Benchmarks and comparison of different bioinformatics tools to detect AMR genes in sequencing datasets, including whole metagenome shotgun experiments. Compare the tools and the databases, and the limitations of different approaches.

Different bioinformatics approaches for AMR gene detection show varying levels of sensitivity and accuracy, with performance heavily dependent on tool type, database selection, and specific experimental context.

Abstract

Benchmarks for antimicrobial resistance gene detection vary substantially with tool type, data source, and database choice. In simulated and real sequencing experiments—including whole metagenome shotgun data—studies report the following:

1. For read-mapping tools: - ABRIcate sensitivity ranged from 0.25 to 0.99, while ResFinder achieved balanced accuracy from 0.40 to 0.92. - KmerResistance reached 99.9% sensitivity for isolated variants but dropped to 76.3% when discriminating closely related variants. - ARIBA and SRST2 recorded 90.4% and 68.5% sensitivity for single variants, yet only 3.6% and 4.3% when variants were similar.

2. For machine learning and model-based approaches: - AMR-meta reported F-scores between 0.2 and 0.9 (median 0.7), with KARGA offering recall of 83–99% over moderate mutation rates. - ROCker models yielded a 0% false positive rate, false negatives under 4%, and demonstrated F1 improvements that were up to 50× higher than competing methods. - Long-read methods such as MegaPath-Nano outperformed comparable tools by achieving 28% higher accuracy relative to ARMA.

Database selection emerged as a critical factor. Multiple studies noted that the choice, curation, and update frequency of databases (e.g., CARD, ResFinder, MEGARes) directly affect allele-level calls and resistance predictions. Furthermore, categorization of tools reveals that k-mer based methods deliver high speed and efficient recall in large-scale metagenomics—with the trade-off of potentially missing novel variants—whereas alignment-based methods provide robust, interpretable results in isolate sequencing applications. Machine learning and species-specific strategies exhibit improved sensitivity in complex resistomes when supplied with appropriate training data.

These findings indicate that performance, resource demands, and limitations of AMR detection approaches are closely tied to the chosen methodology and database, emphasizing the need to match the tool with the specific experimental context.

Paper search

Using your research question "Benchmarks and comparison of different bioinformatics tools to detect AMR genes in sequencing datasets, including whole metagenome shotgun experiments. Compare the tools and the databases, and the limitations of different approaches.", we searched across over 126 million academic papers from the Semantic Scholar corpus. We retrieved the 499 papers most relevant to the query.

Screening

We screened in sources that met these criteria:

- **Comparative Analysis**: Does the study compare multiple (two or more) bioinformatics tools for AMR gene detection with defined performance metrics (sensitivity, specificity, or accuracy)?
- **Data Type**: Does the study analyze sequencing data (such as whole metagenome shotgun data) for AMR gene detection?
- **Dataset Validation**: Does the study evaluate tools using either real datasets with validated results or simulated datasets with known ground truth?
- **Performance Evaluation**: Does the study include quantitative performance evaluation of the tools being compared?
- **Method Type**: Does the study focus on sequencing-based AMR detection methods (rather than non-sequencing methods)?
- **Tool Analysis Scope**: Does the study compare multiple tools (rather than focusing on a single tool)?
- **Database Comparison**: Does the study evaluate different AMR databases or their impact on detection performance?
- **Technical Assessment**: Does the study include analysis of technical limitations or challenges of the different approaches?

We considered all screening questions together and made a holistic judgement about whether to screen in each paper.

Data extraction

We asked a large language model to extract each data column below from each paper. We gave the model the extraction instructions shown below for each column.

- **Bioinformatics Tools Evaluated**:

  List all bioinformatics tools for AMR gene detection that were included in the study.

- Capture the full name of each tool
- Include the version number if specified
- Note the source/repository (e.g., GitHub link)
- If multiple tools were compared, list all of them
- If no specific version is mentioned, write "version not specified"

Example extraction format:

1. Tool Name (Version): [specific details]
2. Tool Name (Version): [specific details]

If tools are not explicitly named, write "No specific tools identified in the text."

- **Databases Used for AMR Gene Detection**:

  Identify and list all databases used for AMR gene detection in the study.

