the resistome may vary across species and storage conditions. However, there is a limited number of studies that focus directly on the effect of temperature on the change in resistome. In our study we focused on the effect of short-term storage (48 h) of sock samples at +4 °C and at room temperature (~ + 25 °C) on the resistome. An increase of temperature did not result in major changes in diversity and abundance of the resistome. Only genes exhibiting resistance against tetracycline showed some slight increase with longer storage at higher temperatures. A similar effect of sample storage temperature increase on tetracycline resistance was shown in the work of Poulsen et al. (2021) in pig fecal samples. However, in his work the total observed abundance of ARGs appeared to be dependent on the storage conditions: cold temperatures (−80 °C and − 20 °C) resulted in a decreased abundance compared to immediate sample processing, while longer storage at RT increased the overall AMR abundance. The difference between our results, and the changes in patterns of resistance in Poulsen's work (Poulsen et al., 2021) could be explained by two factors. Firstly, there is the difference in sample material used: in our study we applied sock sampling, whereas Poulsen's study used pig feces. Secondly, in our study the samples were stored at −80 °C for long-term preservation before DNA extraction and sequencing. This extended storage at low temperatures could reduce any impact that short-term storage at varying

temperatures might have on the results. Another possible explanation is that our samples contained lower overall numbers of resistant bacteria with only tetracycline-resistant fast-growing bacteria in sufficient numbers to affect the results. Tetracycline resistance is commonly observed in indicator as well as clinical isolates of *E. coli* from pigs, although the levels in Sweden are much lower than in most countries (SVA, 2024). Regardless, our results support the strategy to transport samples at low temperature.

4.2. Extraction of DNA from complex environmental samples

In our work we didn't reveal any visible difference in bacterial taxonomy of the most abundant bacteria genera in samples processed with different extraction kits. The EZ kit resulted in a larger number of detected bacterial genera, however it still missed ~5 % of the genera that were recovered by ZM and/or MP. Among the tested DNA isolation kits, the EZ kit was the optimal choice because of its easy usage. In contrast with the two other kits, the quality and quantity of DNA obtained with the EZ kit was good enough to avoid the need for further precleaning of samples, which often leads to loss of at least half of total DNA and is especially needed to avoid pore clogging during ONT sequencing. For additional increase in the quantity of obtained DNA some changes can be made in the EZ protocol: 1) usage of the whole available volume of sample after the bead-beating step (as described in materials and methods); 2) exchange the homogenisation step using vortex to Precellys Evolution homogenizer with bead-beating, however this type of lysis leads to a shorter lengths of the obtained DNA fragments and is therefore not recommended for ONT sequencing.

4.3. Illumina vs ONT sequencing

The comparison between Illumina sequencing and ONT sequencing for taxonomy studies and investigations of the resistome revealed distinct advantages and limitations inherent to each platform. Illumina sequencing, known for its high throughput and accuracy, is particularly effective in generating large volumes of data with low error rates, making it suitable for detailed taxonomic studies and microbial diversity, while long reads produced by the ONT platform allows to resolve repetitive regions and complex genomic structures (Boolchandani et al., 2019; Pillay et al., 2022). In this study, the application of both sequencing methods did not reveal any substantial or consistent effect on bacterial taxonomy. However, Illumina sequencing detected 30 % more bacterial genera compared to ONT sequencing, which might be connected to the significantly larger depth of sequencing obtained with short reads. Therefore, deep sequencing with both the NovaSeq 6000 and MinION sequencing platforms appears adequate for assessment of the major part of the microbial community composition, however deep sequencing with the Illumina approach allows for recovery of more rare bacterial genera. Previous research suggests that Illumina is still the current standard for characterizing complex microbial communities (Stevens et al., 2023), however significant improvement in the quality of long read sequencing with ONT has lead to increase in the accuracy of classification and relative estimates and might be preferred for taxonomy studies (Pearman et al., 2020). However, if the focus is on the resistome, the read length obtained with Illumina sequencing can be regarded as sufficient, as shown in our study.

Moreover, regarding ARG abundance, Illumina sequencing showed better results in comparison with the long reads obtained from ONT. More ARGs and antibiotic classes were found in data from Illumina sequencing, which might be connected to the bigger depth of sequencing. Simulation of sequencing with different depth showed that adequate (>80 %) analysis of the resistome based on ARG families needed at least 25 M of 250 bp PE reads, while to reveal the majority of gene variants the sequencing depth has to be increased to 45 M. This corresponds to previous studies where it was shown that to obtain a stabilised number of ARG families in pig caeca, a sequencing depth of ~60 M 150 bp PE reads per sample is needed (Gweon et al., 2019). Achieving similar sequencing depths with ONT sequencing is not cost-effective. Hence, if the focus is on the resistome and the annotation of the ARGs to their bacterial hosts is less important, Illumina would be the method of choice.

