

SOFTWARE

Open Access



BenchAMRking: a Galaxy-based platform for illustrating the major issues associated with current antimicrobial resistance (AMR) gene prediction workflows

Nikolaos Strepis^{1†}, Dennis Dollee^{2†}, Donny Vrins², Kevin Vanneste³, Bert Bogaerts³, Catherine Carrillo⁴, Amrita Bharat⁵, Kristy Horan⁶, Norelle L. Sherry⁶, Torsten Seemann⁶, Benjamin P. Howden⁶, Saskia Hiltmann⁷, Leonid Chindelevitch^{8*}, Andrew P. Stubbs^{1†} and John P. Hays^{1*†}

Abstract

Background The Joint Programming Initiative on Antimicrobial Resistance (JPIAMR) networks ‘Seq4AMR’ and ‘B2B2B AMR Dx’ were established to promote collaboration between microbial whole genome sequencing (WGS) and antimicrobial resistance (AMR) stakeholders. A key topic discussed was the frequent variability in results obtained between different microbial WGS-related AMR gene prediction workflows. Further, comparative benchmarking studies are difficult to perform due to differences in AMR gene prediction accuracy and a lack of agreement in the naming of AMR genes (semantic conformity) for the results obtained. To illustrate this problem, and as a capacity-building exercise to encourage stakeholder involvement, a comparative Galaxy-based BenchAMRking platform was developed and validated using datasets from bacterial species with PCR-verified AMR gene presence or absence information from abritAMR.

Results The Galaxy-based BenchAMRking platform (<https://erasmusmc-bioinformatics.github.io/benchAMRking/>) specifically focusses on the steps involved in identifying AMR genes from raw reads and sequence assemblies. The platform currently comprises four well-characterised and published workflows that have previously been used to identify AMR genes using WGS data from several different bacterial species. These four workflows, which include the ISO certified abritAMR workflow, make use of different computational tools (or tool versions), and interrogate different AMR gene sequence databases. By utilising their own data, users can investigate potential AMR gene-calling problems associated with their own in silico workflows/protocols, with a potential use case outlined in this publication.

[†]Nikolaos Strepis and Dennis Dollee contributed equally to this publication.

[†]Andrew P. Stubbs and John P. Hays contributed equally to this publication.

*Correspondence:
Leonid Chindelevitch
l.chindelevitch@imperial.ac.uk
John P. Hays
j.hays@erasmusmc.nl

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Conclusions BenchAMRking is a Galaxy-based comparison platform where users can access, visualise, and explore some of the major discrepancies associated with AMR gene prediction from microbial WGS data.

Keywords Antimicrobial resistance, Microbial whole genome sequencing, Benchmarking, Galaxy, Workflows

Background

Antimicrobial resistance (AMR) represents a current global pandemic that detrimentally affects hospitalized patients, community-based care, healthcare system economics and a variety of One Health ecosystems, including foodstuffs, (domestic) animal health and the environment [1]. Furthermore, AMR is facilitated by a variety of factors, including a lack of implementation of infection prevention protocols, inappropriate antibiotic use, the slow development of new (alternative) antimicrobials and the time-consuming detection of AMR phenotypes.

Until recently, the detection of AMR in microbial isolates was almost solely based on phenotypic testing, which tends to provide accurate mechanism-independent results that can be confidently used by clinicians in their antimicrobial prescribing decisions. However, techniques involving mass spectrometry and genotype-to-phenotype AMR gene prediction are gaining in importance [2]. For example, whole genome sequencing of bacteria is frequently incorporated into infectious epidemiology studies and infection prevention programs, as well as in the genotype-to-phenotype prediction of AMR. Although concordance between existing genotype-to-phenotype AMR prediction workflows is generally good, a successful implementation in the clinical setting requires global agreement on standardisation, quality control parameters and validation for genotype-to-phenotype prediction, which begins with the accurate identification of AMR genes from WGS data [3, 4]. This process includes two major steps: (1) the generation of lists of AMR genes from available sequence data and (2) the prediction of AMR phenotypes based on the lists of AMR genes. In this respect, the current number and variety of AMR gene prediction workflows, tools and tool versions is limiting the re-use of both the data and workflows that have previously been published. Therefore, the authors' aim is to provide AMR researchers with easy access to standardised and validated AMR gene prediction workflows, which they could use with confidence when predicting AMR genes in their own One Health ecosystems. The result is BenchAMRking, a reusable Galaxy-based platform for AMR detection workflows that can deliver curated data and ground truth results for use by end users that are not familiar with deploying or using command line applications. The BenchAMRking platform includes a set of Galaxy workflows based on previously published AMR analysis workflows using the associated data and ground truth results to validate these workflows within a single resource. Currently, BenchAMRking allows both

multi-species AMR gene prediction based on abritAMR [5], and species-specific AMR gene prediction originally used for *Escherichia coli* [6] and *Salmonella spp.* found in food [7] and human patients [8], respectively. The workflows represent the ground truth in a comprehensive output format, and their Galaxy versions are available from the Erasmus MC GitHub and Workflow hub. The use of Galaxy and Workflow hub ensures the sustainability, reproducibility and reusability of these tools and associated data, thereby helping mitigate against application obsolescence.

