Homework day 1

Author: Lynn Sanford, 2021

1. Videos for day 1 (if not already done)
2. Videos for day 2
3. Vimtutor
     a. Lessons 1 and 2
     b. Advanced students should do lessons 3, 4, 5
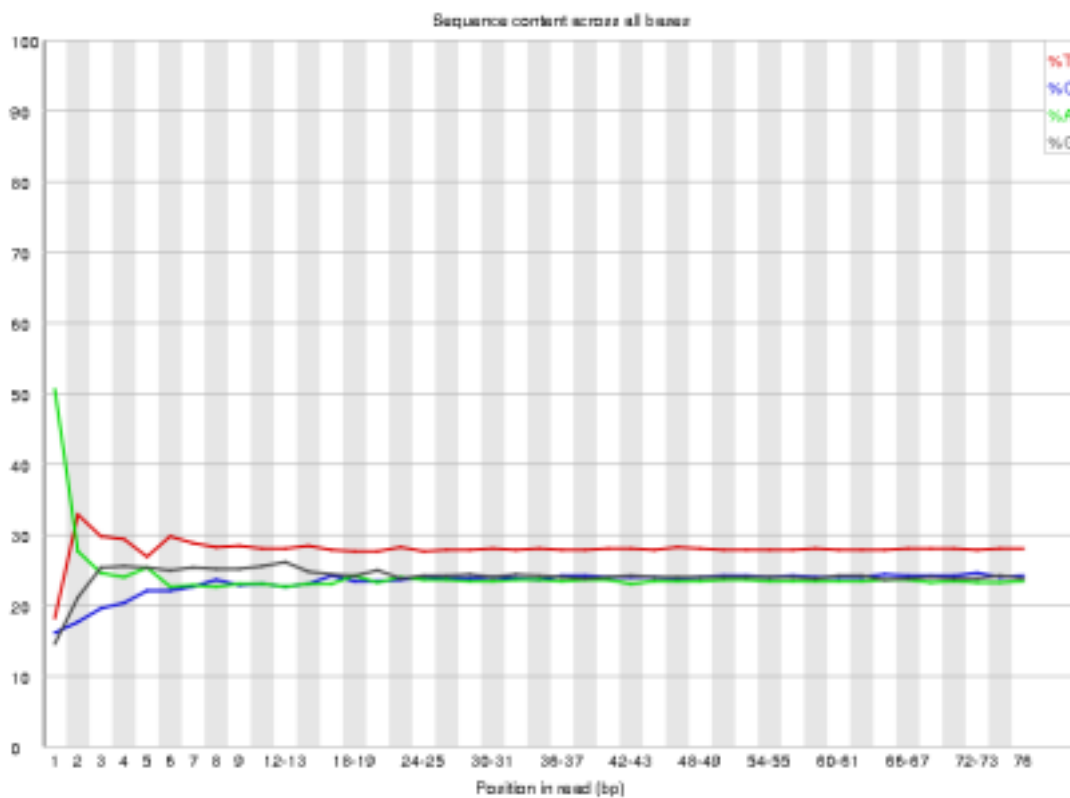4. Library QC challenge (see next page)

Library QC Challenge

In class, we talked about how to read FastQC reports in the context of determining if there are overarching problems with a sequencing library.

