



OPEN ACCESS

EDITED BY

Feng Gao,
Tianjin University, China

REVIEWED BY

John Barlow,
University of Vermont, United States
Natalie Ring,
University of Edinburgh, United Kingdom

*CORRESPONDENCE

Sigrid C. J. De Keersmaecker
✉ sigrid.dekeersmaecker@scienciano.be

RECEIVED 18 April 2025

ACCEPTED 04 July 2025

PUBLISHED 23 July 2025

CORRECTED 29 July 2025

CITATION

Bloemen B, Gand M, Ringenier M, Bogaerts B, Vanneste K, Marchal K, Roosens NHC, Dewulf J, Boyen F and De Keersmaecker SCJ (2025) Overcoming challenges in metagenomic AMR surveillance with nanopore sequencing: a case study on fluoroquinolone resistance. *Front. Microbiol.* 16:1614301. doi: 10.3389/fmicb.2025.1614301

COPYRIGHT

© 2025 Bloemen, Gand, Ringenier, Bogaerts, Vanneste, Marchal, Roosens, Dewulf, Boyen and De Keersmaecker. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Overcoming challenges in metagenomic AMR surveillance with nanopore sequencing: a case study on fluoroquinolone resistance

Bram Bloemen^{1,2}, Mathieu Gand¹, Moniek Ringenier³, Bert Bogaerts¹, Kevin Vanneste¹, Kathleen Marchal^{2,4}, Nancy H. C. Roosens¹, Jeroen Dewulf³, Filip Boyen⁵ and Sigrid C. J. De Keersmaecker^{1*}

¹Transversal Activities in Applied Genomics, Scienciano, Elsene, Belgium, ²Department of Plant Biotechnology and Bioinformatics, Ghent University, Zwijnaarde, Belgium, ³Department of Internal Medicine, Reproduction and population medicine, Faculty of Veterinary Medicine, Ghent University, Merelbeke, Belgium, ⁴Department of Information Technology, Ghent University, Zwijnaarde, Belgium, ⁵Department of Pathobiology, Pharmacology and Zoological Medicine, Faculty of Veterinary Medicine, Ghent University, Merelbeke, Belgium

Introduction: Antimicrobial resistance is an alarming public health problem, and comprehensive surveillance across environments is required to reduce its impact. Phenotypic testing and whole-genome sequencing of isolates are efficient, but culture-free approaches like metagenomic sequencing potentially allow for broader investigation of resistance gene occurrence, evolution and spread. However, technical challenges such as difficulties in associating antimicrobial resistance genes with their bacterial hosts and the collapse of strain-level variation during metagenome assembly, hinder its implementation.

Methods: To illustrate how these challenges can be overcome, we applied Oxford Nanopore Technologies long-read metagenomic sequencing and novel bioinformatic methods to a case study focused on fluoroquinolone resistance in chicken fecal samples.

Results: We demonstrate plasmid-host linking based on detecting common DNA methylation signatures. Additionally, we use new bioinformatic approaches for strain haplotyping, enabling phylogenomic comparison and uncovering fluoroquinolone resistance determining point mutations in metagenomic datasets.

Discussion: We leverage long-read sequencing, including DNA methylation profiling and strain-level haplotyping, to identify antimicrobial resistance gene hosts, link plasmids to their bacterial carriers, and detect resistance-associated point mutations. Although some limitations remain, our work demonstrates how these improvements in metagenomic sequencing can enhance antimicrobial resistance surveillance.

KEYWORDS

metagenomic sequencing, antimicrobial resistance (AMR), nanopore sequencing, DNA methylation, strain-resolved metagenomics, plasmid host prediction

1 Introduction

Antimicrobial resistance (AMR) is an alarming global health challenge and is estimated to lead to around 1.91 million deaths annually by 2050 if no action is taken, according to a recent forecast (GBD 2021 Antimicrobial Resistance Collaborators et al., 2024). Effective responses require both monitoring of antimicrobial usage and surveillance of resistant organisms, and this across human, veterinary, agricultural and natural environments. While conventional AMR surveillance is performed through isolation and phenotypic testing of bacteria to track AMR trends for predefined panels of antibiotics (BELMAP, 2024), genotypic investigation is vital for detailed insights in resistance mechanisms, AMR gene (ARG) transfer and co-occurrence (Anjum et al., 2017; Djordjevic et al., 2024). Currently, short-read whole-genome sequencing (WGS) of isolates is widely used for this purpose, and reliably predicts resistance phenotypes (Zhao et al., 2016; Feldgarden et al., 2019). However, short reads struggle to resolve repetitive regions, a limiting factor in the investigation of plasmids and other mobile genetic elements (MGEs) that can play a role in the transfer of ARGs across bacterial species and their environments (Maguire et al., 2020; Djordjevic et al., 2024). Advances in long-read sequencing, such as Oxford Nanopore Technologies (ONT) R10 flow cells, V14 chemistry, and improvements in basecalling accuracy, now enable complete and high-quality assembly of bacterial genomes and plasmids using long-reads only (Sereika et al., 2022; Agostinho et al., 2024; Sanderson et al., 2024).

While WGS of isolates is efficient for monitoring AMR, there are several limitations: First, WGS is typically only performed on a limited number of species subjected to official control. Second, not all bacteria can be isolated, but even when isolation is successful, culturing and selection further restrict the scope of species and resistances that can be investigated. Finally, only a small fraction of the bacterial population in a sample is isolated and sequenced. This risks missing important isolates, species, and resistance mechanisms, creating large blind spots. Metagenomic sequencing can potentially overcome these issues by sampling all genetic material, revealing a larger diversity of ARGs (i.e., the resistome) and their potential bacterial hosts in a sample. Typically, bioinformatic methods investigate the resistome in a metagenomic dataset either directly from raw sequencing reads (i.e., read-based approaches), or by first assembling the reads into larger contigs (i.e., assembly-based approaches) (Djordjevic et al., 2024). Read-based approaches rapidly detect ARGs by skipping the computationally intensive assembly step, but may suffer from low taxonomic precision (Bloemen et al., 2023; Van Uffelen et al., 2024; Chen et al., 2025) and provide limited or no information on the genomic context of the detected ARGs (Djordjevic et al., 2024). Furthermore, reliably detecting resistance-associated point mutations in single reads is challenging due to sequencing errors, especially with long-read technologies, although their read accuracy is continuously improving (Chen et al., 2025). Assembly-based approaches assemble reads into larger contigs that can be grouped into bins and metagenome-assembled genomes (MAGs) based on features such as coverage, sequence composition or assembly graph information (Agostinho et al.,

2024). Additionally, by aggregating information from multiple reads, more accurate consensus genomes can be generated, while the increased length of contigs offer higher taxonomic resolution and more genetic context on ARGs, enabling linking them to their host replicon. In this approach, long reads generate more contiguous assemblies than short reads, particularly around ARGs and for plasmids (Yorki et al., 2023; Abramova et al., 2024). However, as assembly requires sufficient coverage (typically 3x or more), ARGs at low abundance can be missed (Yorki et al., 2023; Chen et al., 2025). Leveraging the strengths of both read- and assembly-based approaches could potentially maximize ARG detection and allow for more in-depth characterization of their hosts replicons, though frequently, only one method is used.

Metagenomic sequencing allows to probe all ARGs present in a sample, and can reveal their genetic context when assembly is successful. However, linking a plasmid carrying an ARG to its bacterial host based solely on nucleotide sequence remains challenging. Advances in long-read sequencing of native DNA allows for detection of DNA modifications such as N4-methylcytosine (4mC), 5-methylcytosine (5mC), and N6-methyladenine (6mA). In previous studies, this has been used to group plasmids with their hosts based on common methylation patterns (Beaulaurier et al., 2018; Tourancheau et al., 2021). However, with ONT this required establishing a non-methylated baseline signal by sequencing a second methylation-free library, while Pacific Biosciences sequencing has low sensitivity for 5mC motifs (Agostinho et al., 2024; Heidelberg et al., 2024; Soto-Serrano et al., 2024). Recent developments in ONT technology addresses these issues by detecting 4mC, 5mC and 6mA DNA-modifications in native sequencing reads from a single library, and new bioinformatic tools such as MicrobeMod (Crits-Christoph et al., 2023) and NanoMotif (Heidelberg et al., 2024) can use this information for detecting methylation motifs. NanoMotif also applies this information for metagenomic bin improvement and including plasmids in a host bin (Heidelberg et al., 2024). However, this approach has not yet been applied specifically to metagenomic AMR surveillance.

Another remaining issue in metagenomics is the recovery of strain-level genomes and genetic variation, as assemblies typically convolute genetic variation into a single consensus sequence when multiple strains of the same species are present (Nicholls et al., 2021; Agostinho et al., 2024). In the case of AMR, this consensus might mask low-frequency single nucleotide polymorphisms (SNPs) associated with resistance, resulting in undetected AMR genotypes, as some might be caused by point mutations. Furthermore, tracing of outbreak origins often requires genome-wide comparison of bacterial strains. Combining outbreak investigation with AMR surveillance provides a comprehensive picture of pathogen spread, resistance evolution, and public health risks, allowing for faster and more effective interventions. Using metagenomics instead of isolate WGS for this purpose presents challenges as it requires to recover genetic variation that co-occurs within a strain, commonly referred to as *haplotyping* or *phasing*. While short reads accurately detect variants between strains, long-read metagenomics can generate more contiguous haplotypes, and recent tools have been developed to recover strain genomes

(Kazantseva et al., 2024; Shaw et al., 2024). However, these tools have yet to be applied in the context of AMR surveillance, e.g., for strain-level phylogenetic investigation and AMR profiling directly from metagenomic data.

