

Converting FASTQ files to FASTA Files With “SED” at the Command Line

Introduction:

It is often useful to be able to convert fastq files generated from genomic sequencing applications to traditional fasta files. This can be done from the command line using the “SED” command. Sed is a standard command in the Bash terminal program so no need to install any programs. The logic is that fastq files consist of 4 lines per sequence and fasta files consist of two lines per sequence. So you simply need to eliminate the two extra fastq file lines and start the new fasta file with a “>” character. This command should work very quickly from the command line.

Code:

```
sed -n '1~4s/^@/>/p;2~4p' INFILE.fastq > OUTFILE.fasta
```

Obviously, you should replace “INFILE” que the name of the fastq file you wish to convert and “OUTFILE” with the name of the new fasta file. Also, make sure to type the command from within the directory in which the infile is located.