

Bioinformatics Lessons Schedule

Date	Subject
01-07	Process RNA-seq
01-14	Process RNA-seq, continued
01-21	Process RNA-seq, continued
01-28	Analyze RNA-seq
02-04	Analyze RNA-seq, continued
02-11	Analyze RNA-seq, continued
02-18	skipped
02-25	Process RRBS
03-03	Process RRBS, continued
03-10	Process RRBS, continued
03-17	Analyze RRBS
03-24	Analyze RRBS, continued
03-31	???

What are we going to do next?

What are we going to do next?

1. Everyone works through R for Data Science to hone their R skills with help available if they have questions.

What are we going to do next?

1. Everyone works through R for Data Science to hone their R skills with help available if they have questions.
2. Meetup – 1 or a few people show something casually (could be less frequent than every week)
 - Code review
 - Demonstrations of cool code or packages
 - How people have used the code they've learned

What are we going to do next?

1. Everyone works through R for Data Science to hone their R skills with help available if they have questions.
2. Meetup – 1 or a few people show something casually (could be less frequent than every week)
 - Code review
 - Demonstrations of cool code or packages
 - How people have used the code they've learned
3. Nothing

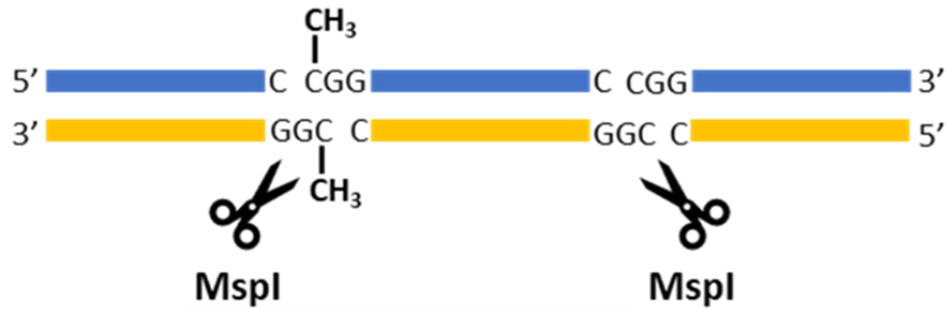
Process Reduced Representation Bisulfite Sequencing (RRBS)

2020-02-24

Reduced Representation Bisulfite Sequencing

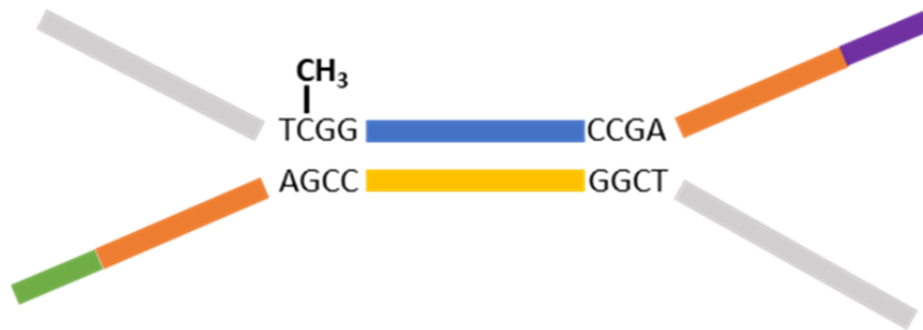
Reduced Representation Bisulfite Sequencing

