

Porechop on HPC cluster:

What is Porechop?

Porechop is a tool for finding and removing adapters from Oxford Nanopore reads. Adapters on the ends of reads are trimmed off, and when a read has an adapter in its middle, it is treated as chimeric and chopped into separate reads. Porechop performs thorough alignments to effectively find adapters, even at low sequence identity.

The full documentation for Porechop is found in the following links:

[GitHub](#)

Versions Available:

- Porechop-v 0.2.3

How to load a version of Porechop?

To load a version of Porechop on the HPC, use the following command:

```
module load bio/porechop
```

Verify by using this command:

```
module list
```

The loaded software and runtime dependencies, python3.6, GDAL2 and GCC will be shown.

How to use Porechop on the cluster?

There are two methods to run Porechop on the cluster.

The Interactive Way:

To run the program interactively, follow the steps:

```
#Open a bash session on compute node
srun -p main --qos main -n 1 -c 12 --mem 10G --pty bash

#Load the module
module load bio/porechop

# Start your commands here
porechop --help
#Follow with commands to execute
```

This method is used for short job runs and debugging since the job terminates after the user exits the current terminal session.

The Script (Preferred):

To run a slurm job, the user must prepare input files. For this example, get input files with,

```
#Download the input files to the test directory
curl -o test.fastq
https://zenodo.org/record/3736457/files/1_control_18S_2019_minq7%20-
%20Copy.fastq?download=1
```

```
#Make a script
touch script.sbatch
```

Use the following script as a template,

```
#!/bin/bash
#SBATCH -p threaded
#SBATCH -q threaded
#SBATCH --mem-per-cpu=4G
#SBATCH -n 1
#SBATCH -c 16

#Load the module
module load bio/porechop

#Go to the test directory
cd $SLURM_SUBMIT_DIR

# Run Porechop
porechop -i test.fastq -o output_reads.fastq --threads 40
```

Schedule the job with the following sbatch command.

```
sbatch script.sbatch
```

All the processed files will be generated in the same directory as the script.

Where to find help?

If you are stuck on some part or need help at any point, please contact OIT at the following address.

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

