

Racon on HPC

What is Racon?

Racon is a software tool used in the field of genomics for improving the accuracy of multiple sequence alignments. It uses a reference genome to guide the iterative correction of the alignment and refinement of the consensus sequence.

Racon is particularly useful for de novo genome assembly and for improving the accuracy of reads in relation to a reference genome. The software takes as input a set of overlapping reads and a reference genome, and produces a polished consensus sequence as output.

Racon operates by using a voting algorithm to correct mismatches and indels in the alignment, resulting in a more accurate representation of the target genome. The process is repeated until convergence, producing a final consensus sequence that is of higher quality compared to the input reads alone.

Links:

[GitHub](#)

Versions Available:

The following versions are available on the cluster:

- Racon v1.3.2

How to load Racon?

To load Racon, use the following commands:

```
#Load the Racon module
module load bio/racon/1.3.2
```

To verify if the module is loaded correctly, use the following command,

```
# List all the module loaded in the environment
module list
```

In a fresh environment, this only load Racon module without any dependencies.

How to use Racon?

To run Racon, you will need to provide the overlapping reads, reference genome, and any necessary parameters as input. The exact command will depend on the specific implementation you are using, so consult the Racon documentation for more information.

```
#See the syntax to run racon by
racon --help
```

Users can use the following sample slurm script to run racon,

```
#!/bin/bash
#SBATCH --job-name=racon_run
#SBATCH --partition=main
#SBATCH --qos main
#SBATCH --nodes=1
#SBATCH --cpus-per-task=12
#SBATCH --mem=32GB
#SBATCH --time=12:00:00
#SBATCH --output=racon_run.out
#SBATCH --error=racon_run.err

# load the necessary modules
module load bio/racon/1.3.2

# define the input and output files
reads_file="reads.fasta"
reference_file="reference.fasta"
```

```
output_file="consensus.fasta"

# run racon in parallel using multiple threads
racon -t $SLURM_CPUS_PER_TASK $reads_file $reference_file $output_file
```

For reference input file, use the following files provided in software github repository

```
#Download reference inputs file
svn checkout https://github.com/isovic/racon/trunk/test
```

Where to find help?

If you are confused or need help at any point, please contact OIT at the following address.

<https://ua-app01.ua.edu/researchComputingPortal/public/oitHelp>