1. MspI digestion

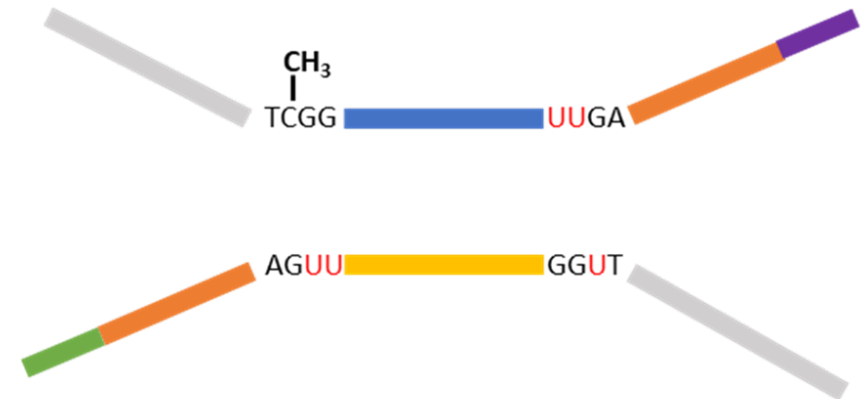


2. End repair

3. Adaptor Ligation and size selection



4. Bisulfite conversion

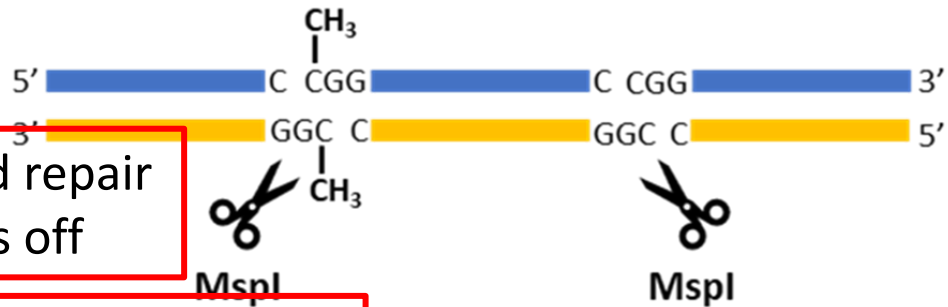


5. PCR amplification

6. Library preparation & sequencing

Reduced Representation Bisulfite Sequencing

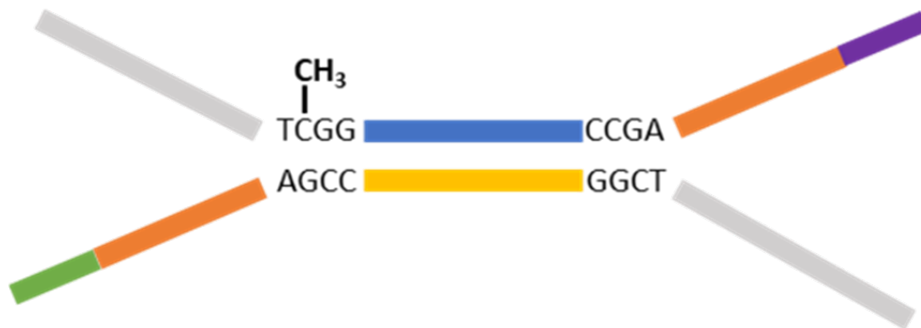
1. MspI digestion