In this study, we applied novel long-read metagenomic methods to detect fluoroquinolone (FQ) resistance in chicken fecal samples as study case. FQs are categorized as highest priority critically important antimicrobials by the WHO, based on evidence of FQ resistant pathogens being transmitted to humans from non-human sources (WHO, 2024). Furthermore, FQ resistance is an interesting proof-of-concept, as both plasmid-mediated ARGs and chromosomal mutations play a role in acquired resistance. For example, FQ resistance genes such as *qnrA*, *qnrB*, *qnrS* and *oqxAB* are often located on plasmids or MGEs that are difficult to assign to a host strain using metagenomics data. Additionally, the predominant FQ resistance mechanism is based on mutations in the coding sequence of the *gyrA* and/or *parC* genes, requiring accurate SNP detection for identification (Bush et al., 2020; Kherroubi et al., 2024). Both plasmid-mediated and SNP-based FQ resistance genes are frequently detected in feces of chickens in Belgian broiler farms, making such samples ideal test cases to detect both resistance determinants in real-life conditions (Ringnier et al., under review)¹. Using this case study, we addressed several current challenges in metagenomic AMR surveillance using ONT sequencing. First, we took advantage of long reads by detecting ARGs and their immediate genetic surroundings in a read-based approach, and used long-read assemblies for a broader view. Next, we applied ONT-based methylation calling and motif detection using Nanomotif (Heidelberg et al., 2024) to link an ARG-carrying plasmid to its host. Finally, we used haplotype phasing to uncover strain-level SNPs involved in FQ resistance, which were not detected at the MAG level. The applied phased reads were then used for a phylogenomic comparison between metagenomics and isolate WGS data. By addressing these challenges, we aimed to demonstrate how long-read metagenomics can be used for improved AMR surveillance.

2 Materials and methods

2.1 Sample collection, processing, and DNA extraction

One chicken (*Gallus gallus*) fecal sample per farm was taken from fresh droppings on each of two Belgian farms, A and B, and are hereby referred to as sample A and B, respectively. The fecal samples were collected using DNA/RNA Shield Fecal Collection Tubes (R1101; Zymo Research, Irvine, United States). Approximately 1 gram of fecal sample was taken and submerged in 9 ml of DNA/RNA shield. Samples were then stored at -80°C until further processing. To extract DNA, 100 mg of fecal material was taken from the collection tube, along with 200 μl of the DNA/RNA Shield supernatant. Next, 7.78 μl of Spike-in Control I (D6320; Zymo Research, Irvine, United States) was added, containing

7.78×10^7 cells of the gram-negative *Imtechella halotolerans* and gram-positive *Allobacillus halotolerans*. DNA extraction was performed according to the three peaks DNA extraction method (Quick, 2019), with the following modifications: Bead-beating was performed on a Bertin Minilys (Bertin Technologies, Montigny-le-Bretonneux, France) at 3,000 rpm. Next, the sample was split in two and divided into two 1.5 ml Eppendorf tubes for the DNA-binding step with Genomic lysis buffer (D3004-1; Zymo Research, Irvine, United States). After the DNA-binding step, the magnetic beads in both tubes were each resuspended in 50 μl DNA elution buffer (D3004-4; Zymo Research, Irvine, United States), and were recombined back into a single 1.5 ml Eppendorf tube. Finally, the elution was done on a thermoshaker at 55°C and 200 rpm for 5 min.

FQ resistant ST10 *E. coli* isolates were obtained from four broiler farms (farms A and B mentioned before, as well as from two additional farms, i.e., C, D) as described in Ringnier et al. (under review)¹, and were grown overnight in Luria-Bertani medium at 37°C and 250 RPM shaking. Isolates were characterized using short-read (Illumina) and long-read (ONT) sequencing. For Illumina sequencing, the GenElute™ Bacterial Genomic DNA Kit from Sigma-Aldrich (Merck, Overijse, Belgium) was used according to the manufacturer's protocol, as described (Ringnier et al., under review)¹. For ONT sequencing, DNA extraction was performed as described for the fecal metagenomic samples, with omission of the bead-beating step.

2.2 Nanopore sequencing and basecalling

Metagenomic nanopore sequencing libraries were generated using the ligation sequencing kit, while isolate libraries were prepared using the native barcoding kit (respectively, SQK-LSK114 and SQK-NBK114.24, Oxford Nanopore Technologies, Oxford, United Kingdom). The manufacturer's instructions were followed for all libraries. The metagenomic samples were sequenced on one MinION R10.4.1 flow cell (FLO-MIN114, Oxford Nanopore Technologies, Oxford, United Kingdom) per sample on a GridION device (Oxford Nanopore Technologies, Oxford, United Kingdom), while the isolates libraries were barcoded and combined on a single R10.4.1 flow cell. For all samples, basecalling was performed using Dorado 0.7.0 and the super high accuracy basecalling model v5 (dna_r10.4.1_e8.2_400bps_sup@v5.0.0), using the option “--modified-bases 4mC_5mC 6mA” to call DNA methylation, with methylation model versions (dna_r10.4.1_e8.2_400bps_sup@v5.0.0_4mC_5mC@v3 and dna_r10.4.1_e8.2_400bps_sup@v5.0.0_6mA@v3). For the barcoded isolate libraries, the option “--barcode-both-ends” was used in dorado for demultiplexing. Reads were filtered for Q-scores ≥ 10 and length ≥ 500 using SeqKit v2.3.0 (Shen et al., 2024).

2.3 Read-mapping based metagenomic analysis

Taxonomic classification of the reads was performed using Kraken2 (Wood et al., 2019) v2.1.1 with default settings to a

¹ Ringnier, M., Boyen, F., De Keersmaecker, S. C. J., Gand, M., Vanneste, K., Bogaerts, B., et al. (under review). Continued presence of fluoroquinolone resistance in *Escherichia coli* in broilers: mapping the spread (under review).

large in-house database containing: complete, chromosome and scaffold NCBI Refseq assembly levels of Archaea, Fungi, Protozoa and Plant; complete Refseq assemblies of Bacteria and Viruses; and complete and chromosome assemblies of Animals (including *Gallus gallus*), supplemented with complete and chromosome assemblies for Plants from GenBank. The databases were accessed on January, 13th 2024 (O'Leary et al., 2016; Bogaerts et al., 2025). Next, reads assigned to Eukaryota were filtered out using the “extract_kraken_reads.py” script from KrakenTools v1.2, with parameter “-t” set to “2759,” and the “-exclude,” “-include-children” and “-fastq-output” parameters (Lu et al., 2022). The remaining reads were then mapped with KMA v1.4.15 (Clausen et al., 2018) to an in-house database consisting of archaea, fungi, plants and protozoa with complete, chromosome and scaffold assembly levels, and bacteria and viruses with complete assembly levels, obtained from the NCBI RefSeq database on February 24th 2023. The following options were used: “-bc” set to 0.7, “-ID” set to 0.0, “-proxi” set to “0.9,” and “-mem_mode,” “-bcNano,” “-ef,” “-l1t1,” “-ca” enabled. KMA v1.4.15 was additionally utilized to map the non-Eukaryotic reads to the ResFinder database version 2.4.0 (Bortolaia et al., 2020), with the same options as for taxonomic mapping, excluding the “-l1t1” setting for reads potentially carrying more than one ARG. ARGs were then filtered for template identity of more than 90%. Reads mapping to both a taxonomic and an ARG template were used to investigate species-ARG links (Bloemen et al., 2023), using a custom python script. Finally, the data were visualized using custom Python (v3.12.3) and R (v4.1.3) scripts. The complete workflow is available as an open-source Snakemake workflow on GitHub: <https://github.com/brambloemen/FLUPOULmeta>.

2.4 Assembly-based metagenomic analysis and detection of methylation motifs

The pipeline for metagenomic assembly, binning, and methylation analysis is also included in the GitHub repository mentioned above (<https://github.com/brambloemen/FLUPOULmeta>), and is further described here. The filtered non-Eukaryotic reads (produced as described in Section 2.3) were assembled using Flye v2.9.4 (Kolmogorov et al., 2020), using the options “-nano-hq -meta -i 2 -read-error 0.03.” Next, a single iteration of polishing was performed with Medaka v2.0.0, using the mini_align, inference and sequence modules with the following basecalling model: dna_r10.4.1_e8.2_400bps_sup@v5.0.0. Contigs were then binned using an ensemble of methods including: MetaBAT2 v2.15 with default options (Kang et al., 2019), graphMB v0.1.5 with the options “-assembly_type flye -vamb -minbin 250000 -mincontig 3000” (Lamurias et al., 2022) and SemiBin v2.1.0 with options “-compression=none -sequencing-type long_read -environment global” (Pan et al., 2023). The outputs from all binning methods were combined using DASTool v1.1.6 with the option “-score_threshold” set to 0 (Sieber et al., 2018). For methylation analysis including methylation-based bin improvement, the raw reads (including methylation tags) were

mapped back to the assembly using Minimap2 v2.28 (Li, 2018), using options “-ax map-ont -y,” and were processed further with SAMtools v1.17 (Danecek et al., 2021). After this, a methylation pileup was made using Modkit v0.3.0. Next, NanoMotif v0.5.6 (Heidelberg et al., 2024) was used for methylation analysis and bin improvement as follows. Motif discovery was carried out using the “motif_discovery” module with the following options: “-min_motif_score” set to 0.2, “-read_level_methylation” enabled, and “-threshold_valid_coverage” set to 1. Bin inclusion was done using the “include” function of NanoMotif with the “-run_detect_contamination” option enabled. Classification of the MAGs corrected with NanoMotif was done using GTDB-Tk v2.3.2 with GTDB database release 214 (Chaumeil et al., 2022), using the command “gtdbtk classify_wf” with option “-skip_ani_screen” enabled. Completeness and contamination of MAGs were checked using the “checkm lineage_wf” command from CheckM v1.2.2 (Parks et al., 2015) with default options. The methylation output reported by NanoMotif in the “motifs-scored.tsv” file were analyzed by visualizing the methylation percentage per motif, using the R packages ComplexHeatmap v2.10.0 (Gu et al., 2016) and UMAP v0.2.10.0 (McInnes et al., 2020). Motifs were only used if they occurred in at least one contig with more than 50% methylation. For each bin, the 10 largest contigs were included, and bins were filtered for >10% completeness. NCBI's AMRFinderPlus v3.12.8 with database version 2024-05-02.2 was used to detect ARGs (Feldgarden et al., 2021), along with KMA and the ResFinder database as described in Section 2.3 for read-based analysis. For *Klebsiella pneumoniae* or *Escherichia coli* MAGs, AMRFinderPlus v3.12.8 was run separately with the “-organism” option set to the corresponding species. AMRFinderPlus and KMA were used in the same way to detect ARGs in the reference genomes of the spike-in species (Refseq accessions NZ_CP117969.2 for *I. halotolerans* and NZ_CP117968.1 for *A. halotolerans*), detecting no ARGs. Plasmids and mobile elements were detected with an ensemble approach: First, genomad v1.8.0 (Camargo et al., 2023) was run to detect plasmid/viral sequences. Next, SCAPP v0.1.4 (Pellow et al., 2021) was used to detect remaining plasmid contigs not reported by genomad (e.g., plasmids consisting of multiple contigs that could only be detected with assembly graph information). MOB-suite v3.1.9 (Robertson et al., 2020) was additionally used to profile plasmids. Finally, mobileOG-db with the mobileOG-pl scripts (<https://github.com/clb21565/mobileOG-db/tree/main> commit 91ee129; Brown et al., 2022) was used to detect MGE-related genes. Proksee was used to illustrate plasmids and mobile elements (Grant et al., 2023).

