Computational Practical 6: Detecting antimicrobial resistance from bacterial genomes using command-line tools

Module Developers: Mr. Collins Kigen and Dr Francesc Coll I Cerezo

Table Of Contents

Introduction

Growing rates of antimicrobial resistance make antibiotic susceptibility testing (AST) increasingly needed to ensure the right antibiotics are prescribed for patients with bacterial infections. Determining antibiotic susceptibility is preferred over empiric therapy, wherein typically broad-spectrum drugs are used without a definitive confirmation of the infectious agent and which antibiotics infectious bacteria are resistant to. Data collected on antibiograms (strains' full susceptibility pattern) can also be used for surveillance purposes and, in turn, inform empiric therapy.

AST is routinely performed using culture-based techniques in clinical diagnostic laboratories, frequently disk diffusion, broth microdilution and gradient diffusion (i.e., E-test). As antibiotic resistance is genetically encoded, i.e. mediated by acquisition of new genes, gene copy number, or mutations in regulatory and coding regions of existing chromosomal genes, molecular tests have been developed to target the detection of such genetic markers.[1] In the last decade, **whole-genome sequencing** has emerged as an alternative technology to both culture and targeted molecular

tests for the detection of AMR as it can, in principle, detect all AMR genetic determinants and predict resistance to all antibiotics in a single experiment. The **accuracy of genotypic predictions** depends on the availability of: (1) accurate databases of AMR genetic determinants, (2) large collections of whole-genome sequenced strains with AST measurements to assess the diagnostic accuracy of such catalogues, and (3) automated genome analysis and interpretation tools.

Mutational (chromosomal) resistance is the main driver of acquired resistance in certain bacterial species, such as *Mycobacterium tuberculosis* and *Helicobacter pylori*, or for particular antibiotics, especially to synthetic agents such as fluoroquinolones and oxazolidinones. Resistance mutations are vertically transmitted, i.e., via clonal reproduction of bacteria, or can be transmitted horizontally via homologous recombination between different strains. **Gene-mediated resistance** is the main driver of acquired resistance in certain bacterial species, particularly in gram-negatives. Resistance genes can be horizontally transmitted (via mobile genetic elements such as plasmids) and vertically transmitted via clonal reproduction of bacteria, particularly stable if integrated into the chromosome. In some bacterial species, chromosomal and gene-mediated resistance are equally common (e.g., *Staphylococcus aureus*). Resistance to the same antibiotic can be conferred by both mutations and acquired genes (e.g., fusidic acid in *Staphylococcus aureus*, colistin resistance in *Escherichia coli*).

Over the years, a number of global studies have identified the genes and mutations that confer resistance to particular antibiotics. There are several databases such as the **Comprehensive Antimicrobial Resistance Database (CARD)** (https://card.mcmaster.ca/), **ResFinder** (https://cge.cbs.dtu.dk/services/ResFinder/)**, AMRFinder** (https://www.ncbi.nlm.nih.gov/pathogens/antimicrobial-resistance/AMRFinder/) **or Pathogenwatch** (https://pathogen.watch/) that contain information about the genes and mutations that confer resistance. The use of these databases and tools depends on the species and mechanisms of resistance one is interested in.

6.1 Bacterial strains to be analysed.

Table 6.1 contains the list of strains to be analysed in this practical. We will use three different command-based tools (AMRFinder, ResFinder and CARD RGI) to identify AMR genetic determinants from whole-genome sequences. These strains were sourced from key studies on the genomic epidemiology of methicillin-resistant *Staphylococcus aureus* (MRSA)[2] and extensively drug-resistant (XDR) *Salmonella typhi*.[3] In later computational practicals we will explore the epidemiology of outbreaks and epidemic clones of these bacteria, and study the strains investigated in this practical in a broader context. We will also identify the genomic context of AMR genes, for example, if they are carried on mobile genetic elements.

Table 6.1

Species	Study and origin	Strain Id	Genome accession	Assembly file name
S. aureus	holden2013, Berlin (Germany), 2007, ST22 EMRSA-15	07-02477	ERR017261	ERR017261.assembly.fa
	holden2013, UK, 2005, ST22 EMRSA-15	HO50960412	HE681097 (GenBank)	HO50960412.fa
S. typhi	klemm2018 (ACT), Pakistan, 2016, 4.3.1 (H58) XDR	BL0006	ERR2093245	ERR2093245.assembly.fa
	klemm2018, Pakistan (2016) – 4.3.1 (H58) pre-XDR	Pak60168	ERR2093329	ERR2093329.assembly.fa

Typhoid outbreak on campus

A total of twenty-four cases of Typhoid fever have been reported on campus. We will collectively investigate the source of this outbreak by collecting epidemiological information about these cases using the **Epicollect5**. We will also investigate the antibiotic susceptibility of these isolates and the AMR genes contained in their genomes. Finally, we will integrate all this data with a phylogenetic tree to contextualise the local outbreak in relation to other Typhoid isolates collected from other South African cities and countries. **In this practical**, remember to **run the commands** below on the genome assembly of your assigned *S. typhi* sample, and take note of the AMR genetic determinants identified by the AMR detections we will use. At the end of this practical, we will input this data into the Epicollect5.

