

# Day 1 overview

- SSH keys and logging into the AWS
- Basic Bash/Intro to Git
- Vim and Vimtutor
- Variables in Bash
- Review library prep and sequencing (non-DS2)

# Logging into the supercomputer

- Open and follow the worksheet **SSH\_AWS**  
([https://github.com/Dowell-Lab/srworkshop/tree/main/day01/worksheets/SSH\\_AWS.md](https://github.com/Dowell-Lab/srworkshop/tree/main/day01/worksheets/SSH_AWS.md))
- You MUST
  1. Be logged into your GitHub account
  2. Have a terminal application open (see worksheet for details)
- After finishing the worksheet, try to log into the AWS:  

```
ssh <github_username>@3.20.14.225
```

**\*\*replace <github\_username>, including brackets, with your username**

**If you need help, flag us with a red sticky note**

**If you are successfully logged in, put up a green sticky note**

# Basic Bash and Git/GitHub

- Open and follow the worksheet **Git\_github\_bash**

([https://github.com/Dowell-Lab/srworkshop/tree/main/day01/worksheets/Git\\_github\\_bash.md](https://github.com/Dowell-Lab/srworkshop/tree/main/day01/worksheets/Git_github_bash.md))

- This worksheet orients you to the class GitHub repository and gets you started with some basic Bash navigation

- Main commands:

<code>pwd</code>	Print working (current) directory
<code>cd</code>	Change directory
<code>ls</code>	List contents

**If you need help, flag us with a red sticky note**

**If you are done, put up a green sticky note**

# Break

If you do not have a working terminal,  
please log into Google Shell now

([.../srworkshop/day01/worksheets/Google\\_shell.md](#))

# Vim and Vimtutor

- What is Vim?
  - Text editor – read, write, and save text files
  - Entirely keyboard-based
  - You CANNOT use your mouse to move the cursor!!!
- Vimtutor is on every Unix/Linux system and teaches you how to use Vim
- Open and follow the worksheet **Vimtutor\_crashes\_colors**

(.../srworkshop/day01/worksheets/Vimtutor\_and\_Vim\_crashes.md)

**If you need help, flag us with a red sticky note**

**If you are done, put up a green sticky note**

# Variables in Bash

- Open and follow the worksheet **Variables\_in\_bash**

(.../srworkshop/day01/worksheets/Variables\_in\_bash.md)

- Reference:

Defining a variable

`variable=value`

Calling a variable

`$variable`

Viewing a variable value

`echo $variable`

Other variable calling notations

`"$variable"`    `${variable}`

**If you need help, flag us with a red sticky note**

**If you are done, put up a green sticky note**

# Day 1 overview

- SSH keys and logging into the AWS
- Basic Bash/Intro to Git
- Vim and Vimtutor
- Variables in Bash
- Review library prep and sequencing (non-DS2)

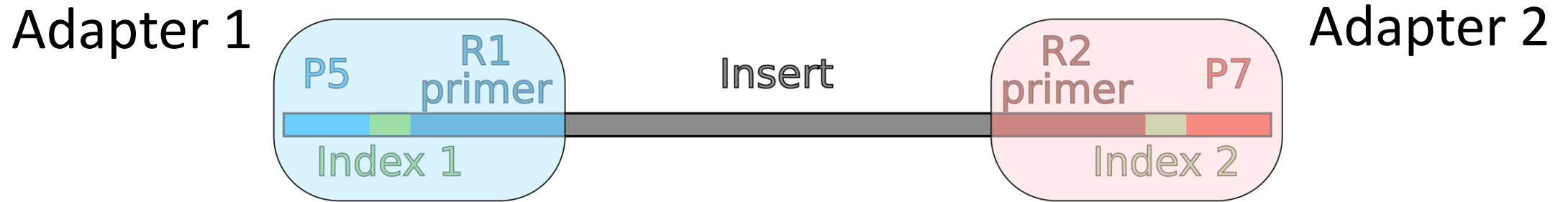
# Homework for day 2

- Videos for day 1 (if not already done)
- Videos for day 2
- Vimtutor
  - Lessons 1 and 2
  - Advanced students should do lessons 3-5
- Day1\_homework.md (library QC challenge)



