Application of hAMRonization Tool in Investigating Anti-microbial Resistance Genes to Generate Harmonized Reports

Katuwal N,[1-3] Tamrakar D,[1,3,4] Shrestha R[1-3,5]

[1]Center for Infectious Disease Research and Surveillance, Dhulikhel Hospital, Kathmandu University Hospital, Dhulikhel, Nepal.

[2]Molecular and Genome Sequencing Research Lab, Dhulikhel Hospital, Kathmandu University Hospital, Dhulikhel, Nepal.

[3]Research and Development Division, Dhulikhel Hospital, Kathmandu University Hospital, Dhulikhel, Nepal.

[4]Department of Community Medicine, Kathmandu University School of Medical Sciences, Dhulikhel, Nepal.

[5]Department of Pharmacology, Kathmandu University School of Medical Sciences, Dhulikhel, Nepal.

Corresponding Author

Rajeev Shrestha

Center for Infectious Disease Research and Surveillance,

Research and Development Division,

Department of Pharmacology,

School of Medical Sciences, Kathmandu University,

Dhulikhel Hospital, Kathmandu University Hospital,

Dhulikhel, Kavre, Nepal.

E-mail: rajeev.shrestha@kusms.edu.np

Citation

Katuwal N, Tamrakar D, Shrestha R. Application of hAMRonization Tool in Investigating Anti-microbial Resistance Genes to Generate Harmonized Reports. *Kathmandu Univ Med J.* 2023;83(3):320-4.

ABSTRACT

Background

Antimicrobial resistance has become a serious threat to global public health, with the potential to cause significant morbidity and mortality. Therefore, the detection and surveillance of antimicrobial resistance genes are critical to monitor the prevalence of resistance and to guide appropriate treatment decisions. Several bioinformatics tools have been developed to facilitate the detection and investigation of antimicrobial resistance genes from sequencing data. However, these tools often require a significant amount of manual curation and provide results which can limit their intra-operability, accessibility and usability. This lack of standardization in the reporting of antimicrobial resistance genes investigation greatly hinders the comparison of results across the public health sector.

Method

This study focuses on the application of hAMRonize tool developed by Public Health Alliance for Genomic Epidemiology. In this study, we have implemented hAMRonization to evaluate antimicrobial resistance genes in Salmonella and compared to the results from other tools (ARIBA, RGI, amrfinderplus and abricate).

Result

We observed that hAMRonize was more comprehensive than other tools in predicting genes, including information on antimicrobials, drug classes, and start and end base pair of genes, among others. The hAMRonize tool also provided results in an interactive Hypertext Markup Language format.

Conclusion

hAMRonization is an innovative tool for investigating and sharing antimicrobial resistance data in the current field of research, where scientists use a plethora of prediction tools that provide different results and subsequent interpretations.

KEY WORDS

Antimicrobial resistance, Gene-prediction, hAMRonize

INTRODUCTION

Antimicrobial resistance (AMR) has become a serious threat to global public health, with the potential to cause significant morbidity and mortality.[1-3] The misuse and overuse of antibiotics in both humans and animals have led to the emergence and spread of antibiotic-resistant bacteria.[4] Therefore, the detection and surveillance of AMR genes are critical to monitor the prevalence of resistance and to guide appropriate treatment decisions. High-throughput sequencing technologies, such as whole-genome sequencing (WGS) and metagenomics, have revolutionized the field of microbiology and made it possible to detect and identify AMR genes in complex microbial communities.[5,6] However, the analysis of these genome data can be challenging due to massive sequencing data sets.

Several tools (ResFinder, ARG-ANNOT and AMRfinder) have already been developed to developed to facilitate the detection and investigation of AMR genes from sequencing data and then, integrate the results of multiple AMR prediction tools, including hAMRonization.[5-10] The data generated from such tools have the potential to guide antibiotic treatment decisions and patient therapy in clinical cases of disease.[11] However, these tools often require a significant amount of manual curation and provide results which can limit their intra-operability, accessibility and usability. Additionally, these tools use specific databases and are marred by limited coverage of AMR genes and lack of regular updates in current detection tools can lead to false-negative results.[12,13] There is also a lack of standardization in the methods and tools used for the detection and annotation of AMR genes.[14] This is because the platforms differ in supported inputs, search algorithm, parametrization, underlying reference databases and output formats. This lack of standardization in the reporting of AMR gene detection greatly hinders the comparison of results across the public health sector. Thus, the myriad of options available for this purpose highlights a grave interoperability problem.

