# Bioinformatics Lab Guide 3:

## *Antimicrobial Resistance Gene Detection*

James B. Herrick, Ph.D.
*James Madison University*

There are essentially two approaches to antimicrobial resistance (AMR) gene detection: assembly-based or read-based. For assembly-based, we'll use *Abricate* in GalaxyTrakr. For read-based, we'll use *KmerResistance* on the Center for Genomic Epidemiology (CGE) website.

## Detecting AMR Genes Using *Abricate* in GalaxyTrakr

1. Find *Abricate* under "NGS: Screening and Prediction".

2. Input your SPAdes assembly file (for example, "SPAdes on data 2 and data 1: contigs (fasta)").

3. **Click on 'Advanced Options.' Try each database one at a time**. Remember how to quickly re-run analyses by clicking on the name of the analysis in the right window, then on the 're-do' icon:
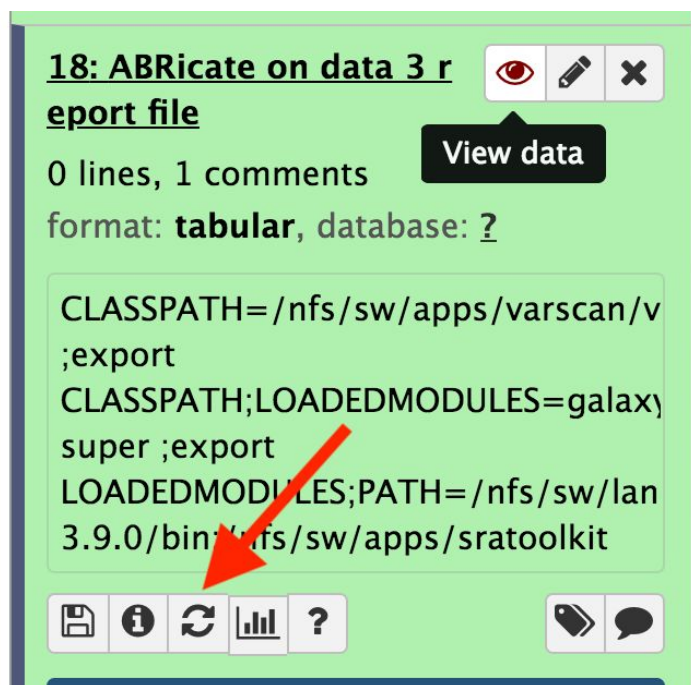
4. ARG-ANNOTT is a good database. The CARD databse will normally return a lot of genes, most of which may or may not be involved in AMR. Try more than one and see what you get. You can and ultimately should [BLAST](#) any results you get just to verify; also read what the genes actually encode.

## Abricate Output (from [here](#)):

## Output

Abricate produces a tap-separated output file with the following columns:

| Column | Example | Description |
| --- | --- | --- |
| FILE | Ecoli.fna | The filename this hit came from |
| SEQUENCE | contig000324 | The sequence in the filename |
| START | 23423 | Start coordinate in the sequence |
| END | 24117 | End coordinate |
| GENE | tet(M) | AMR gene name |
| COVERAGE | 1-1920/1920 | What proportion of the gene is in our sequence |
| COVERAGE_MAP | =============== | A visual represenation |
| GAPS | 1/4 | Openings / gaps in subject and query - possible psuedogene? |
| %COVERAGE | 100.00% | Proportion of gene covered |
| %IDENTITY | 99.95% | Proportion of exact nucleotide matches |
| DATABASE | card | The database this sequence comes from |
| ACCESSION | NC_009632:49744-50476 | The genomic source of the sequence |

| PRODUCT | aminoglycoside O-phosphotransferase APH(3')-IIIa | Gene product (if available) |
|---|---|---|

## Caveats

- Does not find mutational resistance, only acquired genes.
- Gap reporting incomplete
- Sometimes two heavily overlapping genes will be reported for the same locus
- Possible coverage calculation issues

MegaRes - Has good summaries of the activities of, as well as references for, many of the R genes you will find.

# Detecting AMR Genes using *KmerResistance* on CGE

The Center for Genomic Epidemiology has a number of programs which can be useful in microbial genomics, and especially in the study of *Salmonella* and *E. coli.* The site can be slow, though, and sometimes a bit buggy as well. Sometimes it's finicky about file names. Be sure they follow our file-naming suggestions and if in doubt, make them shorter. For CGE analyses, plan to start early and to be patient.

Philip T. L. C. Clausen, Ea Zankari, Frank M. Aarestrup, Ole Lund; Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data, Journal of Antimicrobial Chemotherapy, Volume 71, Issue 9, 1 September 2016, Pages 2484–2488, https://doi.org/10.1093/jac/dkw184. See also this comment.

1. Click on "KmerResistance" in the left menu. Click on the 'Instructions' link at the top and read through (you should do this with any CGE analysis)

2. Leave the defaults as they are. Click on 'isolate file', upload your two short read (not assembly) files, then click 'execute'. This analysis can take a few hours to a day to return. You can input your email address to be notified when your analysis is complete.