

Codes running for assembly and annotation

1. Genomic assembly

software: Unicycler (v0.4.8)

Homepage: <https://github.com/rrwick/Unicycler>

Requirements:

Operating system: Linux or macOS

Python 3.4 or later

C++ compiler with C++14 support:

GCC 4.9.1 or later

Clang 3.5 or later

ICC also works

setuptools (only required for installation of Unicycler)

For short-read or hybrid assembly: SPAdes v3.14.0 or later

For long-read or hybrid assembly: Racon (racon)

For rotating circular contigs: BLAST+ (makeblastdb and tblastn)

Running code:

Hybrid assembly: `unicycler -1 JNDY-SC_1.clean.fq.gz -2 JNDY-SC_2.clean.fq.gz -l JNDY-SC.fq.gz --out assembly --threads 24`

2. Genome annotation

software: prokka 1.13

Homepage: <https://github.com/tseemann/prokka>

Requirements:

Operating system: Linux

CPU: 40 cores

Running code:

`prokka --force --outdir prokka_out --prefix assembly --kingdom Bacteria --gcode 11 --cpus 32 --evaluate 1e-5 --locustag assembly --addgenes --nammer --rfam --gram sig_peptide --compliant assembly.fasta`