This study focuses on the application of hAMRonize tool, which is a standardized data specification to improve AMR data harmonization and interoperability, developed by Public Health Alliance for Genomic Epidemiology (PHA4GE).[9,15] As the unified global platforms require a common ground for the comparison of results from different tools, we have implemented hAMRonize to evaluate AMR genes in Salmonella Typhi and in comparison, to other AMR-predicting bioinformatics tools.[16]

METHODS

In this project, we evaluated the Salmonella whole genomes from European Nucleotide Archive (ENA). The selected genomes were submitted by Surveillance for Enteric Fever in Asia Project (SEAP) from clinical isolates of Nepal (2016-2019). The AMR associated genes were evaluated through Antimicrobial Resistance Identifier by Assembly (ARIBA) version 2.10.0.[16]

hAMRonization package

hAMRonization is conda installable and able to compile the output of many AMR tools into a unified format. Inside the hAMRonization package, biopython-compatible parser and command-line utility automatically transform reports from 14 different species-agnostic AMR gene detection tools into "hAMRonization" compatible reports. Also, validation and programmatic use of the specification is facilitated via the development of JSON and SALAD schemata. Additionally, the validation of the parsing utilities is ensured by unit testing.[9]

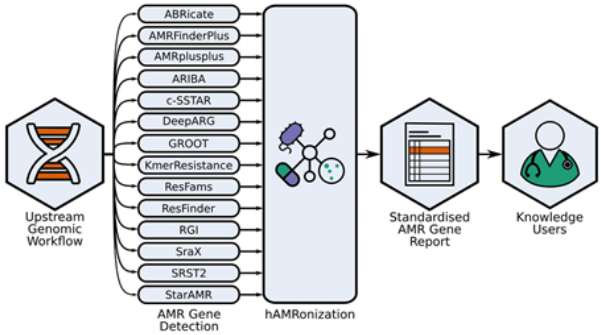

Figure 1. The hAMRonization package automates conversion to a standardized output, as shown by the bioinformatics pipeline.[9]

Installation of hAMRonization package

hAMRonization platform is implemented in a Biopython-compatible parser and command-line utility.[9] The hAMRonization package v1.1.1 (Release Date: 26 September 2022) was installed on Ubuntu 22.04.1 LTS, with prior installation of miniconda. This was followed by subsequent installation of several tools required for the package. This included amr (v3.10.45), fastp (v0.23.2), shovill (v1.1.0), ncbi-amrfinderplus (v3.10.45) packages. Before running the hAMRonize command, we first activated the 'amr' environment.

hAMRonization package workflow

a. Trim Adapter Reads from raw pair ended reads

fastp --in1 /path/Read1.fq.gz --in2 /path/Read2.fq.gz --out1 /path/Read1_trimmed.fq.gz –out2 /path/Read2_trimmed.fq.gz

b. Assemble the reads into contigs

shovill --R1 /path/Read1_trimmed.fq.gz --R2 /path/Read2_trimmed.fq.gz –outdir /path/assembly_samplename

c. Finding amr associated genes

To find the amr associated genes and generate output the result in tsv (excel) format, following command was used for amrfinderplus and abricate respectively:

• amrfinder --nucleotide /path/contigs.fa --output /path/amrfinderplus_results.tsv

• abricate /path/contigs.fa > /path/abricate_results.tsv

d. hAMRonizing the results

The hAMRonization for amrfinderplus and abricate was done separately as:

• hamronize amrfinderplus --analysis_software_version 3.10.16 --reference_database_version 2021-09-30.1 --input_file_name samplename /path/amrfinderplus_results.tsv > /path/hAMRonized_amr_report.tsv

•hamronize abricate /path/abricate_results.tsv --reference_database_version 3.2.5 --analysis_software_version 1.0.0 --format json > /path/hAMRonized_abricate_report.tsv

e. Summarising the result in tsv and interactive html format

The summary was first reported in .tsv format as:

• hamronize summarize -o /path/combined_report.tsv -t tsv /path/hAMRonized_abricate_report.tsv /path/hAMRonized_amr_report.tsv

for interactive html format:

• hamronize summarize --summary_type interactive /path/combined_report.tsv > /path/combined_report.html

Comparison of hAMRonize and other AMR detection tools

After successfully running the hAMRonize command, the output was compared to that of ARIBA (v2.10.0), Resistance Gene Identifier (RGI) (v6.0.1, CARD 3.2.6), amrfinderplus (v3.11.14) and abricate (v0.8.13).

