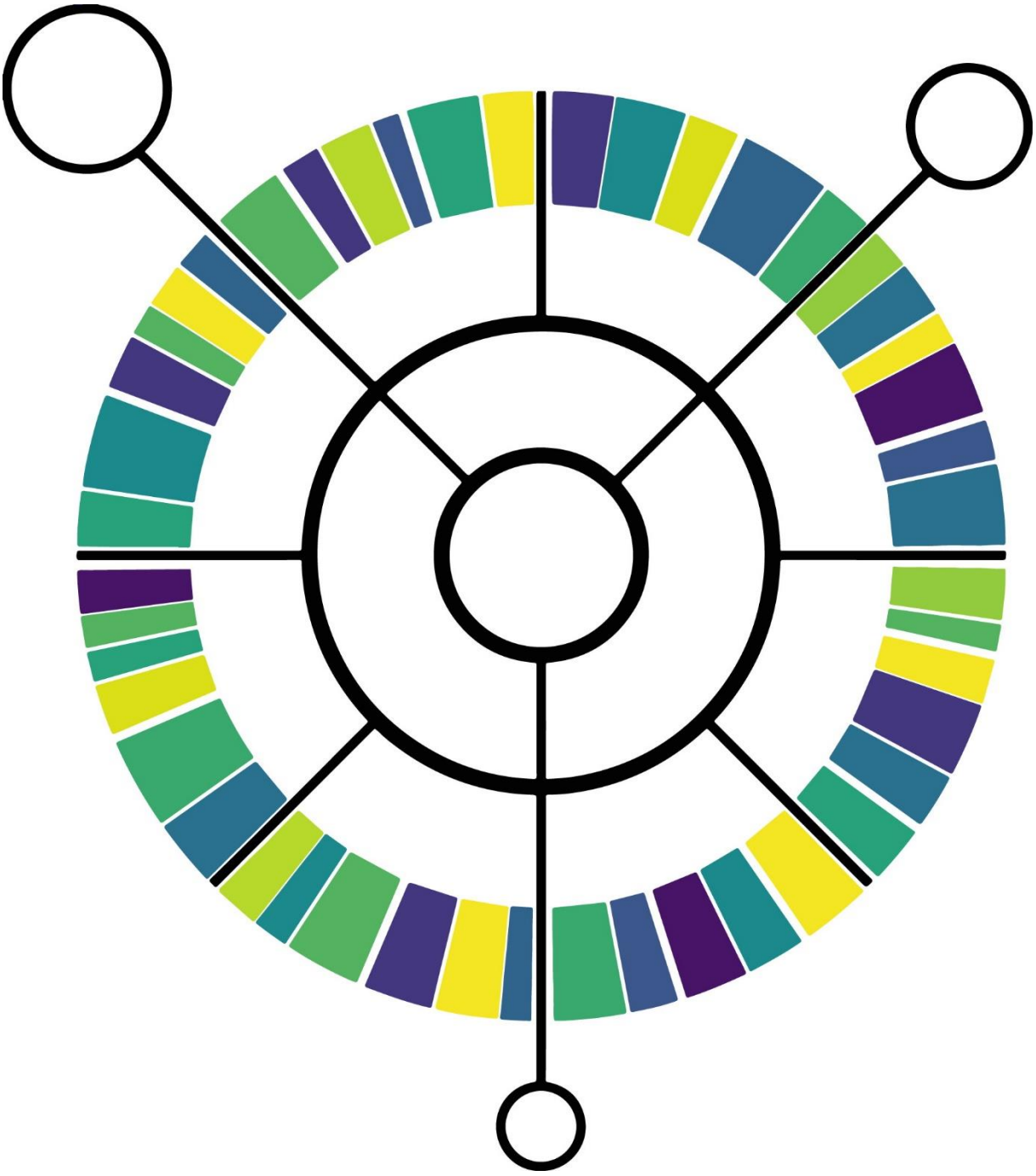
RGI WEB USAGE MANUAL



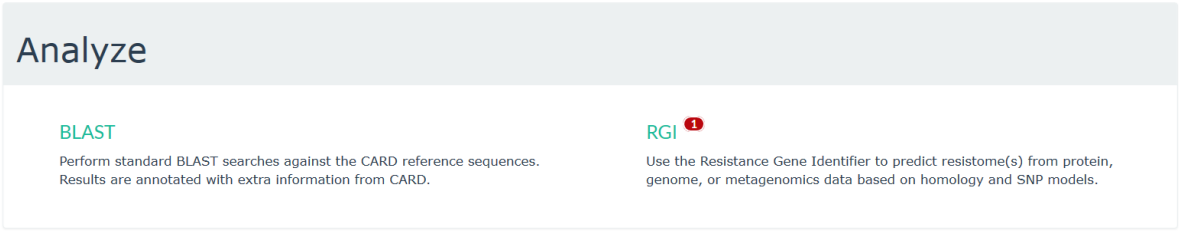
Appendix:

1.0 – How do I use RGI on CARD?

To use RGI on the CARD website. Navigate to the CARD website at https://card.mcmaster.ca/.

Once on the page, click on Analyze at the top right of the webpage. This will present to options:

1. To use BLAST
2. To use RGI

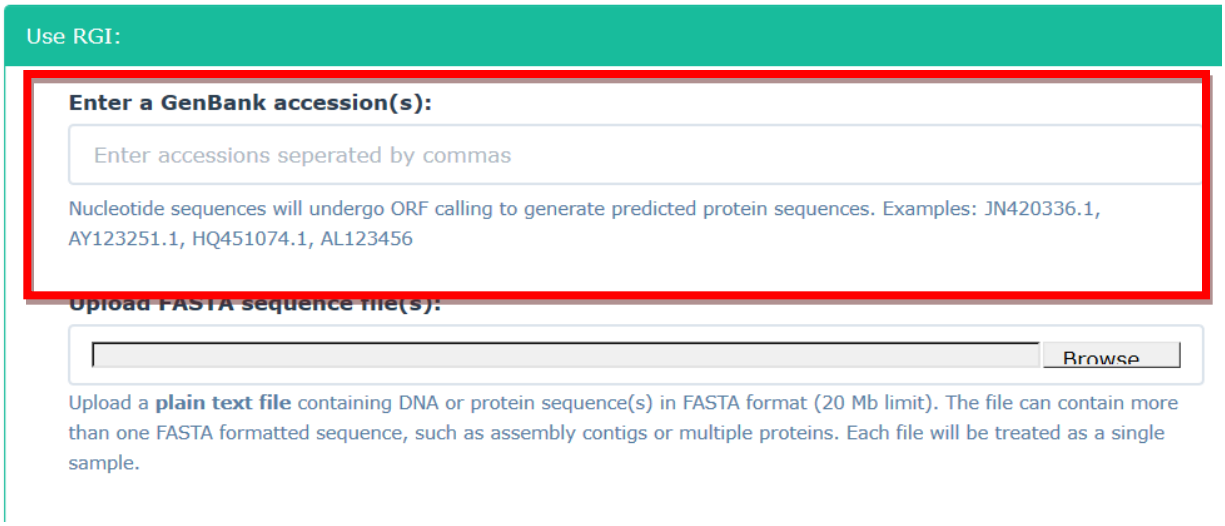


Click on RGI to access the RGI tool on the CARD website.

1.1 – Using GenBank Accessions

On the page, there will be multiple options for input. The first is to input the GenBank accession associated with the bacterial genome. RGI will use the accession to access the genome and using its built-in tools to analyze it for resistance determinants.

Go to GenBank or use the accession you have associated to your genome, copy the accession, and paste it into the input box requesting for an accession. Once complete, you may click submit at the bottom of the page to proceed with your analysis.