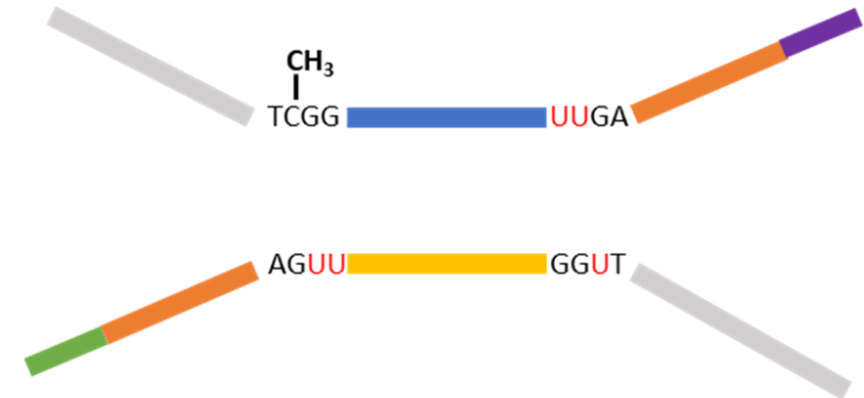
Trim end repair
bases off

2. End repair

3. Adaptor Ligation and size selection



4. Bisulfite conversion

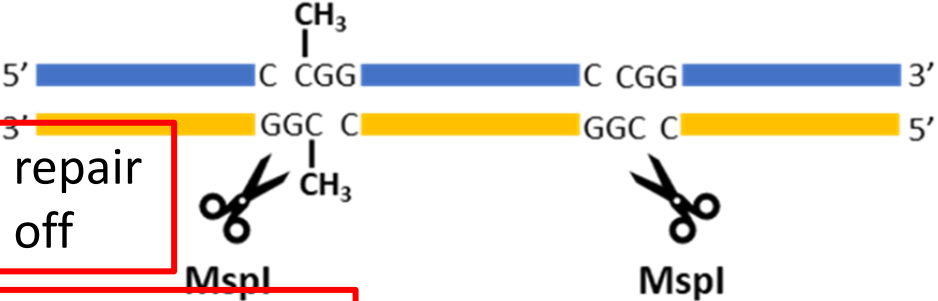


5. PCR amplification

6. Library preparation & sequencing

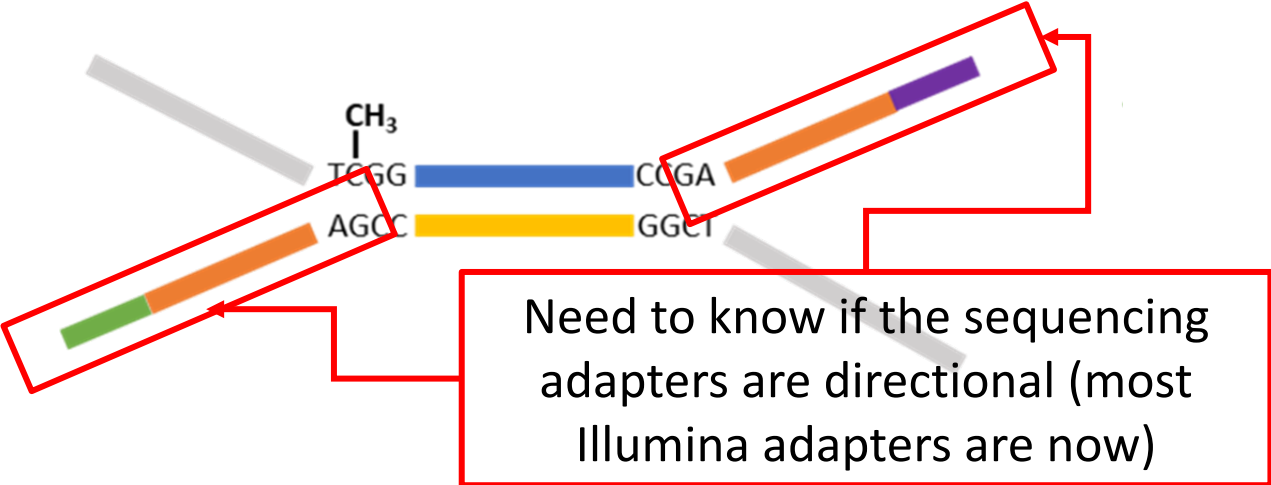
Reduced Representation Bisulfite Sequencing