2.5 Strain phasing and strain-level detection of resistance-associated SNPs

Longshot v1.0.0 (Edge and Bansal, 2019) was utilized to call variants with default settings (as indicated in the Floria documentation), after which Floria (Shaw et al., 2024, git commit 22fb0ab) was used to phase reads into haplotypes. The results were visualized with the script “visualize_vartigs.py” in the Floria github repository. Phased reads were then clustered into strains with

floria-strainer v0.2.1 (<https://github.com/maxibor/floria-strainer>). Each set of clustered reads was then separately assembled using the ONT assembly method for isolates described in Section 2.6. The resulting phased assemblies were checked for quality with CheckM (Parks et al., 2015) as described in Section 2.4, and ARGs and resistance-conferring SNPs were analyzed using AMRFinderPlus as described in Section 2.4, with the additional “-organism” option set to the organism of interest to identify SNPs. Further manual inspection of SNPs was performed with SAMtools v1.17 (Danecek et al., 2021) and IGV (Thorvaldsdóttir et al., 2013).

2.6 Bioinformatic analysis of isolates, phylogenetic comparison and plasmid alignment

Illumina short-read sequencing data for *E. coli* isolates was generated, assembled and sequence typed as described in Ringenier et al. (under review)¹, using a previously published STEC pipeline (v1.0, available at <https://galaxy.sciensano.be>; Bogaerts et al., 2021). ONT assemblies for the isolates were generated with Flye (Kolmogorov et al., 2020) and Medaka as described in Section 2.4, but without the “-meta” option in Flye, and without further binning steps. ARGs were detected as described in Section 2.4. Multi-locus sequence typing (MLST) of the metagenomic strain-phased genome and isolate assemblies was performed on the Sciensano Galaxy instance (Bogaerts et al., 2025), using the seven-gene university of Warwick scheme (Zhou et al., 2020), updated September 1st 2024. For SNP-phylogeny of the metagenomic strain-phased ST10 genome in sample A and ST10 isolate genomes, PACU v0.0.7 (Bogaerts et al., 2024) was applied as follows: First, both long- and short-read data of a single isolate (Farm A Isolate 3) were used with an in-house implemented hybrid assembly pipeline to generate a common reference genome (Bouras et al., 2024). Only the largest contig in this hybrid assembly (4.9 Mb), representing the *E. coli* chromosome, was retained as the reference genome for PACU v0.0.7 (Bogaerts et al., 2025). Next, the web-based PHASTEST (Wishart et al., 2023) was used to identify phage regions in the reference, from which a BED file was created to mask phage regions in the PACU analysis with the “-ref-bed” option. For each isolate, reads (Illumina or ONT) were then mapped to the reference genome with the “PACU_map” script (Bogaerts et al., 2024). Strain-phased reads from the metagenomic data were converted from BAM to FASTQ format with the SAMtools “fastq” command, and were then aligned to the reference genome using the approach described above. Next, the PACU pipeline was used with the reference-mapped read sets to construct a phylogeny, with phage regions masked, and other settings left at default values. To compare the metagenomically detected and linked plasmid, hybrid assemblies for two additional isolates (Farm A Isolate 1 and 2) were generated as described above for the PACU reference genome. Finally, the dnadiff command from Mummer4 v4.0.0rc1 (Marçais et al., 2018) was used to align and compare the metagenomic plasmids and the hybrid assemblies for the isolates (Farm A Isolate 1 and 2) to the reference (Farm A Isolate 3).

3 Results

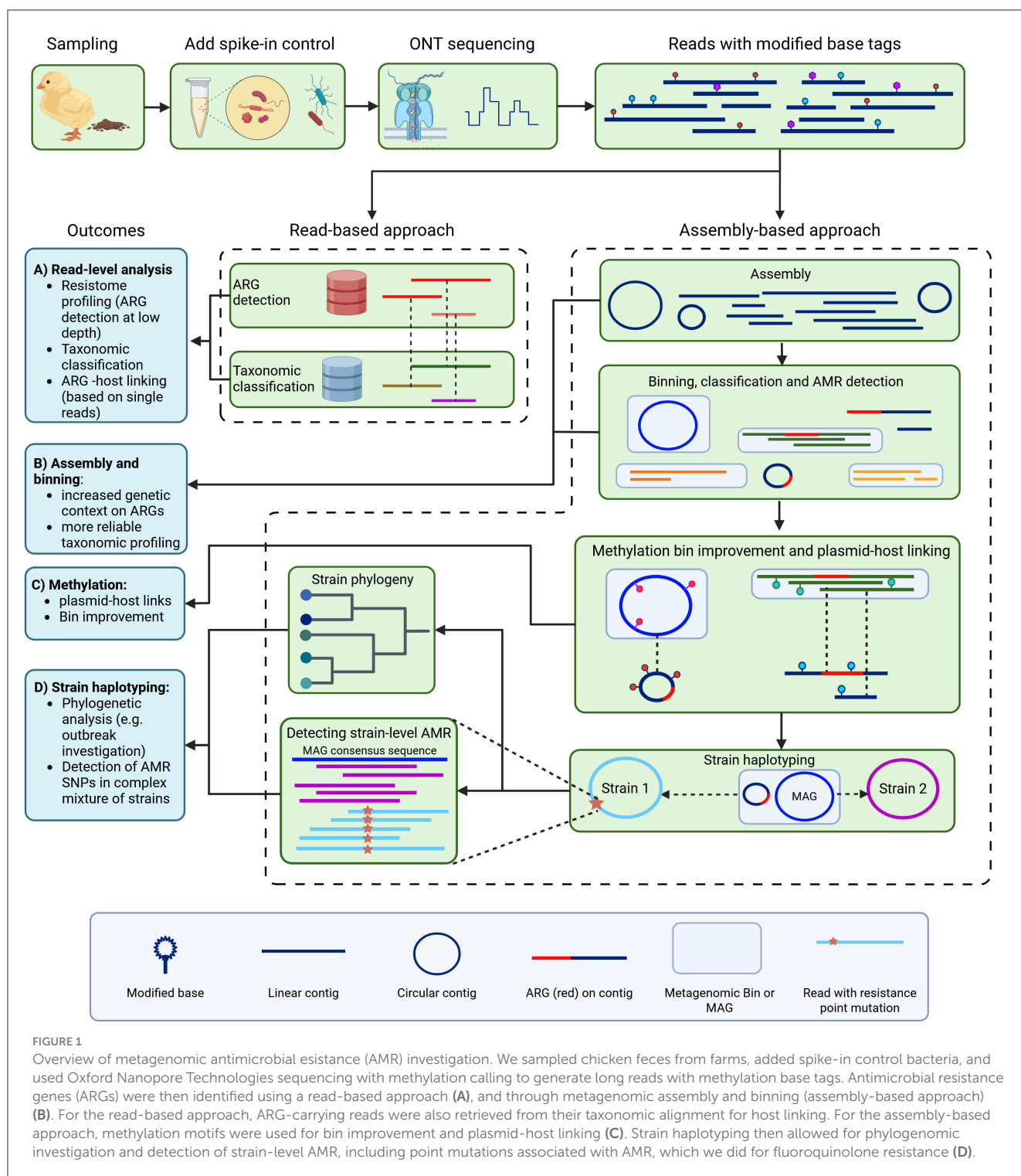
3.1 Read-based resistome analysis of chicken fecal samples

To perform a resistome analysis, we collected chicken fecal samples from two farms (A and B, [Supplementary Table S1](#)), and performed ONT sequencing ([Figure 1](#)). For sample A, this yielded 14.55 Gigabases (Gb) of total DNA on 1.61×10^6 reads, with an N50 of 15.03 kb. For sample B, 15.03 Gb of data was obtained on 3.14×10^6 reads, with an N50 of 9.94 kb. Initially, read-based ARG detection and taxonomic host linking were performed ([Figure 1A](#)). The following FQ resistance genes were detected in sample A: *qnrS1*, *qnrB5*, and *oqxB*, but not the *oqxA* part of the *oqxAB* operon ([Figure 2A](#)). In sample B, the *oqxAB* operon and *qnrS1* gene were found alongside other ARGs ([Figure 2B](#)). To investigate species-ARG associations, we extracted reads that mapped to both ARGs and taxonomic templates. In sample A, the *qnrS1* gene was linked to *Klebsiella pneumoniae*, *Escherichia coli* and a plasmid according to taxonomic mapping ([Figure 2C](#)). The *qnrB5* gene was associated with a plasmid, while no taxonomic match was found for *oqxB*. Surprisingly, the spike-in control strains *I. halotolerans* and *A. halotolerans* were also linked to an ARG, while the reference genomes do not contain any. In sample B, the *oqxAB* operon was linked to *K. pneumoniae*, while *qnrS1* was associated to plasmids ([Figure 2D](#)). Similar to sample A, *A. halotolerans* was also reported as ARG host in sample B. Point mutations were not investigated using the read-based resistome analysis workflow described here.