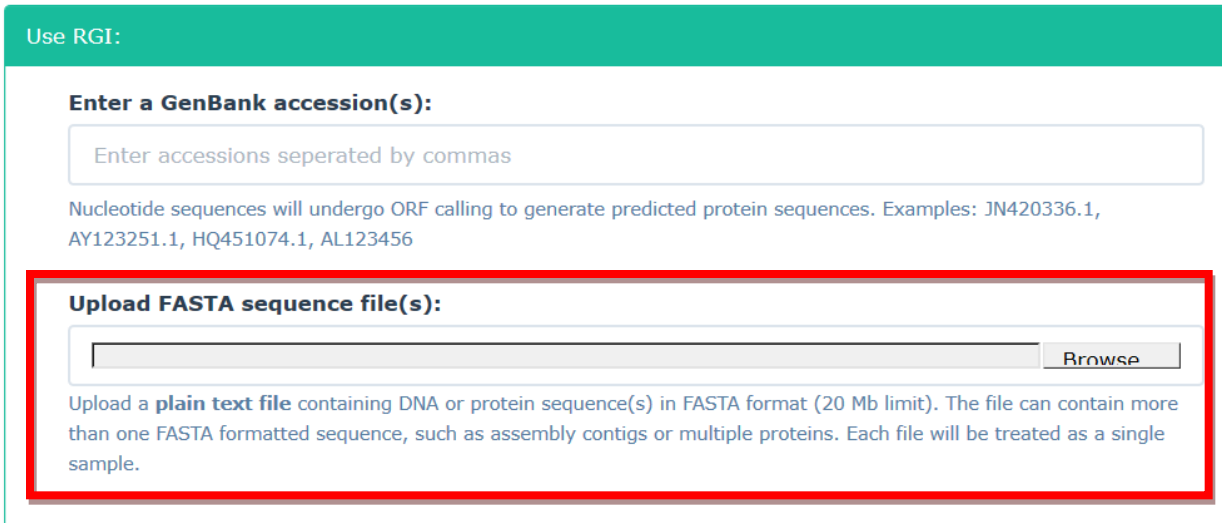
If you would like to apply some options, see section 1.4.

1.2 – Using a FASTA file

A FASTA formatted file is a text file that ends with extensions **.fsa**, **.fa**, or **.fasta**. It contains DNA or Protein sequence information, each sequence belonging to a new line with a header prefixed with ">". Please make sure that your FASTA file is valid, otherwise RGI will not work.

An example of a valid FASTA entry is:

>header
ATCGTGTCGAATGA...

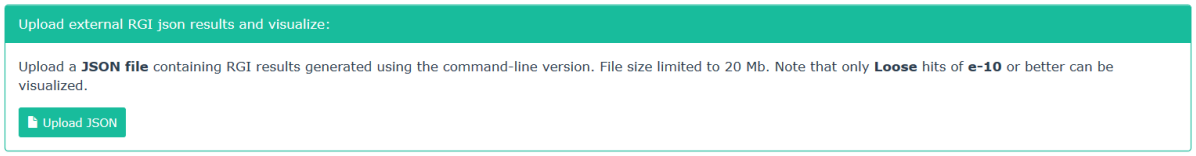


If you have access to a FASTA formatted file containing your desired genome, you may input it into RGI. Under "Upload FASTA sequence file(s)", click on browse and navigate to the

location of your FASTA file. Upload it and press submit to proceed with the analysis. If you would like to apply some options, see section 1.4.

1.3 – Using command-line generated RGI JSON

A JSON file is a JavaScript Object Notation file. JSON files follow a strict format. Please ensure that your JSON file is valid and is from the RGI output or else RGI will not work.

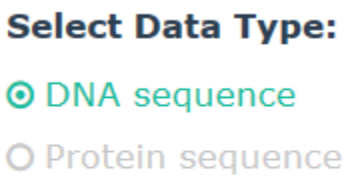


This method is usually used by more advanced users. If you have successfully used the command-line RGI software and would like to navigate your JSON using the webpage interface, you may scroll to the bottom of the page where you may select "Upload JSON". Upon successfully uploading the JSON, you may click submit to proceed with your analysis. If you would like to apply some options, see section 1.4.