6.2 WGS-based prediction of AMR using AMRFinderPlus

6.2.1 Introduction to AMRFinderPlus

To enable accurate assessment of AMR gene content, as part of a multi-agency collaboration, the National Center for Biotechnology Information (NCBI) in the US developed a comprehensive AMR gene database, the Bacterial Antimicrobial Resistance Reference Gene Database, and AMRFinder, an AMR gene detection tool.[9] Recently, NCBI released a new version of AMRFinder, known as AMRFinderPlus that, among several new functionalities, has been expanded to detect point mutations in both protein and nucleotide sequences, and taxon-specific analyses that include, or exclude, certain genes and point mutations for specific taxa. AMRFinderPlus (https://github.com/ncbi/amr) is available on as a command-line tool only. In this section we will run AMRFinderPlus on the same strain genomes analysed with ResFinder and CARD RGI in previous sections.

6.2.2 AMRFinderPlus commands

Navigate to the ‘cp6’:

```
cd ~/course/cp6/
```

The raw sequencing reads and genome assemblies used in this practical can be found in the directory ~/course/cp6/

The only required arguments to run AMRFinderPlus are either -p <protein_fasta> for proteins or -n <nucleotide_fasta> for nucleotides. Use '--help' to see the complete set of options and flags.

```
amrfinder --help
```

Use ‘amrfinder -u’ to download and prepare database for AMRFinderPlus:

```
amrfinder -u
```

First, a local database of the latest the latest AMR database must be download.

```
mkdir amrfinder_db
amrfinder_update -d ./amrfinder_db
```

After making sure the latest AMR database is downloaded, you can run amrfinder on genome assemblies, as showed in the command line below:

```
amrfinder -n HO50960412.fa -O Staphylococcus_aureus -o HO50960412_amrfinder.txt
```

It should take a couple of minutes for this command to finish.

From the command above, note the following chosen options:

- AMRFinder only supports the processing of input nucleotide sequences in FASTA format (with the -n/--nucleotide option), and not the analysis of raw reads in fastq format. This means that raw reads must be *de novo* assembled first.
- The option ‘-o/--output’ allows you to choose the name of the output file.
- One of the strengths of AMRFinfer is the option ‘-O/--organism’ which can be used to get organism-specific results. For those organisms which have been curated, using --organism will get optimized organism-specific results, and it is therefore recommended. AMRFinderPlus uses the --organism for screening for point mutations and to filter out genes that are nearly universal in a group and uninformative.

Use ‘amrfinder -l’ to list the organism options supported by AMRFinder:

```
amrfinder -l
```

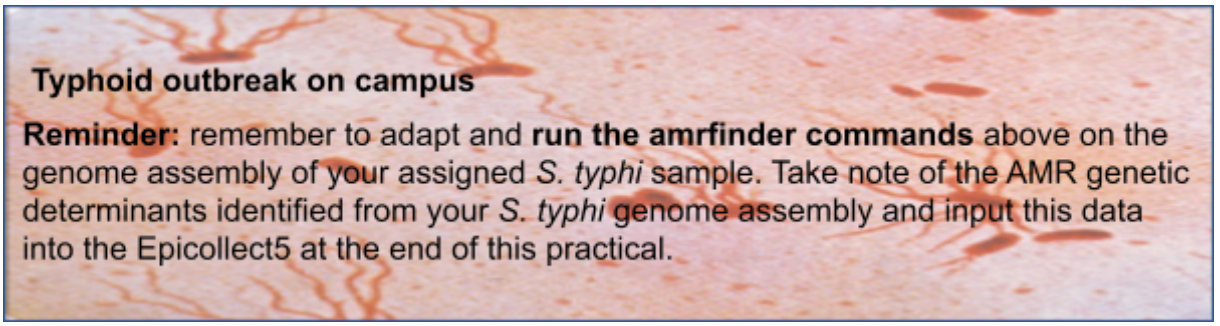
You will find that ‘Staphylococcus_aureus’ and ‘Salmonella’ are included in the list of supported taxa.

The list of commands below will run AMRFinder for the remaining samples:

amrfinder -n ERR017261.assembly.fa -O Staphylococcus_aureus -o ERR017261_amrfinder.txt

amrfinder -n ERR2093245.assembly.fa -O Salmonella -o ERR2093245_amrfinder.txt

amrfinder -n ERR2093329.assembly.fa -O Salmonella -o ERR2093329_amrfinder.txt



6.2.3 Interpreting AMRFinderPlus results

The table below includes some of the columns of the AMRFinderPlus output file HO50960412_amrfinder.txt.