3.2 Metagenome assembly to investigate genetic context and host species of fluoroquinolone resistance genes

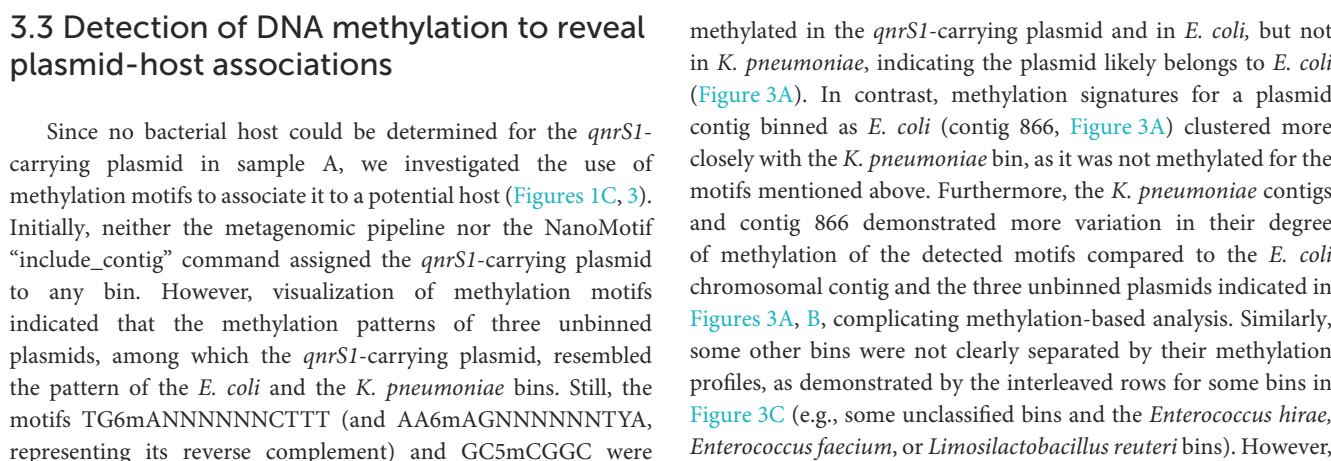
To investigate the broader genetic context of the FQ ARGs, we performed an assembly-based analysis, and binned the resulting contigs ([Figure 1B](#)). In both samples, the spike-in control species were retrieved as high-quality MAGs (>90% completion and <5% contamination), except for *I. halotolerans* in sample A, which had the lowest coverage of 4.39X and was recovered at 87.52% completion ([Supplementary Table S2](#)). Most of these MAGs consisted of multiple contigs (8 to 16 contigs, with N50 ranging from 186 kbp to 1.9 Mbp), except for *A. halotolerans* in sample B, where a single 2.7 Mbp contig was assembled.

Next, the assemblies were screened for the presence of plasmids, MGEs, and for ARGs ([Figure 1B](#)). We compared the detection of FQ resistance genes in the overall metagenome assemblies to their detection in reads ([Table 1](#)). In contrast to the multiple hosts reported for *qnrS1* in sample A, it was found exclusively on a single conjugative IncFII, MOB_P plasmid in the assembly, which carried various other ARGs such as *bla*LAP-2 and *tet*(A) ([Supplementary Figure S1](#)). However, this plasmid could not be attributed to a metagenomic bin. For both *E. coli* and *K. pneumoniae*, MAGs were obtained, although the *K. pneumoniae* MAG was of low completeness ([Table 1](#)). The *E. coli* MAG was highly complete, but the high strain heterogeneity indicates that the MAG represents a collapse of multiple strains. The *qnrB5*



and *oqx*B genes observed in the read-based approach were not detected in the metagenomic assembly of sample A. In sample B, the *oqx*AB operon was located on a high quality *K. pneumoniae* MAG, supporting the read-based analysis. Additionally, the operon was found in proximity to multiple mobile element related genes, as determined by mobileOGdb (Supplementary Figure S2). The *qnr*S1

gene was not detected in the sample B assembly. In both samples, no FQ resistance-associated SNPs were found in the *E. coli* and *K. pneumoniae* bins (Table 1). Finally, in both samples no ARGs were detected in MAGs obtained for the spike-in species *A. halotolerans* and *I. halotolerans*, in contrast to the read-based approach where these species were indicated as ARG hosts.



Since no bacterial host could be determined for the *qnrS1*-carrying plasmid in sample A, we investigated the use of methylation motifs to associate it to a potential host (Figures 1C, 3). Initially, neither the metagenomic pipeline nor the NanoMotifs “include_contig” command assigned the *qnrS1*-carrying plasmid to any bin. However, visualization of methylation motifs indicated that the methylation patterns of three unbinned plasmids, among which the *qnrS1*-carrying plasmid, resembled the pattern of the *E. coli* and the *K. pneumoniae* bins. Still, the motifs TG6mANNNNNNCTTT (and AA6mAGNNNNNNNTYA, representing its reverse complement) and GC5mCGGC were

frontiersin.org

TABLE 1 Overview of fluoroquinolone (FQ) resistance genes.

Sample	FQ resistance gene/mutation	Detected in assembly and reads	Classification	Depth of coverage*	CheckM completeness (%)	CheckM contamination (%)	CheckM strain heterogeneity (%)
A	<i>qnrS1</i>	Yes	IncFII, MOBP plasmid	36.20	n.a.	n.a.	n.a.
	<i>qnrB5</i>	Only in reads	Not detected in assembly, but reads classified as <i>E. coli</i> plasmid.	4.00	n.a.	n.a.	n.a.
	no FQ ARG/mutation found in bin	/	<i>Escherichia coli</i>	32.3	99.97	0.99	77.78
	no FQ ARG/mutation found in bin	/	<i>Klebsiella pneumoniae</i>	3.68	12.07	0.00	0.00
	<i>oqxB</i>	Only in reads	n.a.	2.57	n.a.	n.a.	n.a.
B	<i>oqxAB</i>	Yes	<i>Klebsiella pneumoniae</i>	8.71	98.79	0.04	0.00
	<i>qnrS1</i>	Only in reads	n.a.	1.29	n.a.	n.a.	n.a.

If these were detected in the assembly, metrics for the bin or contig are given. Additionally, if read-level analysis (cfr. Figure 1) identified a species as a potential FQ ARG host, the bin representing these hosts is included here as well. *Depth of coverage: For bins (with or without ARGs) with multiple contigs, the depth of coverage is calculated as the weighted mean of contig depths, weighted by contig length. If detected on an unbinned contig (e.g., *qnrS1* in sample A), the depth of the contig is given. If an ARG was not detected in the assembly but only in reads, the coverage depth of the ARG template is given. n.a., not applicable; /, not detected.

we observed that for other bins, the visualized methylation profiles did correspond to the output of the binning pipeline. This includes the internal control species *I. halotolerans* and *A. halotolerans* (Figures 3C, D).

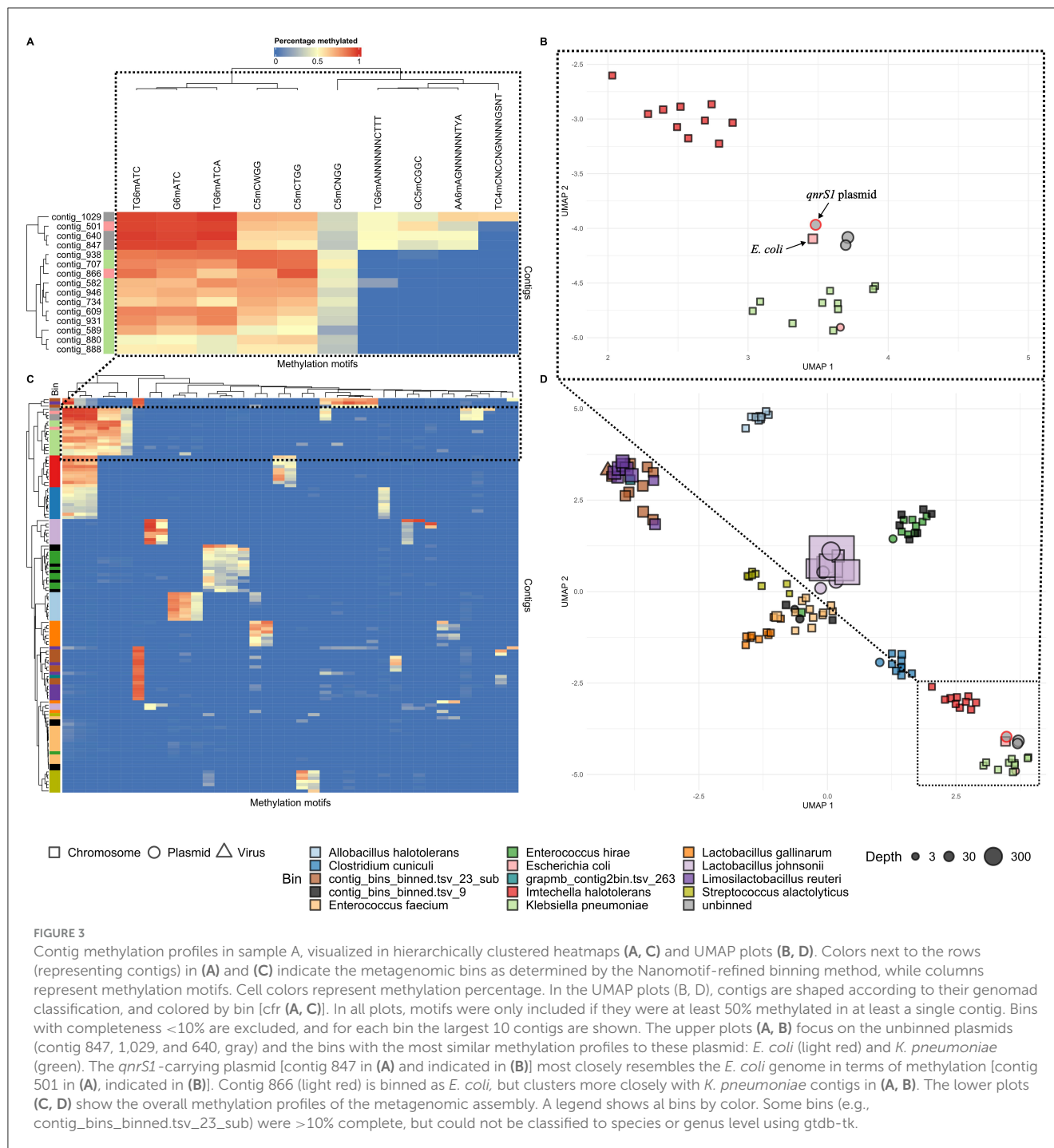
3.4 Long-read haplotyping uncovers fluoroquinolone resistance SNPs and enables MLST