Here, your job is to look at some FastQC reports to determine if a library looks good or whether it has problems (and what those problems might be). For each library, briefly answer the following questions:
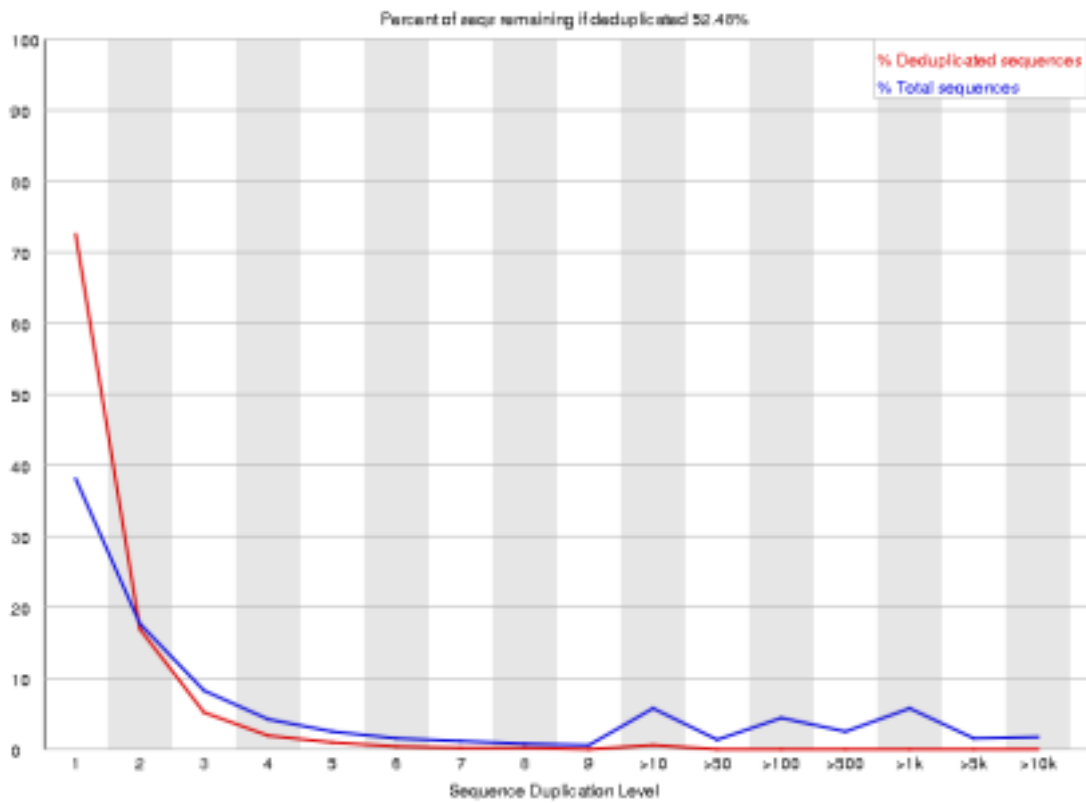
1. Does this library pass general QC?
2. If not:
   a. What QC metric(s) show problems with the library?
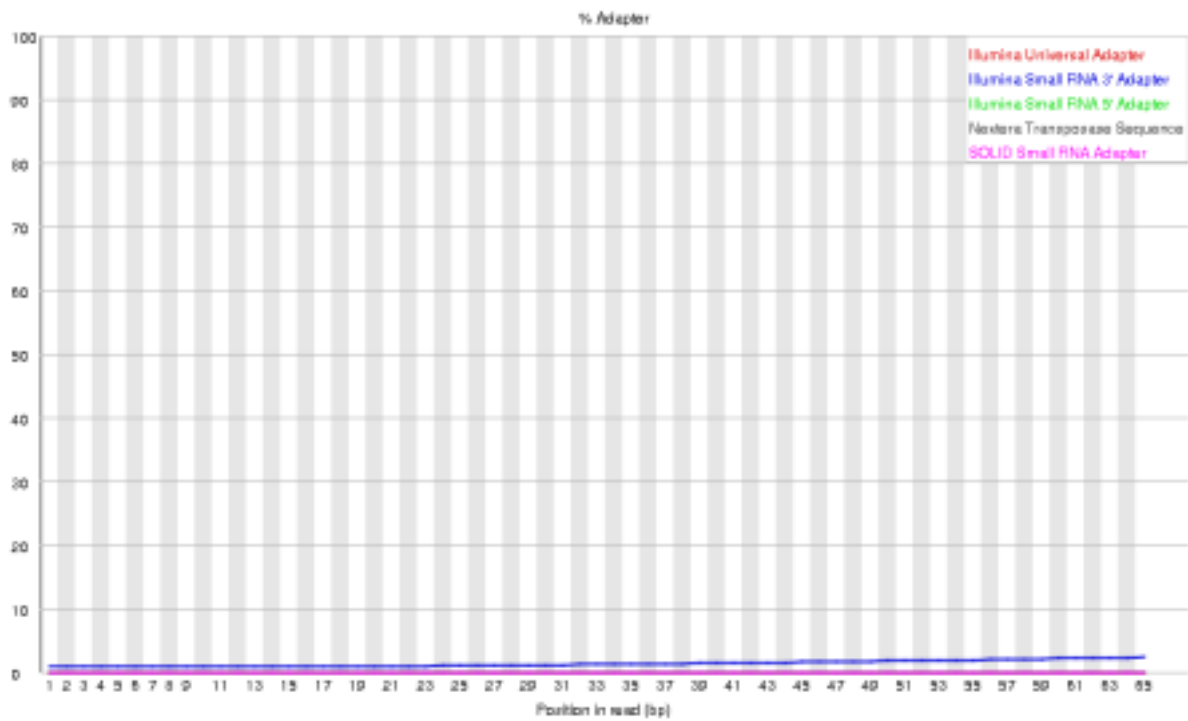   b. What might be the biological or technological source of the problem(s)?

