



# QIB: Data Science

Antimicrobial resistance  
genes' profilers for metagenome

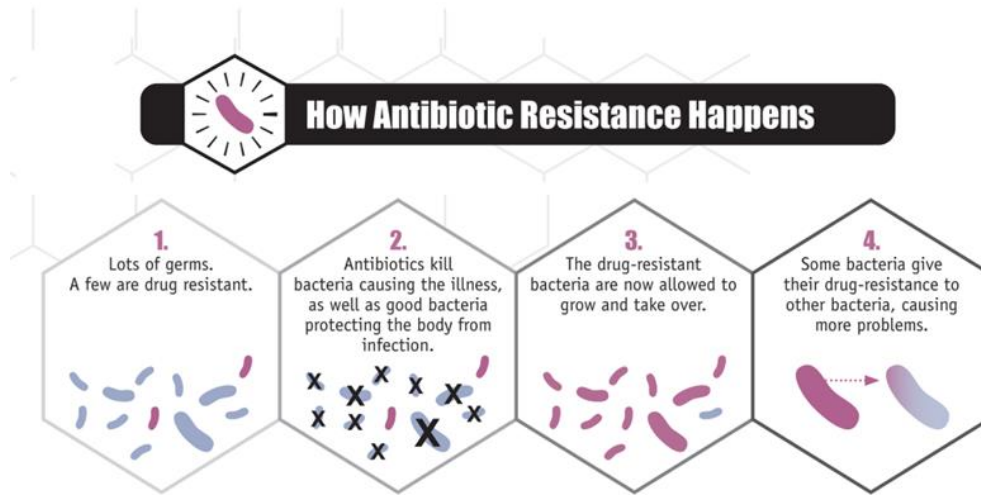
25<sup>th</sup> Feb 2025

Sumeet Tiwari

Thanks to Core-Binfie Team

---

# Introduction:



<https://asm.org/Articles/2024/October/Antimicrobial-Resistance-Pandemic-Breaking-Silence>

Type of drug resistance: Intrinsic and acquired

**hAMRoaster: a tool for comparing performance of AMR gene detection software**

Emily F. Wissel<sup>A</sup>, Brooke M. Talbot<sup>B</sup>, Noriko A. B. Toyosato<sup>C</sup>, Robert A Petit III<sup>D</sup>, Vicki Hertzberg<sup>A</sup>, Anne Dunlop<sup>E</sup>, Timothy D. Read<sup>D,F</sup>

Data Descriptor | [Open access](#) | Published: 15 June 2022

## Datasets for benchmarking antimicrobial resistance genes in bacterial metagenomic and whole genome sequencing

[Amogelang R. Raphenya](#), [James Robertson](#), [Casper Jamin](#), [Leonardo de Oliveira Martins](#), [Finlay Maguire](#), [Andrew G. McArthur](#) & [John P. Hays](#) ✉

[Scientific Data](#) **9**, Article number: 341 (2022) | [Cite this article](#)

**9415** Accesses | **14** Altmetric | [Metrics](#)

# Selection of tools:

## Tools:

- Ariba
- ARGprofiler
- AMR++
- DeepARG
- Groot
- KMA
- KARGA
- Kmerresistance
- MetaCerberus
- RGI
- SRST2
- Shortbred