As SNPs in genes encoding target enzymes are an important mechanism of FQ resistance (Kherroubi et al., 2024), the MAGs in both samples were screened, but initially no FQ resistance SNPs were found. However, as MAGs often represent a species-level consensus, they can mask low-frequency FQ resistant strains. Therefore, we performed a haplotype-phasing analysis (Figure 1D), focused on FQ resistance for several MAGs that were indicated as having high strain heterogeneity (Table 2). In sample A, two phased assemblies were produced for the *E. coli* MAG: a first (phased 1) detected at a coverage of 31.3X, with high completeness and reduced contamination than the original MAG; and a second (phased 2) at a lower coverage of 11.9X, with lower completeness and higher contamination than the original MAG (Table 2, Supplementary Figure S3). However, no FQ resistance SNPs were found in either phased assembly in sample A. Using MLST, the phased 1 assembly could be identified as belonging to the ST10 sequence type. Interestingly, some FQ resistant *E. coli* isolates obtained from four farms, including A and B, were also typed as ST10 by MLST (Ringénier et al., under review)¹. In contrast, no sequence type could be determined for the phased 2 assembly. For the *E. coli* MAG in sample B, three phased assemblies were

generated with varying coverages and CheckM quality measures. One of these (phased 2) contained the S83L mutation in *gyrA*, while the other two (phased 1 and 3) did not. However, a high level of contamination and strain heterogeneity was still observed for the phased 2 assembly, which inhibited MLST analysis. Inspection of the phased reads demonstrated that the majority of reads in the phased 2 set contained the mutation, while a smaller proportion did not (Supplementary Figure S5). In the phased 1 assembly, no FQ-resistance associated mutation was observed. In the phased 3 assembly, no FQ resistance-related SNPs were detected. However, as the assembly was only 53.8% complete and had high strain heterogeneity, it is possible that some FQ resistance-related SNPs were not detected. None of the farm B phased assemblies could be sequence typed.

3.5 Phylogenomic comparison of metagenomics data and isolates provide further evidence for plasmid-host link

To confirm that the *qnrS1*-carrying plasmid in sample A originated from an *E. coli* strain, as suggested by the metagenomic DNA methylation analysis (cf. 3.3) we performed two additional analyses: We sought whether an isolate representing the metagenomic strain could be found, and next if this isolate indeed carried the *qnrS1*-carrying IncFII, MOBP plasmid. To accomplish this, we phylogenomically compared the metagenomic *E. coli* strain to several previously isolated ST10 strains, some of which carried an IncFII, MOBP plasmid and the *qnrS1* gene (Figure 4). A hybrid assembly using ONT long-read and short-read data of one of the isolates (Farm A isolate 3) was used as the reference genome, from which plasmids were removed.



Some other isolates were re-sequenced with ONT, and these data were also included in the phylogenomic analysis. The resulting phylogenetic tree demonstrates that the *qnrS1*-carrying farm A isolates are (nearly) identical to the *E. coli* phased 1 strain (Figure 4). Specifically, the metagenomically detected strain differs by 0 SNPs to isolates 1 and 2 from farm A, and by 3 SNPs to isolate 3 from farm A (Supplementary Table S3), indicating that they all represent the same strain. Additionally, long-read assemblies of the farm A isolates contained the same ARGs as the metagenomically phased strain, and similarly no FQ

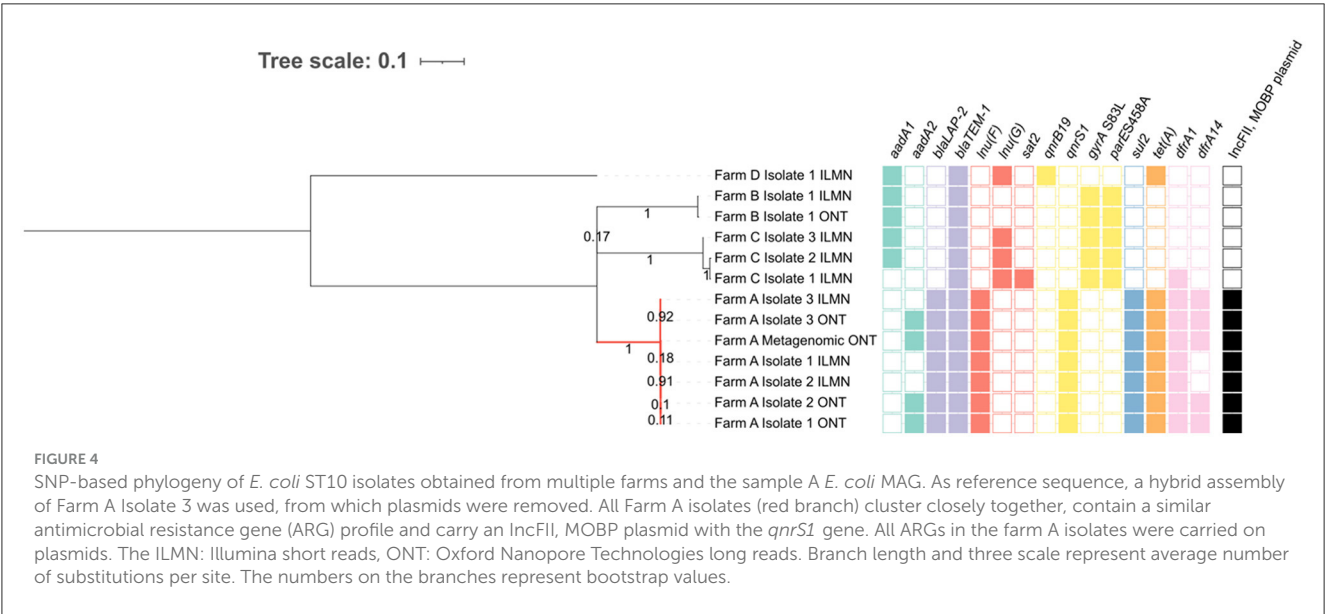
resistance-conferring SNPs. However, the short-read assemblies were missing several ARGs compared to the (metagenomic) long-read assemblies (Figure 4). In both the metagenomic assembly and the isolate hybrid assembly used as reference, these genes were situated on an IncFII, MOB_P plasmid, and were hence not included in phylogenetic analysis.

Next, to confirm that the *qnrS1*-carrying IncFII, MOB_P plasmid in the sample A MAG was identical to the IncFII, MOB_P plasmids in the farm A isolates, we aligned the metagenomically detected plasmids to the plasmids in the reference sequence

TABLE 2 Strain haplotype phasing results for the investigated MAGs, along with fluoroquinolone (FQ) resistance-related point mutations.

Sample	FQ resistance mutation	Classification	Depth of coverage*	CheckM completeness (%)	CheckM contamination (%)	CheckM strain heterogeneity (%)
A	/	<i>Escherichia coli</i> —MAG	32.3	99.97	0.99	77.78
A	/	<i>Escherichia coli</i> —phased 1	31.3	99.97	0.04	0.00
A	/	<i>Escherichia coli</i> —phased 2	11.9	68.65	13.79	100.00
B	/	<i>Escherichia coli</i> —MAG	39.5	95.21	1.53	85.71
B	/	<i>Escherichia coli</i> —phased 1	25.6	91.63	0.30	100.00
B	<i>gyrA</i> S83L	<i>Escherichia coli</i> —phased 2	18.3	99.87	7.75	98.20
B	/	<i>Escherichia coli</i> —phased 3	7.39	53.80	1.60	57.14

Both the original MAG and the strain-phased assemblies are shown. *Coverage depth was calculated as the mean depth over all contigs, weighted by contig length. /, not detected.



(the hybrid assembly of farm A isolate 3, Table 3). Additionally, IncFII, MOBP plasmids from the other two farm A isolates (1 and 2) were also aligned to the reference. Both the IncFII, MOBP plasmid from the metagenomics data, as well as those from the hybrid assemblies were highly similar to the reference plasmid, with coverage of 100% and identity of at least 99.96% (Table 3). Finally, we further investigated the complete plasmid content of the farm A isolates to confirm whether the detection of the other plasmids and their methylation-based association to *E. coli* in the sample A metagenomic data (cfr. 3.3) was correct (Table 3). Aside from the IncFII, MOBP plasmid mentioned earlier, a second and third plasmid, contig 2 and 4 in the reference genome, were also found in all isolates. A fourth plasmid, contig 7 in the reference, was detected in all isolates, but not in the metagenomic assembly. In the reference and isolate 2 it represented a small plasmid of 6,446 bp, but in isolate 1 it was detected as a 3,223 bp plasmid. Interestingly, the plasmid contig that was binned with *E. coli* in the sample A metagenomic results (contig 866 on Figure 3A) did not have a counterpart in the plasmids from the reference genome. However, it partially aligned to the *E. coli* chromosome in the reference,

but only with 21.91% of its sequence (Supplementary Table S4). In summary, from the four plasmids consistently detected in the isolates, we detected three in the sample A metagenomic assembly and correctly attributed them to *E. coli* based on further methylation analysis (cfr 3.3). In contrast, one plasmid associated to the *E. coli* MAG by the binning pipeline was not observed in the isolates.