Library 1
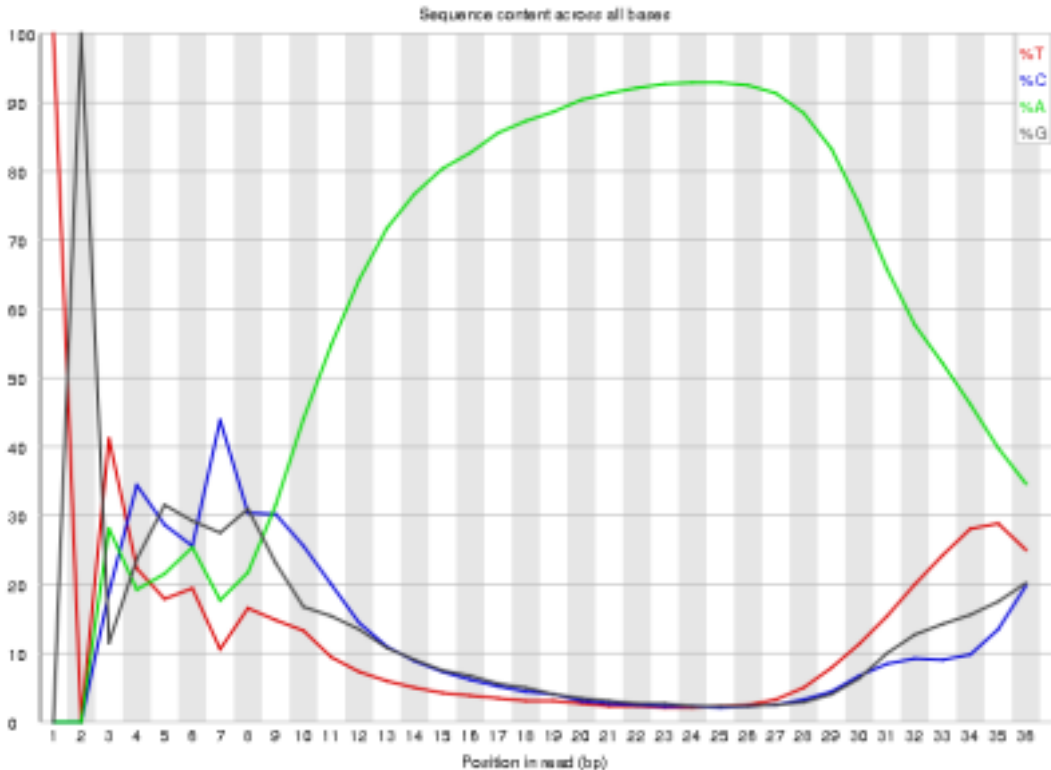
Per base sequence content
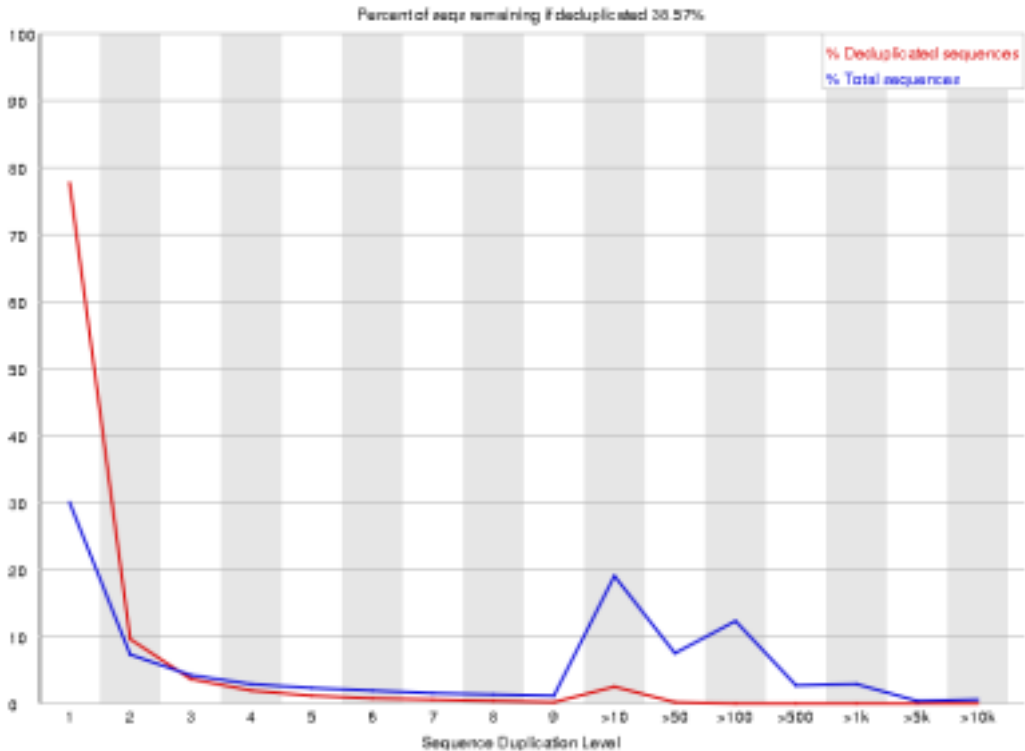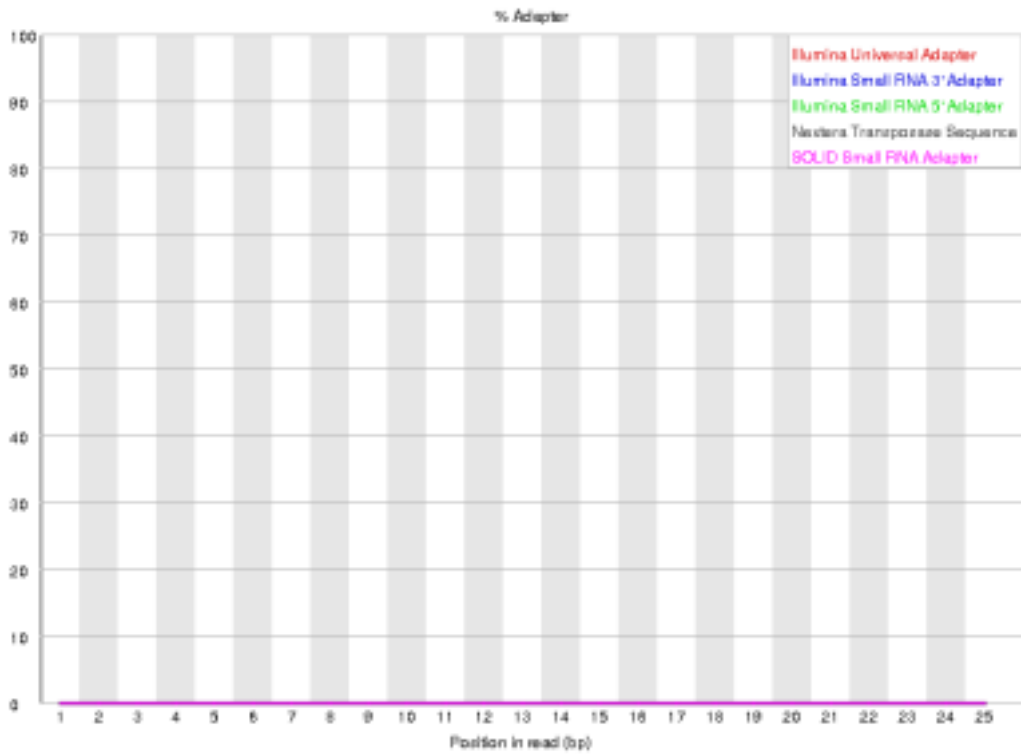
Sequence Duplication Levels