- Capture the full name of each database
- Include version number if specified
- Note any specific characteristics or focus of the database (e.g., resistance gene type)

If multiple databases were used or compared, list all of them. If no specific database is mentioned, write "No databases specified."

Example extraction format:

1. Database Name (Version): [specific details]
2. Database Name (Version): [specific details]

- **Data Type and Source**:

  Describe the type of genomic data used in the study:

- Specify the data type (e.g., whole metagenome shotgun, short-read Illumina)
- Indicate the source of data (e.g., simulated community, clinical samples, sequenced strains)
- Note sample size or number of datasets used
- Capture any specific characteristics of the data (e.g., coverage levels, resistance levels)

If multiple data types were used, describe each separately. If data details are unclear, write "Data details not sufficiently specified."

- **Performance Metrics**:

  List all performance metrics used to evaluate the bioinformatics tools:

- Capture the full name of each metric
- Include the specific value or range of values if reported
- Note the context of the metric (e.g., sensitivity at different coverage levels)

Typical metrics might include:

- Sensitivity
- Specificity
- Precision
- Accuracy
- F1 score
- Cohen's kappa

Example extraction format:

1. Metric Name: [value/range] (context)
2. Metric Name: [value/range] (context)

If no specific metrics are reported, write "No performance metrics specified."

- **Key Comparative Findings**:

  Summarize the main comparative findings about AMR gene detection tools:

- Capture significant differences between tools
- Note any consistent patterns or variations across tools
- Highlight any limitations or challenges identified in the study

Focus on:

- Variations in AMR gene identification
- Performance across different drug classes
- Differences at various genome coverage levels

If findings are complex, use bullet points to break down key observations. If no clear comparative findings are reported, write "No significant comparative findings identified."

Results

Characteristics of Included Studies

Study	Study Design	Tools Evaluated	Data Types	Performance Metrics	Full text retrieved
Wissel et al., 2022	Benchmarking, simulated metagenomes	9 open-source tools (not named)	Simulated Illumina metagenomes (mock communities, various coverage)	Sensitivity, specificity, precision, accuracy	No
"Benchmarking software to predict antibiotic resistance phenotypes," 2022	Systematic benchmarking, simulated data	ABRIcate, fARGene, ResFinder, shortBRED, Resistance Gene Identifier (RGI), AMRFinderPlus, starAMR, sraX, deepARG	Simulated Illumina metagenomes (8 pathogens, 3 coverage levels)	Sensitivity, specificity, F1, precision, accuracy, Cohen's kappa	Yes
Clausen et al., 2016	Method development and benchmarking	KmerResistance, ResFinder, SRST2	Whole genome sequencing (clinical and farm isolates, 339 total)	No mention found	No
Marini et al., 2022	Tool development and benchmarking	AMR-meta, AMRPlusPlus, DeepARG, Meta-MARC	Metagenomic shotgun (semi-synthetic, cross-validation)	F-score, hit rate, run-time	No
"OUP accepted manuscript," 2022	Comparative benchmarking	AMRPlusPlus, DeepARG, KARGA, ResFinder, Meta-MARC	High-throughput sequencing (clinical, 585 isolates)	Balanced accuracy	No

Study	Study Design	Tools Evaluated	Data Types	Performance Metrics	Full text retrieved
Marini et al., 2021	Comparative benchmarking	AMRPlusPlus, DeepARG, KARGA, Meta-MARC, ResFinder	Short-read Illumina (clinical, 585 isolates)	Sensitivity, specificity, balanced accuracy, F1	Yes
Davies et al., 2021	Systematic comparison, simulated and real data	ABRicate, ARIBA, Kmer-Resistance, SRST2	Simulated constructs, whole genome sequencing (Escherichia coli, 1818 isolates)	Accuracy, discrepancy rate, sensitivity at coverage	Yes
Nunez-Garcia et al., 2022	Harmonisation study	GeneFinder, APHA SeqFinder/ABRicate, WBVR BLAST, ResFinder/PointFinder, ARIBA	Illumina whole genome sequencing (clinical/veterinary, 436 Escherichia coli)	Sensitivity, specificity	Yes
Madden et al., 2024	Tool development and benchmarking	ARDaP, abritAMR, AMRFinderPlus, ResFinder, RGI, Comprehensive Antibiotic Resistance Database (CARD)	Whole genome sequencing (Pseudomonas aeruginosa, 1877+102 isolates)	Balanced accuracy, precision, recall	Yes
Abramova et al., 2023	Assembly benchmarking	Velvet, Ray, MEGAHIT, metaSPAdes, Trinity, others	Simulated and real metagenomes (Illumina, PacBio)	Assembled length, contig number, mapping rate, antibiotic resistance gene (ARG) recovery	Yes
Cooper et al., 2020	Evaluation of whole genome sequencing-based antimicrobial resistance prediction	CARD-RGI, ResFinder, SRST2, KMA	Illumina whole genome sequencing (Salmonella, 111 isolates)	Accuracy	Yes