# Library prep and sequencing

# Anatomy of a library



P5/P7

Ends that attach to flow cell

Index 1/2

ID sequences for multiplexing samples

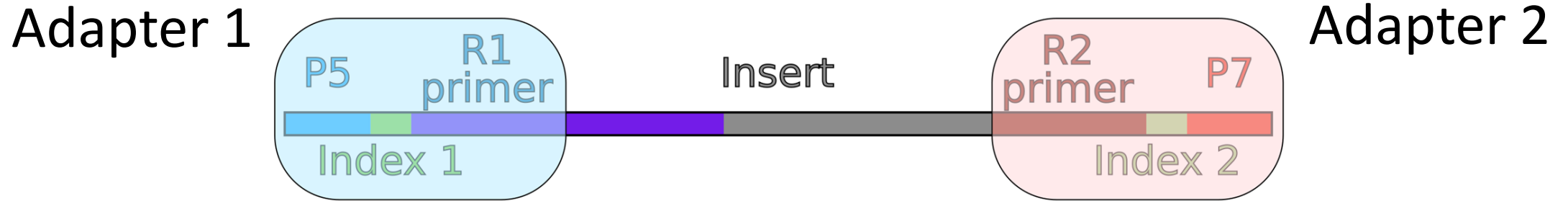
R1/R2 primers

Sequencing primers

Insert

Fragment of sample DNA/cDNA

# Anatomy of a library



P5/P7

Ends that attach to flow cell

Index 1/2

ID sequences for multiplexing samples

R1/R2 primers

