

Finding tRNA genes

For finding tRNA genes we recommend tRNA Scan. This uses a stochastic context free grammar to identify the tRNA genes. By default, tRNA scan tries to find eukaryotic tRNA genes, but you can easily change the specification to bacterial tRNA genes using the -B flag:

```
tRNAscan-SE -B -o trna.out assembly/scaffolds.fasta
```

You should hopefully find one tRNA gene per codon (at least) in your genome!

Note that the Lowe lab at UCSC has an online tRNA Scan server, although you maybe limited to upload limitations.

Note 2: You will need at least version 5 of the SDSU Computational Genomics Image to run tRNAscan-SE. You may need to upgrade your image.

Finding rRNA genes

For rRNA genes, we recommend barrnap for ribosomal RNA predictions.

barrnap has pre-built hidden Markov models for ribosomal RNA genes from Bacteria, Archaea, or Eukarya, and uses **nhmmer**, part of the HMMER 3.1 suite, to compare your sequences to the prebuilt hidden Markov model profiles.

barrnap is already installed on the AWS image and you can easily run it with this command:

```
barrnap assembly/scaffolds.fasta
```

The output will include some information about the program, and then the locations of the ribosomal RNA genes that have been identified. Recall that bacteria should contain 5S, 16S, and 23SrRNA genes. One of the problems that you are likely to run into is that the ribosomal RNA genes are so similar that the assembly algorithms tend to break at rRNA genes. Recall, from sequence assembly that assembly algorithms have a big problem with highly repetitive sequences.

If you are assembling a genome, you should pull out the sequences that map to the ribosomal RNA genes and assemble them separately, and then figure out the organization of the genome using a biological method like PCR or a long range sequencing method like PacBio or Nanopore.