Study	Study Design	Tools Evaluated	Data Types	Performance Metrics	Full text retrieved
Mahé et al., 2019	Large-scale evaluation	TBProfiler, Mykrobe	Illumina whole genome sequencing (Mycobacterium tuberculosis, 6574 genomes)	Sensitivity, specificity, precision, area under the curve (AUC)	Yes
Pesesky et al., 2016	Rules-based vs. machine learning benchmarking	ResFinder, CARD, Resfams, BLAST+, HMMER3, custom scripts	Illumina whole genome sequencing (Enterobacteriaceae, 78 isolates)	Agreement, major/very major error, area under the curve (AUC)	Yes
Zhang et al., 2022	ROCker model benchmarking	ROCker, BLASTx, HMMs, DeepARG, ShortBRED, AMRFinder	Simulated short-read (150/250 base pairs)	False positive rate (FPR), false negative rate (FNR), F1	No
Carroll et al., 2021	Concordance study	ABRicate, AMRFinderPlus, ARIBA, BTyper, SRST2, PATRIC3	Illumina whole genome sequencing (Salmonella, 128 strains)	Accuracy, Cohen's kappa, sensitivity, specificity	Yes
Cooper et al., 2024	Comparative analysis, synthetic metagenomes	KMA, SRST2, CARD-RGI	Simulated Illumina metagenomes (beef/lettuce, 100+50 replicates)	Sensitivity, specificity, accuracy	Yes
Madden et al., 2020	Tool development and benchmarking	ARIBA, CARD, ResFinder, AMRFinderPlus, ARDaP	Whole genome sequencing (Burkholderia pseudomallei, 41 strains)	No mention found	No
Doyle et al., 2019	Inter-lab reproducibility study	ABRicate, rgi, c-SSTAR, Resfinder, ariba, srst2, Genefinder	Illumina whole genome sequencing (carbapenem-resistant, 10 datasets)	Carbapenemase reporting, consensus, sensitivity	Yes

Study	Study Design	Tools Evaluated	Data Types	Performance Metrics	Full text retrieved
Madden et al., 2019	Tool development and benchmarking	ARIBA, CARD, ARDaP	Whole genome sequencing (Burkholderia pseudomallei, 47+3 strains)	No mention found	Yes
McCall and Xagoraraki, 2018	Aligner comparison	Bowtie2, bwa-mem, blastn, blastx	Simulated reads, wastewater metagenomes	Correctly mapped, false positives, multi-reads, partials	No
Prosperi and Marini, 2021	Toolkit development and benchmarking	KARGA, AMRPlusPlus, DeepARG, MetaMARC	Metagenomic short reads (simulated, empirical)	Recall, hit score, run-time	No
Hodges et al., 2021	Whole genome sequencing-based antimicrobial resistance prediction	Sipprverse, KMA, SRST2, GeneSeekr, StarAMR	Illumina whole genome sequencing (Campylobacter, 271 isolates)	Concordance, association, positive predictive value (PPV), sensitivity, specificity	Yes
Rocha et al., 2024	Antimicrobial resistance prediction in Corynebacterium	BV-BRC, ResFinder, RAST, Kmer-Resistance	Whole genome sequencing (Corynebacterium, 107+18 isolates)	Area under the curve (AUC), major error rate (MER), very major error rate (VMER), Matthews correlation coefficient (MCC), F1, accuracy	No
Zhang et al., "Rocker Models for Reliable Detection," 2020	ROCker benchmarking	ROCker, BLASTx, HMMer, DeepARG, ShortBRED, AMRFinder	Simulated short-read (150/250 base pairs)	False positive rate (FPR), false negative rate (FNR), F1, precision, recall, accuracy, sensitivity, specificity	Yes