Sequencing primers

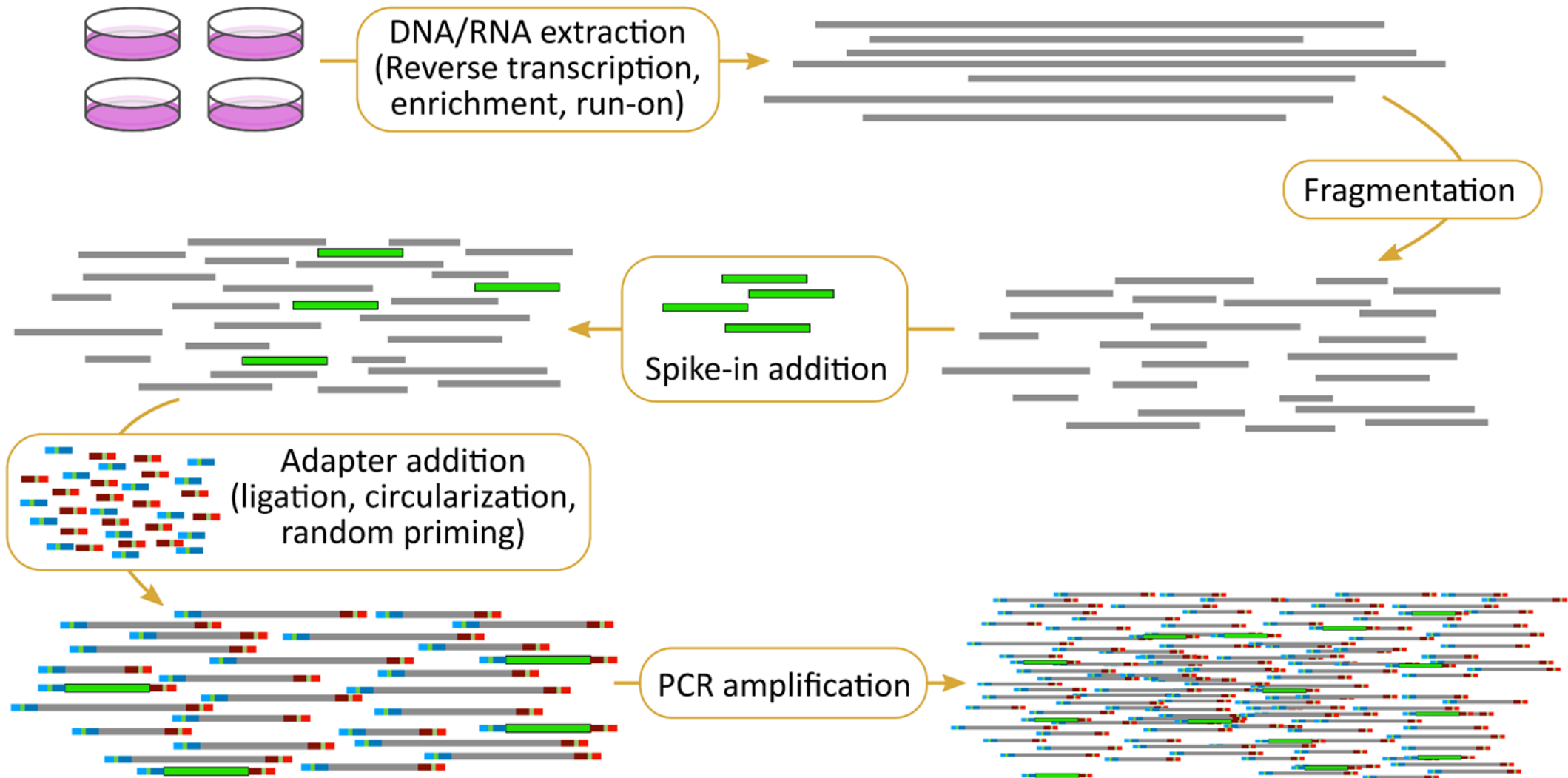
Insert

Fragment of sample DNA/cDNA

Read

Sequenced portion of fragment

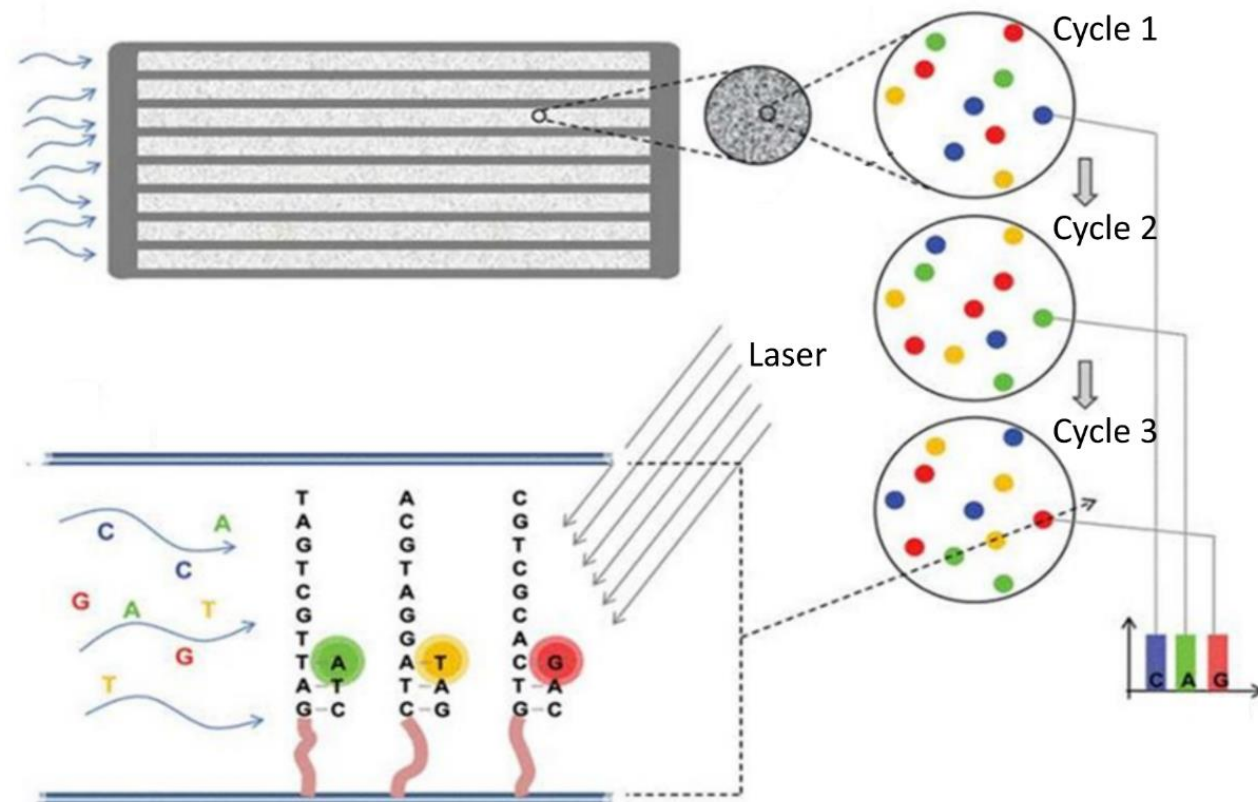
# Creating libraries



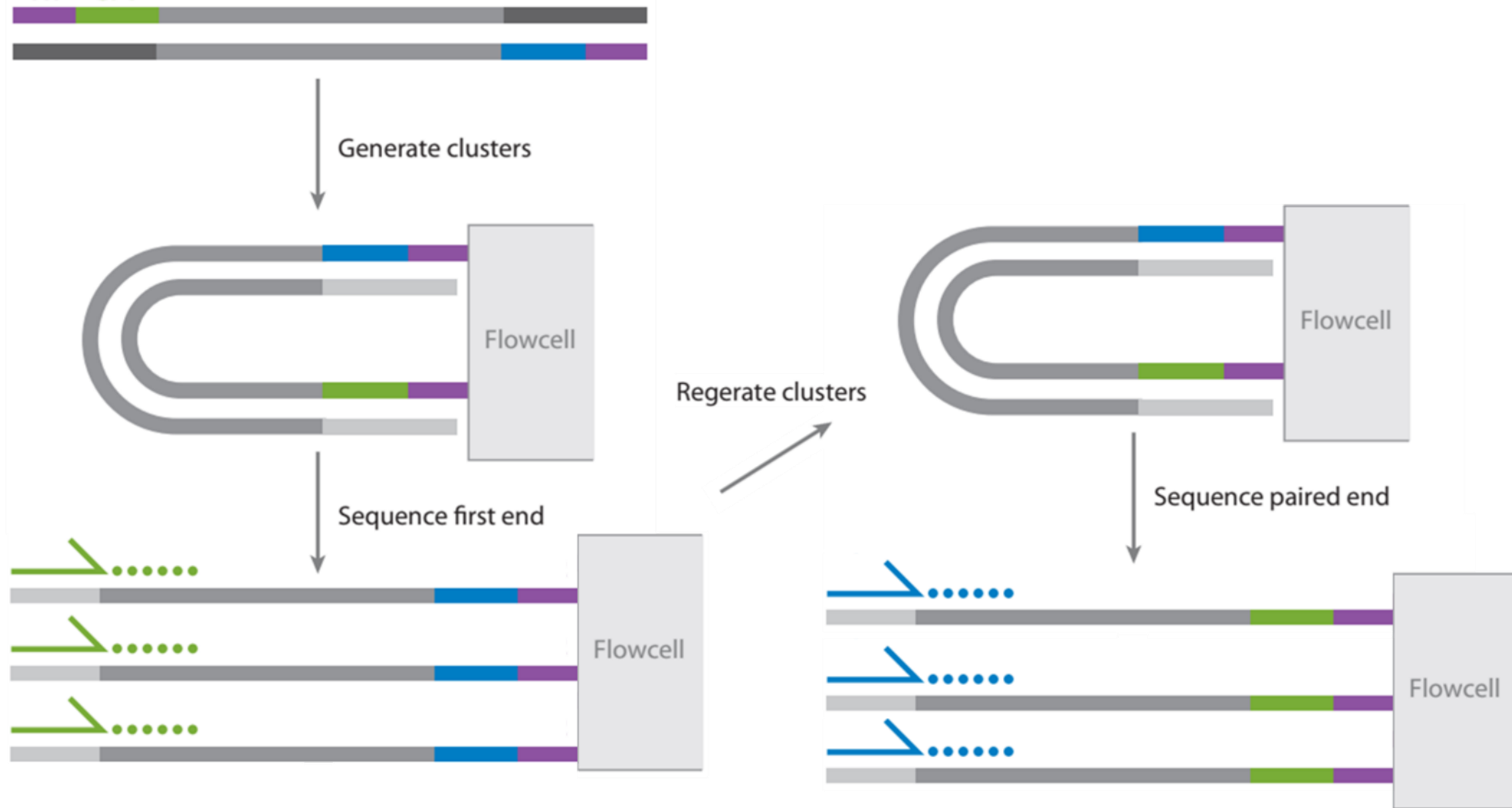
# Illumina sequencing technology

Imaging a slide (flow cell) with millions/billions of DNA clusters by cycling in fluorescent nucleotides