4.4. Bioinformatic analysis

There are several crucial steps in bioinformatic analysis to characterise the resistome: choice of assembler, binning and choice of ARG database. Although some authors have described the pro's and con's of different choices (Boolchandani et al., 2019; Pillay et al., 2022; Lee et al., 2023), these must be optimised for the specific purpose of each study, scientific question and complexity of investigated samples. To rely on previously published studies in the choice of methods may seem a robust strategy but, as our study shows, there are many context-specific aspects that must be understood and addressed. The typical pipeline for ARG identification from metagenome data is based either on read-mapping or assembly, which can be further used for binning or an annotation step (Pillay et al., 2022). Every additional step in the pipeline for ARG search brings the risk of additional loss in diversity and abundance of identified ARGs. Therefore, depending on the research question, the assembly and binning step may or may not be performed. A search of ARGs on the basis of raw reads may be recommended if the focus of the research is quantity and variety of different ARGs, but it increases the risk of false-positive detections (Abramova et al., 2024). However, if the focus of the research is the fluctuation of ARGs in regard to specific bacteria, the step of assembly becomes crucial (Abramova et al., 2024). Recent studies suggest that the use of ONT-based long reads, either alone or in combination with Illumina sequencing, can enhance ARG detection and characterization by overcoming the challenge of plasmid reconstruction and the tendency of assemblies to break around ARGs due to surrounding repetitive regions (Berbers et al., 2020; Zhang et al., 2023; Abramova et al., 2024). Taking into account the factors discussed above, we evaluated the performance of three distinct assembly techniques: hybrid, short-read, and long-read assemblies, focusing on their impact on assembly quality. The implications of the resulting assemblies for biological interpretation were assessed by examining the contextualization of ARGs which was the primary focus of this study. We compared seven different assemblers that were seen as most promising for complex environmental metagenome analysis in each type of assembly technique, including two short reads assemblers (metaSPAdes and MEGAHIT) (Brown et al., 2021; Zhang et al., 2023; Abramova et al., 2024), two long reads assemblers (CN and metaFlye) (Latorre-Perez et al., 2020; Brown et al., 2021; Zhang et al., 2023), two tools based on hybrid assembly (hybridSPAdes and OPERA-MS) (Brown et al., 2021; Zhang et al., 2023) and metaplasmidSPAdes to focus on plasmid-located ARGs. In our work, among long-read assemblers metaFlye performed markedly better, while Canu collapsed our data. Similar results with larger size of both total assembly and largest contig were shown in the work of Brown et al. (2021). Previous research showed contradicting results in comparison of MEGAHIT and metaSPAdes. In the work of Abramova et al. (2024) metaSPAdes was suggested as a better option for complex sample scenario, while other studies showed that MEGAHIT outperforms metaSPAdes in generating assemblies from deeply sequenced datasets, which is probably connected to its optimised algorithms to analyze large datasets (Zhang et al., 2023). Similar to the latter, in our work MEGAHIT performed better and preserved the biggest total aligned length among all assemblers. Addition of long reads to short reads increased the length of obtained contigs in hybridSPAdes, however they were still smaller than the ones obtained with metaFlye assemblies, which is expected as short-reads were mainly used to reduce misassemblies rather than fill gaps between contigs (Zhang et al., 2023). Contradictory to previous research (Brown et al., 2021; Zhang et al., 2023), OPERA-MS assembler did not perform well on our data and was excluded from further analysis together with Canu.

Comparison of assemblies based on ARG identification resulted in higher diversity and abundance of ARGs in short and hybrid read assemblies, which is similar to the results of Brown et al. (2021) Overall, among all tested assemblers MEGAHIT showed the best performance in regard of both assembly quality and effectiveness of ARG search.

As stated above, one of the possible steps in ARG search is binning. On one hand, binning may allow to reconstruct genomes of unknown or uncultured bacteria (Pillay et al., 2022), while on the other hand it results in loss of some information. Binning of MEGAHIT contigs reduced the variety of recovered ARGs by 56 %, which is markedly bigger than in the work of Maguire which was performed on a simulated metagenome from 30 chosen genomes and resulted in up to 15 % loss in ARG recovery (Maguire et al., 2020).