Study	Study Design	Tools Evaluated	Data Types	Performance Metrics	Full text retrieved
Lui et al., 2020	Tool development and benchmarking	MegaPath-Nano, WIMP, Kraken2, Bracken, MetaMaps, ARGpore, ARMA, CARD, ResFinder, AMRfinder, MEGARes, CBMAR	Oxford Nanopore Technologies (ONT) long-read metagenomics (Zymo, 5 patient isolates)	Precision, sensitivity, recall, accuracy, F1	Yes

Study design:

- Benchmarking or comparative evaluation (including tool development, systematic, ROCker, and toolkit benchmarking) was reported in 16 studies.
- Assembly benchmarking was reported in 1 study, harmonisation in 1, inter-laboratory reproducibility in 1, aligner comparison in 1, concordance in 1, whole genome sequencing-based antimicrobial resistance prediction in 2, large-scale evaluation in 1, and antimicrobial resistance prediction in a specific taxon in 1.

Tools evaluated:

- ResFinder was evaluated in 12 studies, the most of any tool.
- SRST2 appeared in 7 studies; DeepARG, ARIBA, and CARD each appeared in 5–6 studies; ABRicate, AMRFinderPlus, Meta-MARC, and KARGA each appeared in 3–5 studies.
- Over 30 other tools were evaluated in 1–2 studies each.
- We did not find tool names specified in 1 study.

Data types:

- Illumina or high-throughput whole genome sequencing of isolates was used in 15 studies.
- Simulated metagenomes, short reads, or constructs were used in 8 studies.
- Metagenomic shotgun or short reads (empirical or semi-synthetic) were used in 3 studies.
- Real metagenomes (environmental, wastewater) were used in 3 studies.
- Oxford Nanopore Technologies (ONT) long-read metagenomics was used in 1 study.

Performance metrics:

- Sensitivity was reported in 12 studies, specificity in 9, and accuracy in 9.
- F1/F-score was reported in 7 studies, precision in 6, recall in 4, and balanced accuracy in 3.
- Area under the curve (AUC) and Cohen's kappa were each reported in 2–3 studies.
- Other metrics (e.g., discrepancy rate, hit rate, run-time, agreement, major/very major error, mapping rate, carbapenemase reporting, concordance, association, positive predictive value (PPV), major error rate (MER), very major error rate (VMER), Matthews correlation coefficient (MCC), assembled length, contig number, false positive rate (FPR), false negative rate (FNR)) were each reported in 1–2 studies.

- We did not find mention of performance metrics in 3 studies.

Tool Performance Analysis

Detection Accuracy

Tool Name	Sensitivity Range	Specificity Range	Resource Requirements
ABRIcate	0.25–0.99 (simulated data)	0.2–1 (simulated data)	No mention found
fARGene	No mention found	No mention found	No mention found
ResFinder	0.40–0.92 (balanced accuracy)	No mention found	No mention found
shortBRED	No mention found	No mention found	No mention found
Resistance Gene Identifier (CARD)	No mention found	No mention found	No mention found
AMRFinderPlus	No mention found	No mention found	No mention found
starAMR	No mention found	No mention found	No mention found
sraX	No mention found	No mention found	No mention found
deepARG	No mention found	No mention found	No mention found
KmerResistance	99.9% (simulated, single variant), 76.3% (closely related variants)	No mention found	No mention found
ARIBA	90.4% (simulated, single variant), 3.6% (closely related variants)	No mention found	No mention found
SRST2	68.5% (simulated, single variant), 4.3% (closely related variants)	No mention found	No mention found
AMR-meta	F-score 0.2–0.9 (median 0.7)	No mention found	3x faster than DeepARG, 30x faster than Meta-MARC
KARGA	Recall 83–99% (mutation rates 10–25%)	No mention found	2x faster than AMRPlusPlus, 7x than DeepARG, 100x than MetaMARC
ARDaP	Balanced accuracy 81–85% (Pseudomonas aeruginosa)	No mention found	No mention found
ROCker	False positive rate 0%, false negative rate 0–4%, F1 up to 50x higher than others	No mention found	No mention found