## Criteria:

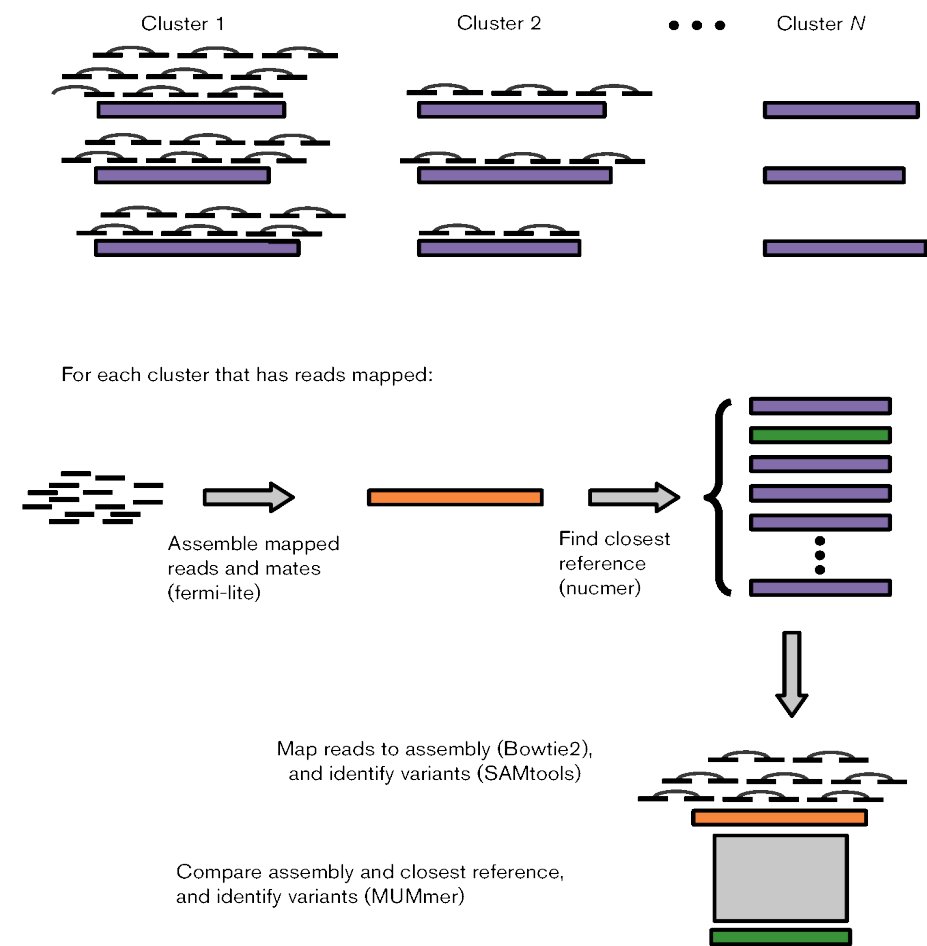
- FASTQ file as input
- Custom AMR gene database not protein.
- Executable via command line
- Only-local assembly is allowed.
- Open-source