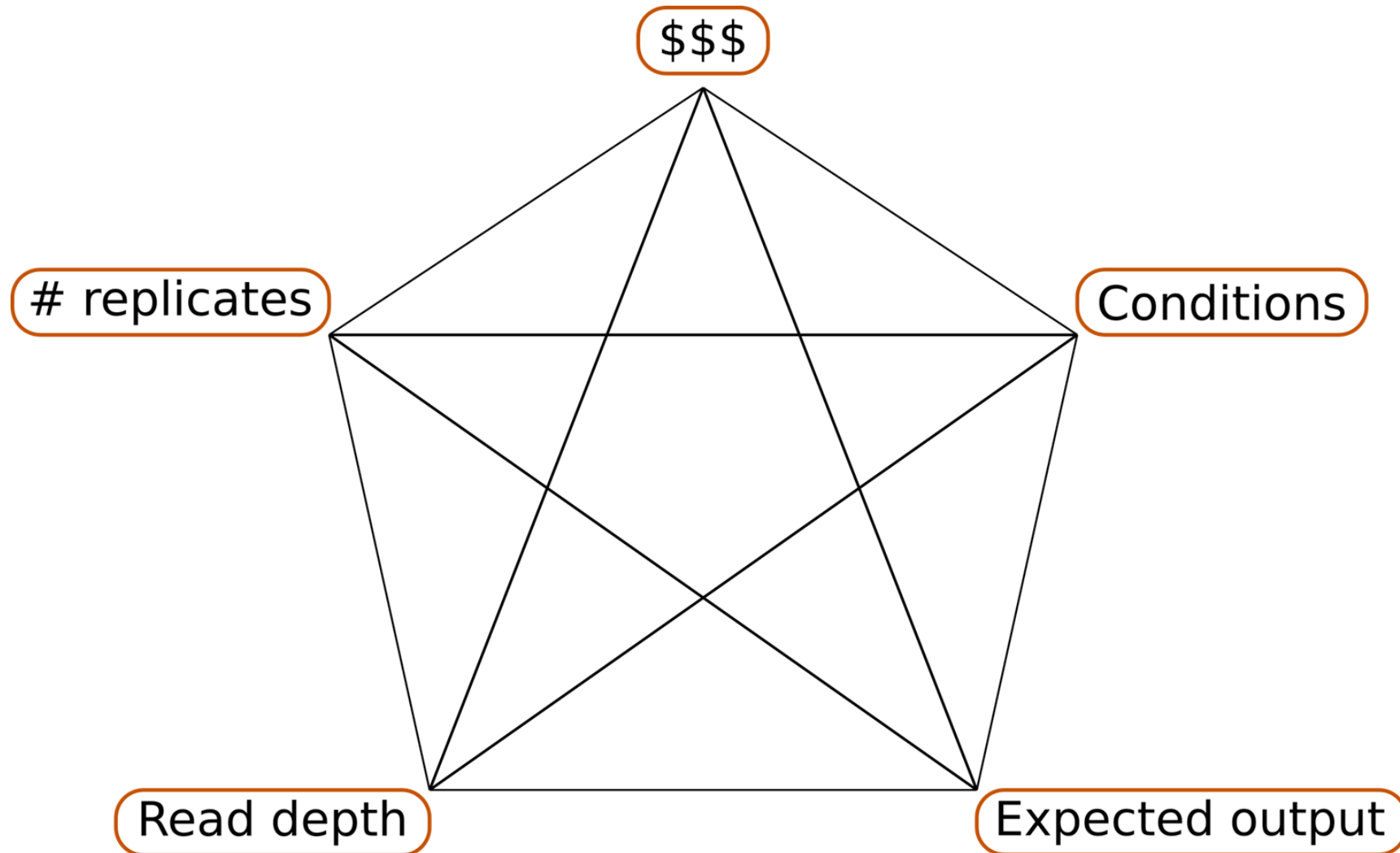
Sequencing:



# Single/paired end sequencing



# Designing a sequencing experiment



Library QC

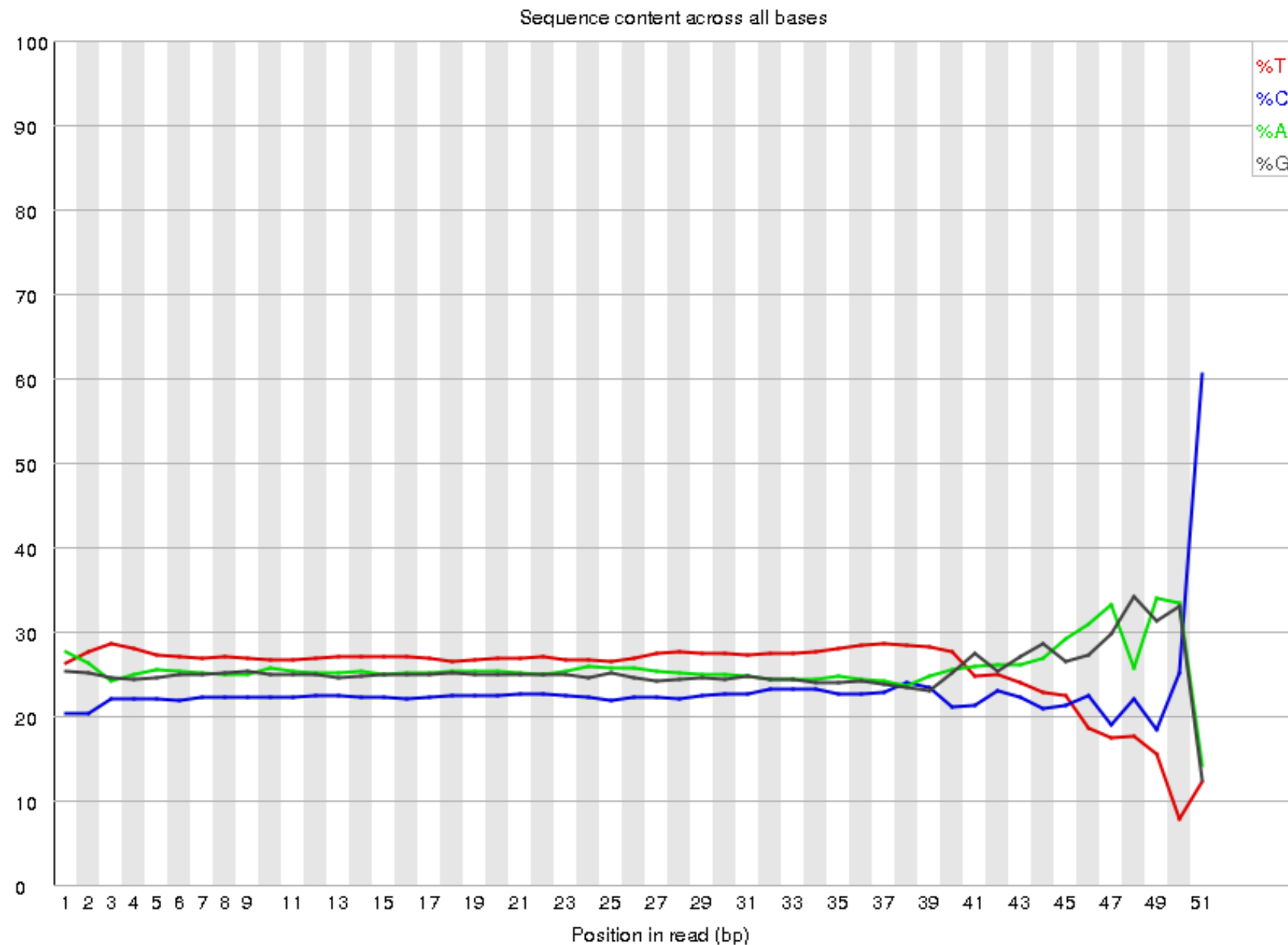


# FastQC

Base diversity

Complexity

## Per base sequence content