Gene symbol	Sequence name	Element subtype	Class	Subclass
gyrA_S84L	Staphylococcus aureus quinolone resistant GyrA	POINT	QUINOLONE	QUINOLONE
mecA	PBP2a family beta-lactam-resistant peptidoglycan transpeptidase MecA	AMR	BETA-LACTAM	METHICILLIN
mecR1	beta-lactam sensor/signal transducer MecR1	AMR	BETA-LACTAM	METHICILLIN
tet(38)	tetracycline efflux MFS transporter Tet(38)	AMR	TETRACYCLINE	TETRACYCLINE
23S_C2220T	Staphylococcus aureus linezolid resistant 23S	POINT	OXAZOLIDINONE	LINEZOLID
23S_C2220T	Staphylococcus aureus linezolid resistant 23S	POINT	OXAZOLIDINONE	LINEZOLID
parC_S80F	Staphylococcus aureus quinolone resistant ParC	POINT	QUINOLONE	QUINOLONE
blaI	penicillinase repressor BlaI	AMR	BETA-LACTAM	BETA-LACTAM
blaZ	penicillin-hydrolyzing class A beta-lactamase BlaZ	AMR	BETA-LACTAM	BETA-LACTAM

The column 'Gene symbol' indicates the genetic determinant (either acquired gene or point mutation) associated with phenotypic resistance, the latter indicated in the column 'Subclass'.

Based on AMRFinderPlus output files, fill in the tables below to facilitate comparison of WGS-predicted antibiograms between strains.

Summary of AMRFinderPlus results for *S. aureus* strains

Antibiotic	HO50960412	pAST	07-02477 (ERR017261)	pAST
Amikacin		ND		ND
Gentamicin		S		S
Tobramycin		ND		ND
Kanamycin		ND		ND
Oxacillin		R		R
Methicillin/Cefoxitin		ND		ND
Penicillin		R		R
Fusidic Acid		S		S
Vancomycin		S		S
Clindamycin		S		R
Erythromycin		R		R
Mupirocin		S		S
Linezolid		S		R
Tetracycline		S		S
Trimethoprim		ND		ND
Co-trimoxazole		S		S
Daptomycin		S		S
Tigecycline		S		S
Rifampicin		S		S
Ciprofloxacin (quinolone)		R		R
Moxifloxacin (quinolone)		R		R
Teicoplanin		S		S
Phosphomycin		S		S

Phenotypic AST extracted from holden2013 Supplementary Table 1
Co-Trimoxazole: Trimethoprim + sulfamethoxazole

Summary of AMRFinderPlus results for *S. typhi* strains

Antibiotic	BL0006 (ERR2093245)	pAST	Pak60168 (ERR2093329)	pAST
Tobramycin		ND		ND
Amikacin		ND		ND
Streptomycin		ND		ND
Ampicillin		R		S
3GC (cefotaxime, ceftazidime)		R		S
Chloramphenicol		R		R
Ciprofloxacin (quinolone)		R		I
Sulfamethoxazole		ND		ND

(sulfonamides)				
Trimethoprim		ND		ND
Co-Trimoxazole		R		R
Tetracycline		ND		ND
Azithromycin		S		S
Colistin		ND		ND
Meropenem		S		S

3GC: 3rd Generation cephalosporins
Co-Trimoxazole: Trimethoprim + sulfamethoxazole
Phenotypic AST abbreviations: ND, not determined; R, resistant; I, intermediate; S, susceptible.

6.3 WGS-based prediction of AMR using ResFinder

6.3.1 Introduction to ResFinder

ResFinder, developed by Center for Genomic Epidemiology at the Technical University of Denmark (http://www.genomicepidemiology.org/), is a freely accessible tool to identify acquired genes and/or chromosomal mutations mediating antimicrobial resistance in total or partial DNA sequence of bacteria. Published in 2012 for the first time[4], ResFinder was the first web-based bioinformatics tool developed to provide detection of AMR genes in WGS, aimed at users without specialized bioinformatic skills. A command-line (https://bitbucket.org/genomicepidemiology/resfinder/) version was later developed which allows the automation of ResFinder analyses within bioinformatic scripts. The authors claim ResFinder (web-based version, September 2021) has been executed more than 800,000 times from more than 61,000 different users in over 171 countries.[5]

ResFinder, originally developed to detect acquired AMR genes, was later expanded with PointFinder[6], a tool that detects chromosomal point mutations mediating resistance to selected antimicrobial agents. Recently, additional databases were developed to link each AMR determinant with phenotypic resistance to specific antimicrobial compounds, and species-specific panels for *in silico* antibiograms. ResFinder 4.0 was validated for several bacterial species including *Salmonella spp*. and *Staphylococcus aureus* strains with a diversity of AST profiles, human and animal sources and geographical origins.[7]

6.3.1 ResFinder commands

Navigate to the ‘cp6’ directory:
cd ~/course/cp6/

The raw sequencing reads and genome assemblies used in this practical can be found in the directory ~/course/cp6/

Because the genomes of our strains are available as raw sequencing reads or genome assemblies, we will need to indicate ResFinder the format of input files. ResFinder can analyse both paired-end Illumina reads in fastq.gz format and genome assemblies in FASTA format.