The analytical and interpretation-based problems associated with predicting AMR phenotypes from AMR gene-based data are not addressed by BenchAMRking, as this subject requires an additional level of complexity. Further, if the correct identification of AMR genes is challenging, then those challenges will also be likely to affect the downstream prediction of AMR phenotypes.

Implementation

Tools

We have integrated a diverse collection of four previously published AMR gene prediction workflows into Galaxy for comparative benchmarking via the BenchAMRking platform (Fig. 1; Tables 1 and 2). The platform can be found at <https://erasmusmc-bioinformatics.github.io/benchAMRking/>, including brief instructions on its use. The output of the BenchAMRking platform may be visualised using the R-based Confusion Matrix and Heatmap scripts available from the BenchAMRking website. The workflows included in the BenchAMRking platform enable non-bioinformatics-trained researchers to perform extensive genomics analysis using short read sequence data, without the need for any coding. All workflows and their dependencies are installed on Galaxy and are managed by the Bioconda framework for dependency management. BenchAMRking workflows and their dependencies are available from the Bioconda Conda channel. The Galaxy wrappers were developed in GitHub for testing and have been made available on the Galaxy ToolShed.

Workflows

We have integrated four published WGS-AMR genotype prediction workflows (WF1-WF4) into Galaxy (Table 3). These workflows utilise a variety of bioinformatics applications copied from the original publications in which they were defined i.e., references WF1 [5]; WF2 [6]; WF3 [7] and WF4 [8]. In this publication, the ground truth

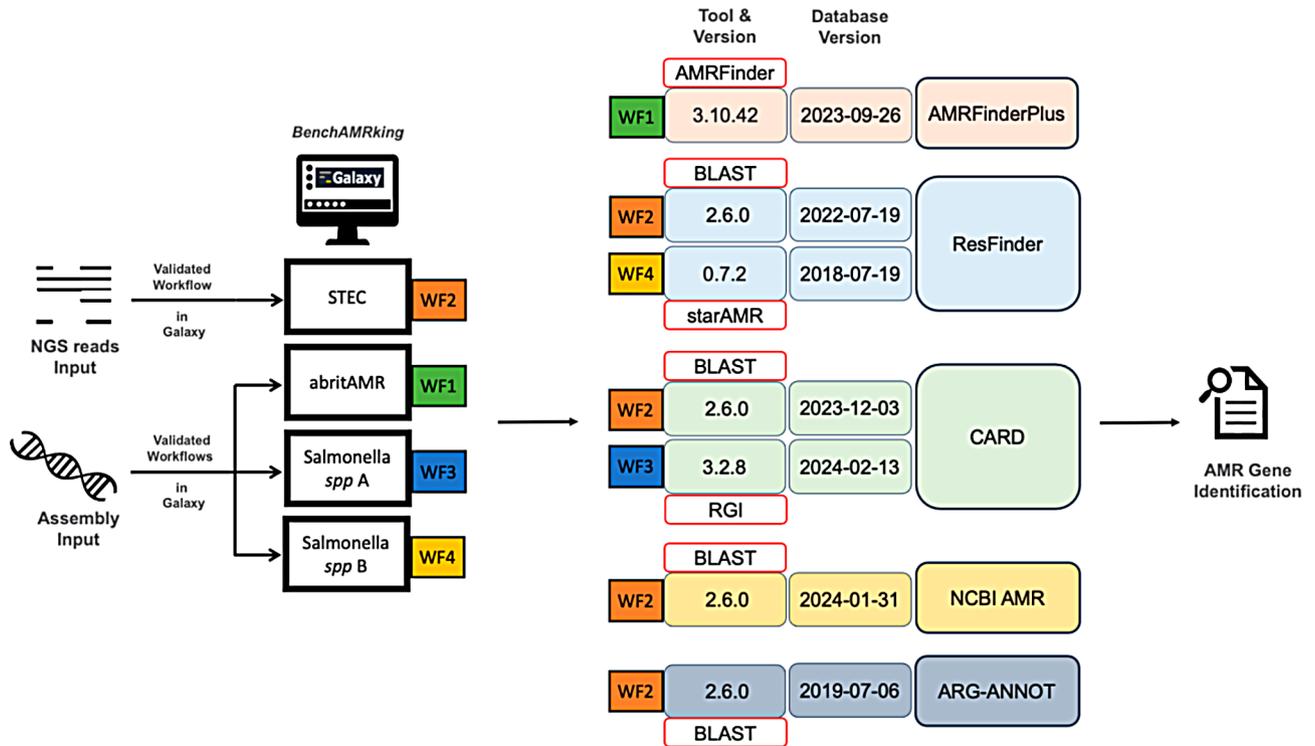


Fig. 1 Overview of BenchAMRking platform and workflows. Selected AMR gene prediction workflows (WFs) are translated into Galaxy workflows and stored in Workflow Hub. Researchers can load them into a Galaxy instance of their choice and either use the published data to reproduce the results or analyse their own data. Published *Salmonella* spp A WF3 (from broiler chickens) and published *Salmonella* spp B WF4 (from human infections) represent different workflows

