

Reads Trimming (Trim Galore)

[Trim Galore](#) is a Perl wrapper around Cutadapt and FastQC to consistently apply adapter and quality trimming to FastQ files. We will use this tool for quality trimming, adapter trimming, and removing short sequences.

Install Trim Galore

Before installation, ensure that [Cutadapt](#) and [FastQC](#) are already installed.

```
# Check the version of cutadapt
$ cutadapt --version

# Check the version of FastQC
$ fastqc -v
```

Install the latest version of Trim Galore from [Github](#) or [project website](#):

```
# Install Trim Galore
$ curl -fsSL https://github.com/FelixKrueger/TrimGalore/archive/0.6.6.tar.gz -o trim_galore.tar.gz
$ tar xvf trim_galore.tar.gz

# Verify installation
$ trim_galore -v
```

Adaptive Quality and Adapter Trimming

In this procedure, first, low-quality base calls are trimmed off from the 3' end of the reads before adapter removal. Next, adapter sequences from the 3' end of reads are detected and removed by cutadapt. Lastly, trimmed short sequences (default: < 20bp) are filtered.

```
#!/bin/bash
trimGalore(){
  trim_galore -o /path/to/trim_galore_result/ /path/to/raw_data/homo/$1.fastq
}
export -f trimGalore
for s in SRR5978827 SRR5978828 SRR5978829 SRR5978834 SRR5978835 SRR5978836
SRR5978869 SRR5978870 SRR5978871 SRR5179446 SRR5179447 SRR5179448
do
  trimGalore ${s}
done
```

```
#!/bin/bash
trimGalore(){
trim_galore -o /path/to/trim_galore_result/ /path/to/raw_data/mm10/$1.fastq
}
export -f trimGalore
for s in SRR866997 SRR866998 SRR866999 SRR867000 SRR867001 SRR867002 SRR866991
SRR866992 SRR866993 SRR866994 SRR866995 SRR866996
do
trimGalore ${s}
done
```

For trimming paired-end data, you need to add a `--paired` option in the `trim_galore` command .

```
trim_galore --paired -o /path/to/trim_galore_result/ *_1.fastq *_2.fastq
```

Outputs

Trim Galore produced two output files for each FastQ file: one text file ("SRR5978869.fastq_trimming_report.txt") and a trimmed FastQ file ("SRR5978869_trimmed.fq").

1. The text file

The text file provides a summary of running parameters.

```
# To see the first few lines of the text file
$ head SRR5978869.fastq_trimming_report.txt
```

```
SUMMARISING RUN PARAMETERS
=====
Input filename: SRR5978869.fastq
Trimming mode: single-endw
Trim Galore version: 0.6.4_dev
Cutadapt version: 3.2
Number of cores used for trimming: 1
Quality Phred score cutoff: 20
Quality encoding type selected: ASCII+33
```

2. The trimmed FastQ

The trimmed FastQ file can be used for further analysis.