## Final-tools:

- Ariba
- ARGprofiler
- Groot
- SRST2
- KARGA

# ARIBA



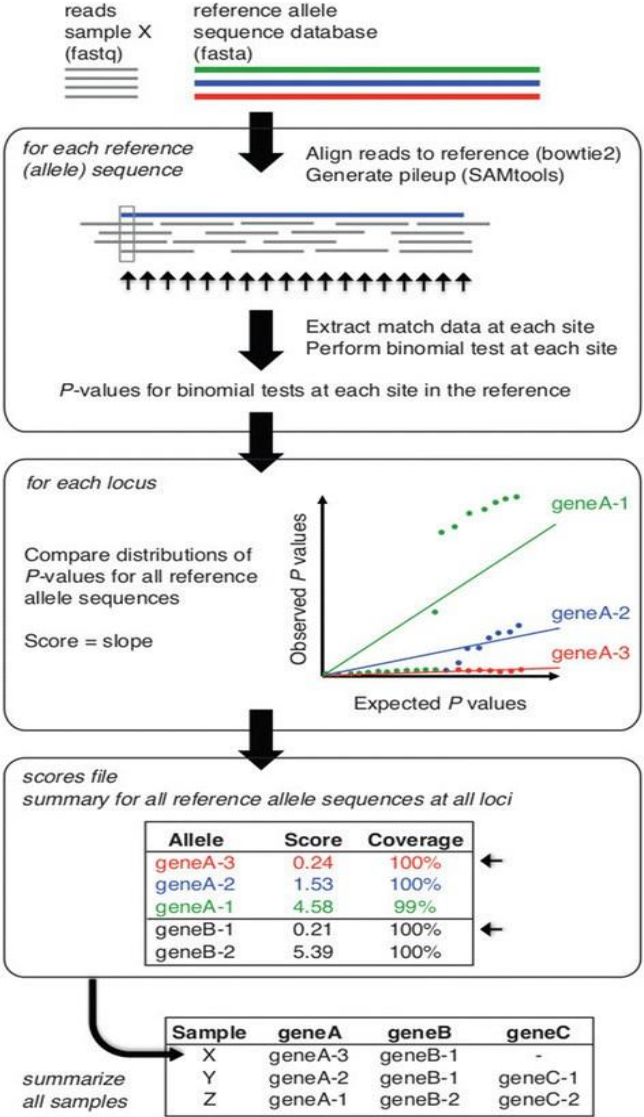
Research Article | Open Access

## ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads

<sup>1</sup> Martin Hunt<sup>1</sup>, Alison E Mather<sup>1,2</sup>, Leonor Sánchez-Busó<sup>1</sup>, Andrew J Page<sup>1</sup>, Julian Parkhill<sup>1</sup>, Jacqueline A Keane<sup>1</sup> and Simon R Harris<sup>1</sup>

 View Affiliations

# SRST2



# Groot

JOURNAL ARTICLE

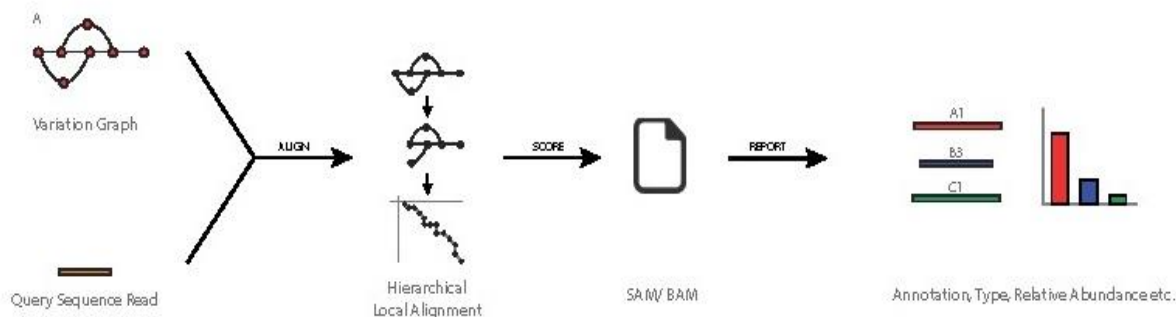
## Indexed variation graphs for efficient and accurate resistome profiling