Table 1 Version and licence information for the different workflow tools used in the BenchAMRking platform

Tool	Version	Licence
abritAMR	1.0.14	Creative Commons Attribution 4.0 International
RGI	5.2.1	custom license - free academic, government, non-profit use
SeqSero2	1.2.1	GNU GPL v2.0
BBTools	39.01	MIT License
SRST2	0.2.0	BSD License
hamronize	1.0.3	GNU LGPL v3.0
SPAdes	v3.15.5	GNU GPL v2.0
SKESA	3.0.0	Public Domain
pilon	1.1.0	GNU GPL v2.0
sistr	1.1.1	Apache-2.0 license
MOB-Recon	3.0.3	Apache-2.0 license
Shovill	1.0.4	GPL-3.0 license
staramr	0.8.0	Apache-2.0 license

for accurate AMR gene prediction is taken as the results obtained from the ISO certified abritAMR workflow (WF1). The replicated workflow data and results are all accessible at the BenchAMRking website (<https://erasm.usmc-bioinformatics.github.io/benchAMRking/>) via WorkflowHub (<https://workflowhub.eu/>). We note that a systematic review of available workflows was not performed when choosing the workflows used in BenchAMRking.

Table 2 Database version information for the different workflow tools used in the BenchAMRking platform

Workflow	Database	Database version
WF1	AMRFinderPlus	2023-09-26
WF2	ResFinder	2022-07-19
WF2	CARD	2023-12-03
WF2	NCBI AMR	2024-01-31
WF2	ARG-ANNOT	2019-07-06
WF3	CARD	2024-02-13
WF4	ResFinder	2018-07-19

Table 3 Workflow availability

Workflow	WorkflowHub ID
WF1 : abritAMR Multi-species [5]	https://workflowhub.eu/workflows/634
WF2 : Sciensano <i>Escherichia coli</i> [6]	https://workflowhub.eu/workflows/644
WF3 : CFIA Food <i>Salmonella</i> spp. [7]	https://workflowhub.eu/workflows/407
WF4 : Staramr Human Health <i>Salmonella</i> spp. [8]	https://workflowhub.eu/workflows/470
WF: workflow	

Instead, a simple search of existing literature was made for workflows that met the criteria mentioned above and described below (WF1 and WF4). WF2 and WF3 are workflows that also meet these criteria and are currently in use by one or more partners of the 'Seq4AMR' and 'B2B2B AMR Dx' networks. Additionally, publications validated their workflows with clinical or surveillance isolates and the workflows contain most known AMR genes.

The current workflows are supported by publications that include validated datasets. The tool and versions used in this publication are shown in Fig. 1 and provided in a machine-readable format in Tables 1 and 2, respectively. Users should be aware that the results of the workflows might change when newer versions of tools or databases are implemented in future version of the original workflows (we used the versions listed in the relevant publications). Descriptions of the individual workflows are given below.

WF1: ISO abritAMR

An AMR detection and reporting workflow (certified to ISO standards in the originating laboratory), based on the AMRFinderPlus tool and further optimized for clinical use. AMR prediction is based on the AMRFinderPlus database, and reports customized for clinical and public health microbiology applications are generated with an enhanced database to classify AMR mechanisms, and reports filtered to contain the most relevant AMR mechanisms. An additional module provides inferred phenotype reports for *Salmonella* spp. An extensive validation dataset is provided including PCR data and synthetic genomic data across 42 species. The workflow was validated with 1,184 bacterial isolates (42 species) [5].

WF2: Sciensano

This workflow uses multiple tools that perform read trimming, genome assembly, contamination checks, quality control of reads, plasmid detection, sequence typing, serotype determination, virulence factors identification and AMR characterization (against the NCBI NDARO, ResFinder and PointFinder databases). The database used for AMR gene prediction are ResFinder, CARD, ARGannot and NDARO. The data for this workflow included 137 Shiga toxin-producing *E. coli* isolates from human

faeces and various food matrices that were tested with disc diffusion or PCR-based methods [6].