Tool Name	Sensitivity Range	Specificity Range	Resource Requirements
MegaPath-Nano	No mention found	No mention found	Outperformed others in precision, sensitivity, recall; 28% higher accuracy than ARMA
Others (see full table)	Variable	Variable	Variable

- Sensitivity or related performance metrics were reported for 9 tools (ABRIcate, ResFinder, KmerResistance, ARIBA, SRST2, AMR-meta, KARGA, ARDaP, ROCker).
  - ABRIcate: 0.25–0.99 (simulated data)
  - ResFinder: 0.40–0.92 (balanced accuracy)
  - KmerResistance: 99.9% (simulated, single variant), 76.3% (closely related variants)
  - ARIBA: 90.4% (simulated, single variant), 3.6% (closely related variants)
  - SRST2: 68.5% (simulated, single variant), 4.3% (closely related variants)
  - AMR-meta: F-score 0.2–0.9 (median 0.7)
  - KARGA: Recall 83–99% (mutation rates 10–25%)
  - ARDaP: Balanced accuracy 81–85% (Pseudomonas aeruginosa)
  - ROCker: False positive rate 0%, false negative rate 0–4%, F1 up to 50x higher than others
- We did not find sensitivity or related metrics for 8 tools (fARGene, shortBRED, Resistance Gene Identifier (CARD), AMRFinderPlus, starAMR, sraX, deepARG, MegaPath-Nano). For 1 tool (Others), sensitivity was described as "variable."
- Specificity was reported for only 1 tool (ABRIcate: 0.2–1, simulated data), was described as "variable" for 1 tool (Others), and we did not find specificity data for the remaining 16 tools.
- Resource requirements were described for 3 tools in terms of relative speed or performance:
  - AMR-meta: 3x faster than DeepARG, 30x faster than Meta-MARC
  - KARGA: 2x faster than AMRPlusPlus, 7x than DeepARG, 100x than MetaMARC
  - MegaPath-Nano: Outperformed others in precision, sensitivity, recall; 28% higher accuracy than ARMA
- We did not find resource requirement information for 14 tools. For 1 tool (Others), resource requirements were described as "variable."

Thematic Analysis

Database Impact on Detection

- Several studies reported that database choice (for example, Comprehensive Antibiotic Resistance Database (CARD), ResFinder, MEGARes, Antibiotic Resistance Genes Database (ARDB), or custom species-specific databases) affected detection results, with some tools failing to detect clinically relevant antimicrobial resistance due to database limitations.
- Database curation, update frequency, and ontology standardization were recurring challenges across studies.

- Discrepancies in allele-level calls and resistance predictions were often attributed to differences in database content and structure.

Metagenomic-Specific Challenges

- Metagenomic and low-coverage data presented unique challenges, with assembly-based approaches often losing information compared to direct read mapping.
- Some tools (such as KmerResistance, ARDaP, and ROCker) were reported to better handle complex or low-abundance scenarios.
- The presence of background microbiota and sample complexity affected detection accuracy, with some tools more prone to false positives.

Implementation Considerations

Tool Category	Key Strengths	Major Limitations	Best Use Cases
k-mer based (KARGA, KmerResistance)	High speed, robust to genome rearrangements, good recall	May miss novel or rare variants, database-dependent	Large-scale metagenomics, rapid screening
Alignment-based (AMRPlusPlus, ResFinder)	Well-established, interpretable	Slower, may lose information in assembly, database-dependent	Whole genome sequencing of isolates, clinical diagnostics
Machine learning (AMR-meta, DeepARG)	Improved sensitivity and specificity, handles negative results	Requires training data, may not generalize	Metagenomics, complex resistomes
Custom/Species-specific (ARDaP)	High accuracy for target species, detects chromosomal variants	Not generalizable, requires curation	Pathogen-specific surveillance
Assembly-based (metaSPAdes, MEGAHIT, Trinity)	Contextual information, taxonomic assignment	Fragmentation, underestimation, resource-intensive	Contextualization, taxonomic origin studies
ROCker models	High precision and recall for target genes	Requires model building, limited to modeled genes	Beta-lactamase surveillance, targeted detection
Oxford Nanopore Technologies/Long-read (MegaPath-Nano)	Handles long reads, drug-level prediction	Error-prone reads, tool maturity	Rapid clinical metagenomics