Another crucial step in resistome investigation is the choice of ARG database. There are a range of different databases and tools for ARG search that have been described and compared in multiple reviews (Boolchandani et al., 2019; de Abreu et al., 2020; Papp and Solymosi, 2022; Pillay et al., 2022). The databases can be specialised (focused on species specific information) or general, with focus on ARGs from all bacterial hosts, and the tools can be divided into those that work on reads or contigs. In our study we used ABRicate as it allows to standardize and compare outcome of search with several ARG databases. As the focus of the study was the overall resistome of the farm environment, we compared the performance of five general databases: ARG-ANNOT, CARD, MEGARes, NCBI AMRFinderPlus and ResFinder. The choice of databases was based on availability of them in ABRicate and how actively they are curated. Previous research suggested that CARD should be the number one data resource (Papp and Solymosi, 2022), while a lot of environmental and veterinary-based resistome studies are conducted based on the highly cited database ResFinder (Macedo et al., 2021; Ekhlas et al., 2023), however it contains only acquired ARGs (Zankari et al., 2012). In our study, most of the resistance was identified by MEGARes which is based on data from all other databases and had a big number of gene variants (Macedo et al., 2021; Papp and Solymosi, 2022), but it recovered only 84 % of all ARGs found in the study. The ARG-ANNOT database is based on the Lahey Clinic β-lactamase database (Gupta et al., 2014) and therefore had the second best result in detection of ARGs against beta-lactams. Only CARD and MEGARes revealed ARGs conveying resistance to aminocoumarins, mupirocin, fluoroquinolone and peptide antibiotics. The other plus of these two databases is that it is possible to identify biocide and/or metal resistant genes. NCBI AMRFinder Plus recovered the greatest number of ARGs against aminoglygosides, folic acid synthesis inhibitors, glycopeptides, MLS and tetracyclines. One of the minuses of ResFinder identified in this work was that the database has 52 pairs of different ARGs with the same accession number which might lead to misinterpretation of the outcome. Overall, our findings indicate that achieving higher diversity in antimicrobial resistance gene (ARG) detection requires simultaneous searches across multiple databases.

5. Limitations of the study

In this study, we addressed several key factors that impact the reproducibility of ARG identification in complex farm-derived samples. Nonetheless, some limitations remain that future research should address. Firstly, we did not compare the effectiveness of sock sampling to individual fecal samples, leaving unresolved whether the sock sampling method may miss microbial diversity or low-abundance ARGs. Secondly, sequencing data from fresh samples would be valuable to assess potential effects of long-term storage at −80 °C on bacterial and ARG diversity and abundance. Additionaly, this study did not compare ARG detection between assembled and non-assembled data, as quantifying ARGs in non-assembled Illumina short reads presents challenges. Such analysis could clarify the magnitude of potential losses in ARG abundance or diversity connected to the assembly step. The use of samples from several different pens, including the only one where the pigs had been treated with antibiotics, was to obtain some variation and a sufficient number of samples for our purposes. However, as the pigs in Swedish farms are managed batchwise all pigs in the same batch should be regarded as one epidemiological unit.

6. Conclusion

We have developed an approach for reproducible ARG detection in complex samples from the pig farm environment with low levels of antibiotic usage, based on sock sampling as a novel method for detecting AMR. E.Z.N.A Universal Pathogen (Omega Bio-tek, USA) was the optimal choice among the tested kits for metagenome analysis. Choice of sequencing technique didn't drastically affect taxonomy identification; however deep Illumina sequencing allows to recover approximately 30 % more of low-abundance bacterial genera. The optimal sequence depth for ARG detection from farm environment with low levels of antibiotic usage is within 25 to 45 M of 250 bp reads. Storage of sock samples at room temperature up to two days did not affect diversity or abundance of ARGs, however choices made during bioinformatical analysis can markedly change the outcome. In this work, we propose computational analysis based on MEGAHIT assembly as a compromise between need to identify bacteria which carry ARGs and possible loss in diversity and/or abundance of resistance genes. Our research suggested that search against multiple ARG databases is needed for the detection of the highest diversity of ARGs.

Declaration of generative AI and AI-assisted technologies in the writing process

Generative Artificial Intelligence (AI) or AI-assisted technologies were not used in writing or generation of any of the sections of this manuscript.

Funding information

Internal funding from the Swedish University of Agricultural Sciences.

Ethics statement

This work did not involve the use of animals or human participants.