WF3: CFIA

This workflow is based on multiple tools that include quality control and read trimming, genome assembly, plasmid prediction and serotype prediction for *Salmonella* spp. genomes. The AMR prediction database is based on the CARD database. All results are subsequently standardised using the hAMRronization tool. This workflow was validated using phenotypic verification of AMR in isolates, which were performed using broth micro dilutions [7].

WF4: Staramr

This workflow is based on staramr, a tool for genotypic AMR prediction based on the Centre for Genomic Epidemiology's ResFinder, PointFinder, and PlasmidFinder databases as well as PubMLST databases. Validation of the workflow was based on AMR phenotypic broth micro dilution of 1,321 *Salmonella enterica* isolates from the Canadian Integrated Program for Antimicrobial Resistance Surveillance (CIPARS) [8].

Experiments

The output of each WF was generated in a tabular format. For the visualization of the WF results, the output was concatenated and grouped using Python scripts (available in the BenchAMRking Github (Github repository of Erasmus/donny)). R scripts to visualize the output data from the WFs - as shown in Figs. 2 and 3a and b - are also available in the BenchAMRking GitHub (<https://github.com/ErasmusMC-Bioinformatics/BenchAMRking-script>).

RO-Crate FAIR digital objects

RO-Crate, or Research Object Crate, is a format for storing research related files, datasets, and documents in a FAIR way [9]. For findability, RO-Crate contains metadata such as title, authors, date of creation, and other ID's relevant for findability. All data, resources, and metadata are contained within the RO-Crate, ensuring accessibility. Interoperability and reusability were achieved by using the JSON-LD format, which is widely supported by most (bio-)informatics systems. The RO-Crates for each workflow are located in the corresponding Workflow Hub.

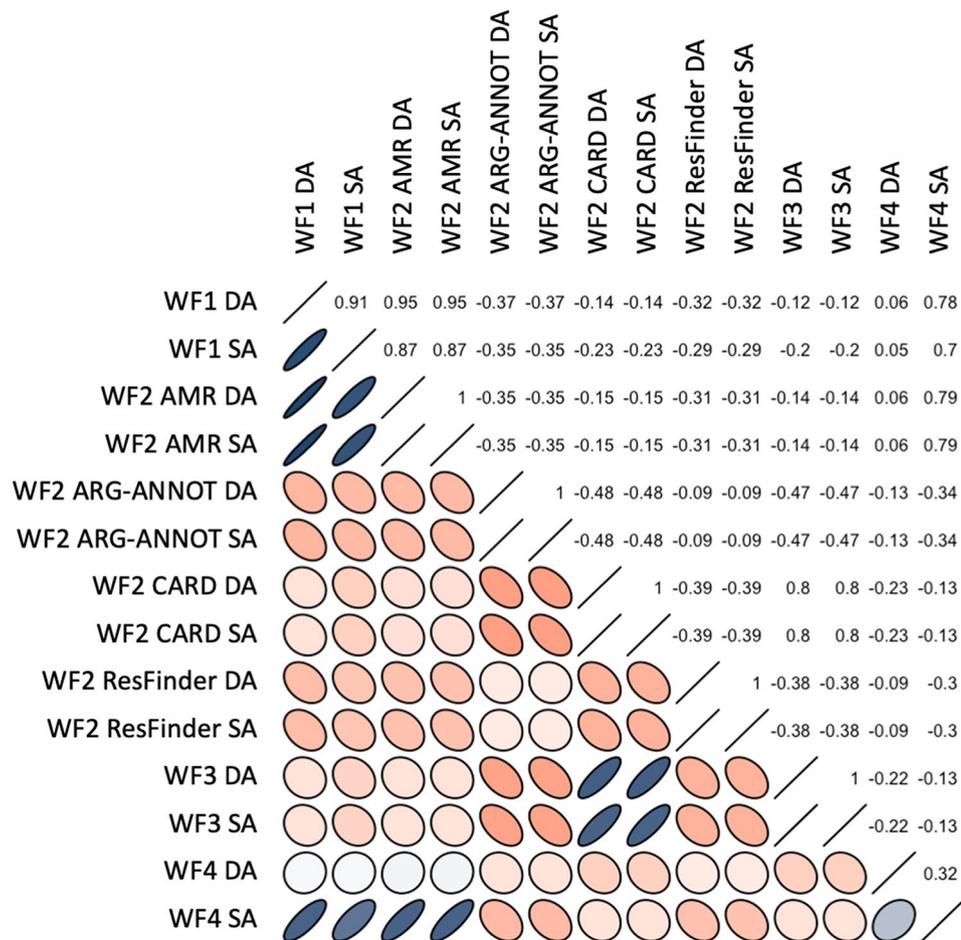


Fig. 2 Correlation matrix of AMR gene presence/absence vectors among different workflows included in BenchAMRking. WF1 - AbridAMR; WF2 - Sciensano; WF3 - CFIA; WF4 - Staram. Numbers on the top right indicate the correlation among workflows. Colour indicates a positive (red) or negative (blue) correlation, and shape indicates the strength of correlation. The more circular the shape, the stronger the correlation; the more oval the shape, the weaker the correlation. SA: same assembler; DA: different assembler (part of AMR identification and input of BenchAMRking). The supplemental data for the heatmaps are both the binary and identity excel files in the scripts repository at <https://github.com/ErasmusMC-Bioinformatics/BenchAMRking-scripts/tree/main>