- Seven distinct tool categories for antimicrobial resistance detection were identified, each represented by one or more example tools.
- Key strengths were unique to each tool category, with no single strength reported for more than one category. The most common strengths included high speed, improved sensitivity and specificity, high accuracy for target species, and contextual information.
- Major limitations were also mostly unique, except for "database-dependence," which was reported for both k-mer based and alignment-based tools. All other limitations were reported for only one category each.

- Best use cases were specific to each tool category, with no overlap between categories. No tool category was recommended for more than one primary use case.
- We did not find missing information for any tool category in the table.

References

A. Cooper, Andrew J Low, Adam Koziol, Matthew C. Thomas, D. Leclair, S. Tamber, A. Wong, B. Blais, and C. Carrillo. "Systematic Evaluation of Whole Genome Sequence-Based Predictions of Salmonella Serotype and Antimicrobial Resistance." *Frontiers in Microbiology*, 2020.

A. Cooper, Andrew Low, Alex Wong, Sandeep Tamber, B. Blais, and C. Carrillo. "Modeling the Limits of Detection for Antimicrobial Resistance Genes in Agri-Food Samples: A Comparative Analysis of Bioinformatics Tools." *BMC Microbiology*, 2024.

Anna Abramova, Antti Karkman, and J. Bengtsson-Palme. "Metagenomic Assemblies Tend to Break Around Antibiotic Resistance Genes." *bioRxiv*, 2023.

"Benchmarking Software to Predict Antibiotic Resistance Phenotypes in Shotgun Metagenomes Using Simulated Data," 2022.

Camille McCall, and I. Xagoraraki. "Comparative Study of Sequence Aligners for Detecting Antibiotic Resistance in Bacterial Metagenomes." *Letters in Applied Microbiology*, 2018.

Danielle E. Madden, Jessica R. Webb, E. Steinig, B. Currie, E. Price, and D. Sarovich. "Taking the Next-Gen Step: Comprehensive Antimicrobial Resistance Detection from Burkholderia Pseudomallei." *EBioMedicine*, 2020.

———. "Taking the Next-Gen Step: Comprehensive Antimicrobial Resistance Detection from Burkholderia Pseudomallei Genomes." *bioRxiv*, 2019.

Danielle E. Madden, Timothy Baird, Scott C. Bell, Kate L. McCarthy, Erin P. Price, and D. Sarovich. "Keeping up with the Pathogens: Improved Antimicrobial Resistance Detection and Prediction from Pseudomonas Aeruginosa Genomes." *Genome Medicine*, 2024.

Danilo J P G Rocha, Carolina S. Silva, Hendor N. R. Jesus, Felipe G. Sacoda, João V. O. Cruz, Carina S. Pinheiro, Eric R.G.R. Aguiar, et al. "Suboptimal Bioinformatic Predictions of Antimicrobial Resistance from Whole-Genome Sequences in Multidrug-Resistant Corynebacterium Isolates." *Journal of Global Antimicrobial Resistance*, 2024.

E. Wissel, B. Talbot, Noriko A. B. Toyosato, R. Petit, V. Hertzberg, A. Dunlop, and T. Read. "hAMRoaster: A Tool for Comparing Performance of AMR Gene Detection Software." *bioRxiv*, 2022.

J. Nunez-Garcia, M. AbuOun, N. Storey, M. Brouwer, J. Delgado-Blas, S. S. Mo, N. Ellaby, et al. "Harmonisation of in-Silico Next-Generation Sequencing Based Methods for Diagnostics and Surveillance." *Scientific Reports*, 2022.