1.4 – Using the options

Select Data Type:

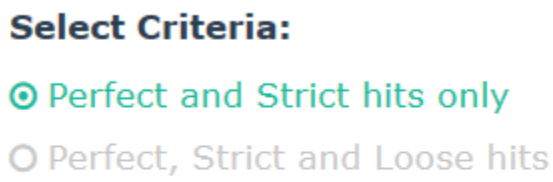
You may select either Protein or DNA sequence. Please select the option that applies to your sequence. If you are submitting a genome or a gene, select DNA Sequence. If you are submitting a protein sequence, select Protein Sequence.



Select Criteria:

You may select criterions for detection. Perfect and Strict ensures that the results from RGI are either Perfect matches or passed the curated bit-score, respectively.

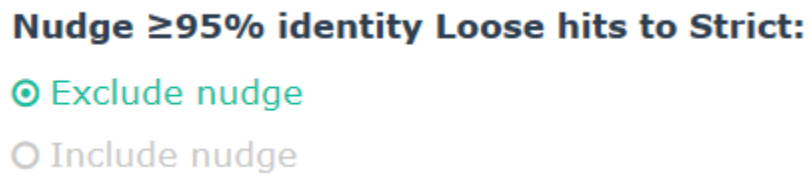
Perfect, Strict and Loose option functions the same as before, however selecting this option extends wait time slightly to present you with results that are less than the curated bit-score.



Nudge ≥95% identity Loose hits to Strict:

Exclude Nudge would cause RGI to ignore results that are less than the curated bit-score despite being more than 95% identity.

Include Nudge would cause RGI to allow results that are less than the curated bit-score if it is more than 95% identity.

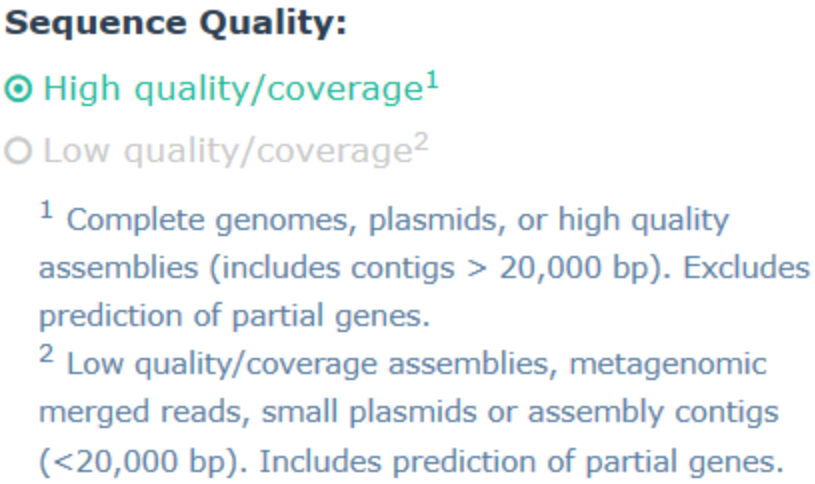


Sequence Quality:

You may select criterions based on the quality of your sequences. Depending on the lengths of your sequences and the identity of your sequence (small plasmids/partial sequences), you may opt to use High Quality coverage or Low Quality coverage.

High Quality coverage is recommended for contigs that are greater than 20,000 base pairs and for complete genomes, plasmids, or high quality assemblies.

Low Quality coverage is recommended for contigs that are less than 20,000 base pairs and for low quality/coverage assemblies, metagenomic merged reads, small plasmids or assembly contigs and partial sequences.



2.0 – How do I interpret the RGI web results?

Once you have found your desired options, press submit to proceed with your analysis. The analysis will take variable amounts of time depending on the size of your file and the options chosen.

2.1 – Reading the results table

If your analysis has completed successfully, you will be moved to a new page with your file at the top and a download link. Underneath you will see your table with multiple columns. You may sort this table using the ascending and descending arrows in each column and you may search this table using the search bar.