Results

We have developed BenchAMRking (Fig. 1) to provide end-users and bioinformaticians with a suite of standardised AMR gene prediction workflows that have been replicated for use in the Galaxy environment. We have implemented four workflows: WF1 is an ISO certified AMR gene prediction workflow; WF2 – WF3 are examples of workflows developed by partners in the JPIMAR Seq4AMR and B2B2B networks, while WF4 is a well-characterised workflow for *Salmonella* spp. in human patients. All workflows were chosen to be representative of standardised AMR gene prediction analysis

methodologies for multiple pathogens and for single pathogenic species. Furthermore, all the selected workflows were demonstrated to function properly using validation data sets. In the following sections, we outline the tools incorporated into the Galaxy toolshed and the steps in these individual workflows (WF1-4). The user can access and use all workflows and retrieve all FASTQ files (both primary data and contigs). The underlying code in our GitHub repository is accessible from the BenchAMRking landing page (<https://erasmusmc-bioinformatics.github.io/benchAMRking/>).

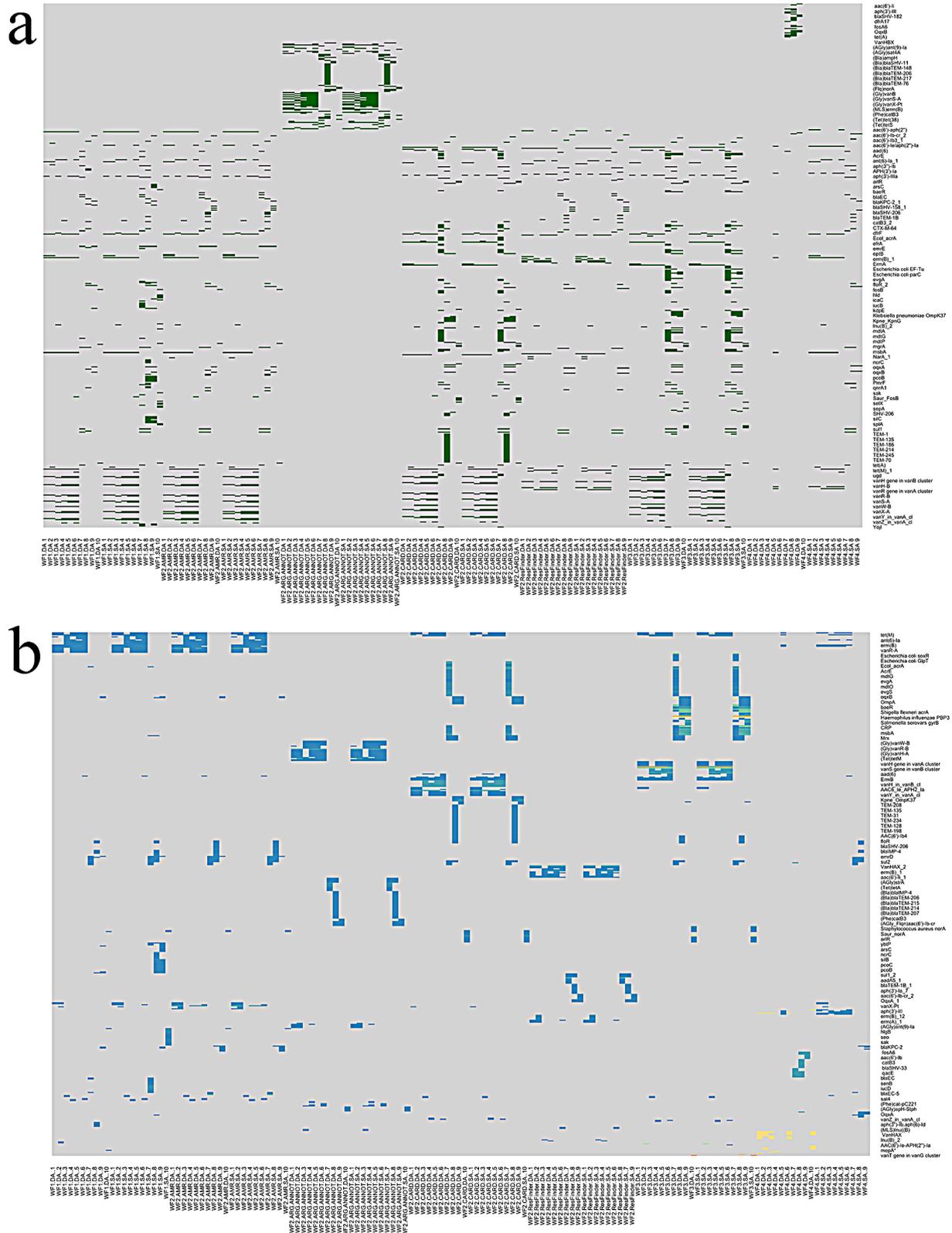


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 (a) Heatmap representation of the relationships of AMR genes detected in the workflows included in BenchAMRking. Green colour represents gene presence/absence. AMR genes are clustered based on identification by different workflows. SA: same assembler; DA: different assembler. WF – Workflow number SA: same assembler; DA: different assembler. Samples are numbered in the order shown in Table 4. The supplemental data for the heatmaps are both the Binary and Identity excel files in the scripts repository at <https://github.com/ErasmusMC-Bioinformatics/BenchAMRking-scripts/tree/main>. **(b)** Heatmap representation of the identity of AMR genes detected in the workflows included in BenchAMRking. Colours represent different values of AMR gene identity between the different workflows. SA: same assembler; DA: different assembler. Samples are numbered in the order shown in Table 4. The supplemental data for the heatmaps are both the Binary and Identity excel files in the scripts repository at <https://github.com/ErasmusMC-Bioinformatics/BenchAMRking-scripts/tree/main>

To illustrate the differences in AMR gene calling generated by the four different BenchAMRking workflows, a pilot study was performed using ten whole genome sequences obtained from abritAMR's validation dataset [5], in two experiments (see Table 5). Accessions for the whole genome sequences are shown in Table 4. The output of the WFs is generated in a tabular format. For the visualization of the results, the output was concatenated and grouped using