1. MspI digestion

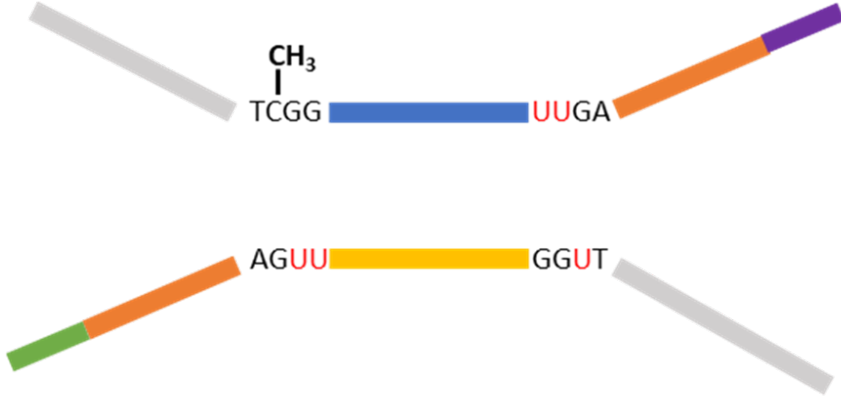


2. End repair

3. Adaptor Ligation and size selection



4. Bisulfite conversion

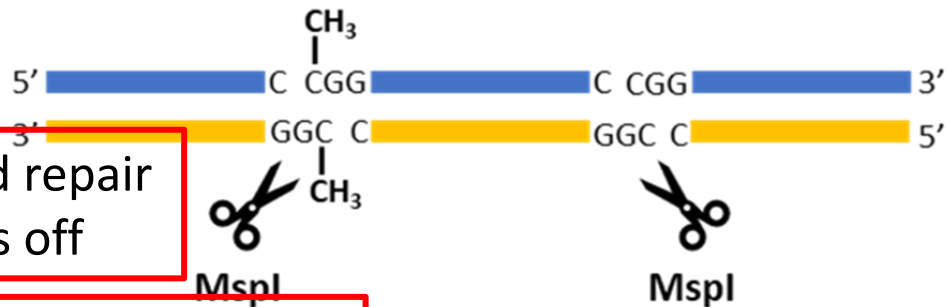


5. PCR amplification

6. Library preparation & sequencing

Reduced Representation Bisulfite Sequencing

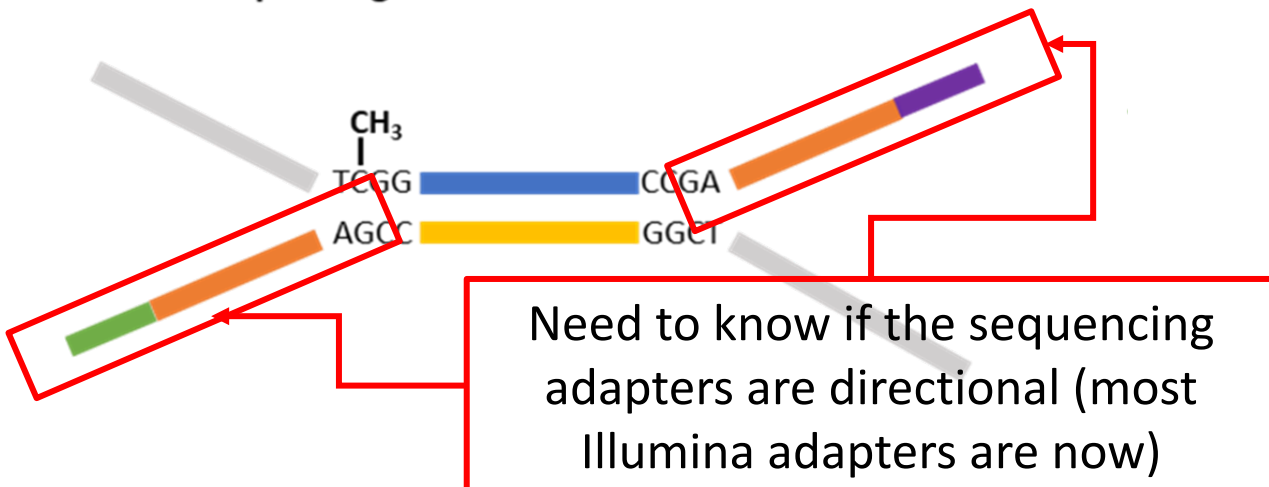
1. MspI digestion



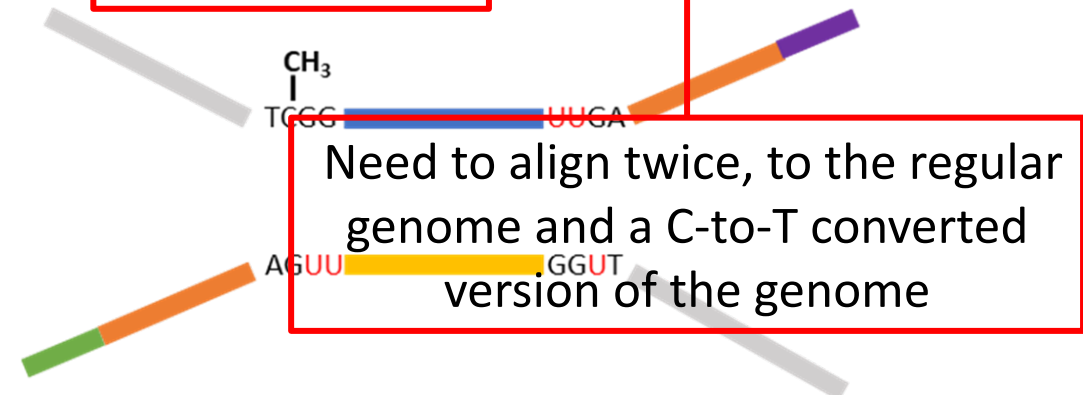
Trim end repair bases off

2. End repair

3. Adaptor Ligation and size selection



4. Bisulfite conversion



5. PCR amplification

6. Library preparation & sequencing

The Data

- Blood from healthy controls from the National Institute of Neurological Disorders and Stroke (NINDS) repository here at Coriell
- Woonbok did the RRBS as part of a larger study to identify CpGs that are good biomarkers of aging (we've talked about it at length in lab meetings)
- I subsampled 10,000 reads from every file so we could run the commands quickly

File Prefix	Sex	Age
F22_sub10000	female	22
F24_sub10000	female	24
F25_sub10000	female	25
F77_sub10000	female	77
F78_sub10000	female	78
F80_sub10000	female	80

