

Basic quality control with FastQC

June 2021

Differential Gene Expression Analysis Workflow

Fastq file format

Fastq file format - Headers

Fastq file format - Sequences

Fastq file format - Quality Scores

QC is important

Check for any problems before we put time and effort into analysing potentially bad data

- ▶ Start with FastQC
 - ▶ Quick
 - ▶ Outputs an easy to read html report

We run fastQC from the terminal with the command

fastqc <fastq>

but there are lots of other parameters which you can find to tailor your QC by typing

fastqc -h

Per base sequence quality

Good Data

Bad Data

Per base sequence content

Good Data

Bad Data

Per sequence GC content

Good Data

Bad Data

Adaptor content

Good Data

Bad Data

And now onto the exercise. . .

- ▶ Go to: <https://ushers.bio.cam.ac.uk/guacamole2>
- ▶ Log on with YOUR credentials that were emailed to you