python scripts, with R scripts being used to visualize the output data (see <https://github.com/ErasmusMC-Bioinformatics/BenchAMRking-scripts>). Comparison of the results from the two experiments indicated limited concordance in the prediction of the AMR genes between the four WFs (Figs. 2 and 3a and b, and 4). Many of the discrepancies obtained were associated with different names (spelling variants) between AMR genes and different databases.

Table 4 Isolates and their online location for the ten whole genome bacterial sequences used in the pilot comparison of four BenchAMRking workflows and obtained from abritAMR [5]

N, Sample ID	Species	Reads Accession link and ID	Project Accession link and ID
1. SRR13803595	<i>E. coli</i>	https://www.ncbi.nlm.nih.gov/sra/SRS5393392	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA565795
2. SRR13803509	<i>E. faecium</i>	https://www.ncbi.nlm.nih.gov/sra/SRS8336102	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA565795
3. SRR6980278	<i>E. faecium</i>	https://www.ncbi.nlm.nih.gov/sra/SRS3155711	https://www.ncbi.nlm.nih.gov/bioproject/PRJEB11390
4. SRR14673537	<i>E. faecium</i>	https://www.ncbi.nlm.nih.gov/sra/SRR14673537	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA856406
5. SRR14673481	<i>E. faecium</i>	https://www.ncbi.nlm.nih.gov/sra/SRS9084754	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA565795
6. SRR14673267	<i>E. faecium</i>	https://www.ncbi.nlm.nih.gov/sra/SRS9084968	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA565795
7. SRR10127028	<i>E. coli</i>	https://www.ncbi.nlm.nih.gov/sra/SRS5393392	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA565795
8. SRR9734617	<i>K. pneumoniae</i>	https://www.ncbi.nlm.nih.gov/sra/SRR9734617	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA529744
9. SRR15097991	<i>K. pneumoniae</i>	https://www.ncbi.nlm.nih.gov/sra/SRS9451555	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA529744
10. SRR13803536	<i>S. aureus</i>	https://www.ncbi.nlm.nih.gov/sra/SRS8336075	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA565795

Table 5 Experiments performed using BenchAMRking platform

EXPERIMENT 1

Aim - To check WF AMR predictions based on the use of the same assembler (SA)

Methodology - WF2 Assemblies were used as an input of the four WF using ten samples (Table 4)

Tools - Spades v3.15.5 / BenchAMRking

Samples – 10 samples from abritAMR (Table 4)

EXPERIMENT 2

Aim - To check WF AMR predictions based on using a different assembler (DA)

Methodology - WF2 and shovill assembler for 10 samples (Table 4)

Tools - Spades v3.15.5 / shovill v1.0.4 / BenchAMRking

Samples – 10 samples from abritAMR (Table 4)