Will P M Rowe , Martyn D Winn

*Bioinformatics*, Volume 34, Issue 21, November 2018, Pages 3601–3608,

<https://doi.org/10.1093/bioinformatics/bty387>

**Published:** 14 May 2018 **Article history** ▼



# ARGprofiler

JOURNAL ARTICLE

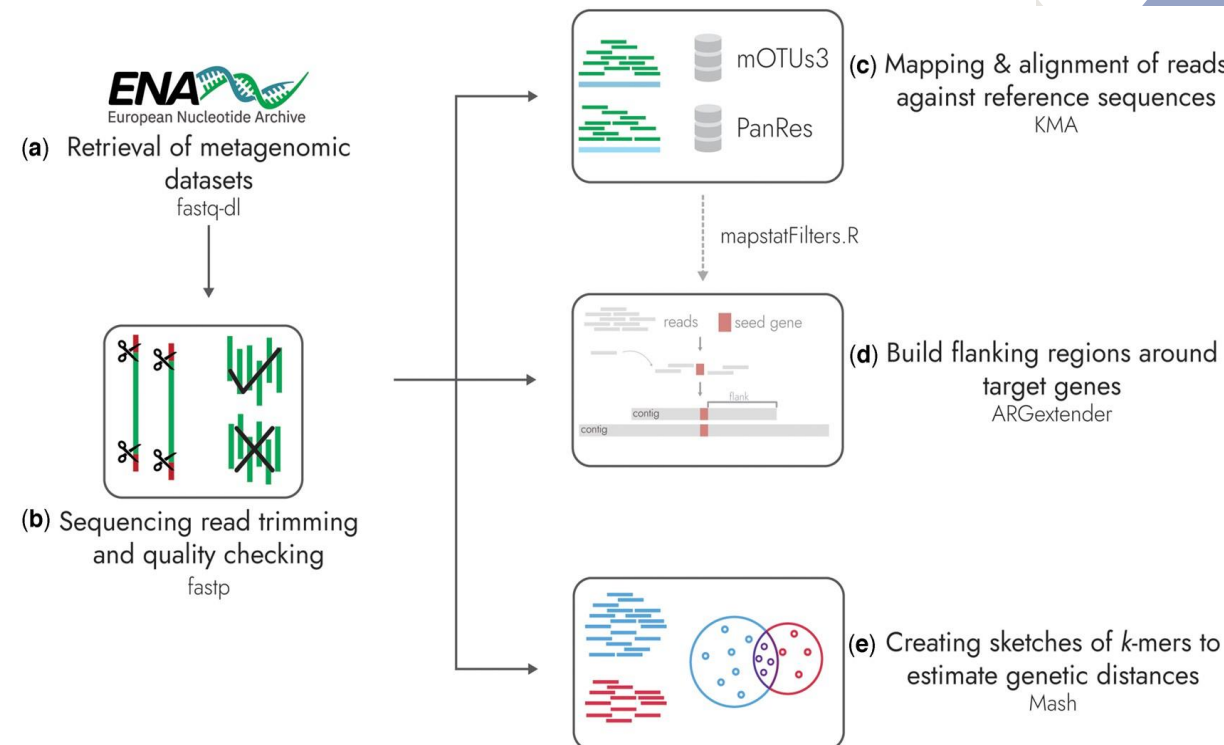
## ARGprofiler—a pipeline for large-scale analysis of antimicrobial resistance genes and their flanking regions in metagenomic datasets

Hannah-Marie Martiny , Nikiforos Pyrounakis, Thomas N Petersen, Oksana Lukjančenko, Frank M Aarestrup, Philip T L C Clausen, Patrick Munk

*Bioinformatics*, Volume 40, Issue 3, March 2024, btae086,

<https://doi.org/10.1093/bioinformatics/btae086>

**Published:** 20 February 2024 **Article history** ▼



# AMR databases:

AMRfinderPlus (v3.12)  
 ARG-ANNOT (v6, July 2019)  
 BacMet (v1.1)  
 CARD (v3.3.0)  
 MegaRes (v3.0)  
 ResFinderFG (v2.0)  
 ResFinder

Remove genes  
 required SNP  
 confirmation

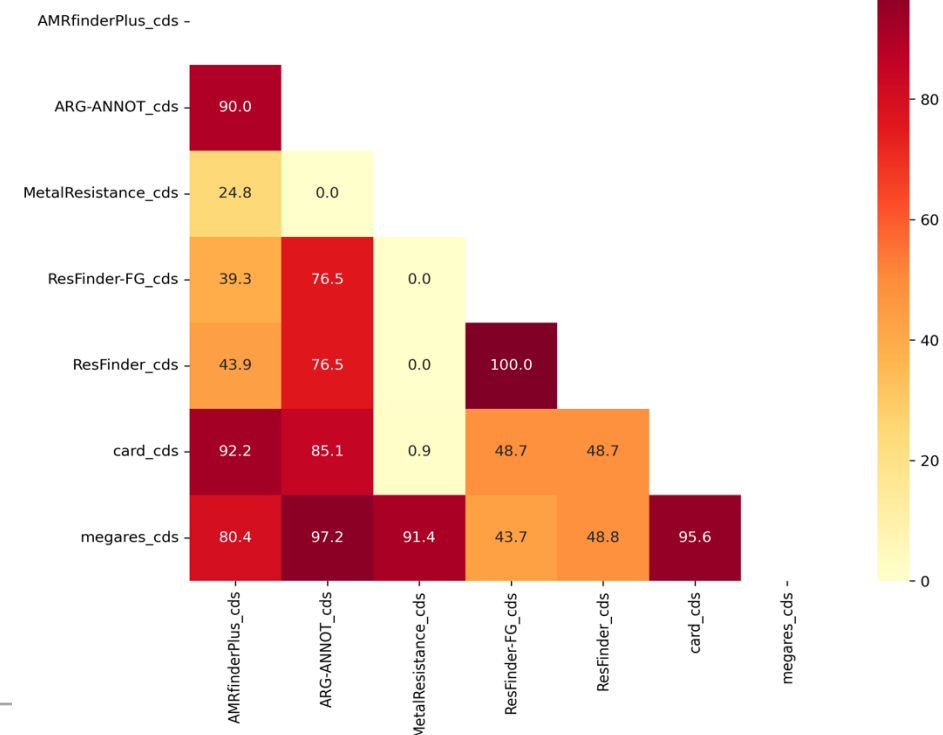
Remove genes  
 without start and  
 stop codons

