Multiple ways to download public sequencing data

Lingzi Li

April 27, 2019

Option 1: Direct download of fastq files

Download fastq files from ENA

Copy the "FASTQ files (FTP)" link of the sequencing data on https://www.ebi.ac.uk/ena

```
wget ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR300/001/SRR3000761/SRR3000761.fastq.gz
```

Download fastq files from NCBI (only with SRR accession number, no path required)

Find out the SRR accession number of the sequencing data on https://www.ncbi.nlm.nih.gov/sra

```
fastq-dump <SRR accession>

# As "fastq-dump"" is time-consuming, it is better to add "nohup &" to the command line. Like this, eve
nohup fastq-dump <SRR accession> &
```

Option 2: Download SRA files first and then convert to fastq files

Download SRA files from NCBI website

Find out the SRR accession number of the sequencing data on https://www.ncbi.nlm.nih.gov/sra

```
wget ftp://ftp-trace.ncbi.nih.gov/sra/sra-instant/reads/ByRun/sra/SRR/SRR300/SRR3000759/SRR3000759.sra
```

Download SRA files from ENA

Find out the SRR accession number of the sequencing data on https://www.ebi.ac.uk/ena

```
wget ftp://ftp.sra.ebi.ac.uk/vol1/srr/SRR300/000/SRR3000760
```

Convert SRA file to fastq file

For a single SRA file

```
nohup fastq-dump SRR3000759.sra &
```

For multiple SRA files

```
nohup for i in SRR*; do fastq-dump -A $i; done &
```

Further readings

To check out features of the command line, type:

```
nohup --help
```

To check out the related NCBI documentation of the SRA Toolkit, see: Downloading SRA data using command line utilities, and Using the SRA Toolkit to convert SRA files into other formats