The first column indicates whether your match was Perfect, Strict or Loose. Please revise the definitions of these criterions in section 1.4 if you feel lost at any point.

The second column are your ARO terms; your gene names. These allow you to identify the genes that were detected in your input.

The third column are your SNPs/Single Nucleotide Polymorphisms. These will show you the mutations present in the gene if they are confirmed to confer resistance.

The fourth column is your detection criteria allowing you to identify the method by which the gene was found. Please refer to the CARD model ontologies for each detection criterion at https://card.mcmaster.ca/ontology/40323.

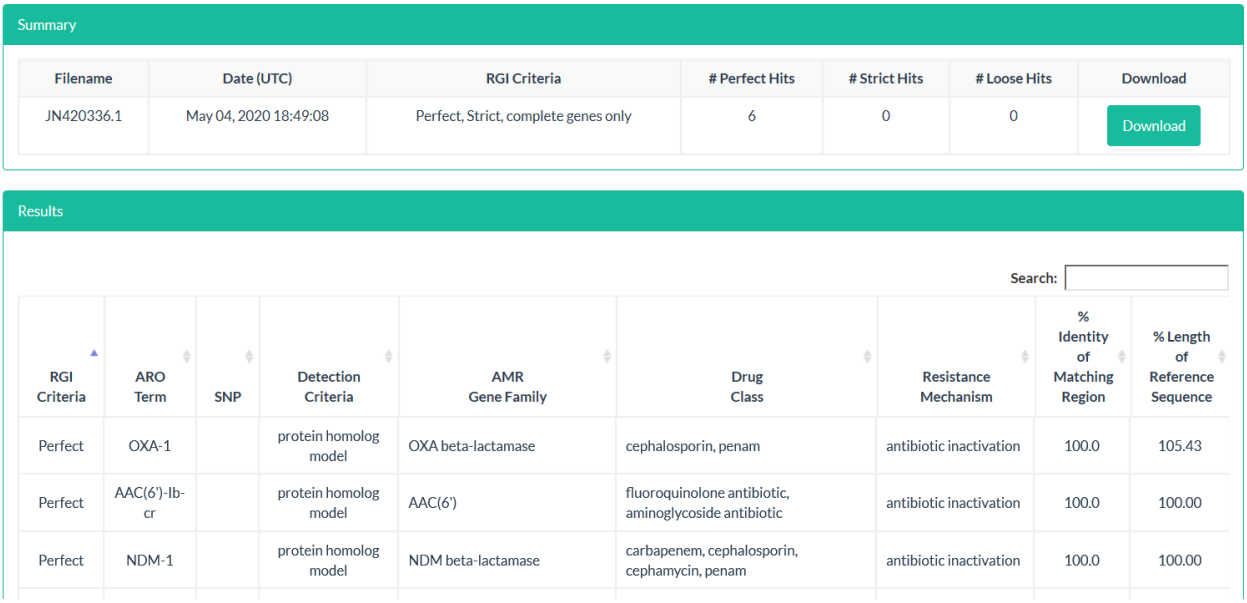
The fifth column is the AMR Gene Family, which show you the gene family your gene belongs to. The gene family provides information on the mechanisms and identity of your gene.

The sixth column is Drug Class, showing you the drug(s) that the gene confers resistance to.

The seventh column is the Resistance Mechanism, indicating the method by which the gene introduces drug resistance.

The eighth column is the percent identity indicating the degree to which your inputs matches with the ARO in CARD.

The ninth column is the percent length, indicating the length of your input relative to its matched ARO term.