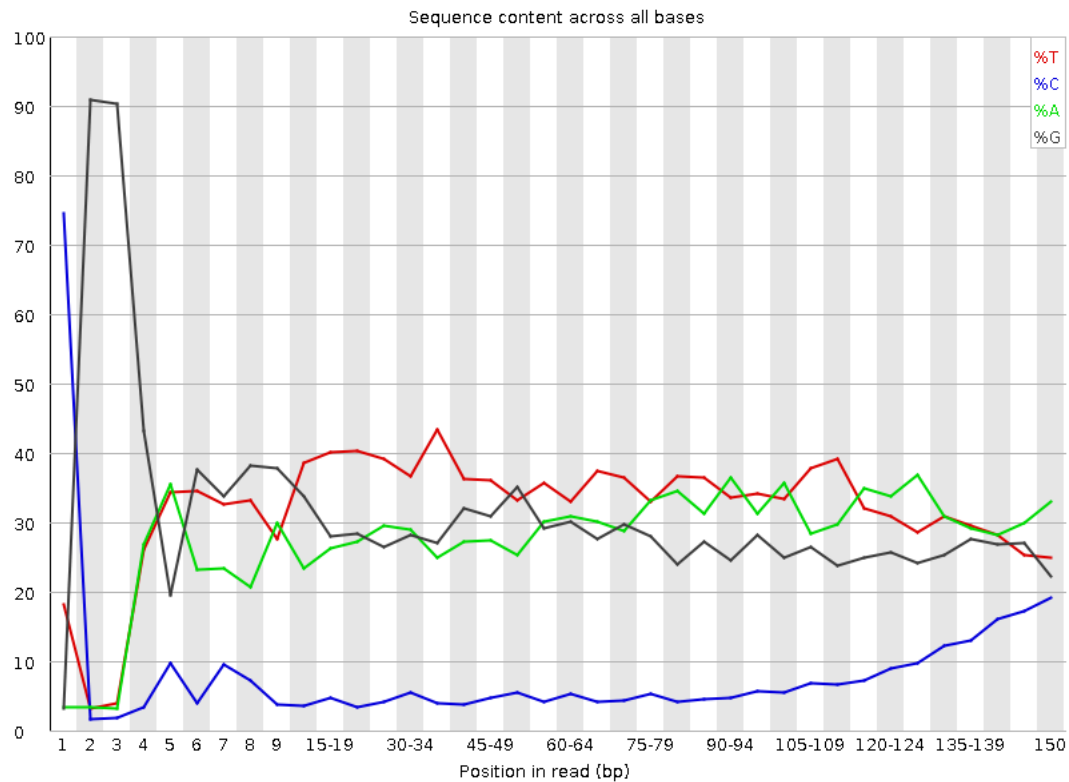
Step 1: Quality Check w/ FastQC

```
fastqc *.fq.gz -o fastqc/
```