Percent of seqs remaining if deduplicated 52.46%

% Deduplicated sequences
% Total sequences

Sequence Duplication Level

Adapter Content

% Adapter

Illumina Universal Adapter
Illumina Small RNA 3' Adapter
Illumina Small RNA 5' Adapter
Nextera Transposase Sequence
SOLID Small RNA Adapter

Position in read (bp)

Library 2

Per base sequence content

Sequence content across all bases

Sequence Duplication Levels

Percent of seqs remaining if deduplicated 38.57%
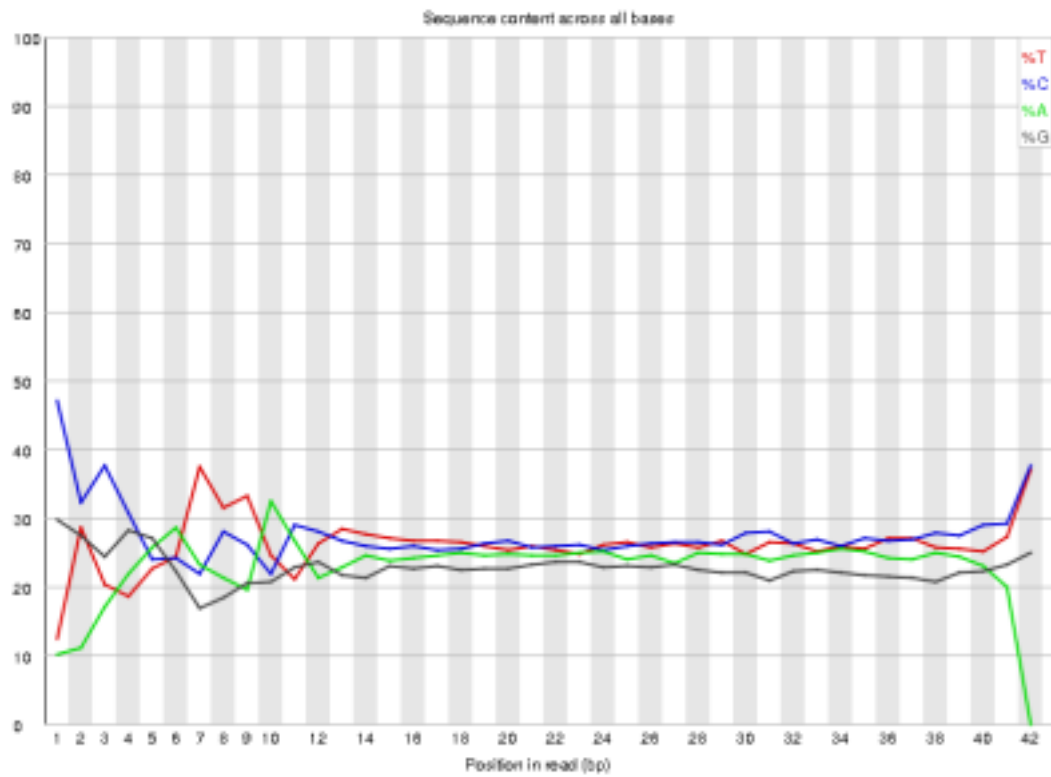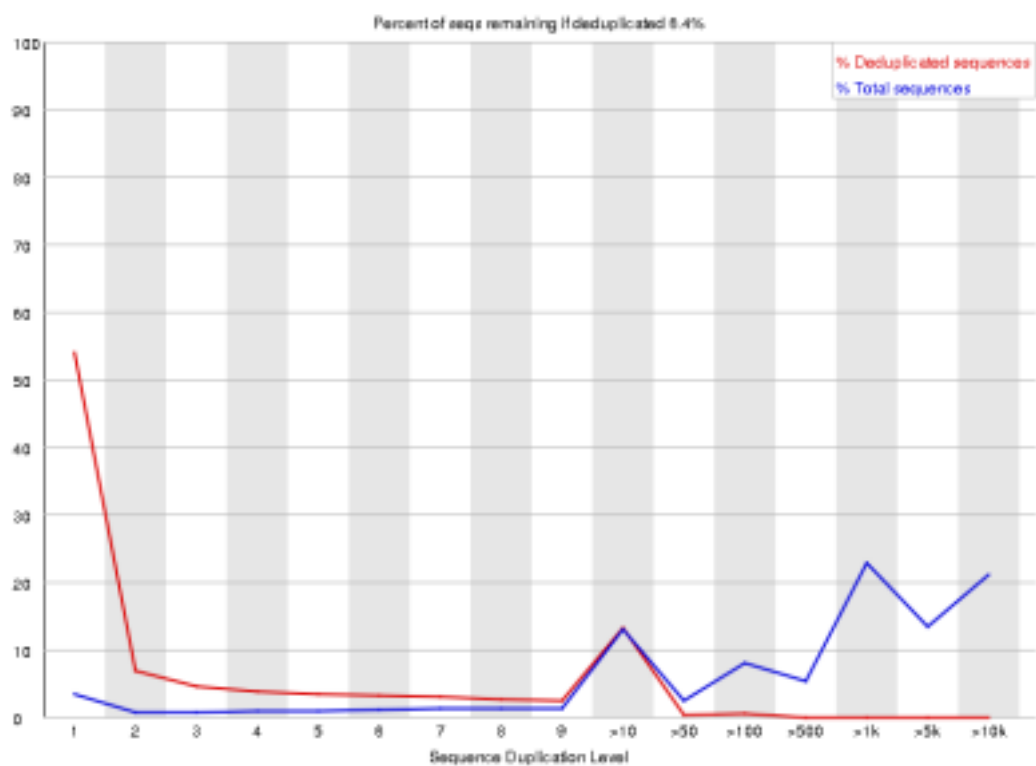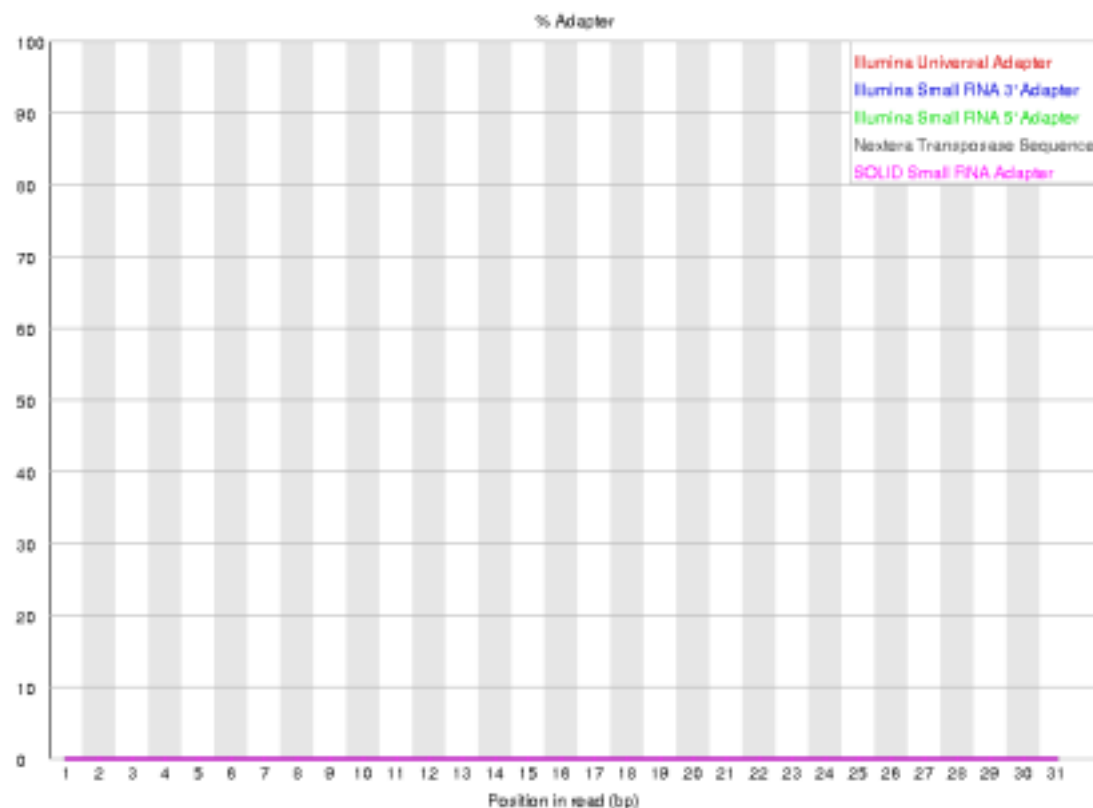
Adapter Content