RESULTS

For this study, five Salmonella genomes were randomly selected. After evaluation of the genomes by ARIBA, RGI, amrfinderplus, abricate and hAMRonize, information on AMR genes were observed (Table 1).

With evaluation from RGI, details including AMR Genes, gene family, drug class, resistance mechanism among others were observed, while from abricate details on AMR genes, gene position, resistant antibiotic, among others were observed. Similarly, from amrfinderplus, details on AMR genes, drug class, gene position, resistance method were observed, while hAMRonize tool combined all information and provided details on AMR gene, gene name, analysis software details, antimicrobial agent, drug class, gene location, protein length, resistance mechanism and others. As observed in table 1, hAMRonize tool also predicted higher number of antibiotic resistance genes than other tools.

Table 1. AMR genes shown by ARIBA, RGI, amrfinderplus, abricate and hAMRonize for selected Salmonella genomes.

Accession	ARIBA	RGI	Amrfinderplus	abricate	hAMRonize
ERR6294433	blaTEM-1, catA1, dfrA7, sul1, sul2	catI, TEM-1, sul1, qacEdelta1	CatA1, dfrA7, qacEdelta1, sul1, sul2, aph(3'')-Ib, aph(6)-Id, blaTEM-1	catA1, dfrA7, sul1, sul2, aph(3'')-Ib, aph(6)-Id, blaTEM-1	aph(3'')-Ib, aph(6)-Id, blaTEM-1, catA1, dfrA7, qacEdelta1, sul1, sul2
ERR6294437	blaTEM-1, catA1,dfrA7,sul1,sul2	catI, TEM-1, sul1, qacEdelta1	CatA1, dfrA7, qacEdelta1, sul1, sul2, aph(3'')-Ib, aph(6)-Id, blaTEM-1	catA1, dfrA7, sul1, sul2, aph(3'')-Ib, aph(6)-Id, blaTEM-1	aph(3'')-Ib, aph(6)-Id, blaTEM-1, catA1, dfrA7, qacEdelta1, sul1, sul2
ERR6294558	blaTEM-1, sul2	N/A	N/A	N/A	N/A
ERR6294572	blaTEM-1, catA1, dfrA7, sul1, sul2	catI, TEM-1, sul1, qacEdelta1	CatA1, dfrA7, qacEdelta1, sul1, sul2, aph(3'')-Ib, aph(6)-Id, blaTEM-1	catA1, dfrA7, sul1, sul2, aph(3'')-Ib, aph(6)-Id, blaTEM-1	aph(3'')-Ib, aph(6)-Id, blaTEM-1, catA1, dfrA7, qacEdelta1, sul1, sul2
ERR6294680	blaTEM-1, catA1, dfrA7, sul1, sul2	catI, TEM-1, sul1, qacEdelta1	CatA1, dfrA7, qacEdelta1, sul1, sul2, aph(3'')-Ib, aph(6)-Id, blaTEM-1	catA1, dfrA7, sul1, sul2, aph(3'')-Ib, aph(6)-Id, blaTEM-1	aph(3'')-Ib, aph(6)-Id, blaTEM-1, catA1, dfrA7, qacEdelta1, sul1, sul2

BlaTEM-1: broad-spectrum class A beta-lactamase TEM-1, catA1: type A-1 chloramphenicol O-acetyltransferase, dfrA7: trimethoprim-resistant dihydrofolate reductase DfrA7, sul1: sulfonamide-resistant dihydropteroate synthase Sul1, sul2: sulfonamide-resistant dihydropteroate synthase Sul2; aph(6)-Id: aminoglycoside O-phosphotransferase APH(6)-Id; aph(3'')-Ib : aminoglycoside O-phosphotransferase APH(3'')-Ib; qacEdelta1: quaternary ammonium compound efflux SMR transporter QacE delta 1

DISCUSSION

In this study, the utility of hAMRonize tool was compared to other widely used bioinformatics tools for predicting antimicrobial resistance genes (ARIBA, RGI, amrfinderplus and abricate).[8,17,18] In comparison, hAMRonize combined the information and provided more comprehensive data, including information on antimicrobials, drug classes, resistance mechanism, gene position, among others. hAMRonize disseminated results in a single consistent format, allowing not only the comparison of tools and databases (AMRFinderPlus, RGI, Resfinder, ARIBA), but the validation of results through multiple detection algorithms.[19]