CRediT authorship contribution statement

Valeriia Ladyhina: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Elisabeth Rajala:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition, Conceptualization. **Susanna Sternberg-Lewerin:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Leila Nasirzadeh:** Writing – review & editing, Methodology, Formal analysis, Data curation, Conceptualization. **Erik Bongcam-Rudloff:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition, Conceptualization. **Johan Dicksved:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

None.

Acknowledgements

Sequencing was performed by the SNP&SEQ Technology Platform in Uppsala. The facility is part of the National Genomics Infrastructure (NGI) Sweden and Science for Life Laboratory.

The SNP&SEQ Platform is also supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation.

Support by SLUBI (SLUs bioinformatics infrastructure) is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mimet.2025.107103.

Data availability

The data supporting this study are available from the corresponding author upon reasonable request. The data will not be published as open access to protect the privacy and confidentiality of the farm.

References

Abramova, A., Karkman, A., Bengtsson-Palme, J., 2024. Metagenomic assemblies tend to break around antibiotic resistance genes. BMC Genomics 25, 959. https://doi.org/10.1186/s12864-024-10876-0.

Andersen, V.D., Lv, D.E.K., Munk, P., Jensen, M.S., Agerso, Y., Aarestrup, F.M., Vigre, H., 2017. The association between measurements of antimicrobial use and resistance in the faeces microbiota of finisher batches. Epidemiol. Infect. 145, 2827–2837. https://doi.org/10.1017/S0950268817001285.

Andersen, V.D., Jensen, M.S., Munk, P., Vigre, H., 2021. Robustness in quantifying the abundance of antimicrobial resistance genes in pooled faeces samples from batches of slaughter pigs using metagenomics analysis. J. Glob. Antimicrob. Resist. 24, 398–402. https://doi.org/10.1016/j.jgar.2021.02.005.

Antipov, D., Korobeynikov, A., McLean, J.S., Pevzner, P.A., 2016. hybridSPAdes: an algorithm for hybrid assembly of short and long reads. Bioinformatics 32, 1009–1015. https://doi.org/10.1093/bioinformatics/btv688.

Antipov, D., Raiko, M., Lapidus, A., Pevzner, P.A., 2019. Plasmid detection and assembly in genomic and metagenomic data sets. Genome Res. 29, 961–968. https://doi.org/10.1101/gr.241299.118.

Berbers, B., Saltykova, A., Garcia-Graells, C., Philipp, P., Arella, F., Marchal, K., Winand, R., Vanneste, K., Roosens, N.H.C., De Keersmaecker, S.C.J., 2020. Combining short and long read sequencing to characterize antimicrobial resistance genes on plasmids applied to an unauthorized genetically modified Bacillus. Sci. Rep. 10, 4310. https://doi.org/10.1038/s41598-020-61158-0.

Bertrand, D., Shaw, J., Kalathiyappan, M., Ng, A.H.Q., Kumar, M.S., Li, C., Dvornicic, M., Soldo, J.P., Koh, J.Y., Tong, C., Ng, O.T., Barkham, T., Young, B., Marimuthu, K., Chng, K.R., Sikic, M., Nagarajan, N., 2019. Hybrid metagenomic assembly enables high-resolution analysis of resistance determinants and mobile elements in human microbiomes. Nat. Biotechnol. 37, 937–944. https://doi.org/10.1038/s41587-019-0191-2.

Boolchandani, M., D'Souza, A.W., Dantas, G., 2019. Sequencing-based methods and resources to study antimicrobial resistance. Nat. Rev. Genet. 20, 356–370. https://doi.org/10.1038/s41576-019-0108-4.

Brown, C.L., Keenum, I.M., Dai, D., Zhang, L., Vikesland, P.J., Pruden, A., 2021. Critical evaluation of short, long, and hybrid assembly for contextual analysis of antibiotic resistance genes in complex environmental metagenomes. Sci. Rep. 11, 3753. https://doi.org/10.1038/s41598-021-83081-8.

Buhr, R.J., Richardson, L.J., Cason, J.A., Cox, N.A., Fairchild, B.D., 2007. Comparison of four sampling methods for the detection of Salmonella in broiler litter. Poult. Sci. 86, 21–25. https://doi.org/10.1093/ps/86.1.21.

Chen, S., 2023. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. iMeta 2 (2). https://doi.org/10.1002/imt2.107.

Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

Chen, C., Zhou, Y., Fu, H., Xiong, X., Fang, S., Jiang, H., Wu, J., Yang, H., Gao, J., Huang, L., 2021. Expanded catalog of microbial genes and metagenome-assembled genomes from the pig gut microbiome. Nat. Commun. 12, 1106. https://doi.org/10.1038/s41467-021-21295-0.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of SAMtools and BCFtools. Gigascience 10. https://doi.org/10.1093/gigascience/giab008.

de Abreu, V.A.C., Perdigao, J., Almeida, S., 2020. Metagenomic approaches to analyze antimicrobial resistance: an overview. Front. Genet. 11, 575592. https://doi.org/10.3389/fgene.2020.575592.

Doster, E., Lakin, S.M., Dean, C.J., Wolfe, C., Young, J.G., Boucher, C., Belk, K.E., Noyes, N.R., Morley, P.S., 2020. MEGARes 2.0: a database for classification of antimicrobial drug, biocide and metal resistance determinants in metagenomic sequence data. Nucleic Acids Res. 48, D561–D569. https://doi.org/10.1093/nar/gkz1010.

Ekhlas, D., Cobo Diaz, J.F., Cabrera-Rubio, R., Alexa, E., Ortiz Sanjuan, J.M., Garcia Manzanilla, E., Crispie, F., Cotter, P.D., Leonard, F.C., Arguello, H., Burgess, C.M., 2023. Metagenomic comparison of the faecal and environmental resistome on Irish commercial pig farms with and without zinc oxide and antimicrobial usage. Anim. Microbiome 5, 62. https://doi.org/10.1186/s42523-023-00283-3.

Feldgarden, M., Brover, V., Haft, D.H., Prasad, A.B., Slotta, D.J., Tolstoy, I., Tyson, G.H., Zhao, S., Hsu, C.H., McDermott, P.F., Tadesse, D.A., Morales, C., Simmons, M., Tillman, G., Wasilenko, J., Folster, J.P., Klimke, W., 2019. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. Antimicrob. Agents Chemother. 63. https://doi.org/10.1128/AAC.00483-19.

Graesboll, K., Larsen, I., Clasen, J., Birkegard, A.C., Nielsen, J.P., Christiansen, L.E., Olsen, J.E., Angen, O., Folkesson, A., 2019. Effect of tetracycline treatment regimens on antibiotic resistance gene selection over time in nursery pigs. BMC Microbiol. 19, 269. https://doi.org/10.1186/s12866-019-1619-z.

Gupta, S.K., Padmanabhan, B.R., Diene, S.M., Lopez-Rojas, R., Kempf, M., Landraud, L., Rolain, J.M., 2014. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob. Agents Chemother. 58, 212–220. https://doi.org/10.1128/AAC.01310-13.

Gweon, H.S., Shaw, L.P., Swann, J., De Maio, N., AbuOun, M., Niehus, R., Hubbard, A.T. M., Bowes, M.J., Bailey, M.J., Peto, T.E.A., Hoosdally, S.J., Walker, A.S., Sebra, R.P., Crook, D.W., Anjum, M.F., Read, D.S., Stoesser, N., consortium, R., 2019. The impact of sequencing depth on the inferred taxonomic composition and AMR gene content of metagenomic samples. Environ. Microbiome 14, 7. https://doi.org/10.1186/s40793-019-0347-1.

Horie, M., Yang, D., Joosten, P., Munk, P., Wadepohl, K., Chauvin, C., Moyano, G., Skarzynska, M., Dewulf, J., Aarestrup, F.M., Blaha, T., Sanders, P., Gonzalez-Zorn, B., Wasyl, D., Wagenaar, J.A., Heederik, D., Mevius, D., Schmitt, H., Smit, L.A. M., Van Gompel, L., 2021. Risk factors for antimicrobial resistance in Turkey farms: a cross-sectional study in three European countries. Antibiotics 10. https://doi.org/10.3390/antibiotics10070820.

Jia, B., Raphenya, A.R., Alcock, B., Waglechner, N., Guo, P., Tsang, K.K., Lago, B.A., Dave, B.M., Pereira, S., Sharma, A.N., Doshi, S., Courtot, M., Lo, R., Williams, L.E., Frye, J.G., Elsayegh, T., Sardar, D., Westman, E.L., Pawlowski, A.C., Johnson, T.A., Brinkman, F.S., Wright, G.D., McArthur, A.G., 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res. 45, D566–D573. https://doi.org/10.1093/nar/gkw1004.