4 Discussion

4.1 Comparison of ARG taxonomic host detection from long reads and long-read assemblies

Long-read metagenomics holds great promise as a novel tool to detect and monitor AMR. It can complement current methods by providing a fast and comprehensive approach to detect and investigate the genetic factors behind AMR occurrence and spread. In this study, we applied recent developments in Oxford nanopore sequencing to overcome some of the key challenges

TABLE 3 Alignment results of the metagenomic plasmids and the plasmids from isolate hybrid assemblies to plasmids in the Farm A Isolate 3 reference.

Sample	Identity (%)	Reference length	Query length	Reference coverage	Query coverage	Reference contig	Query contig
Farm A metagenomic	99.99	326,267	326,266	100.00	100.00	contig_2	contig_640
	99.42	108,180	106,863	98.32	99.56	contig_4	contig_1029
	99.99	81,767	81,767	100.00	100.00	contig_9	contig_847
Farm A Isolate 1 hybrid assembly	99.99	326,267	326,268	100.00	100.00	contig_2	contig_6
	98.98	108,180	10,8217	99.02	98.99	contig_4	contig_1
	100.00	6446	3,223	50.00	100.00	contig_7	contig_13
	99.99	81,767	8,1767	100.00	100.00	contig_9	contig_2
Farm A Isolate 2 hybrid assembly	99.99	326,267	325,905	99.89	100.00	contig_2	contig_9
	99.97	108,180	108,165	100.00	100.00	contig_4	contig_4
	100.00	6,446	6,446	100.00	100.00	contig_7	contig_8
	99.96	81,767	81,733	100.00	100.00	contig_9	contig_5

Alignments of the *qnrS1*-carrying plasmid are highlighted in bold.

in metagenomics for AMR surveillance. Specifically, we focus on attributing ARGs and plasmids to their bacterial hosts and resolving strain-level variation associated to AMR, using FQ resistance in poultry as a case study.

Apart from detecting ARGs in metagenomic data, it is also crucial to determine their taxonomic origins and genetic surroundings. However, short-read sequencing often fails to identify sufficient genomic context to do so (Abramova et al., 2024). Here, we used long reads to address this issue, and investigated FQ ARGs and their hosts with both read-based and assembly-based approaches. With read-based profiling, several FQ resistance genes were found that went undetected in assemblies, such as *qnrB5* and *oqxB* in sample A and *qnrS1* in sample B. However, these were all observed with low read counts, and their bacterial host reported by the read-based approach could not be confirmed by the assembly-based approach. Additionally, the low sequencing coverage of these ARGs likely impeded their assembly into larger contigs, explaining why they were not detected in the assemblies (Yorki et al., 2023). For example, detection of *oqxB* but not *oqxA* in sample A reads could indicate incomplete coverage of the *oqxB* surroundings, as these normally form a single operon (Li et al., 2019).

In contrast, FQ ARGs retrieved by both read- and assembly-based approaches were found at higher read counts than those exclusively observed in reads. For example, in the case of *oqxAB* in sample B, both analysis methods corresponded, and identified *K. pneumoniae* as the host. In sample A however, read-based ARG-host linking attributed the *qnrS1* gene to multiple host replicons, while it was unambiguously situated on a single plasmid in the metagenomic assembly. Moreover, read-based analysis identified *A. halotolerans* and *I. halotolerans* as carrying an ARG, while their respective MAGs did not contain any. The latter observation corresponds to the reference genomes, and illustrates how spike-in controls can be used to assess metagenomic results (Yorki et al., 2023). Furthermore, the assembled *qnrS1*-carrying plasmid in sample A contained several other types of ARGs, illustrating the need for sufficient genetic context to detect mechanisms of AMR co-selection and multi-drug resistance.

We also compared the metagenomically detected *E. coli* strain in sample A to isolates of the same strain, revealing that metagenomics retrieved three plasmids also found in the isolates, among which the *qnrS1*-carrying plasmid. When comparing the plasmid assemblies, the data from isolates and metagenomics corresponded well. However, a fourth small plasmid which was present in all isolates was not found in the metagenomic assembly, and differed in length between the isolate assemblies. This could be explained by the finding by Johnson et al. (2023) that long-read assemblers struggle with small plasmids.

Taken together, our analysis of both read- and assembly-based approaches highlight that although long reads can capture ARGs within their genetic context, determining their bacterial hosts directly from reads suffers from low precision, especially when carried on plasmids (Bloemen et al., 2023; Lou et al., 2024). On the other hand, assembly enables more accurate linking of ARGs to a replicon, but can be less sensitive at recovering ARGs at low coverages, and struggles with small plasmids (Johnson et al., 2023; Yorki et al., 2023). Novel methods exploiting the benefits of both methods, such as Argo (which was released after finalization of our study), could be tested in the future to allow for rapid, sensitive and precise ARG detection (Chen et al., 2025).

4.2 Methylation-based plasmid-host attribution

Despite enhancing recovery of MGEs and ARGs compared to short-read methods, long-read metagenomics still faces difficulties in associating them to a bacterial host, especially for extrachromosomal elements such as plasmids. In our case, a plasmid carrying *qnrS1* in sample A could be recovered, but was not attributed to a host species by current binning methods. We addressed this problem by applying DNA methylation calling with Oxford nanopore sequencing and methylation binning

with NanoMotif (Heidelberg et al., 2024). On its own, the bin improvement module of NanoMotif did not place the plasmid in the *E. coli* MAG. However, closer manual inspection of the methylation patterns uncovered high similarity between the *qnrS1*-carrying plasmid and the *E. coli* MAG, indicating that the plasmid most likely belonged to *E. coli*. Conversely, methylation and coverage data of another plasmid (contig 866) that was binned as *E. coli* showed that it was possibly incorrectly binned. Both of these findings were supported by isolate data where the *qnrS1*-plasmid was found, and the other plasmid was not observed.

The inability of NanoMotif to correctly bin these plasmids could have multiple causes. First, erroneous basecalling or false positive methylation calls near methylated sites complicate correct methylation motif calling (Kulkarni et al., 2024; Galeone et al., 2025). Second, more complex motifs require higher coverage and motif frequencies for correct detection, which can be challenging in smaller contigs such as plasmids (Heidelberg et al., 2024). A third issue is specific to plasmids, as some have undergone selection for lower numbers of restriction-modification recognition sites (Shaw et al., 2023; Dimitriu et al., 2024). Therefore, our example indicates that methylation data alone can be insufficient to identify a plasmid host, especially when multiple closely related species with conserved methyltransferases are present in the same sample. As an example, the *E. hirae* and *L. reuteri* bins detected in sample A presented similar methylation patterns to other, unclassified bins. Conversely, the spike-in species *I. halotolerans* and *A. halotolerans* had more distinct methylation profiles, and could be clearly separated from all other bins, again highlighting the benefit of adding spike-in controls. Increased coverage of the incompletely separated bins, for example by using ONT adaptive sampling (which allows to accept or reject DNA molecules for sequencing based on analysis of an initial part; Martin et al., 2022), could have better distinguished their genomes by uncovering more DNA methylation motifs. Also, higher coverage and more accurate methylation calling models would likely improve methylation detection and the methylation-assisted binning process in general, as NanoMotif requires sufficient coverage and motif occurrences for reliable motif detection, depending on the type of motif (Heidelberg et al., 2024; Galeone et al., 2025). In our case-study, methylation could be used to point out a likely plasmid-host link, but isolate data were needed to reduce uncertainty. Other novel culture-independent technologies such as Hi-C or single-cell sequencing methods could likely overcome many of the limitations of methylation-based plasmid binning, but are more costly (Brito, 2021; Yang et al., 2025). In contrast, DNA methylation can provide valuable insights in a more cost-effective manner, despite the lower resolution in determining plasmid-host links.

4.3 Strain haplotyping reveals strain-level resistance-associated point mutations

Although long-read metagenomic assembly performs well in recovering MAGs, these often represent a species-level consensus, concealing strain-level variation such as resistance-associated SNPs, and inhibiting strain-level phylogenomic analysis. Therefore, we further investigated MAGs suspected of representing multiple

subpopulations with a strain-haplotyping approach. In sample B, this uncovered a FQ resistance point mutation that was suppressed at the MAG level. In sample A, strain-level phylogenomic comparison related the metagenomic data to nearly identical isolates obtained from the same farm. However, many of the resulting phased assemblies in both samples remained incomplete or contaminated, or represented multiple strains. This was especially the case at low coverages, indicating insufficient strain separation. Novel ONT adaptive sampling approaches such as BOSS-RUNS could enhance strain haplotyping by enriching for regions with low coverage or high variation (Weilguny et al., 2023). This could improve the strain-level resolution of metagenomics, allowing for applications beyond AMR surveillance, such as outbreak investigation, e.g., when no isolate can be obtained or in time-critical situations (Buytaers et al., 2021).

5 Conclusion

In conclusion, we applied new developments in long-read metagenomic data analysis to address several current limitations in metagenomic AMR surveillance. While we analyzed only two samples, these provided sufficient examples to illustrate these limitations and how to overcome them. We demonstrated that read-based approaches are complementary to assembly-based approaches, with the former detecting ARGs at lower sequencing depths, while the latter can better determine their bacterial hosts. We also highlighted the value of DNA methylation detection for AMR surveillance, as this allowed association of a plasmid with its host species. Furthermore, we demonstrated the power of strain-resolved phylogenomic comparison to link metagenomically detected strains to related isolates, and to perform strain-level analysis to unveil resistance-conferring SNPs. However, several challenges remain: DNA methylation might be unable to differentiate organisms with similar methylation profiles, and insufficient coverage can limit detection of DNA methylation motifs and low frequency SNPs, inhibiting both plasmid-host association and strain deconvolution. Future work could exploit ongoing improvements in nanopore sequencing such as increases in throughput, novel ONT adaptive sampling strategies, and improved methylation calling to address these issues, enabling deeper new insights in the occurrence, spread and evolution of AMR.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: <https://www.ebi.ac.uk/ena>, PRJEB86979.