Summary

Filename	Date (UTC)	RGI Criteria	# Perfect Hits	# Strict Hits	# Loose Hits	Download
JN420336.1	May 04, 2020 18:49:08	Perfect, Strict, complete genes only	6	0	0	Download

Results

Search:

RGI Criteria	ARO Term	SNP	Detection Criteria	AMR Gene Family	Drug Class	Resistance Mechanism	% Identity of Matching Region	% Length of Reference Sequence
Perfect	OXA-1		protein homolog model	OXA beta-lactamase	cephalosporin, penam	antibiotic inactivation	100.0	105.43
Perfect	AAC(6')-Ib-cr		protein homolog model	AAC(6')	fluoroquinolone antibiotic, aminoglycoside antibiotic	antibiotic inactivation	100.0	100.00
Perfect	NDM-1		protein homolog model	NDM beta-lactamase	carbapenem, cephalosporin, cephamycin, penam	antibiotic inactivation	100.0	100.00

2.2 – Using the wheel

i). AMR Gene wheel

You can find the RGI results sorted in the form of an interactive wheel for AMR Genes. The results are sorted by RGI Criteria (Perfect, Strict, Loose). Pressing any of them will reveal the genes under that criteria. To move back on the wheel, press the center.

ii). AMR Gene Family wheel

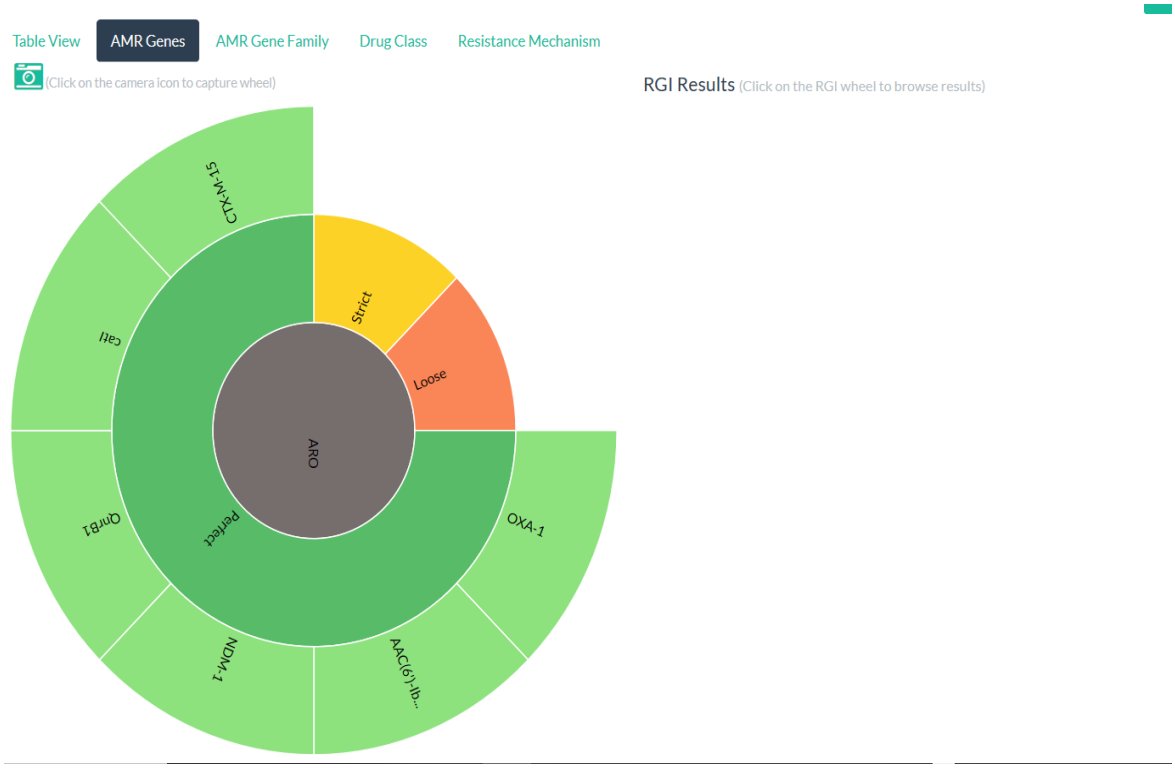
You can find the RGI results sorted in the form of an interactive wheel for AMR Gene Families. The results are sorted by RGI Criteria (Perfect, Strict, Loose). Pressing any of them will reveal the gene families under that criteria. Pressing the gene family will reveal the AMR genes under that family. To move back on the wheel, press the center.

iii). Drug Class wheel

You can find the RGI results sorted in the form of an interactive wheel for Drug Classes. The results are sorted by RGI Criteria (Perfect, Strict, Loose). Pressing any of them will reveal the Drug Classes under that criteria. Pressing the drug classes will reveal the AMR genes that confer resistance to the drug. To move back on the wheel, press the center.