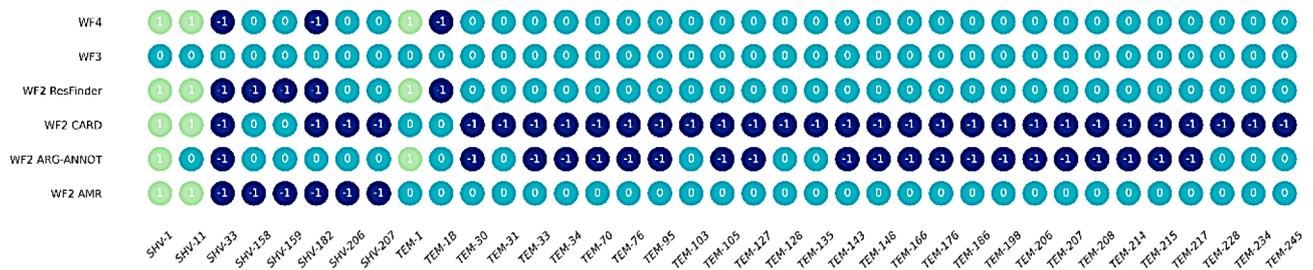


Fig. 4 A comparison of the results obtained by WF1 (abritAMR) with those of WF2-4 via BenchAMRking. The AMR genes identified by both WF1 and WF2-4 are shown in light blue; those identified only by WF1 are shown in dark blue; those identified only by WF2-4 are shown in green

Discussion and conclusions

BenchAMRking delivers easy and FAIR (Findable, Accessible, Interoperable, Reusable) access to both input data and standardised state-of-the-art AMR gene prediction workflows (WF1-4) in Galaxy. The workflows may be used in research and in diagnostic microbiology laboratories. Whilst BenchAMRking workflows are designed to be executed on Galaxy, the use of the Workflow Hub to generate a RO-Crate for each workflow ensures that they can also be executed in non-Galaxy based workflow applications. Thus, BenchAMRking goes beyond delivering reproducible workflow results that are comparable and available to the broader research community and the clinical field. The platform can contribute to the epidemiology and treatment of global AMR using WGS. As BenchAMRking is an open source and freely available platform, the authors hope that collaborations with interested colleagues will facilitate additional workflows and adaptations of the code and content beyond its current version. Our aim is to help democratize and promote a more comprehensive, standardised, and validated series of bioinformatics workflows for AMR gene prediction to help combat the current AMR pandemic. Finally, BenchAMRking is a tool whose feasibility is shown and described in this publication, examining over 500 AMR genes within 10 samples and 20 assemblies. More extensive studies using a broader and deeper range of international sequence data are currently being performed. Additional contributions and suggestions from international stakeholders interested in AMR, bioinformatics, workflow development and policy are welcome with the final goal of generating internationally agreed standards for gene sequence to AMR phenotype prediction workflows.

Project link and requirements

- Project name: BenchAMRking.
- Project home page: <https://erasmusmc-bioinformatics.github.io/benchAMRking/>.
- Operating system: Windows and Linux.
- Programming Language: Python and R.
- License: GNU GPL.

Abbreviations

AMR	Antimicrobial resistance
B2B2B AMRDx	Bench to Bedside to Business and Beyond: innovative solutions for AMR diagnostics
ISO	International Organisation for Standardisation
JPIAMR	Joint Programming Initiative on Antimicrobial Resistance
NGS	Next Generation Sequencing
Seq4AMR	JPIAMR Network for Integrating Microbial Sequencing and Applications for Antimicrobial Resistance
WF	Workflows
WGS	Whole genome sequencing
FAIR	Findable, Accessible, Interoperable, Reusable
RO-Crate	Research Object Crate

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-024-11158-5>.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Acknowledgements

The authors would like to acknowledge the expert discussion and guidance of the Seq4AMR and B2B2B AMRDx JPIAMR networks in helping generate the background for the current publication. The authors would like to also acknowledge Gary van Domselaar of the National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, MB R3E 3R2, Canada.

Author contributions

SH, DD, and NS implemented the tools and the workflows and carried out the analysis. AS and JH designed the study and supervised the development of the tools and analysis. AS, JH, NS and DD conceptualized the methodology and wrote the manuscript. NS, DD, DV, KV, BB, CC, AB, KH, NS, TS, BH, SH, LC, AS and JH read, edited, and approved the final version of the manuscript.

Funding

This work was made possible and supported by a collaboration between two Joint Programming Initiative on Antimicrobial Resistance (JPIAMR) networks: the Network for Integrating Microbial Sequencing and Platforms for AMR (Seq4AMR), funded via a Network Plus 2020 grant (ZonMW 549010001), and the Bench, Bedside, Business and Beyond: innovative solutions for AMR diagnostics network (B2B2B AMRDx), funded via a 2022 Diagnostics and Surveillance Networks grant (MRC MR/X036936/1). LC acknowledges funding from the MRC Centre for Global Infectious Disease Analysis (reference MR/X020258/1), funded by the UK Medical Research Council (MRC). This UK funded award is carried out in the frame of the Global Health EDCTP3 Joint Undertaking.