Ethics statement

Ethical approval was not required for the study involving animals in accordance with the local legislation and institutional requirements because Ethical approval was not required for the study, involving animals in accordance with the local legislation

and institutional requirements because only animal fecal material was used and collected from environment.

Author contributions

BBI: Formal analysis, Writing – review & editing, Software, Methodology, Visualization, Investigation, Conceptualization, Data curation, Writing – original draft. MG: Methodology, Writing – review & editing, Investigation, Data curation, Resources. MR: Resources, Writing – review & editing, Investigation. BBo: Software, Writing – review & editing, Formal analysis, Data curation, Resources. KV: Resources, Data curation, Writing – review & editing, Software. KM: Writing – review & editing, Supervision. NR: Writing – review & editing, Resources, Funding acquisition. JD: Writing – review & editing, Resources. FB: Resources, Writing – review & editing. SD: Writing – original draft, Funding acquisition, Resources, Conceptualization, Methodology, Writing – review & editing, Project administration, Validation, Supervision.

Funding

The authors declare that financial support was received for the research and/or publication of this article. The research that yielded these results was funded by Sciensano through contract RP FARMED-internal.

Acknowledgments

We thank the technicians of the service transversal activities in Applied Genomics at Sciensano, Belgium for conducting the Illumina runs, and Stefan Hoffman for the assistance with the Oxford Nanopore Technologies devices.

References

- Abramova, A., Karkman, A., and Bengtsson-Palme, J. (2024). Metagenomic assemblies tend to break around antibiotic resistance genes. *BMC Genomics* 25:959. doi: 10.1186/s12864-024-10876-0
- Agustinho, D. P., Fu, Y., Menon, V. K., Metcalf, G. A., Treangen, T. J., and Sedlazeck, F. J. (2024). Unveiling microbial diversity: harnessing long-read sequencing technology. *Nat. Methods* 21, 954–966. doi: 10.1038/s41592-024-02262-1
- Anjum, M. F., Zankari, E., and Hasman, H. (2017). Molecular methods for detection of antimicrobial resistance. *Microbiol. Spectr.* 5. doi: 10.1128/microbiolspec.ARBA-0011-2017
- Beaulaurier, J., Zhu, S., Deikus, G., Mogno, I., Zhang, X.-S., Davis-Richardson, A., et al. (2018). Metagenomic binning and association of plasmids with bacterial host genomes using DNA methylation. *Nat. Biotechnol.* 36, 61–69. doi: 10.1038/nbt.4037
- BELMAP. (2024). *BELMAP 2024: One Health Report of Antimicrobial Consumption and Resistance in Belgium*. Brussels: BELMAP. doi: 10.25608/v1x6-1e19
- Bloemen, B., Gand, M., Vanneste, K., Marchal, K., Roosens, N. H. C., and De Keersmaecker, S. C. J. (2023). Development of a portable on-site applicable metagenomic data generation workflow for enhanced pathogen and antimicrobial resistance surveillance. *Sci. Rep.* 13:19656. doi: 10.1038/s41598-023-46771-z
- Bogaerts, B., Nouws, S., Verhaegen, B., Denayer, S., Van Braekel, J., Winand, R., et al. (2021). Validation strategy of a bioinformatics whole genome sequencing workflow for Shiga toxin-producing *Escherichia coli* using a reference collection extensively characterized with conventional methods. *Microbial Genomics* 7:000531. doi: 10.1099/mgen.0.000531
- Bogaerts, B., Van Braekel, J., Van Uffelen, A., D'aes, J., Godfroid, M., Delcourt, T., et al. (2025). Galaxy @Sciensano: a comprehensive bioinformatics portal for genomics-based microbial typing, characterization, and outbreak detection. *BMC Genomics* 26:20. doi: 10.1186/s12864-024-11182-5
- Bogaerts, B., Van den Bossche, A., Verhaegen, B., Delbrassinne, L., Mattheus, W., Nouws, S., et al. (2024). Closing the gap: Oxford Nanopore Technologies R10 sequencing allows comparable results to Illumina sequencing for SNP-based outbreak investigation of bacterial pathogens. *J. Clin. Microbiol.* 62:e01523. doi: 10.1128/jcm.01576-23
- Bortolaia, V., Kaas, R. S., Ruppe, E., Roberts, M. C., Schwarz, S., Cattoir, V., et al. (2020). ResFinder 4.0 for predictions of phenotypes from genotypes. *J. Antimicrob. Chemother.* 75, 3491–3500. doi: 10.1093/jac/dkaa345
- Bouras, G., Houtak, G., Wick, R. R., Mallawaarachchi, V., Roach, M. J., Papudeshi, B., et al. (2024). Hybracter: enabling scalable, automated, complete and accurate bacterial genome assemblies. *Microb. Genomics* 10:001244. doi: 10.1099/mgen.0.001244
- Brito, I. L. (2021). Examining horizontal gene transfer in microbial communities. *Nat. Rev. Microbiol.* 19, 442–453. doi: 10.1038/s41579-021-00534-7
- Brown, C. L., Mullet, J., Hindi, F., Stoll, J. E., Gupta, S., Choi, M., et al. (2022). mobileOG-db: a Manually curated database of protein families mediating the life cycle

Conflict of interest

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

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Correction note

This article has been corrected with minor changes. These changes do not impact the scientific content of the article.

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2025.1614301/full#supplementary-material>