Execute the command below to display all ResFinder arguments and options:
resfinder --help

IMPORTANT NOTE: if resfinder could not be found, you can install it yourself on your virtual machine running the command below:

mamba install resfinder

Next, if not already available on ~/course/cp6/ directory, download a local copy of the latest ResFinder databases:
git clone https://bitbucket.org/genomicepidemiology/resfinder_db/
git clone https://bitbucket.org/genomicepidemiology/pointfinder_db/
git clone https://bitbucket.org/genomicepidemiology/disinfinder_db/

Local databases need to be indexed using kma:
cd resfinder_db
python3 INSTALL.py kma_index
cd ..
cd pointfinder_db
python3 INSTALL.py kma_index
cd ..
cd disinfinder_db
python3 INSTALL.py kma_index

Set approximate environment bash variables for ResFinder executable to locate these databases.
export CGE_RESFINDER_RESGENE_DB="/home/manager/course/cp6/resfinder_db";
export CGE_RESFINDER_RESPOINT_DB="/home/manager/course/cp6/pointfinder_db";
export CGE_DISINFINDER_DB="/home/manager/course/cp6/disinfinder_db";

Remember to set these variables in any new terminal window. Otherwise ResFinder will exist with the error: ‘Could not locate ResFinder database path’.

Now everything is set to run ResFinder on your terminal screen as shown in the command below:

resfinder -ifa HO50960412.fa -s "Staphylococcus aureus" --acquired --point --outputPath HO50960412_resfinder

IMPORTANT NOTE: if ResFinder database could not be found ('Could not locate ResFinder database path', you can use the option '--db_path_res' to indicate where the directory of such database is:

resfinder -ifa HO50960412.fa -s "Staphylococcus aureus" --acquired --point --outputPath HO50960412_resfinder --db_path_res ./resfinder_db

The command line above was used to run ResFinder on the genome assembly of *S. aureus* HO50960412 strain (Table 1). Note the following parameters:

- the option '-ifa' is used to indicate that the input genome is provided in FASTA format, following by the path to the genome assembly file we want to analyse.
- the option '-s' is used to indicate the bacterial species in the same. This is important for ResFinder to use the antimicrobial panel specific to each bacterial species.
- the option '--acquired' is chosen to detected acquired resistance genes, and
- the option '--point' to scan for AMR chromosomal mutations.
- the option '--outputPath' allows you to specify the name of the output directory where ResFinder files will be stored

Next, we will run ResFinder on the raw sequencing reads of strain 07-02477 (accession number ERR017261).

fastq.gz were previously downloaded directly from the ENA using their FTP links (https://www.ebi.ac.uk/ena/browser/view/ERR017261) so they are already available on the working directory.

optional
wget ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR017/ERR017261/ERR017261_1.fastq.gz
wget ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR017/ERR017261/ERR017261_2.fastq.gz

resfinder -ifq ERR017261_1.fastq.gz ERR017261_2.fastq.gz -s "Staphylococcus aureus" --acquired --point --outputPath ERR017261_resfinder --db_path_res ./resfinder_db

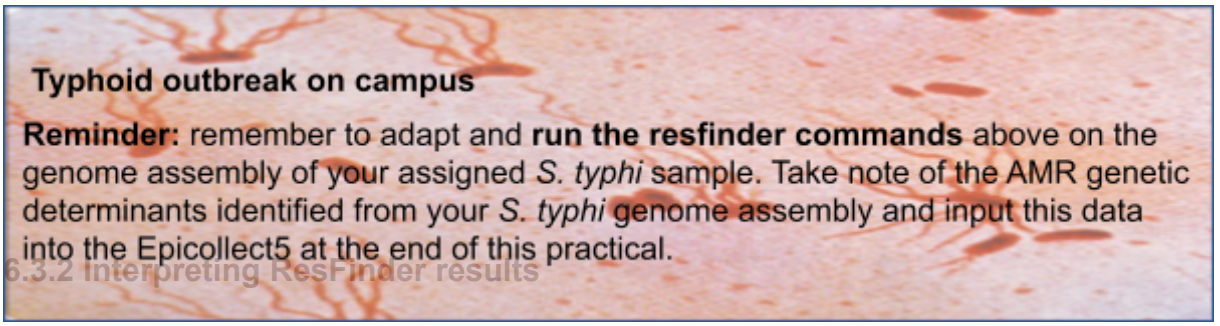
In the ResFinder command above note we used the same options as for sample HO50960412 except for '-ifq', used here to specify input fastq file(s). ResFinder assumes the input to be single-end fastq if only one file is provided after '-ifq', and to be paired-end data if two files are provided instead.