There is a challenge in developing prediction software to generate standardized reports on AMR genes and to facilitate comparison of results across the public health sector.[14] In addition, the AMR associated data are also crucial for understanding the mechanism of resistance,

developing primers for detection, and devising intervention approaches. Furthermore, the classification of AMR genes by their corresponding drug class (es) is necessary to match them to known phenotypically resistant drug classes, which facilitates appropriate intervention strategies.[20] This is because researchers and epidemiologists often focus on a narrow range of AMR phenotypes in practice, making tool selection crucial, particularly when resistance to a specific drug class is of interest.[21] The hAMRonize tool subdued intricacy of such results by providing comprehensive information, combining the results of most widely used tools and also provided results in an interactive HTML format.[22] These depictions are consistent with previous studies that have highlighted the utility of hAMRonization in compiling the results of 12 different AMR tools.[22-24] Additionally, this tool had ability to integrate a large number of genomes using automatic workflows that are compatible with workflow managers such as Galaxy, Snakemake, and NextFlow.

One hurdle in translating genomic data to clinically actionable information for AMR is the use of different databases that classify genes differently. The hAMRonize tool addressed this issue by providing a drug class column and filling in values from the Chemical Entities of Biological Interest (ChEBI) ontology when available.[25] This standardized approach enables researchers and clinicians to easily compare and interpret AMR gene data, facilitating the development of targeted interventions for specific drug-resistant infections.[22,25,26]

The hAMRonize tool is designed as command line input therefore, there is need of utlising locally stored large-sized genomic data which always might not be the case. Therefore, there is an opportunity of modification of hAMRonize tool where genomic data can directly be accessed from databases (ENA, NCBI) through accession numbers.

CONCLUSION

hAMRonize is an innovative bioinformatics tool for investigating and sharing AMR data in the current field of research, where scientists use a plethora of AMR prediction tools that provide different results and subsequent interpretations. This standardized approach, of AMR description by hAMRonize, enables researchers and clinicians to easily compare and interpret AMR gene data, facilitating the development of targeted interventions for specific drug-resistant infections.

ACKNOWLEDGMENTS

We are thankful to the researchers associated PHA4GE for resources associated with hAMRonization tool and PHA4GE sub-grant for supporting this study. We are appreciative of Jason R. Andrews, Kesia Esther da Silva and the SEAP team for providing accessions numbers (European Nucleotide Archive) for the Salmonella isolates.

REFERENCES

1. O'Neill, J. Tackling drug-resistant infections globally: final report and recommendations. The review on antimicrobial resistance. The Review on Antimicrobial Resistance. 2016.
2. World Health Organization. Antimicrobial resistance: global report on surveillance 2014. World Health Organization; 2014.
3. World Health Organization. Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics. World Health Organization; 2017.
4. Van Boeckel TP, Brower C, Gilbert M, Grenfell BT, Levin SA, Robinson TP, et al. Global trends in antimicrobial use in food animals. Proc Natl Acad Sci USA. 2015;112(18):5649-54
5. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother*. 2012 Nov;67(11):2640-4. doi: 10.1093/jac/dks261. Epub 2012 Jul 10. PMID: 22782487; PMCID: PMC3468078.
6. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol. 2017 Sep 12;35(9):833-44. doi: 10.1038/nbt.3935. Erratum in: Nat Biotechnol. 2017 Dec 8;35(12):1211. PMID: 28898207.
7. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother.* 2014;58(1):212-20. doi: 10.1128/AAC.01310-13. Epub 2013 Oct 21. PMID: 24145532; PMCID: PMC3910750.
8. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, et al. Validating the AMRFinder Tool and Resistance Gene Database by Using Antimicrobial Resistance Genotype-Phenotype Correlations in a Collection of Isolates. *Antimicrob Agents Chemother*. 2019 Oct 22;63(11):e00483-19. doi: 10.1128/AAC.00483-19. Erratum in: Antimicrob Agents Chemother. 2020 Mar 24;64(4): PMID: 31427293; PMCID: PMC6811410.
9. PHA4GE. hAMRonization. Github; Accessed: 24 April 2023. Available at: https://github.com/pha4ge/hAMRonization
10. Sakagianni A, Koufopoulou C, Feretzakis G, Kalles D, Verykios VS, Myrianthefs P, et al. Using Machine Learning to Predict Antimicrobial Resistance-A Literature Review. *Antibiotics (Basel)*. 2023 Feb 24;12(3):452. doi: 10.3390/antibiotics12030452. PMID: 36978319; PMCID: PMC10044642.
11. Zhou H, Larkin PMK, Zhao D, Ma Q, Yao Y, Wu X, et al. Clinical Impact of Metagenomic Next-Generation Sequencing of Bronchoalveolar Lavage in the Diagnosis and Management of Pneumonia: A Multicenter Prospective Observational Study. *J Mol Diagn.* 2021 Oct;23(10):1259-68. doi: 10.1016/j.jmoldx.2021.06.007. Epub 2021 Jun 29. PMID: 34197923.
12. Zankari E, Hasman H, Kaas RS, Seyfarth AM, Agersø Y, Lund O, et al. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. *J Antimicrob Chemother.* 2013 Apr;68(4):771-7. doi: 10.1093/jac/dks496. Epub 2012 Dec 11. PMID: 23233485.

13. McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, et al. The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother.* 2013 Jul;57(7):3348-57. doi: 10.1128/AAC.00419-13. Epub 2013 May 6. PMID: 23650175; PMCID: PMC3697360.

14. Bioinformatics approaches to the study of antimicrobial resistance. *Rev Esp Quimioter.* 2021 Sep;34 Suppl 1(Suppl1):15-17. doi: 10.37201/req/s01.04.2021. Epub 2021 Sep 30. PMID: 34598416; PMCID: PMC8683011.

15. Maguire F. AMR Genomics: Reads to Reports. AMR Workshop. 2021

16. da Silva KE, Tanmoy AM, Pragasam AK, Iqbal J, Sajib MSI, Mutreja A, et al. The international and intercontinental spread and expansion of antimicrobial-resistant Salmonella Typhi: a genomic epidemiology study. Lancet Microbe. 2022 Aug;3(8):e567-e577. doi: 10.1016/S2666-5247(22)00093-3. Epub 2022 Jun 21. PMID: 35750070; PMCID: PMC9329132.

17. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR, Wlodarski MA, et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res.* 2023 Jan 6;51(D1):D690-D699. doi: 10.1093/nar/gkac920. PMID: 36263822; PMCID: PMC9825576.

18. Hunt M, Mather AE, Sánchez-Busó L, Page AJ, Parkhill J, Keane JA, et al. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. Microb Genom. 2017 Sep 4;3(10):e000131

19. Ines Mendes. Comparing and reporting AMR results using hAMRonization. AMR Workshop. 2021

20. Noman SM, Zeeshan M, Arshad J, Deressa Amentie M, Shafiq M, Yuan Y, et al. Machine Learning Techniques for Antimicrobial Resistance Prediction of Pseudomonas Aeruginosa from Whole Genome Sequence Data. Comput Intell Neurosci. 2023 Mar 1;2023:5236168. doi: 10.1155/2023/5236168. PMID: 36909968; PMCID: PMC9995192.

21. Sabat AJ, Budimir A, Nashev D, Sá-Leão R, van Dijl Jm, Laurent F, et al. ESCMID Study Group of Epidemiological Markers (ESGEM). Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Euro Surveill.* 2013 Jan 24;18(4):20380. doi: 10.2807/ese.18.04.20380-en. PMID: 23369389.

22. Raphenya AR, Robertson J, Jamin C, de Oliveira Martins L, Maguire F, McArthur AG, Hays JP. Datasets for benchmarking antimicrobial resistance genes in bacterial metagenomic and whole genome sequencing. *Sci Data*. 2022 Jun 15;9(1):341. doi: 10.1038/s41597-022-01463-7. PMID: 35705638; PMCID: PMC9200708.

23. Wissel EF, Talbot BM, Toyosato NAB, Petit III RA, Read TD. hAMRoaster: a tool for comparing performance of AMR gene detection software. *MedRxiv. bioRxiv* 2022.01.13.476279; doi: https://doi.org/10.1101/2022.01.13.476279

24. Ayorinde O Afolayan and others, Overcoming Data Bottlenecks in Genomic Pathogen Surveillance, Clinical Infectious Diseases. 2021; 73(supp4): pS267–S274

25. Hastings J, Owen G, Dekker A, Ennis M, Kale N, Muthukrishnan V, et al. ChEBI in 2016: Improved services and an expanding collection of metabolites. Nucleic Acids Res. 2016 Jan 4;44(D1):D1214-9. doi: 10.1093/nar/gkv1031. Epub 2015 Oct 13. PMID: 26467479; PMCID: PMC4702775.

26. Veloso M, Acosta J, Arros P, Pasten CB, Rojas R, Marcoleta AE. Population genomics, resistance, pathogenic potential, and mobile genetic elements 1 of carbapenem-resistant Klebsiella pneumoniae causing infections in Chile. *medRxiv.* 2022. doi: https://doi.org/10.1101/2022.11.28.517876