FastQC: Per base sequence content

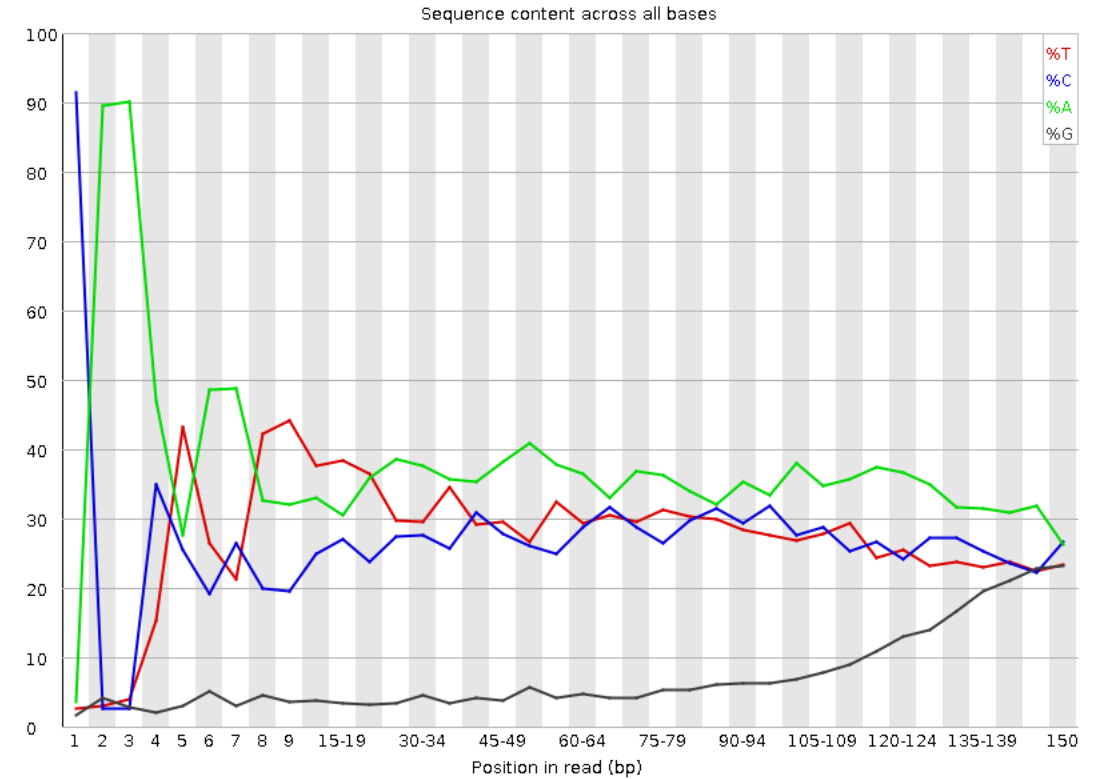
Read 1

✖ Per base sequence content



Read 2

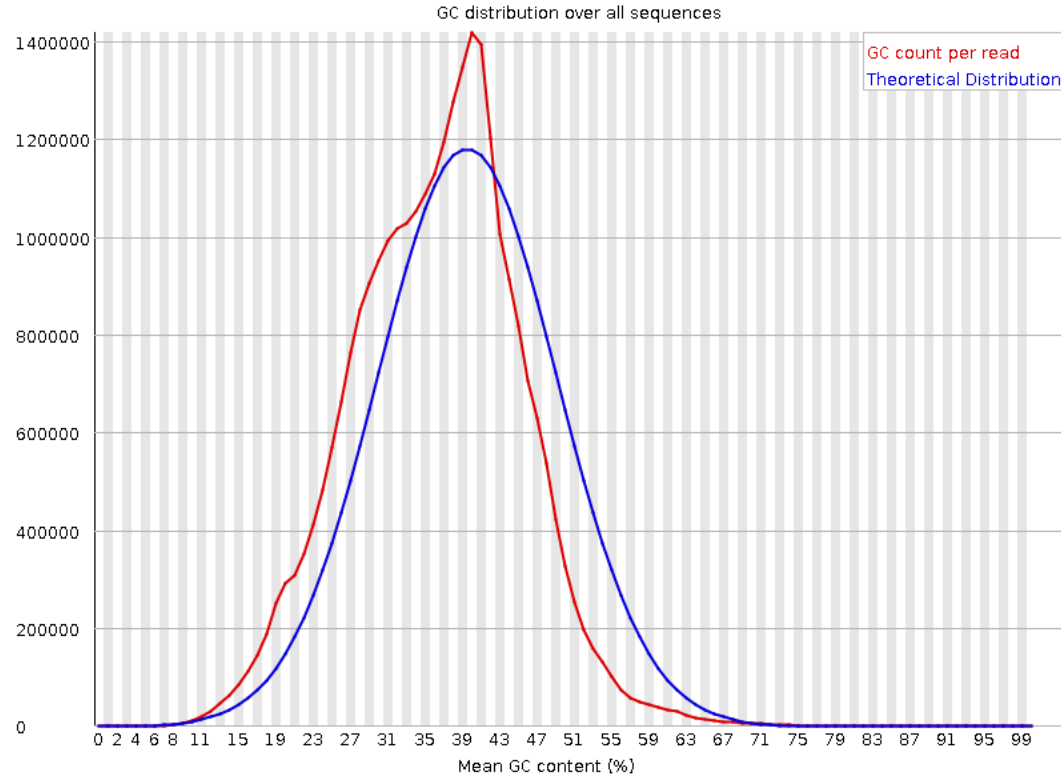
✖ Per base sequence content



FastQC: Per sequence GC content