Kang, D.D., Froula, J., Egan, R., Wang, Z., 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ 3, e1165. https://doi.org/10.7717/peerj.1165.

Kolmogorov, M., Bickhart, D.M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S.B., Kuhn, K., Yuan, J., Polevikov, E., Smith, T.P.L., Pevzner, P.A., 2020. metaFlye: scalable long-read metagenome assembly using repeat graphs. Nat. Methods 17, 1103–1110. https://doi.org/10.1038/s41592-020-00971-x.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., Phillippy, A.M., 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 27, 722–736. https://doi.org/10.1101/gr.215087.116.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923.

Latorre-Perez, A., Villalba-Bermell, P., Pascual, J., Vilanova, C., 2020. Assembly methods for nanopore-based metagenomic sequencing: a comparative study. Sci. Rep. 10, 13588. https://doi.org/10.1038/s41598-020-70491-3.

Lee, C., Polo, R.O., Zaheer, R., Van Domselaar, G., Zovoilis, A., McAllister, T.A., 2023. Evaluation of metagenomic assembly methods for the detection and characterization of antimicrobial resistance determinants and associated mobilizable elements. J. Microbiol. Methods 213, 106815. https://doi.org/10.1016/j.mimet.2023.106815.

Li, H., 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34, 3094–3100. https://doi.org/10.1093/bioinformatics/bty191.

Li, H., 2021. New strategies to improve minimap2 alignment accuracy. Bioinformatics 37, 4572–4574. https://doi.org/10.1093/bioinformatics/btab705.

Li, D., Liu, C.M., Luo, R., Sadakane, K., Lam, T.W., 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31, 1674–1676. https://doi.org/10.1093/bioinformatics/btv033.

Li, X., Rensing, C., Vestergaard, G., Arumugam, M., Nesme, J., Gupta, S., Brejnrod, A.D., Sorensen, S.J., 2022. Metagenomic evidence for co-occurrence of antibiotic, biocide and metal resistance genes in pigs. Environ. Int. 158, 106899. https://doi.org/10.1016/j.envint.2021.106899.

Lillie-Jaschniski, K., Wahner, C., Viehmann, M., Hauf, S., Gale, C., Rohde, J., Hennig-Pauka, I., 2023. Sock and environmental swabs as an efficient, non-invasive tool to assess the Salmonella status of sow farms. Animals 13. https://doi.org/10.3390/ani13061031.

Luiken, R.E.C., Van Gompel, L., Munk, P., Sarrazin, S., Joosten, P., Dorado-Garcia, A., Borup Hansen, R., Knudsen, B.E., Bossers, A., Wagenaar, J.A., Aarestrup, F.M., Dewulf, J., Mevius, D.J., Heederik, D.J.J., Smit, L.A.M., Schmitt, H., Consortium, E., 2019. Associations between antimicrobial use and the faecal resistome on broiler farms from nine European countries. J. Antimicrob. Chemother. 74, 2596–2604. https://doi.org/10.1093/jac/dkz235.

Luiken, R.E.C., Heederik, D.J., Scherpenisse, P., Van Gompel, L., van Heijnsbergen, E., Greve, G.D., Jongerius-Gortemaker, B.G., Tersteeg-Zijderveld, M.H., Fischer, J., Juraschek, K., Skarzynska, M., Zajac, M., Wasyl, D., Group, E, Wagenaar, J.A., Smit, L.A., Wouters, I.M., Mevius, D.J., Schmitt, H., 2022. Determinants for antimicrobial resistance genes in farm dust on 333 poultry and pig farms in nine European countries. Environ. Res. 208, 112715. https://doi.org/10.1016/j.envres.2022.112715.

Macedo, G., van Veelen, H.P.J., Hernandez-Leal, L., van der Maas, P., Heederik, D., Mevius, D., Bossers, A., Schmitt, H., 2021. Targeted metagenomics reveals inferior resilience of farm soil resistome compared to soil microbiome after manure