If you could not download the fastq.gz files of sample ERR017261, remember you can run ResFinder on its assembly file:

resfinder -ifa ERR017261.assembly.fa -s "Staphylococcus aureus" --acquired --point --outputPath ERR017261_resfinder --db_path_res ./resfinder_db

You can use the commands below to run ResFinder for the *S. typhi* samples:
resfinder -ifa ERR2093245.assembly.fa -s "Salmonella enterica" --acquired --point --outputPath ERR2093245_resfinder --db_path_res ./resfinder_db

resfinder -ifa ERR2093329.assembly.fa -s "Salmonella enterica" --acquired --point --outputPath ERR2093329_resfinder --db_path_res ./resfinder_db



Based on the ResFinder commands run in the previous section, you should have obtained a '_resfinder' output directory for each of the samples analysed. Out of the various output files, the following ones are key for interpretation:

- ResFinder_results_table.txt: summary of detected acquired AMR genes by antibiotic class, including BLAST statistics such as percentage of nucleotide identity or percentage of gene length covered.
- PointFinder_table.txt: summary of chromosomal genes scanned for AMR point mutations for the chosen bacterial species.
- PointFinder_results.txt: detected AMR chromosomal point mutations in your sample, and associated phenotypic resistance.
- pheno_table_[species].txt: ResFinder (including PointFinder) WGS-predicted phenotypes for the bacterial species chosen, including the detected genetic determinants supporting such prediction.
- pheno_table.txt: ResFinder (including PointFinder) WGS-predicted phenotypes for all antibiotics included in this database. These results should be interpreted with caution, and the file pheno_table_[species].txt prioritised for reporting.

IMPORTANT: Out of all these output files, we should focus on pheno_table_[species].txt, as it contains WGS-predicted phenotypes that are specific for the chosen bacterial species. Also, look at file PointFinder_results.txt to extract AMR point mutations.

Based on the ResFinder output files, fill in the tables below to facilitate comparison of WGS-predicted antibiograms between strains.

Summary of ResFinder results for *S. aureus* strains

Antibiotic	HO50960412	pAST	07-02477 (ERR017261)	pAST
Amikacin		ND		ND
Gentamicin		S		S
Tobramycin		ND		ND
Kanamycin		ND		ND
Oxacillin		R		R
Methicillin/Cefoxitin		ND		ND
Penicillin		R		R
Fusidic Acid		S		S
Vancomycin		S		S
Clindamycin		S		R
Erythromycin		R		R
Mupirocin		S		S
Linezolid		S		R
Tetracycline		S		S
Trimethoprim		ND		ND
Co-trimoxazole		S		S
Daptomycin		S		S
Tigecycline		S		S
Rifampicin		S		S
Ciprofloxacin (quinolone)		R		R
Moxifloxacin (quinolone)		R		R
Teicoplanin		S		S
Phosphomycin		S		S

Summary of ResFinder results for *S. typhi* strains

Antibiotic	BL0006 (ERR2093245)	pAST	Pak60168 (ERR2093329)	pAST
Tobramycin		ND		ND
Amikacin		ND		ND
Streptomycin		ND		ND
Ampicillin		R		S
3GC (cefotaxime, ceftazidime)		R		S
Chloramphenicol		R		R
Ciprofloxacin (quinolone)		R		I
Sulfamethoxazole (sulfonamides)		ND		ND
Trimethoprim		ND		ND
Co-Trimoxazole		R		R
Tetracycline		ND		ND
Azithromycin		S		S
Colistin		ND		ND
Meropenem		S		S

3GC: 3rd Generation cephalosporins
Co-Trimoxazole: Trimethoprim + sulfamethoxazole

6.4 Adding *in silico* antibiogram of *S. typhi* outbreak strains into EpiCollect

Typhoid outbreak on campus

A total of twenty-four cases of Typhoid fever have been reported on campus. We will collectively investigate the source of this outbreak by collecting epidemiological information about these cases using the **Epicollect5**. We will also investigate the antibiotic susceptibility of these isolates and the AMR genes contained in their genomes. Finally, we will integrate all this data with a phylogenetic tree to contextualise the local outbreak in relation to other Typhoid isolates collected from other South African cities and countries. **In this practical**, after running the commands on the genome assembly of your assigned *S. typhi* sample, make sure the AMR genetic determinants identified by the AMR detection tools are included into Epicollect5.