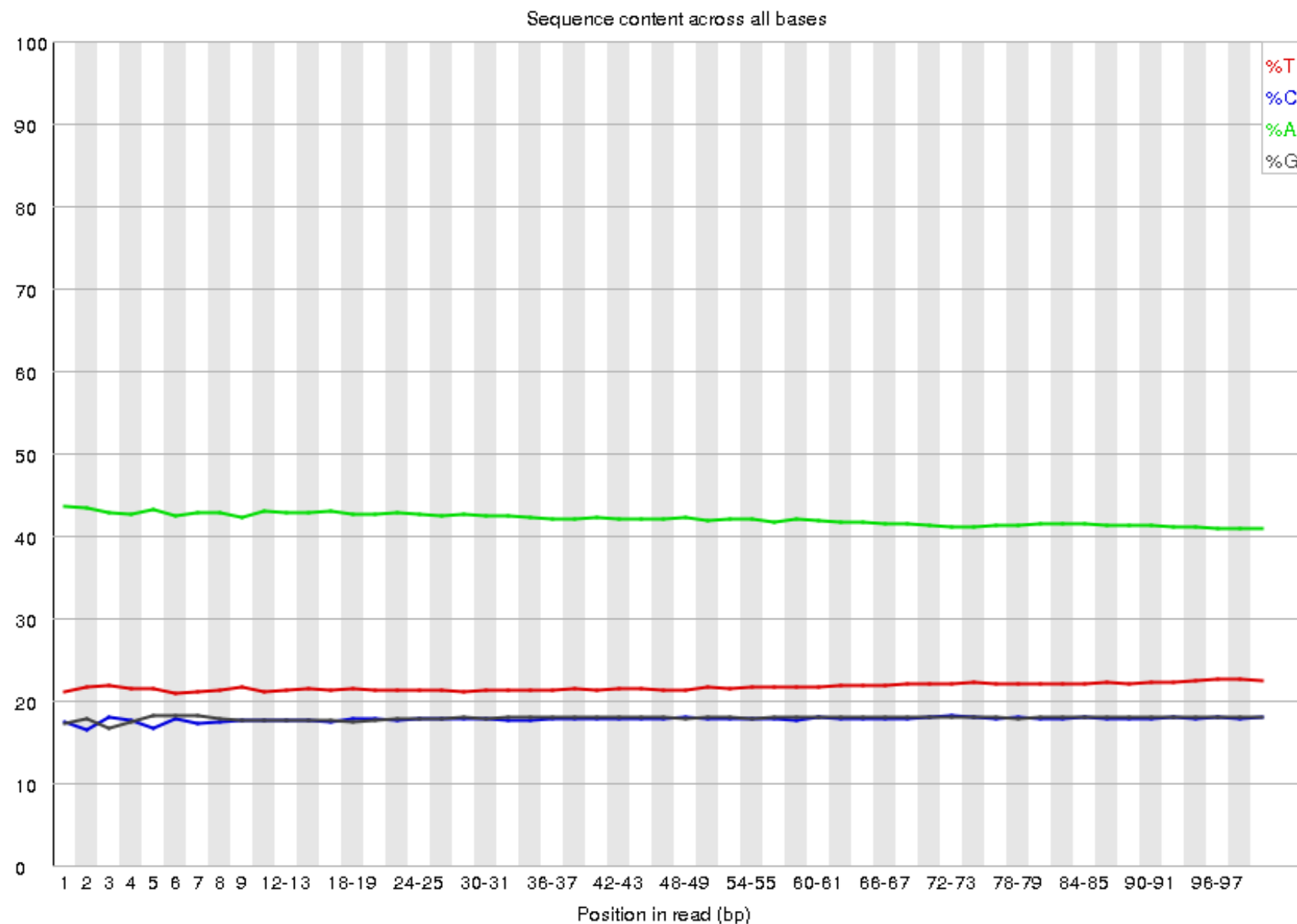


# FastQC

Base diversity

Complexity

## Per base sequence content

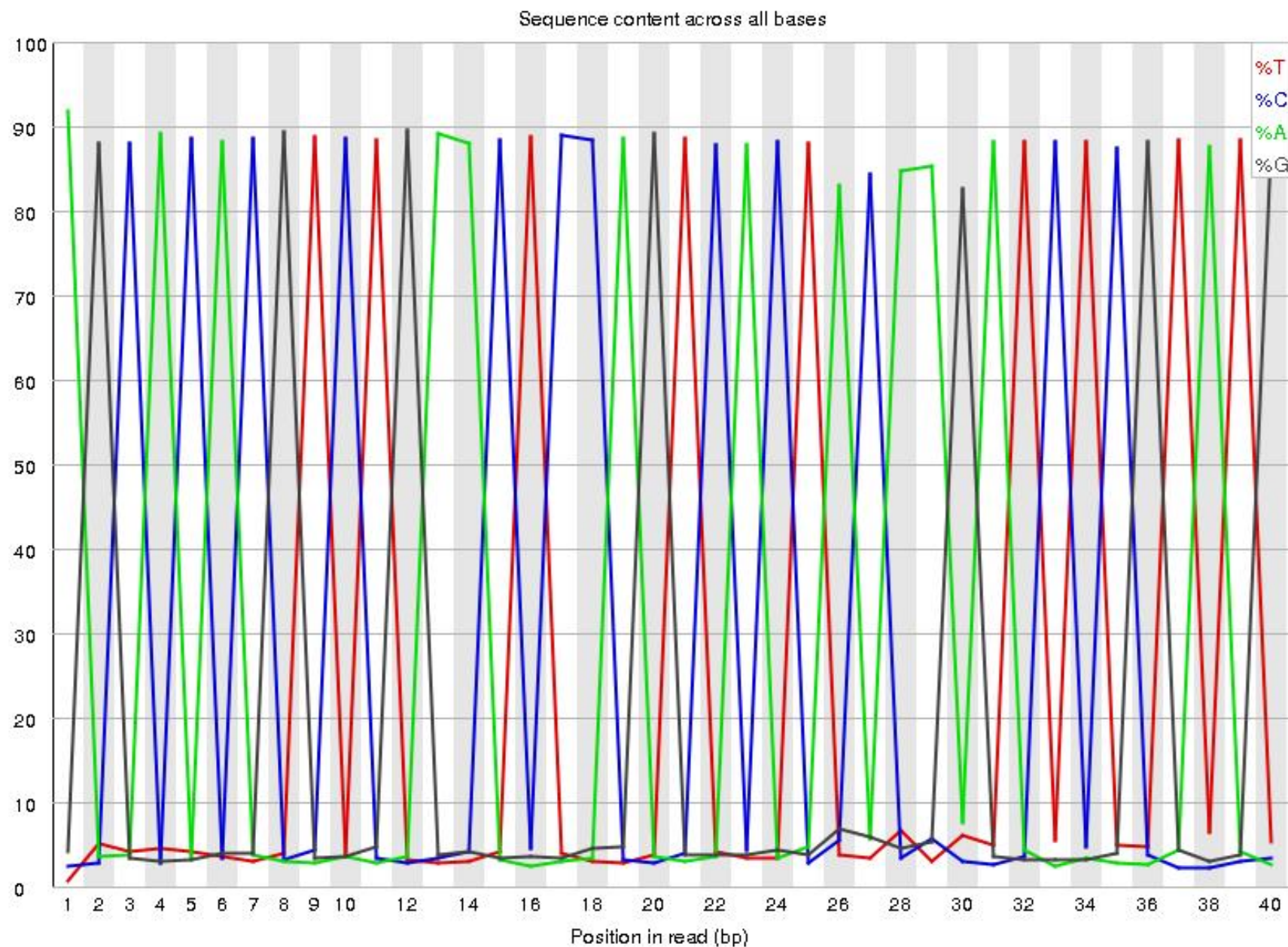


# FastQC

Base diversity

Complexity