application. Sci. Total Environ. 770, 145399. https://doi.org/10.1016/j.scitotenv.2021.145399.
Maguire, F., Jia, B., Gray, K.L., Lau, W.Y.V., Beiko, R.G., Brinkman, F.S.L., 2020. Metagenome-assembled genome binning methods with short reads disproportionately fail for plasmids and genomic islands. Microb. Genom. 6. https://doi.org/10.1099/mgen.0.000436.
Martin de Bustamante, M., Plummer, C., MacNicol, J., Gomez, D., 2021. Impact of ambient temperature sample storage on the equine fecal microbiota. Animals 11. https://doi.org/10.3390/ani11030819.
Mencia-Ares, O., Cabrera-Rubio, R., Cobo-Diaz, J.F., Alvarez-Ordonez, A., Gomez-Garcia, M., Puente, H., Cotter, P.D., Crispie, F., Carvajal, A., Rubio, P., Arguello, H., 2020. Antimicrobial use and production system shape the fecal, environmental, and slurry resistomes of pig farms. Microbiome 8, 164. https://doi.org/10.1186/s40168-020-00941-7.
Menzel, P., Ng, K.L., Krogh, A., 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat. Commun. 7, 11257. https://doi.org/10.1038/ncomms11257.
Mikheenko, A., Saveliev, V., Gurevich, A., 2016. MetaQUAST: evaluation of metagenome assemblies. Bioinformatics 32, 1088–1090. https://doi.org/10.1093/bioinformatics/btv697.
Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D., Gurevich, A., 2018. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics 34, i142–i150. https://doi.org/10.1093/bioinformatics/bty266.
Munk, P., Andersen, V.D., de Knegt, L., Jensen, M.S., Knudsen, B.E., Lukjancenko, O., Mordhorst, H., Clasen, J., Agerso, Y., Folkesson, A., Pamp, S.J., Vigre, H., Aarestrup, F.M., 2017. A sampling and metagenomic sequencing-based methodology for monitoring antimicrobial resistance in swine herds. J. Antimicrob. Chemother. 72, 385–392. https://doi.org/10.1093/jac/dkw415.
Murray, Ch.J.L., et al., 2022. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet 399, 629–655. https://doi.org/10.1016/s0140-6736(21)02724-0.
Noyes, N.R., Yang, X., Linke, L.M., Magnuson, R.J., Dettenwanger, A., Cook, S., Georgnaras, I., Woerner, D.E., Gow, S.P., McAllister, T.A., Yang, H., Ruiz, J., Jones, K. L., Boucher, C.A., Morley, P.S., Belk, K.E., 2016. Resistome diversity in cattle and the environment decreases during beef production. Elife 5, e13195. https://doi.org/10.7554/eLife.13195.
Nurk, S., Meleshko, D., Korobeynikov, A., Pevzner, P.A., 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 27, 824–834. https://doi.org/10.1101/gr.213959.116.
Papp, M., Solymosi, N., 2022. Review and comparison of antimicrobial resistance gene databases. Antibiotics 11. https://doi.org/10.3390/antibiotics11030339.
Pearman, W.S., Freed, N.E., Silander, O.K., 2020. Testing the advantages and disadvantages of short- and long- read eukaryotic metagenomics using simulated reads. BMC Bioinf. 21, 220. https://doi.org/10.1186/s12859-020-3528-4.
Pedersen, K.S., Okholm, E., Johansen, M., Angen, O., Jorsal, S.E., Nielsen, J.P., Baekbo, P., 2015. Clinical utility and performance of sock sampling in weaner pig diarrhoea. Prev. Vet. Med. 120, 313–320. https://doi.org/10.1016/j.prevetmed.2015.04.015.
Pillay, S., Calderon-Franco, D., Urhan, A., Abeel, T., 2022. Metagenomic-based surveillance systems for antibiotic resistance in non-clinical settings. Front. Microbiol. 13, 1066995. https://doi.org/10.3389/fmicb.2022.1066995.
Poulsen, C.S., Kaas, R.S., Aarestrup, F.M., Pamp, S.J., 2021. Standard sample storage conditions have an impact on inferred microbiome composition and antimicrobial resistance patterns. Microbiol. Spectr. 9, e0138721. https://doi.org/10.1128/Spectrum.01387-21.
R Core Team, 2013. R: A Language and Environment for Statistical Computing.
Seemann, T.. *Abricate*.(1.0.1). https://github.com/tseemann/abricate.
Sierra, R., Roch, M., Moraz, M., Prados, J., Vuilleumier, N., Emonet, S., Andrey, D.O., 2024. Contributions of long-Read sequencing for the detection of antimicrobial resistance. Pathogens 13. https://doi.org/10.3390/pathogens13090730.