De-replication and  
 clustering of genes (at  
 90% identity and  
 coverage)

panARG

- Hannah-Marie Martiny, Nikiforos Pyrounakis, Thomas N Petersen, Oksana Lukjančenko, Frank M Aarestrup, Philip T L C Clausen, Patrick Munk, ARGprofiler – a pipeline for large-scale analysis of antimicrobial resistance genes and their flanking regions in metagenomic datasets, Bioinformatics, Volume 40, Issue 3, March 2024, btae086, <https://doi.org/10.1093/bioinformatics/btae086>
- <https://github.com/sanger-pathogens/ariba/wiki/Task:-prepareref#task-prepareref>

Database Gene Sharing (Percentages)



# Preparation of Mock-Community:

Research | [Open access](#) | Published: 31 July 2021

## HumGut: a comprehensive human gut prokaryotic genomes collection filtered by metagenome data





[Pranvera Hiseni](#) , [Knut Rudi](#), [Robert C. Wilson](#), [Finn Terje Hegge](#) & [Lars Snipen](#)

[Microbiome](#) **9**, Article number: 165 (2021) | [Cite this article](#)

**14k** Accesses | **51** Altmetric | [Metrics](#)

381,779 genomes clustered at 97.5% identity resulted in 30,691 genomes.

## hAMRoaster: a tool for comparing performance of AMR gene detection software

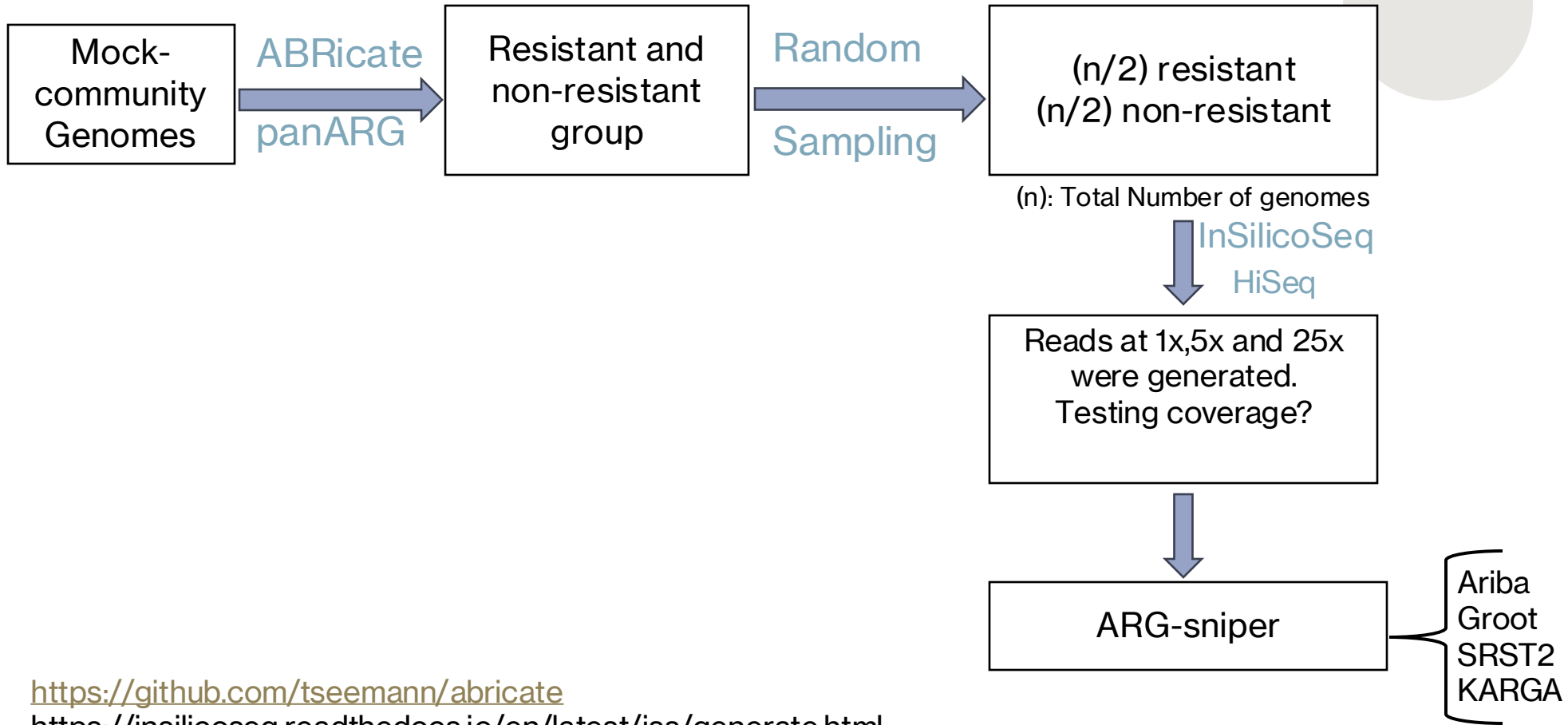
 Emily F. Wissel, Brooke M. Talbot,  Noriko A. B. Toyosato,  Robert A. Petit III, Vicki Hertzberg, Anne Dunlop,  Timothy D. Read