Following on from the EpiCollect practical in which you collected epidemiological information about twenty-four Typhoid cases reported on campus, you will now need to input the AMR genes found in your Typhoid genomes into EpiCollect. First, use the same mobile phone you used in the previous EpiCollect practical. Make sure WIFI is turned on and connect to the WIFI network using the credentials provided.

1. Open the Epicollect5 app on your phone and select the project named 'Typhoid XDR outbreak' (Figure 1)

Figure 1: Click on the 'Typhoid XDR outbreak' project

2. Then click on the entry you want to edit (Figure 2).

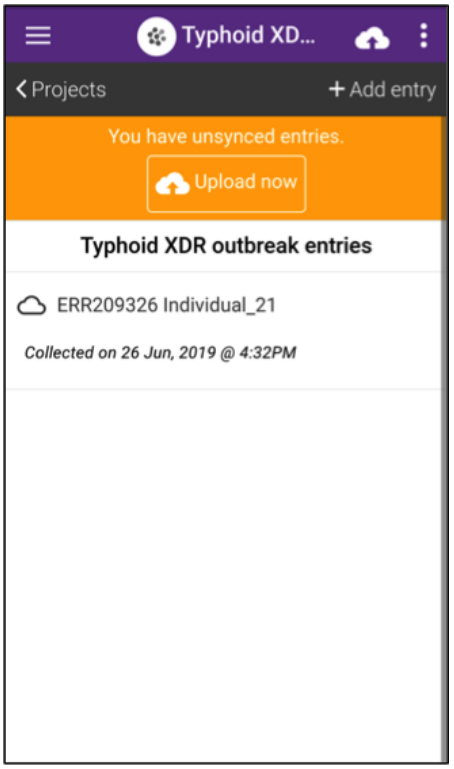


Figure 2: Click on an existing entry

3. Scroll down to the field ‘Antibiotic resistance genes’ and select the AMR genes/mutations detected in your Typhoid isolates (Figure 3 and Figure 4).

