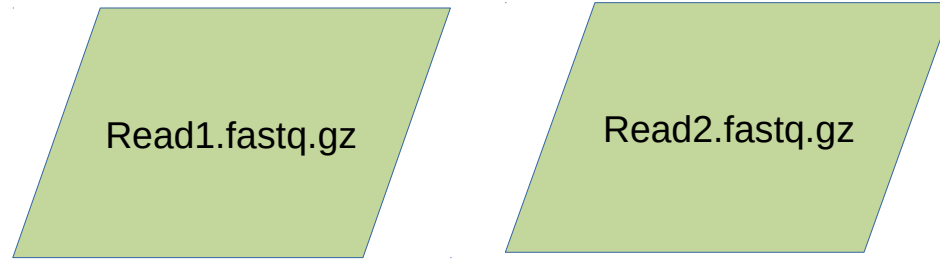
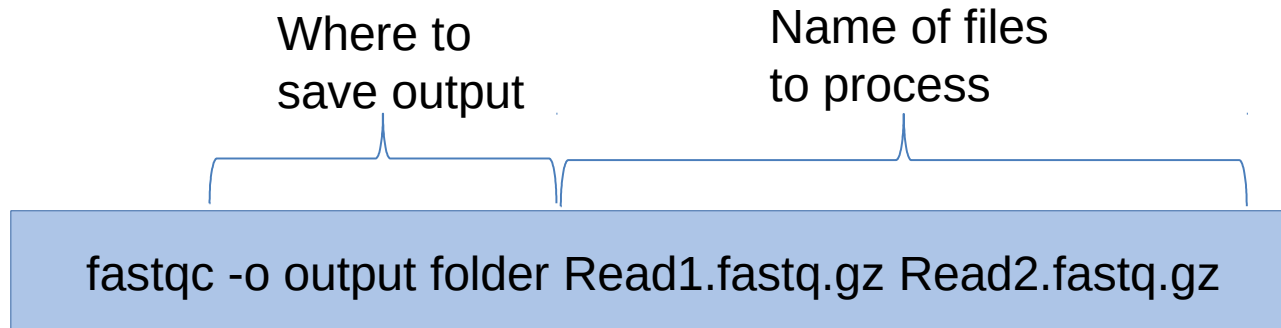


Input:



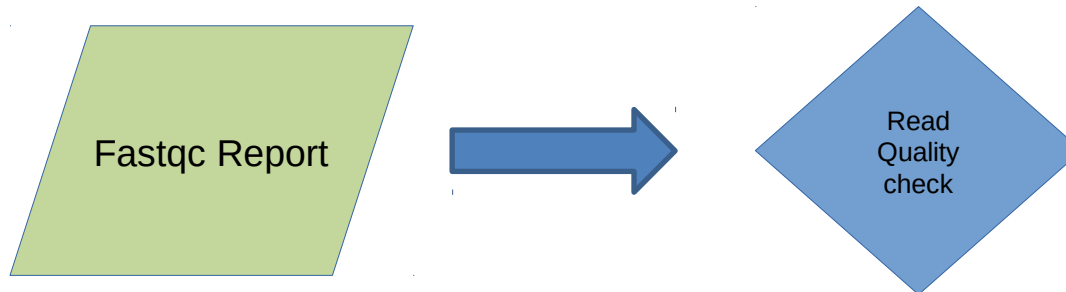
Raw data from the sequencing machine, in fastq format and zipped.

Command Line:



- Start with raw sequencing data, in fastq format and zipped.
- Run Fastqc, a program that summarizes the quality of reads. Also outputs a number of useful metrics. Generates an html report.

Output:



Are the reads of sufficient quality to proceed?

Summary

- ✓ [Basic Statistics](#)
- ✓ [Per base sequence quality](#)
- ✓ [Per tile sequence quality](#)
- ✓ [Per sequence quality scores](#)
- ✗ [Per base sequence content](#)
- ✗ [Per sequence GC content](#)
- ✓ [Per base N content](#)
- ✓ [Sequence Length Distribution](#)
- ✗ [Sequence Duplication Levels](#)
- ! [Overrepresented sequences](#)
- ✗ [Adapter Content](#)
- ✗ [Kmer Content](#)

Links to other reports

✓ Basic Statistics

Measure	Value
Filename	293-412-5-12-16-16-B-R6_S20_L001_R1_001.fastq.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	849986
Sequences flagged as poor quality	0
Sequence length	151
%GC	46

Fastqc Report

Preliminary information, number of sequences in file, average sequence length, etc.

✓ Per base sequence quality