- of bacterial mobile genetic elements. *Appl. Environ. Microbiol.* 88, e00991–e00922. doi: 10.1128/aem.00991-22
- Bush, N. G., Diez-Santos, I., Abbott, L. R., and Maxwell, A. (2020). Quinolones: mechanism, lethality and their contributions to antibiotic resistance. *Molecules* 25:5662. doi: 10.3390/molecules25235662
- Buytaers, F. E., Saltykova, A., Mattheus, W., Verhaegen, B., Roosens, N. H. C., Vanneste, K., et al. (2021). Application of a strain-level shotgun metagenomics approach on food samples: resolution of the source of a *Salmonella* food-borne outbreak. *Microbial Genomics* 7:000547. doi: 10.1099/mgen.0.000547
- Camargo, A. P., Roux, S., Schulz, F., Babinski, M., Xu, Y., Hu, B., et al. (2023). Identification of mobile genetic elements with geNomad. *Nat. Biotechnol.* 42, 1303–1312. doi: 10.1038/s41587-023-01953-y
- Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2022). GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 38, 5315–5316. doi: 10.1093/bioinformatics/btac672
- Chen, X., Yin, X., Xu, X., and Zhang, T. (2025). Species-resolved profiling of antibiotic resistance genes in complex metagenomes through long-read overlapping with Argo. *Nat. Commun.* 16:1744. doi: 10.1038/s41467-025-57088-y
- Clausen, P. T. L. C., Aarestrup, F. M., and Lund, O. (2018). Rapid and precise alignment of raw reads against redundant databases with KMA. *BMC Bioinformatics* 19:307. doi: 10.1186/s12859-018-2336-6
- Crits-Christoph, A., Kang, S. C., Lee, H. H., and Ostrov, N. (2023). MicrobeMod: A computational toolkit for identifying prokaryotic methylation and restriction-modification with nanopore sequencing. *bioRxiv* 2023.11.13.566931. doi: 10.1101/2023.11.13.566931
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *Gigascience* 10:giab008. doi: 10.1093/gigascience/giab008
- Dimitriu, T., Szczelkun, M. D., and Westra, E. R. (2024). Various plasmid strategies limit the effect of bacterial restriction–modification systems against conjugation. *Nucleic Acids Res.* 52, 12976–12986. doi: 10.1093/nar/gkac896
- Djordjevic, S. P., Jarocki, V. M., Seemann, T., Cummins, M. L., Watt, A. E., Drigo, B., et al. (2024). Genomic surveillance for antimicrobial resistance — a One Health perspective. *Nat. Rev. Genet.* 25, 142–157. doi: 10.1038/s41576-023-00649-y
- Edge, P., and Bansal, V. (2019). Longshot enables accurate variant calling in diploid genomes from single-molecule long read sequencing. *Nat. Commun.* 10:4660. doi: 10.1038/s41467-019-12493-y
- Feldgarden, M., Brover, V., Gonzalez-Escalona, N., Frye, J. G., Haendiges, J., Haft, D. H., et al. (2021). AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci. Rep.* 11:12728. doi: 10.1038/s41598-021-91456-0
- Feldgarden, M., Brover, V., Haft, D. H., Prasad, A. B., Slotta, D. J., Tolstoy, I., et al. (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:e00483-19. doi: 10.1128/AAC.00483-19
- Galeone, V., Dabernig-Heinz, J., Lohde, M., Brandt, C., Kohler, C., Wagner, G. E., et al. (2025). Decoding bacterial methylomes in four public health-relevant microbial species: nanopore sequencing enables reproducible analysis of DNA modifications. *BMC Genomics* 26:394. doi: 10.1186/s12864-025-11592-z
- GBD 2021 Antimicrobial Resistance Collaborators, Swetschinski, L. R., Gray, A. P., Wool, E. E., Aguilar, G. R., Mestrovic, T., et al. (2024). Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. *Lancet* 404, 1199–1226. doi: 10.1016/S0140-6736(24)01867-1
- Grant, J. R., Enns, E., Marinier, E., Mandal, A., Herman, E. K., Chen, C., et al. (2023). Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic Acids Res.* 51, W484–W492. doi: 10.1093/nar/gkad326
- Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32, 2847–2849. doi: 10.1093/bioinformatics/btw313
- Heidelberg, S., Dall, S. M., Bojer, J. S., Nissen, J., van der Maas, L. N. L., Sereika, M., et al. (2024). Nanomotif: identification and exploitation of DNA methylation motifs in metagenomes using oxford nanopore sequencing. *bioRxiv* 2024.04.29.591623. doi: 10.1101/2024.04.29.591623
- Johnson, J., Soehnlen, M., and Blankenship, H. M. (2023). Long read genome assemblers struggle with small plasmids. *Microbial Genomics* 9:001024. doi: 10.1099/mgen.0.001024
- Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *Peer J.* 7:e7359. doi: 10.7717/peerj.7359
- Kazantseva, E., Donmez, A., Frolova, M., Pop, M., and Kolmogorov, M. (2024). Strainy: phasing and assembly of strain haplotypes from long-read metagenome sequencing. *Nat. Methods* 21, 2034–2043. doi: 10.1038/s41592-024-02424-1
- Kherroubi, L., Bacon, J., and Rahman, K. M. (2024). Navigating fluoroquinolone resistance in Gram-negative bacteria: a comprehensive evaluation. *JAC-Antimicrob. Resist.* 6:dlae127. doi: 10.1093/jacamr/dlae127
- Kolmogorov, M., Bickhart, D. M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S. B., et al. (2020). metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat. Methods* 17, 1103–1110. doi: 10.1038/s41592-020-00971-x
- Kulkarni, O., Mathew, R. J., Zaveri, L., Jana, R., Singh, N. K., Ara, S., et al. (2024). Comprehensive benchmarking of tools for nanopore-based detection of DNA methylation. *bioRxiv* 2024.11.09.622763. doi: 10.1101/2024.11.09.622763
- Lamurias, A., Sereika, M., Albertsen, M., Hose, K., and Nielsen, T. D. (2022). Metagenomic binning with assembly graph embeddings. *Bioinformatics* 38, 4481–4487. doi: 10.1093/bioinformatics/btac557
- Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi: 10.1093/bioinformatics/bty191
- Li, J., Zhang, H., Ning, J., Sajid, A., Cheng, G., Yuan, Z., et al. (2019). The nature and epidemiology of OqxAB, a multidrug efflux pump. *Antimicrob. Resist. Infect. Control* 8:44. doi: 10.1186/s13756-019-0489-3
- Lou, E. G., Fu, Y., Wang, Q., Treangen, T. J., and Stadler, L. B. (2024). Sensitivity and consistency of long- and short-read metagenomics and epicPCR for the detection of antibiotic resistance genes and their bacterial hosts in wastewater. *J. Hazard. Mater.* 469:133939. doi: 10.1016/j.jhazmat.2024.133939
- Lu, J., Rincon, N., Wood, D. E., Breitwieser, F. P., Pockrandt, C., Langmead, B., et al. (2022). Metagenome analysis using the Kraken software suite. *Nat. Protoc.* 17, 2815–2839. doi: 10.1038/s41596-022-00738-y
- Maguire, F., Jia, B., Gray, K. L., Lau, W. Y. V., Beiko, R. G., and Brinkman, F. S. L. (2020). Metagenome-assembled genome binning methods with short reads disproportionately fail for plasmids and genomic Islands. *Microbial Genomics* 6:e000436. doi: 10.1099/mgen.0.000436
- Marçais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., and Zimin, A. (2018). MUMmer4: A fast and versatile genome alignment system. *PLoS Comput. Biol.* 14:e1005944. doi: 10.1371/journal.pcbi.1005944
- Martin, S., Heavens, D., Lan, Y., Horsfield, S., Clark, M. D., and Leggett, R. M. (2022). Nanopore adaptive sampling: a tool for enrichment of low abundance species in metagenomic samples. *Genome Biol.* 23:11. doi: 10.1186/s13059-021-02582-x
- McInnes, L., Healy, J., and Melville, J. (2020). UMAP: uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*. Available online at: <https://arxiv.org/abs/1802.03426>
- Nicholls, S. M., Aubrey, W., De Grave, K., Schietgat, L., Creevey, C. J., and Clare, A. (2021). On the complexity of haplotyping a microbial community. *Bioinformatics* 37, 1360–1366. doi: 10.1093/bioinformatics/btaa977
- O’Leary, N. A., Wright, M. W., Brister, J. R., Ciufu, S., Haddad, D., McVeigh, R., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44, D733–D745. doi: 10.1093/nar/gkv1189
- Pan, S., Zhao, X.-M., and Coelho, L. P. (2023). SemiBin2: self-supervised contrastive learning leads to better MAGs for short- and long-read sequencing. *Bioinformatics* 39, i21–i29. doi: 10.1093/bioinformatics/btad209
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114
- Pellow, D., Zorea, A., Probst, M., Furman, O., Segal, A., Mizrahi, I., et al. (2021). SCAPP: an algorithm for improved plasmid assembly in metagenomes. *Microbiome* 9:144. doi: 10.1186/s40168-021-01068-z
- Quick, J. (2019). *The “Three Peaks” Faecal DNA Extraction Method for Long-Read Sequencing*. Available online at: <https://www.protocols.io/view/the-39-three-peaks-39-faecal-dna-extraction-method-7rsh6e> (Accessed January 29, 2025). doi: 10.17504/protocols.io.584g9yw
- Robertson, J., Bessonov, K., Schonfeld, J., and Nash, J. H. E. (2020). Universal whole-sequence-based plasmid typing and its utility to prediction of host range and epidemiological surveillance. *Microbial Genomics* 6:e000435. doi: 10.1099/mgen.0.000435
- Sanderson, N. D., Hopkins, K. M. V., Colpus, M., Parker, M., Lipworth, S., Crook, D., et al. (2024). Evaluation of the accuracy of bacterial genome reconstruction with Oxford Nanopore R10.4.1 long-read-only sequencing. *Microbial Genomics* 10:001246. doi: 10.1099/mgen.0.001246
- Sereika, M., Kirkegaard, R. H., Karst, S. M., Michaelsen, T. Y., Sørensen, E. A., Wollenberg, R. D., et al. (2022). Oxford Nanopore R10.4 long-read sequencing enables the generation of near-finished bacterial genomes from pure cultures and metagenomes without short-read or reference polishing. *Nat. Methods* 19, 823–826. doi: 10.1038/s41592-022-01539-7
- Shaw, J., Gounot, J.-S., Chen, H., Nagarajan, N., and Yu, Y. W. (2024). Floria: fast and accurate strain haplotyping in metagenomes. *Bioinformatics* 40, i30–i38. doi: 10.1093/bioinformatics/btae252

- Shaw, L. P., Rocha, E. P. C., and MacLean, R. C. (2023). Restriction-modification systems have shaped the evolution and distribution of plasmids across bacteria. *Nucleic Acids Res.* 51, 6806–6818. doi: 10.1093/nar/gkad452
- Shen, W., Sipos, B., and Zhao, L. (2024). SeqKit2: A Swiss army knife for sequence and alignment processing. *iMeta* 3:e191. doi: 10.1002/imt2.191
- Sieber, C. M. K., Probst, A. J., Sharrar, A., Thomas, B. C., Hess, M., Tringe, S. G., et al. (2018). Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat. Microbiol.* 3, 836–843. doi: 10.1038/s41564-018-0171-1
- Soto-Serrano, A., Li, W., Panah, F. M., Hui, Y., Atienza, P., Fomenkov, A., et al. (2024). Matching excellence: Oxford Nano pore Technologies' rise to parity with Pacific Biosciences in genome reconstruction of non-model bacterium with high G+C content. *Microb. Genomics* 10:001316. doi: 10.1099/mgen.0.001316
- Thorvaldsdóttir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinformatics* 14, 178–192. doi: 10.1093/bib/bbs017
- Tourancheau, A., Mead, E. A., Zhang, X.-S., and Fang, G. (2021). Discovering multiple types of DNA methylation from bacteria and microbiome using nanopore sequencing. *Nat. Methods* 18, 491–498. doi: 10.1038/s41592-021-01109-3
- Van Uffelen, A., Posadas, A., Roosens, N. H. C., Marchal, K., De Keersmaecker, S. C. J., and Vanneste, K. (2024). Benchmarking bacterial taxonomic classification using nanopore metagenomics data of several mock communities. *Sci Data* 11:864. doi: 10.1038/s41597-024-03672-8
- Weilguny, L., De Maio, N., Munro, R., Manser, C., Birney, E., Loose, M., et al. (2023). Dynamic, adaptive sampling during nanopore sequencing using Bayesian experimental design. *Nat. Biotechnol.* 41, 1018–1025. doi: 10.1038/s41587-022-01580-z
- WHO (2024). *WHO's List of Medically Important Antimicrobials: A Risk Management Tool for Mitigating Antimicrobial Resistance due to Non-human Use*. Geneva, Switzerland: World Health Organization. Available online at: https://cdn.who.int/media/docs/default-source/gcp/who-mia-list-2024-lv.pdf?sfvrsn=3320dd3d_2 (Accessed November 4, 2024).
- Wishart, D. S., Han, S., Saha, S., Oler, E., Peters, H., Grant, J. R., et al. (2023). PHASTEST: faster than PHASTER, better than PHAST. *Nucleic Acids Res.* 51, W443–W450. doi: 10.1093/nar/gkad382
- Wood, D. E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biol.* 20:257. doi: 10.1186/s13059-019-1891-0
- Yang, Q. E., Gao, J. T., Zhou, S. G., and Walsh, T. R. (2025). Cutting-edge tools for unveiling the dynamics of plasmid–host interactions. *Trends Microbiol.* 33, 496–509. doi: 10.1016/j.tim.2024.12.013
- Yorki, S., Shea, T., Cuomo, C. A., Walker, B. J., LaRocque, R. C., Manson, A. L., et al. (2023). Comparison of long- and short-read metagenomic assembly for low-abundance species and resistance genes. *Briefings Bioinformatics* 24:bbad050. doi: 10.1093/bib/bbad050
- Zhao, S., Tyson, G. H., Chen, Y., Li, C., Mukherjee, S., Young, S., et al. (2016). Whole-genome sequencing analysis accurately predicts antimicrobial resistance phenotypes in *Campylobacter* spp. *Appl. Environ. Microbiol.* 82, 459–466. doi: 10.1128/AEM.02873-15
- Zhou, Z., Alikhan, N.-F., Mohamed, K., Fan, Y., and Group, the, A. S., Achtman, M., et al. (2020). The Enterobase user's guide, with case studies on *Salmonella* transmissions, *Yersinia pestis* phylogeny, and *Escherichia* core genomic diversity. *Genome Res.* 30, 138–152. doi: 10.1101/gr.251678.119