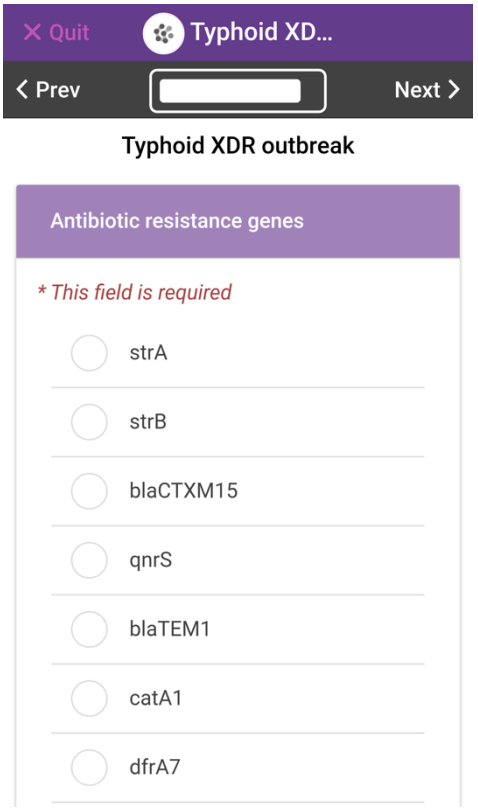


Figure 3

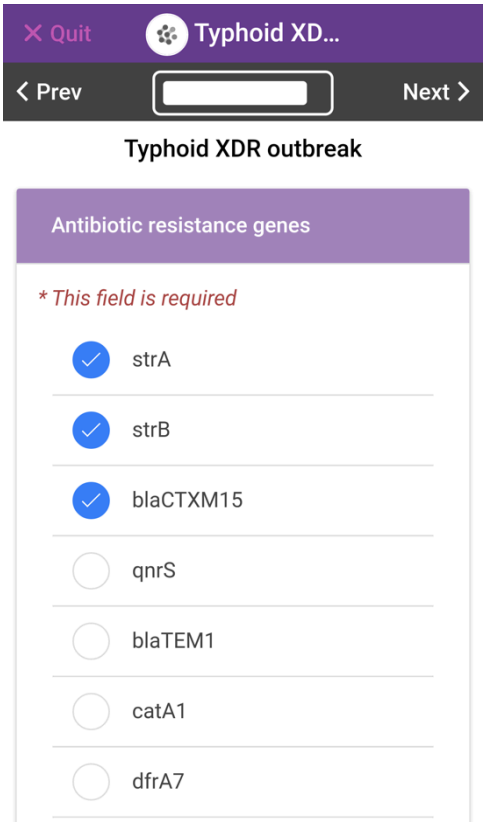


Figure 4

4. Do not forget to save changes in your entry by clicking on ‘Save entry’ at the end of the form (Figure 5).

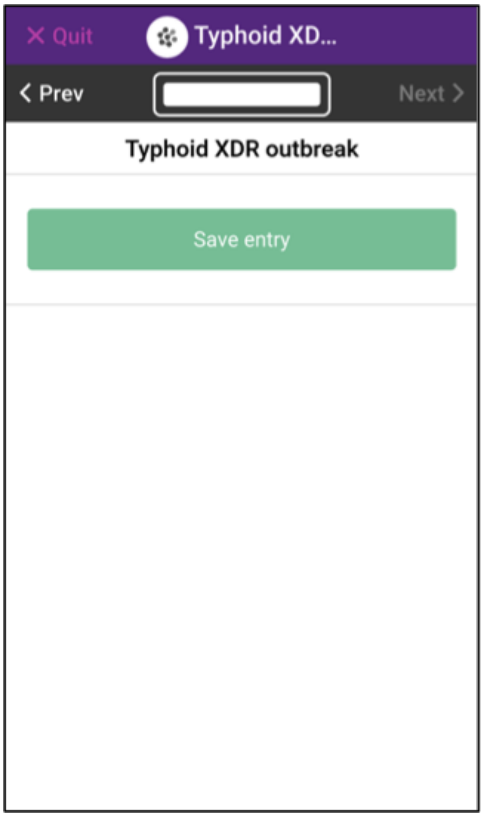


Figure 5: Save changes by clicking on 'Save entry' at the end of the form

6.5 WGS-based prediction of AMR using CARD RGI (Optional)

6.5.1 Introduction to CARD and RGI

The Comprehensive Antibiotic Resistance Database (CARD) is a very commonly queried database of resistance genes and associated AMR phenotypes.[8] CARD is a curated resource (https://card.mcmaster.ca) providing reference DNA and protein sequences, detection models and bioinformatics tools on the molecular basis of bacterial antimicrobial resistance (AMR). CARD focuses on providing high-quality reference data and molecular sequences within a controlled vocabulary, the Antibiotic Resistance Ontology (ARO), designed by the CARD team.

CARD's Resistance Gene Identifier (RGI) is a bioinformatic tool developed to predict resistomes from genomic and metagenomic data using the CARD database as a source of AMR genes and curated genotype-phenotype relationships. Analyses can be performed via this web portal or via the command line (https://github.com/arpcard/rgi), which can be used to automate the analysis of multiple samples in bioinformatic scripts. Briefly, RGI algorithmically predicts AMR genes and mutations from submitted genomes using a combination of open reading frame prediction with Prodigal, sequence alignment with BLAST or DIAMOND, and curated resistance mutations included with the AMR detection model.

6.5.1 CARD RGI commands

Navigate to the 'cp6' directory:
cd ~/course/cp6/

Activate 'rgi' conda environment:
conda activate rgi

The raw sequencing reads and genome assemblies used in this practical can be found in the directory /cps/data/.

Because the genomes of our strains are available as raw sequencing reads or genome assemblies, we will need to indicate RGI the format of input files. RGI can analyse both paired-end Illumina reads in fastq.gz format with the tool *rgi bwt*, and genome assemblies in FASTA format with the tool *rgi main*.

Execute the commands below to display the arguments and options of each RGI tool:
rgi main -h
rgi bwt -h

Next, download a local copy of the latest CARD database:
wget https://card.mcmaster.ca/latest/data --no-check-certificate
tar -xvf data ./card.json
rm data
rgi load --card_json card.json --local
rgi database --local --version

You can use '*rgi main*' to detect AMR genes from genome assemblies as showed in the commands below:
rgi main --input_sequence HO50960412.fa --output_file HO50960412_rgi --input_type contig --clean --local

rgi main --input_sequence ERR017261.assembly.fa --output_file ERR017261_rgi --input_type contig --clean --local

rgi main --input_sequence ERR2093245.assembly.fa --output_file ERR2093245_rgi --input_type contig --clean --local

rgi main --input_sequence ERR2093329.assembly.fa --output_file ERR2093329_rgi --input_type contig --clean --local

In the '*rgi main*' commands above, note the following arguments:

- --input_sequence: file name of the input genome assembly to analyse
- --output_file: name of RGI output file
- --input_type: type of input data, contig (DNA) or protein. Because the default option for this argument is 'contig', '--input_type contig' can be omitted.
- --clean: option to remove temporary files created by RGI
- --local: use local CARD database (the one we just downloaded)

6.5.3 Interpreting CARD RGI results

You can use the 'cut' command below to extract specific fields (i.e. columns) from RGI output files: Cut_Off, ARO, SNPs_in_Best_Hit_ARO, Drug Class, Resistance Mechanism, and AMR Gene Family.

cat HO50960412_rgi.txt | cut -d$'\t' -f6,11,13,15,16,17 > HO50960412_rgi.short.txt

In RGI output files pay particular attention to the fields 'Drug Class', 'Resistance Mechanism', and 'AMR Gene Family' which you can use to derive to what antibiotic each CARD-identified genetic determinant is conferring resistance to. Based on CARD RGI results, fill in the tables below to facilitate the comparison of genotypic antibiograms.