Data availability

The datasets generated and/or analysed during the current study are available in the following repositories. 1) For the pilot comparison study of the four BenchAMRking workflows we used ten whole genome bacterial sequences obtained from abritAMR, and their accession details are listed in Table 4) All

Galaxy wrappers developed are available for installation from the Galaxy Tool Shed (<https://toolshed.g2.bx.psu.edu/>). The workflows described in this publication are publicly available from the European Galaxy server, including published Galaxy histories (Table 3). All scripts used to generate the figures are available on GitHub (<https://github.com/ErasmusMC-Bioinformatics/BenchAMRking-scripts>).

Declarations

Ethical approval

No ethical approval was required for this publication.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Medical Microbiology and Infectious Diseases, Erasmus University Medical Centre (Erasmus MC), Rotterdam, The Netherlands

²Department of Pathology and Clinical Bioinformatics, Erasmus University Medical Centre (Erasmus MC), Rotterdam, The Netherlands

³Transversal activities in Applied Genomics, Sciensano, Brussels, Belgium

⁴Canadian Food Inspection Agency, Ottawa, ON K1A 0Y9, Canada

⁵National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, MB R3E 3R2, Canada

⁶Microbiological Diagnostic Unit Public Health Laboratory (MDU-PHL), Department of Microbiology & Immunology, University of Melbourne at the Peter Doherty Institute for Infection & Immunity, Melbourne, VIC, Australia

⁷Institute of Pharmaceutical Sciences, Faculty of Chemistry and Pharmacy, University of Freiburg, 79104 Freiburg, Germany

⁸MRC Centre for Global Infectious Disease Analysis, Imperial College London, London, England, UK

Received: 22 July 2024 / Accepted: 13 December 2024

Published online: 10 January 2025

References

- Mitchell J, O'Neill AJ, King R. Creating a framework to align antimicrobial resistance (AMR) research with the global guidance: a viewpoint. *J Antimicrob Chemother.* 2022;77(9):2315–20. <https://doi.org/10.1093/jac/dkac205>.
- Gajic I, Kabic J, Kekic D, Jovicevic M, Milenkovic M, Mitic Culafic D, et al. Antimicrobial Susceptibility Testing: A Comprehensive Review of Currently Used Methods. *Antibiot (Basel).* 2022;11(4):427. <https://doi.org/10.3390/antibiotics11040427>.
- Hazards EPB, Koutsoumanis K, Allende A, Alvarez-Ordenez A, Bolton D, Bover-Cid S, et al. Whole genome sequencing and metagenomics for outbreak investigation, source attribution and risk assessment of food-borne microorganisms. *EFSA J.* 2019;17(12):e05898. <https://doi.org/10.2903/j.efsa.2019.5898>.
- Petrillo M, Fabbri M, Kagkli DM, Querci M, Van den Eede G, Alm E, et al. A roadmap for the generation of benchmarking resources for antimicrobial resistance detection using next generation sequencing. *F1000Res.* 2021;10:80. <https://doi.org/10.12688/f1000research.39214.2>.
- Sherry NL, Horan KA, Ballard SA, Gonçalves da Silva A, Gorrie CL, Schultz MB, et al. An ISO-certified genomics workflow for identification and surveillance of antimicrobial resistance. *Nat Commun.* 2023;14(1):60. <https://doi.org/10.1038/s41467-022-35713-4>.
- Bogaerts B, Nouws S, Verhaegen B, Denayer S, Van Braekel J, Winand R, et al. Validation strategy of a bioinformatics whole genome sequencing workflow for Shiga toxin-producing *Escherichia coli* using a reference collection extensively characterized with conventional methods. *Microb Genomics.* 2021;7(3):1–21. <https://doi.org/10.1099/mgen.0.000531>.
- Cooper AL, Low AJ, Koziol AG, Thomas MC, Leclair D, Tamber S, et al. Systematic Evaluation of Whole Genome Sequence-Based Predictions of *Salmonella* Serotype and Antimicrobial Resistance. *Front Microbiol.* 2020;11:549. <https://doi.org/10.3389/fmicb.2020.00549>.
- Bharat A, Petkau A, Avery BP, Chen JC, Folster JP, Carson CA, et al. Correlation between Phenotypic and In Silico Detection of Antimicrobial Resistance in Canada Using Staramr. *Microorganisms.* 2022;10(2):292. <https://doi.org/10.3390/microorganisms10020292>.
- Carragáin EO, Goble C, Sefton P, Soiland-Reyes S. A lightweight approach to research object data packaging. 2019. <https://doi.org/10.5281/ZENODO.3250687>

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.