Library 3

Per base sequence content

Sequence Duplication Levels

Percent of seqs remaining if deduplicated 8.4%

Adapter Content

% Adapter

Library 4

Per base sequence content

Sequence content across all bases

Sequence Duplication Levels

Percent of seqs remaining if deduplicated 55.93%
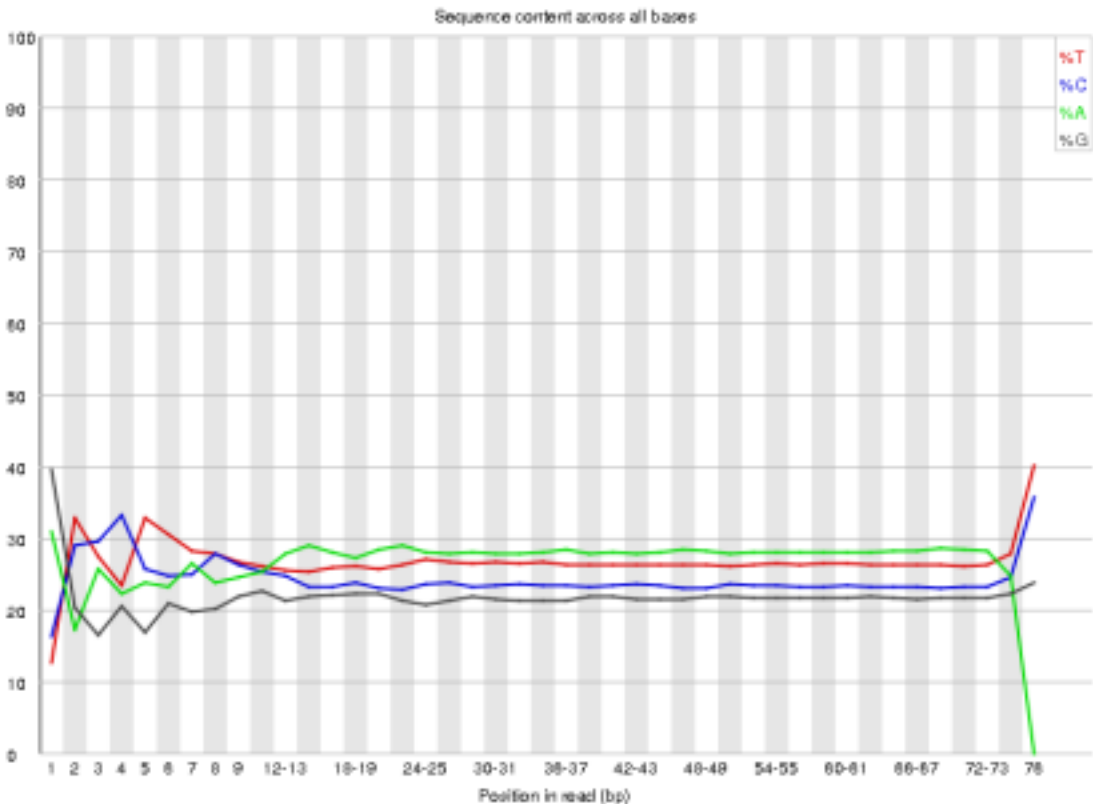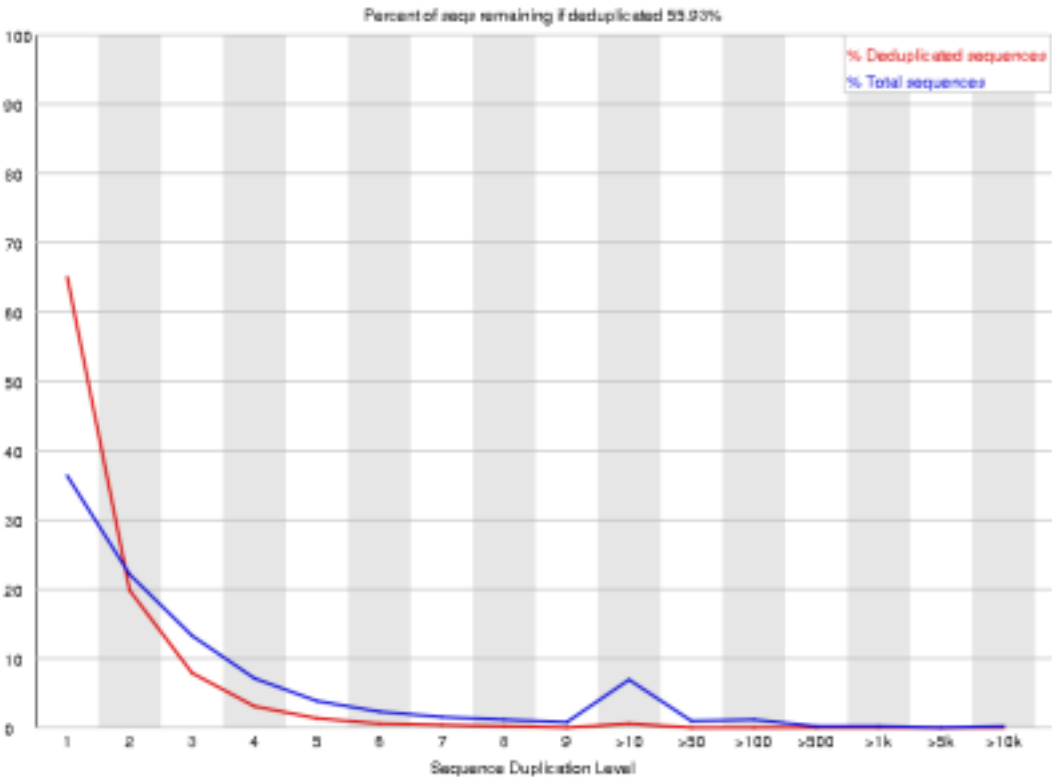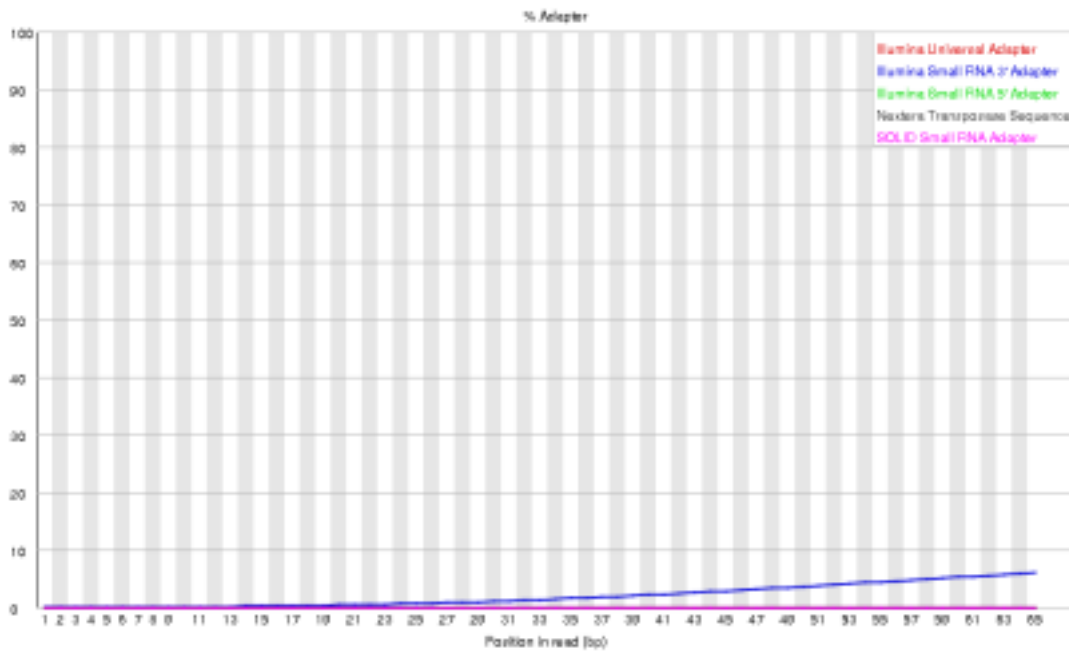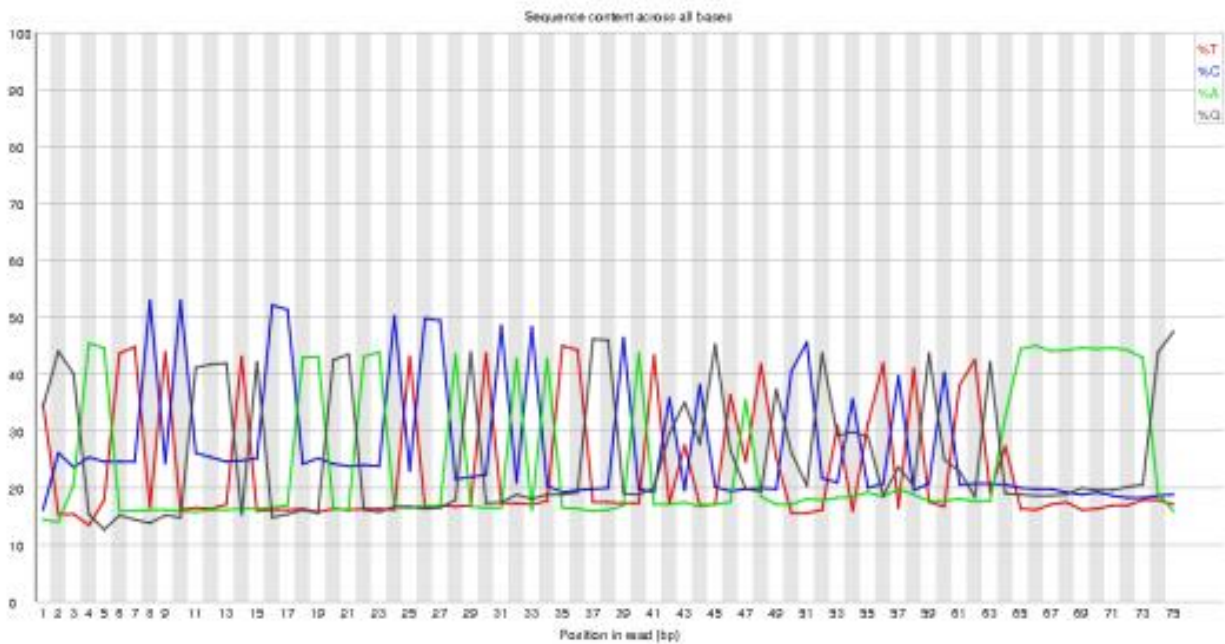
Adapter Content