Summary of CARD RGI results for *S. aureus* strains

Antibiotic	HO50960412	07-02477 (ERR017261)
aminoglycoside antibiotic		
carbapenem		
cephalosporin		
fluoroquinolone antibiotic		
glycopeptide antibiotic		
lincosamide antibiotic		
macrolide antibiotic		
monobactam		
nitrofuran antibiotic		
oxazolidinone antibiotic		
penam		
penem		
phenicol antibiotic		
rifamycin antibiotic		
streptogramin antibiotic		
sulfonamide antibiotic		
tetracycline antibiotic		

Summary of CARD RGI results for *S. typhi* strains

Antibiotic	BL0006 (ERR2093245)	Pak60168 (ERR2093329)
aminoglycoside antibiotic		
carbapenem		
cephalosporin		
fluoroquinolone antibiotic		
glycopeptide antibiotic		
lincosamide antibiotic		
macrolide antibiotic		
monobactam		
nitrofuran antibiotic		
oxazolidinone antibiotic		
penam		
penem		
phenicol antibiotic		
rifamycin antibiotic		
streptogramin antibiotic		
sulfonamide antibiotic		
tetracycline antibiotic		

6.6. Group project questions - Detecting resistance and metagenomics

Considering the results obtained in this practical, discuss the following questions with the members of your group:

Based on the AMRFinder results for *S. aureus* HO50960412 strain:
- How can we tell which AMR genetic determinants are point mutations and which ones are acquired AMR genes?
- What beta-lactams do *blaI* and *blaZ* genes confer resistance to?
- Why is 23S_C2220T genetic determinant detected twice?
- Is this a methicillin-resistant *S. aureus* (MRSA) strain? If so, what genetic determinant would render this strain MRSA and why?

Based on the AMRFinder results for *S. typhi* BL0006 (ERR2093245) strain:
- How can we tell which AMR genetic determinants are point mutations and which ones are acquired AMR genes?
- What beta-lactams do *blaTEM-1* confer resistance to? To all beta-lactams?
- What specific quinolones do *gyrA*_S83F and *qnrS1* confer resistance to?
When answering the last two questions consider the differences between antibiotic and antibiotic class, and how genotypic resistance is reported by each tool.

Based on the results obtained by ResFinder and AMRFinder:
- How do ResFinder results differ from AMRFinder's for these strains in terms of AMR genetic determinants reported and predicted phenotypes?

Regarding the functionality and parameters of each tool, some tools gave the option to indicate specifically what bacterial species to analyse, what implications may this have on AMR genotypic predictions?

References

1. Geneva: World Health Organization. *Molecular methods for antimicrobial resistance (AMR) diagnostics to enhance the Global Antimicrobial Resistance Surveillance System*. (2019).
2. Holden, M. T. G. *et al.* A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant Staphylococcus aureus pandemic. *Genome Research* **23**, 653–664 (2013).
3. Klemm, E. J. *et al.* Emergence of an Extensively Drug-Resistant Salmonella enterica Serovar Typhi Clone Harboring a Promiscuous Plasmid Encoding Resistance to Fluoroquinolones and Third-Generation Cephalosporins. *mBio* **9**, e00105-18 (2018).
4. Zankari, E. *et al.* Identification of acquired antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy* **67**, 2640–2644 (2012).

5. Florensa, A. F., Kaas, R. S., Clausen, P. T. L. C., Aytan-Aktug, D. & Aarestrup, F. M. ResFinder - an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microbial genomics* **8**, 1–10 (2022).
6. Zankari, E. *et al.* PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. *Journal of Antimicrobial Chemotherapy* **72**, 2764–2768 (2017).
7. Bortolaia, V. *et al.* ResFinder 4.0 for predictions of phenotypes from genotypes. *Journal of Antimicrobial Chemotherapy* **75**, 3491–3500 (2020).
8. Alcock, B. P. *et al.* CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Research* **48**, 517–525 (2019).
9. Feldgarden, M. *et al.* Validating the AMRFinder Tool and Resistance Gene Database by Using Antimicrobial Resistance Genotype-Phenotype Correlations in a Collection of Isolates. *Antimicrobial Agents and Chemotherapy* **63**, (2019).