Genome annotation

Now, we will annotate all the genomes. We need to run prokka on the genome, which we can reasonably do one-at-a-time at this scale.

The basic command is (for a given genome file, which here is GENOME.fna, and a given number of CPUs, which here is 2)

prokka --compliant --centre Famnit_name --cpus 2 --outdir annotations/GENOME --prefix GENOME GENOME.fna

 On my laptop, running one genome with a single core takes about 25 minutes. With 4 cpus, around 10 minutes.

However, since it's not easy to install prokka environment and it results time-consuming we will run it on https://usegalaxy.org/

change Force GenBank/ENA/DDJB compliance -> Yes

1) **Provide BUSCO results**
2) **How many genes were retrieved in the gff3 of PROKKA output. How can you count that? Provide some of the results.**