**Downloading SRA datasets and Using Bowtie2**

fastq-dump [SRA\_Run\_ID] – Downloads an SRA dataset using its run Id (SRR, ERR, or DRR)

[This command has lots of options](https://edwards.sdsu.edu/research/fastq-dump/). Some of which are mandatory therefore it is recommended you use it with the following:

fastq-dump --outdir ~/my\_dir/ --skip-technical --readids --dumpbase --clip SRR3403834

--outdir [my\_folder] This is the folder where the file is stored.

--skip-technical Don’t output data about each reads spot. We won’t need it.

--readids Output the ID corresponding with each read, you need this.

--dumpbase Make sure to use only A, C, G and T for sequence data.

--clip Removes any tags from amplification

--split-files Output’s left and right paired end reads to different files

-X [number] Stop at read [number]

-N [number] Start at read [number]

bowtie2 [fastq\_file] – searches a fastq file for the DNA specified and stores matches in .sam files.

Bowtie2 is a read mapping program. For each read in your file, bowtie2 will scan it looking for a mapping to the DNA you are looking for. You will need to construct a bowtie2 index of the DNA you’re looking for prior to running this command.

bowtie2 -x my\_bt2\_index --no-unal -S SRR3403834.sam -q SRR3403834.fastq

-x [bowtie2\_index] This tells bowtie2 what to look for. Don’t include any “.1.bt2” or “.rev.1.bt2” here, just use the base name.

--no-unal If it’s not a match, don’t save it.

-S [output\_file.sam] Specify the name of the output file and where it saves to.

-q The input is going to be a fastq file.

bowtie2-build [input\_file.fasta] [index\_name] – builds a bowtie2 index of a fasta file

Bowtie needs to do billions of searches for the same set of DNA. To speed up these searchers, it needs a highly organized data structure. The details are quite complicated, but you only need to input a fasta file of the DNA you are looking for and this command will do everything for you:

bowtie2-build e\_coli\_genome.fasta ecoli\_index

This command will build the database for you. In this case, the argument “ecoli\_index” will be the input with -x in the bowtie2 command despite there being several different files generated.

In the next class we will cover how to work with SAM files (the plain text output of bowtie2).

**DON’T FORGET TO SHUT OFF YOUR INSTANCES WHEN YOU’RE DONE!**