Slizovskiy, I.B., Bonin, N., Bravo, J.E., Ferm, P.M., Singer, J., Boucher, C., Noyes, N.R., 2024. Factors impacting target-enriched long-read sequencing of resistomes and mobilomes. Genome Res. 34, 2048–2060. https://doi.org/10.1101/gr.279226.124.
Stevens, B.M., Creed, T.B., Reardon, C.L., Manter, D.K., 2023. Comparison of Oxford Nanopore technologies and Illumina MiSeq sequencing with mock communities and agricultural soil. Sci. Rep. 13, 9323. https://doi.org/10.1038/s41598-023-36101-8.
SVA SvarmIT, 2024. Interactive Resistance Monitoring Tool. https://www.sva.se/en/what-we-do/antibiotics/svarm-resistance-monitoring/svarmit-interactive-resistance-monitoring-tool/. (Accessed 10 December 2024).
Tang, K.L., Caffrey, N.P., Nobrega, D.B., Cork, S.C., Ronksley, P.E., Barkema, H.W., Polachek, A.J., Ganshorn, H., Sharma, N., Kellner, J.D., Ghali, W.A., 2017. Restricting the use of antibiotics in food-producing animals and its associations with antibiotic resistance in food-producing animals and human beings: a systematic review and meta-analysis. Lancet Planet Health 1, e316–e327. https://doi.org/10.1016/S2542-5196(17)30141-9.
Van Boeckel, T.P., Brower, C., Gilbert, M., Grenfell, B.T., Levin, S.A., Robinson, T.P., Teillant, A., Laxminarayan, R., 2015. Global trends in antimicrobial use in food animals. Proc. Natl. Acad. Sci. USA 112, 5649–5654. https://doi.org/10.1073/pnas.1503141112.
Van Gompel, L., Luiken, R.E.C., Sarrazin, S., Munk, P., Knudsen, B.E., Hansen, R.B., Bossers, A., Aarestrup, F.M., Dewulf, J., Wagenaar, J.A., Mevius, D.J., Schmitt, H., Heederik, D.J.J., Dorado-Garcia, A., Smit, L.A.M., consortium, E., 2019. The antimicrobial resistome in relation to antimicrobial use and biosecurity in pig farming, a metagenome-wide association study in nine European countries. J. Antimicrob. Chemother. 74, 865–876. https://doi.org/10.1093/jac/dky518.
Weinmaier, T., Conzemius, R., Bergman, Y., Lewis, S., Jacobs, E.B., Tamma, P.D., Materna, A., Weinberger, J., Beisken, S., Simner, P.J., 2023. Validation and application of long-Read whole-genome sequencing for antimicrobial resistance gene detection and antimicrobial susceptibility testing. Antimicrob. Agents Chemother. 67, e0107222. https://doi.org/10.1128/aac.01072-22.
Yamamoto, T., Hayama, Y., Hidano, A., Kobayashi, S., Muroga, N., Ishikawa, K., Ogura, A., Tsutsui, T., 2014. Sampling strategies in antimicrobial resistance monitoring: evaluating how precision and sensitivity vary with the number of animals sampled per farm. PLoS ONE 9, e87147. https://doi.org/10.1371/journal.pone.0087147.
Yang, D., Heederik, D.J.J., Mevius, D.J., Scherpenisse, P., Luiken, R.E.C., Van Gompel, L., Skarzynska, M., Wadepohl, K., Chauvin, C., Van Heijnsbergen, E., Wouters, I.M., Greve, G.D., Jongerius-Gortemaker, B.G.M., Tersteeg-Zijderveld, M., Zajac, M., Wasyl, D., Juraschek, K., Fischer, J., Wagenaar, J.A., Smit, L.A.M., Schmitt, H., consortium, E., 2022. Risk factors for the abundance of antimicrobial resistance genes aph(3′)-III, erm(B), sul2 and tet(W) in pig and broiler faeces in nine European countries. J. Antimicrob. Chemother. 77, 969–978. https://doi.org/10.1093/jac/dkac002.
Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M., Larsen, M.V., 2012. Identification of acquired antimicrobial resistance genes. J. Antimicrob. Chemother. 67, 2640–2644. https://doi.org/10.1093/jac/dks261.
Zhang, Z., Yang, C., Veldsman, W.P., Fang, X., Zhang, L., 2023. Benchmarking genome assembly methods on metagenomic sequencing data. Brief. Bioinform. 24. https://doi.org/10.1093/bib/bbad087.
Zhao, W., Zeng, W., Pang, B., Luo, M., Peng, Y., Xu, J., Kan, B., Li, Z., Lu, X., 2023. Oxford nanopore long-read sequencing enables the generation of complete bacterial and plasmid genomes without short-read sequencing. Front. Microbiol. 14, 1179966. https://doi.org/10.3389/fmicb.2023.1179966.