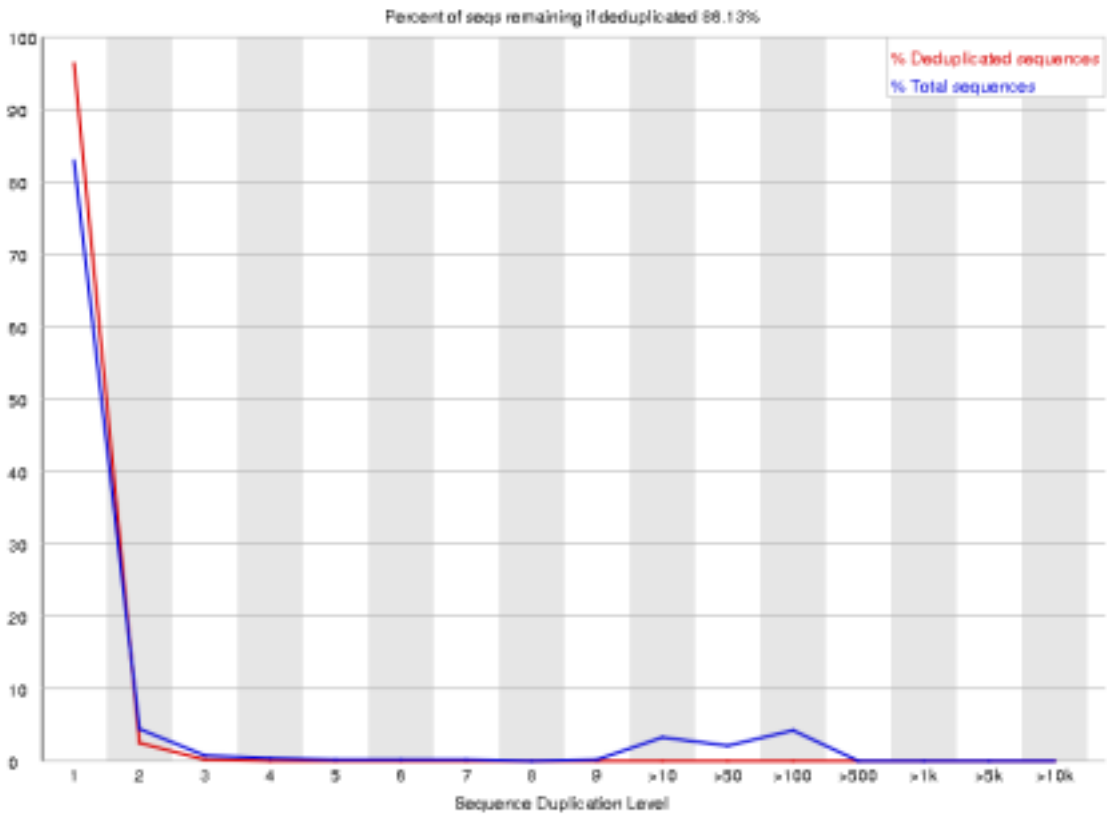# Library 5

Per base sequence content

Sequence Duplication Levels

Percent of seqs remaining if deduplicated 86.13%

Legend:
% Deduplicated sequences
% Total sequences

X-axis: Sequence Duplication Level
(1, 2, 3, 4, 5, 6, 7, 8, 9, >10, >50, >100, >500, >1k, >5k, >10k)

Adapter Content

% Adapter

Legend:
Illumina Universal Adapter
Illumina Small RNA 3' Adapter
Illumina Small RNA 5' Adapter
Nextera Transposase Sequence
SOLID Small RNA Adapter

X-axis: Position in read (bp)