**doi:** <https://doi.org/10.1101/2022.01.13.476279>

Few high resistance and clinically relevant bacterial genomes.



# Simulation of reads:

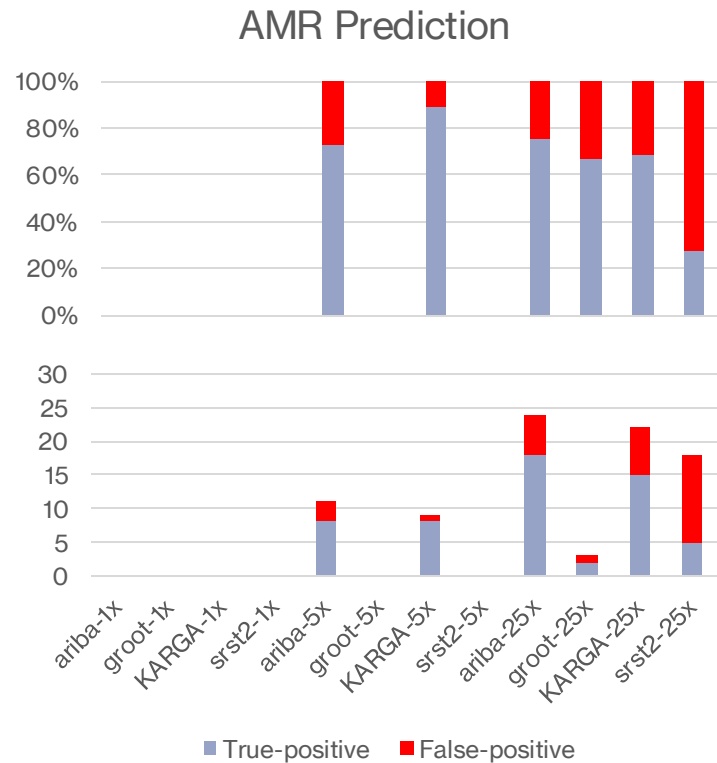


- <https://github.com/tseemann/abricate>
- <https://insilicoseq.readthedocs.io/en/latest/iss/generate.html>



# Preliminary results:

Total AMR genes predicted by ABRICATE: 28



# Thing to be done:

- Back-mapping of simulated-reads to reference.
- Test AMR prediction at 50x or 100x coverage.
- Run ARG-profiler separately.
- Test the effect of complexity on AMR prediction.
- Test the prediction on clinically relevant and phenotypical tested high resistance community of species.
- Characterizing the drug class for AMR-genes.

Thank You!