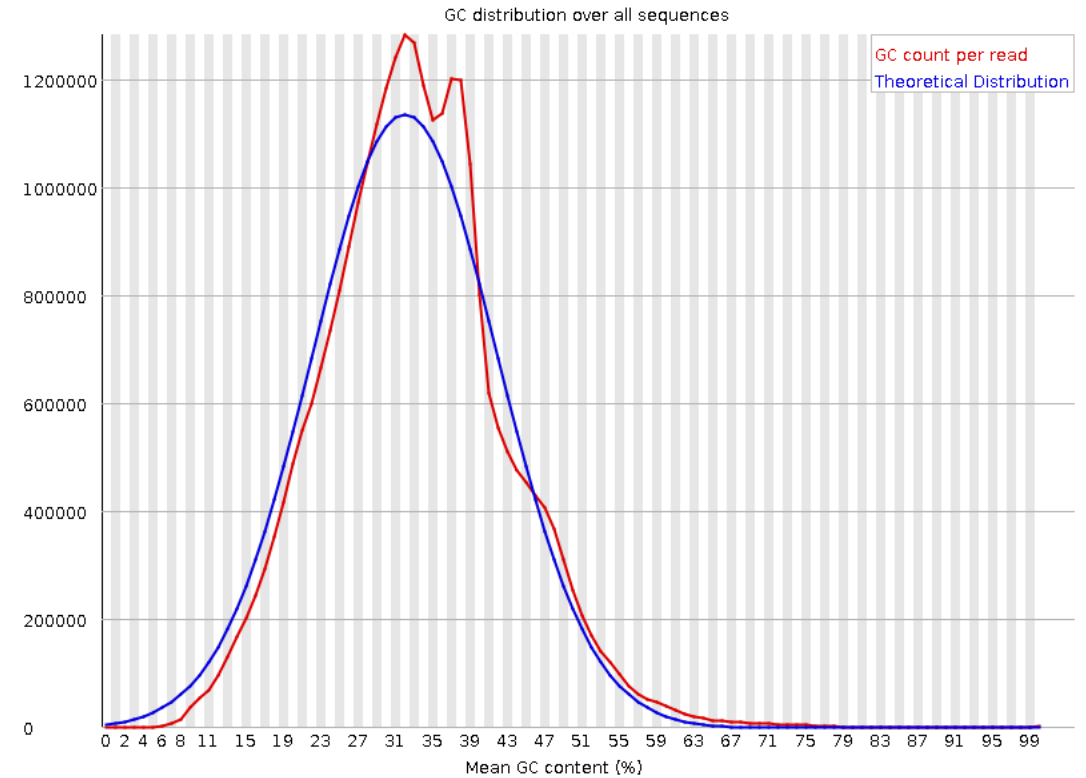
Read 1

! Per sequence GC content



Read 2

✓ Per sequence GC content

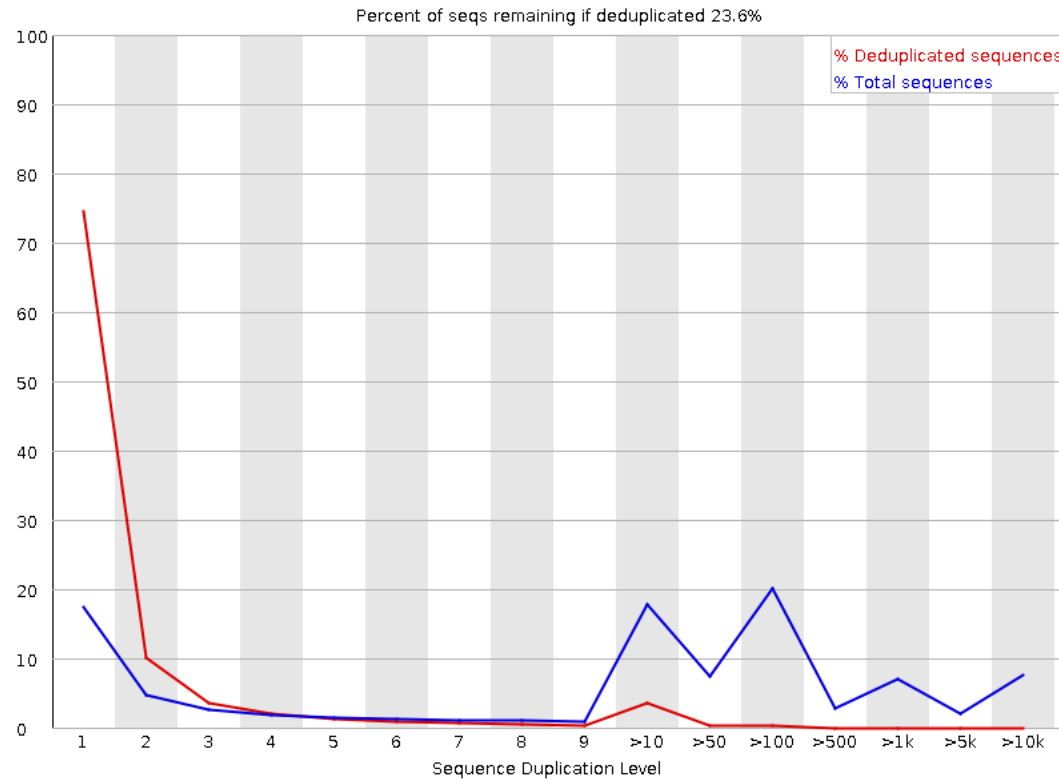


FastQC: Per sequence GC content

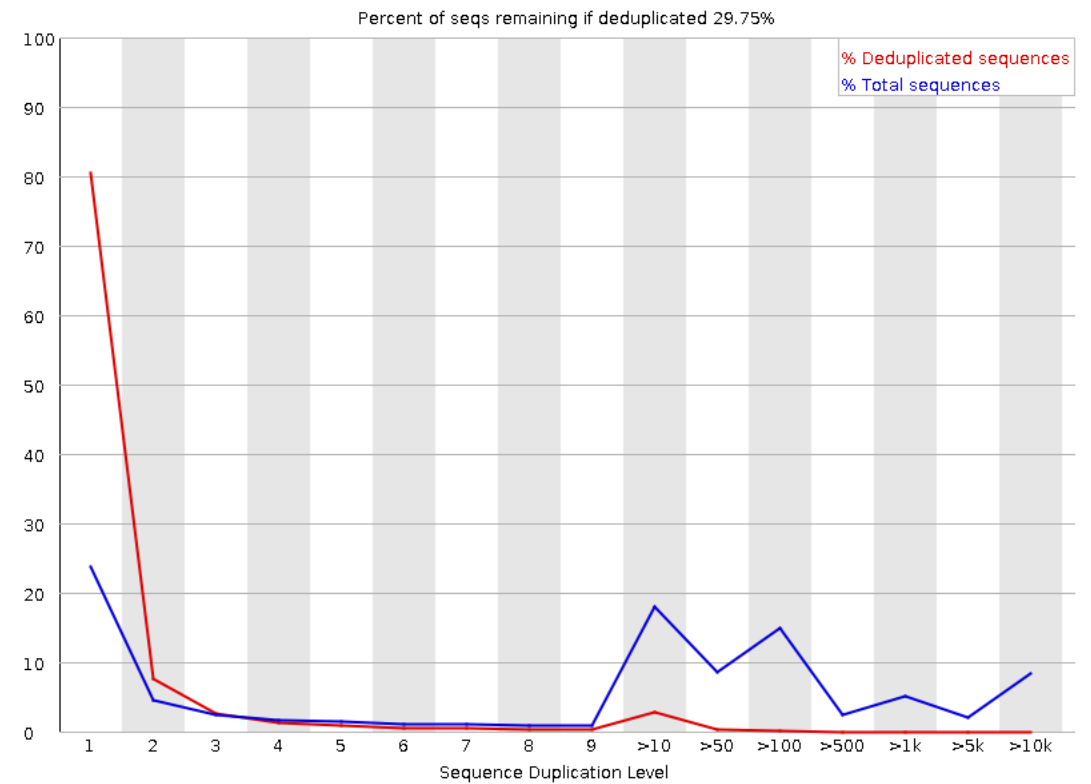