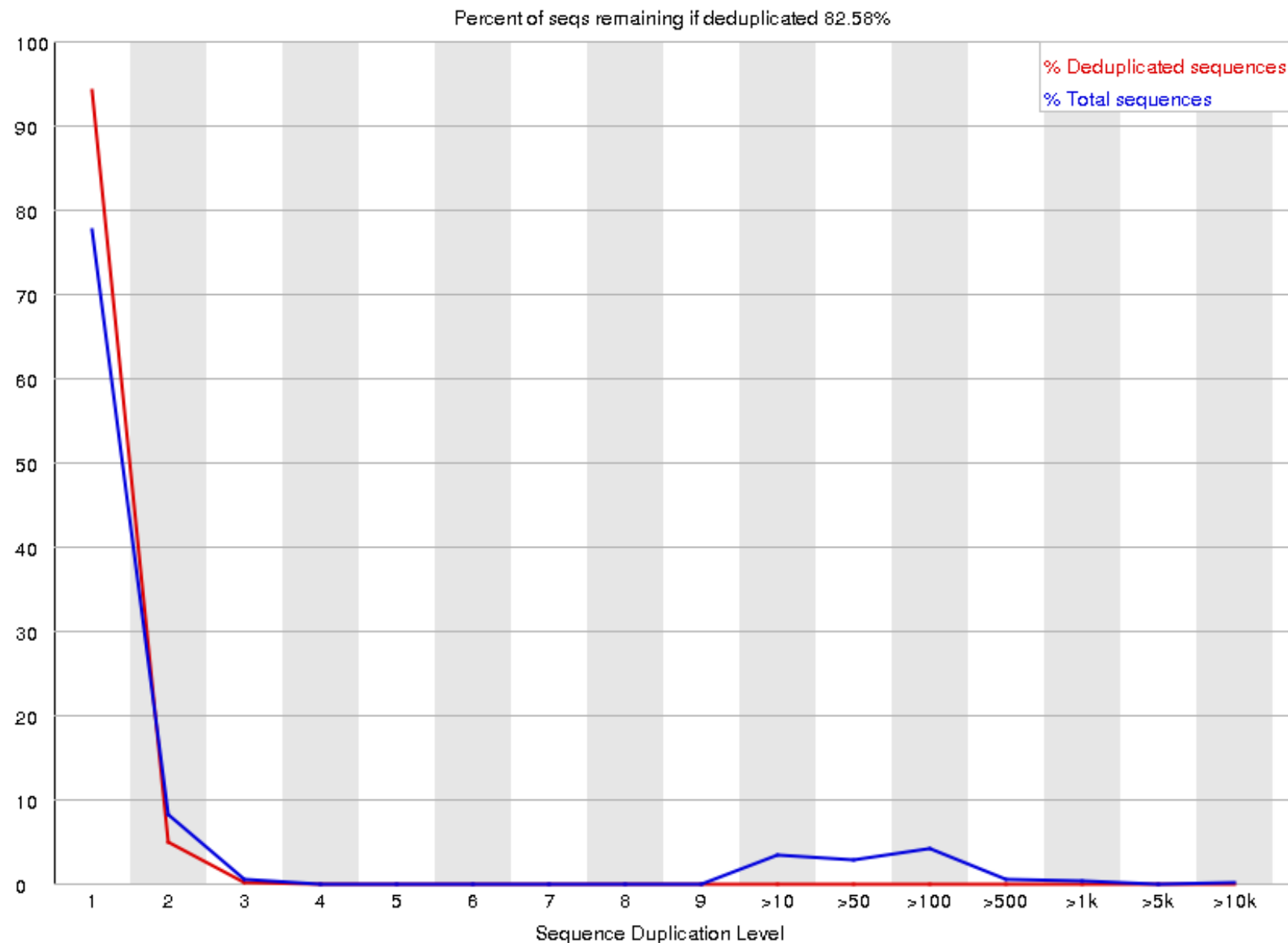
## Per base sequence content



# FastQC

## Complexity Duplication

### Sequence Duplication Levels

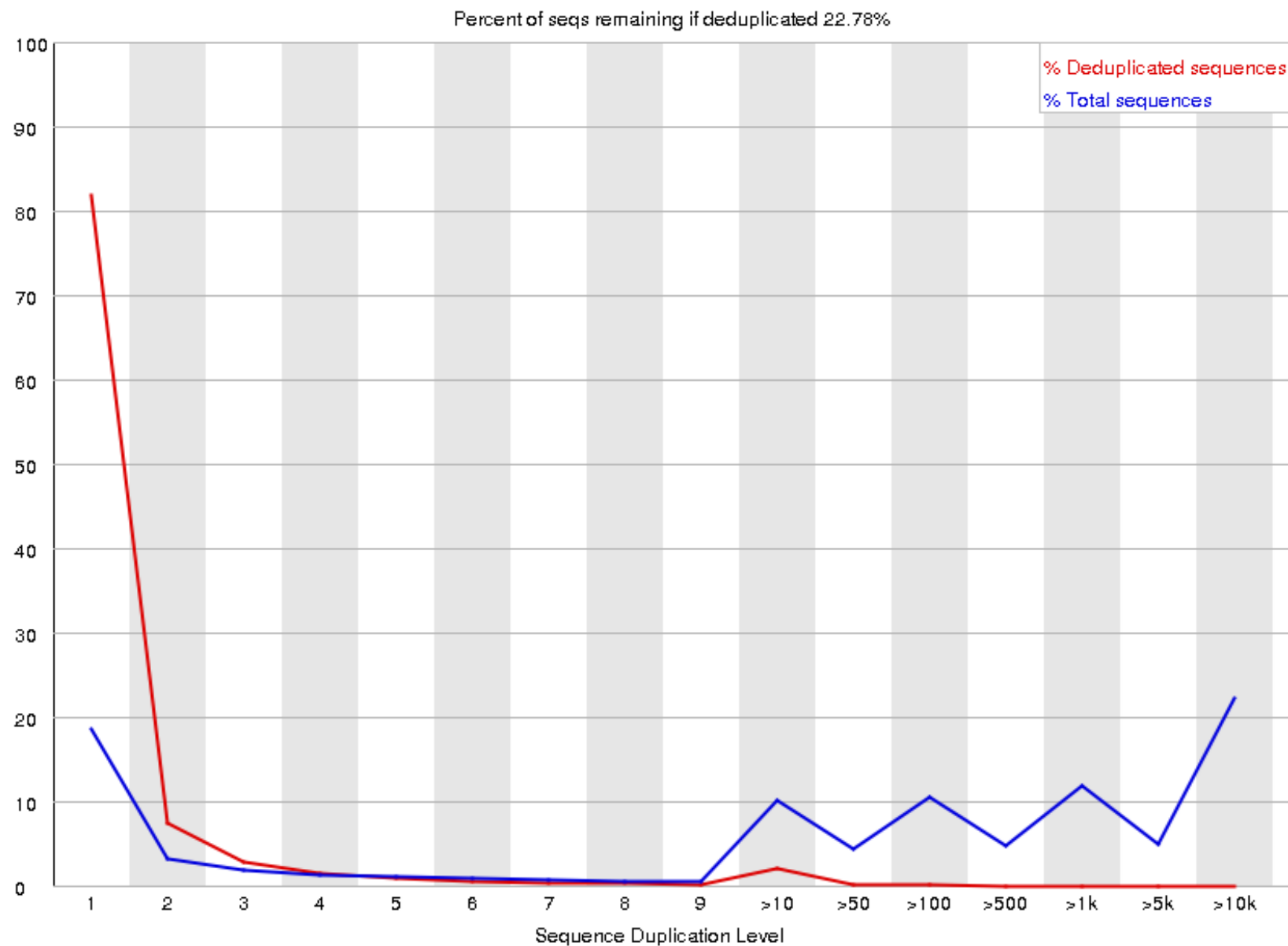


# FastQC

## Complexity

