

Fastx Toolkit

The FASTX-toolkit (http://hannonlab.cshl.edu/fastx_toolkit) is a set of command line tools for processing of FASTA/FASTQ data files. It can be useful for any preprocessing of data, such as filtering by quality, base trimming etc.

FASTX-toolkit is installed on fourierseq, and is also available at TACC as a module: `module load fastx_toolkit`

For more usage information, check the documentation [here](#)

```
<fastx_command> -h will display the help screen for that command.  
Eg: fastx_trimmer -h
```

Note: When using FASTX-toolkit with Illumina HiSeq data from Casava 1.8, make sure to add -Q33 option. FASTX-toolkit assumes quality scores with ASCII offset 64, but the data is Sanger encoded (offset 33), so this needs to be specified explicitly by adding the -Q33 flag. This will apply to most recent GSAF data, and FASTX-toolkit will throw an error when it tries to read a FASTQ file unless the -Q33 option is specified.