Read 1

Read 2

Sequence Duplication Levels



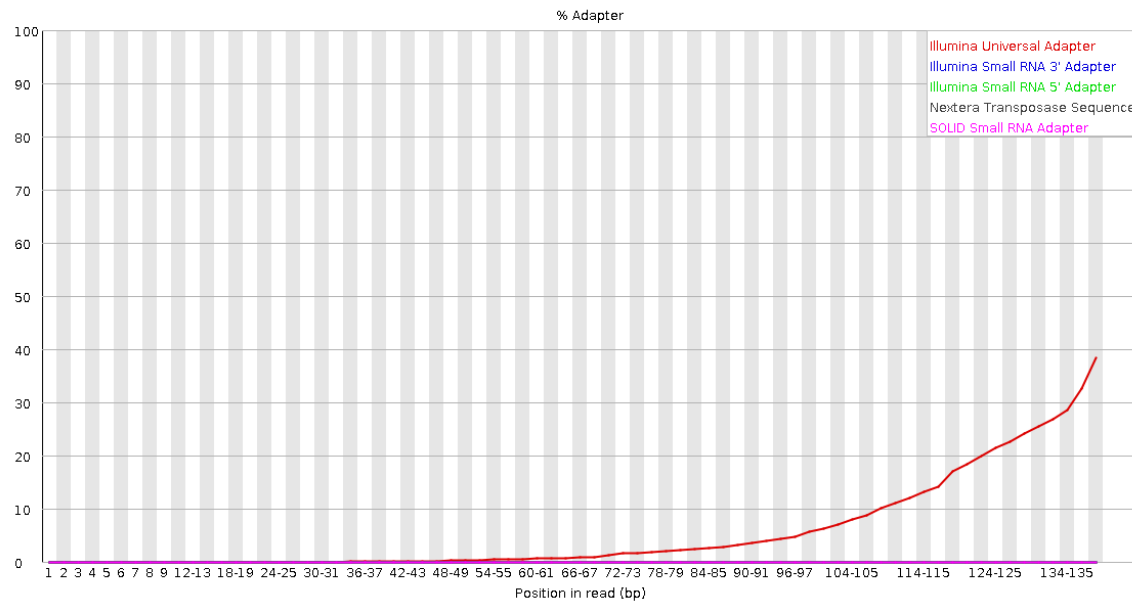
Sequence Duplication Levels



FastQC: Adapter Content

Read 1

✖ Adapter Content



Read 2

✖ Adapter Content