## Duplication

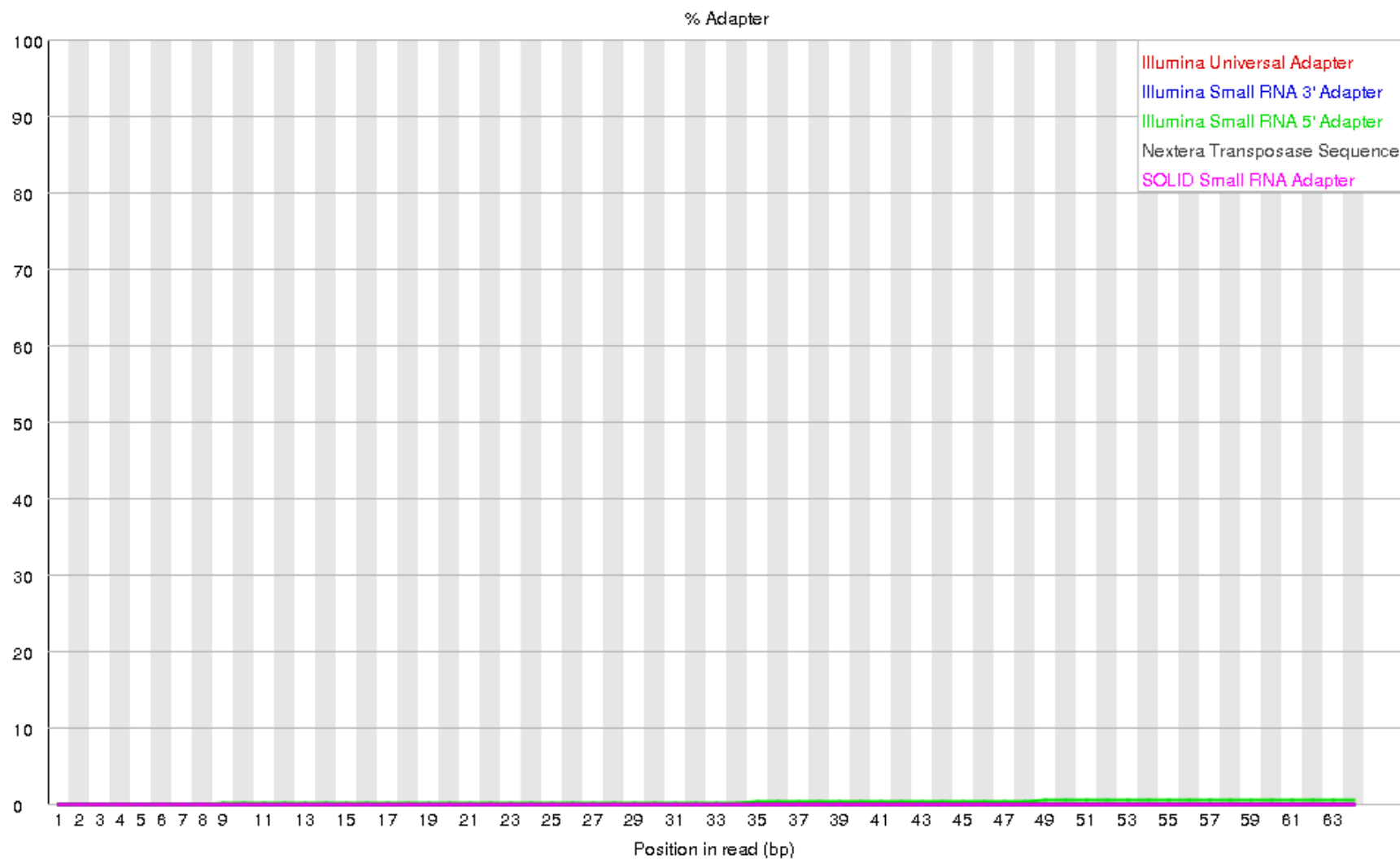
### Sequence Duplication Levels



# FastQC

## Adapter Content

## Adapter Contamination



# FastQC

## Adapter Contamination

### Adapter Content