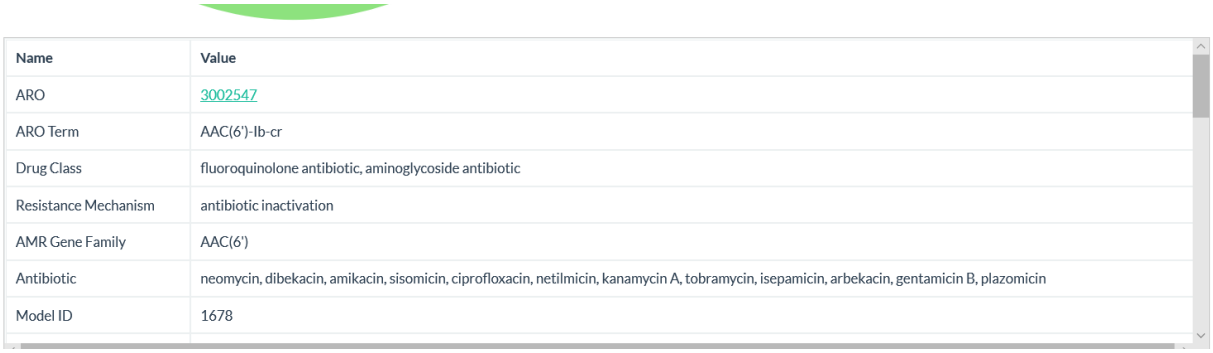
iv). Resistance Mechanism wheel

You can find the RGI results sorted in the form of an interactive wheel for Resistance Mechanisms. The results are sorted by RGI Criteria (Perfect, Strict, Loose). Pressing any of them will reveal the resistance mechanisms under that criteria. Pressing the resistance mechanism will reveal the AMR genes that function through that mechanism. To move back on the wheel, press the center.



2.3 – Reading the wheel results

Once you have decided which wheel you want to use, you may select an AMR gene of interest. On the right side of the wheel is a table where you see the AMR gene. Pressing on the AMR gene will open a dropdown table at the bottom of the page. This will reveal information regarding that gene such as model id, gene name, resistance mechanism, sequence, reference sequence and alignment.



Name	Value
ARO	3002547
ARO Term	AAC(6')-Ib-cr
Drug Class	fluoroquinolone antibiotic, aminoglycoside antibiotic
Resistance Mechanism	antibiotic inactivation
AMR Gene Family	AAC(6')
Antibiotic	neomycin, dibekacin, amikacin, sisomicin, ciprofloxacin, netilmicin, kanamycin A, tobramycin, isepamicin, arbekacin, gentamicin B, plazomicin
Model ID	1678

3.0 – Saving your data

After the analysis, should you wish to save your results, you may navigate to the Table View located to the left of the AMR Gene wheel tab.

At the top of the page, you will find your accession and a download button. Pressing the download button will begin the download.

3.1 – File Type and Access

i). Decompressing

The file downloaded will be double compressed as a gzip (.gz) and as a tarball (.tar). This ensures that the file downloaded is small and easily downloadable. To access this file, please use a decompression tool. The standard is usually 7zip downloadable from https://www.7-zip.org/.

Once decompressed, you will have access to a folder containing your data. Your data will be available as either a JSON file or a TEXT file. The JSON you may use on the RGI webtool to visualize the wheel. The TEXT file is tab delimited so you may open it in excel or your spreadsheet software of choice.