I have picked a paper entitled “[Critical Assessment of Metagenome Interpretation—a benchmark of metagenomics software](https://www.nature.com/articles/nmeth.4458)”. This paper addresses the lack of consensus about assessing the methods for assembly, taxonomic profiling and binning of metagenomic data. The plasmid assemblies, raw data and metadata in this publication have been deposited in the European Nucleotide Archive (ENA) under accession number PRJEB20380. As seen in the figure below, the project summary included 102 experiments that produced 598 nucleotide.

Graphical user interface, application

Description automatically generated

The commands below were run successfully:

# Get the sequencing run information

esearch -db sra -query PRJEB20380| efetch -format runinfo > runinfo.csv

# Download the assembled genome information

esearch -db nucleotide -query PRJEB20380| efetch -format fasta > genomes.fa

The first line of runinfo.csv is observed below:

Text

Description automatically generated

**esearch -db pubmed -query "critical assessment of metagenome interpretation-a benchmark of metagenomics software"| efetch**

Text

Description automatically generated

Graphical user interface, text

Description automatically generated

Running **esearch -db pubmed -query PMC5903868 | elink -target sra** did not give us any meaningful results.

Text

Description automatically generated

As you see in the screenshot below, **csvcut** is not found. I replaced it with **cut -d , -f 1.**

Text

Description automatically generated

I faced the error stating that fastq-dump command is not found. I solved this issue by running command **mamba install sra-tools==2.10.1**

Text

Description automatically generated

Finally, I ran **seqkit stats** on the fast file of one particular sequencing as shown below:

Text

Description automatically generated