L. Carroll, A. Buehler, A. Gaballa, J. Siler, K. Cummings, R. Cheng, and M. Wiedmann. "Monitoring the Antimicrobial Resistance Dynamics of Salmonella Enterica in Healthy Dairy Cattle Populations at the Individual Farm Level Using Whole-Genome Sequencing." *bioRxiv*, 2021.

Lisa M Hodges, Eduardo N. Taboada, Adam Koziol, Steven Mutschall, B. Blais, G. Inglis, D. Leclair, and C. Carrillo. "Systematic Evaluation of Whole-Genome Sequencing Based Prediction of Antimicrobial Resistance in Campylobacter Jejuni and C. Coli." *Frontiers in Microbiology*, 2021.

M. Pesesky, T. Hussain, Meghan A. Wallace, Sanket Patel, S. Andleeb, C. Burnham, and G. Dantas. "Evaluation of Machine Learning and Rules-Based Approaches for Predicting Antimicrobial Resistance Profiles in Gram-Negative Bacilli from Whole Genome Sequence Data." *Frontiers in Microbiology*, 2016.

M. Prosperi, and S. Marini. "KARGA: Multi-Platform Toolkit for k-Mer-Based Antibiotic Resistance Gene Analysis of High-Throughput Sequencing Data." *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, 2021.

"OUP Accepted Manuscript." *Briefings in Bioinformatics*, 2022.

P. Clausen, E. Zankari, F. Aarestrup, and O. Lund. "Benchmarking of Methods for Identification of Antimicrobial Resistance Genes in Bacterial Whole Genome Data." *Journal of Antimicrobial Chemotherapy*, 2016.

P. Mahé, Meriem El Azami, P. Barlas, and Maud Tournoud. "A Large Scale Evaluation of TBProfiler and Mykrobe for Antibiotic Resistance Prediction in Mycobacterium Tuberculosis." *PeerJ*, 2019.

R. Doyle, D. O'Sullivan, Sean Aller, Sebastian Bruchmann, T. Clark, A. Pelegrin, M. Cormican, et al. "Discordant Bioinformatic Predictions of Antimicrobial Resistance from Whole-Genome Sequencing Data of Bacterial Isolates: An Inter-Laboratory Study." *bioRxiv*, 2019.

S. Marini, Marco Oliva, Ilya B Slizovskiy, Rishabh A Das, N. Noyes, Tamer Kahveci, C. Boucher, and M. Prosperi. "AMR-Meta: A k-Mer and Metafeature Approach to Classify Antimicrobial Resistance from High-Throughput Short-Read Metagenomics Data." *GigaScience*, 2022.

S. Marini, Rodrigo A. Mora, C. Boucher, N. Noyes, and M. Prosperi. "Towards Routine Employment of Computational Tools for Antimicrobial Resistance Determination via High-Throughput Sequencing." *bioRxiv*, 2021.

Sirui Zhang, Brittany Suttner, Luis M. Rodriguez-R, Luis H. Orellana, R. Conrad, Fang Liu, J. Rowell, et al. "ROCker Models for Reliable Detection and Typing of Short-Read Sequences Carrying -Lactamase Genes." *mSystems*, 2022.

Si-Yu Zhang, Brittany Suttner, Luis Rodriguez-R, Luis Orellana, Jessica Rowell, Hattie Webb, A. Williams-Newkirk, Andrew D. Huang, and Konstantinos T. Konstantinidis. "Rocker Models for Reliable Detection and Typing of Short Read Sequences Carrying -Lactamases," 2020.

T. J. Davies, J. Swann, A. Sheppard, H. Pickford, S. Lipworth, M. AbuOun, M. Ellington, et al. "Discordance Between Different Bioinformatic Methods for Identifying Resistance Genes from Short-Read Genomic Data, with a Focus on Escherichia Coli." *bioRxiv*, 2021.

W. Lui, A. W. Leung, Henry C. M. Leung, Yan Xin, J. L. Teng, P. Woo, T. Lam, and Ruibang Luo. "MegaPath-Nano: Accurate Compositional Analysis and Drug-Level Antimicrobial Resistance Detection Software for Oxford Nanopore Long-Read Metagenomics." *IEEE International Conference on Bioinformatics and Biomedicine*, 2020.