



A modular tool to aggregate results from bioinformatics analyses across many samples into a single report.

Report generated on 2025-02-18, 17:58 EST based on data in:  
/courses/BINF6400.202530/students/shukla.ar/fastqc\_results

✦ Summarize report

# General Statistics

📄 Copy table

⚙️ Configure columns

📊 Scatter plot

📈 Violin plot

📄 Export as CSV...

Showing <sup>5</sup>/<sub>5</sub> rows and <sup>4</sup>/<sub>6</sub>

columns.

✦ Summarize table

Sample Name	Dups	GC	Median len	Seqs
fastqc_results   base   SRR2589044_1	21.8 %	50.0 %	150 bp	1.1 M
fastqc_results   base   SRR2589044_2	19.0 %	50.0 %	150 bp	1.1 M
fastqc_results   decon   Ecoli	41.6 %	48.0 %	102 bp	0.5 M
fastqc_results   dedupe   Ecoli	41.6 %	48.0 %	102 bp	0.5 M
fastqc_results   trimmed   Ecoli	83.1 %	58.0 %	90 bp	0.2 M

# FastQC

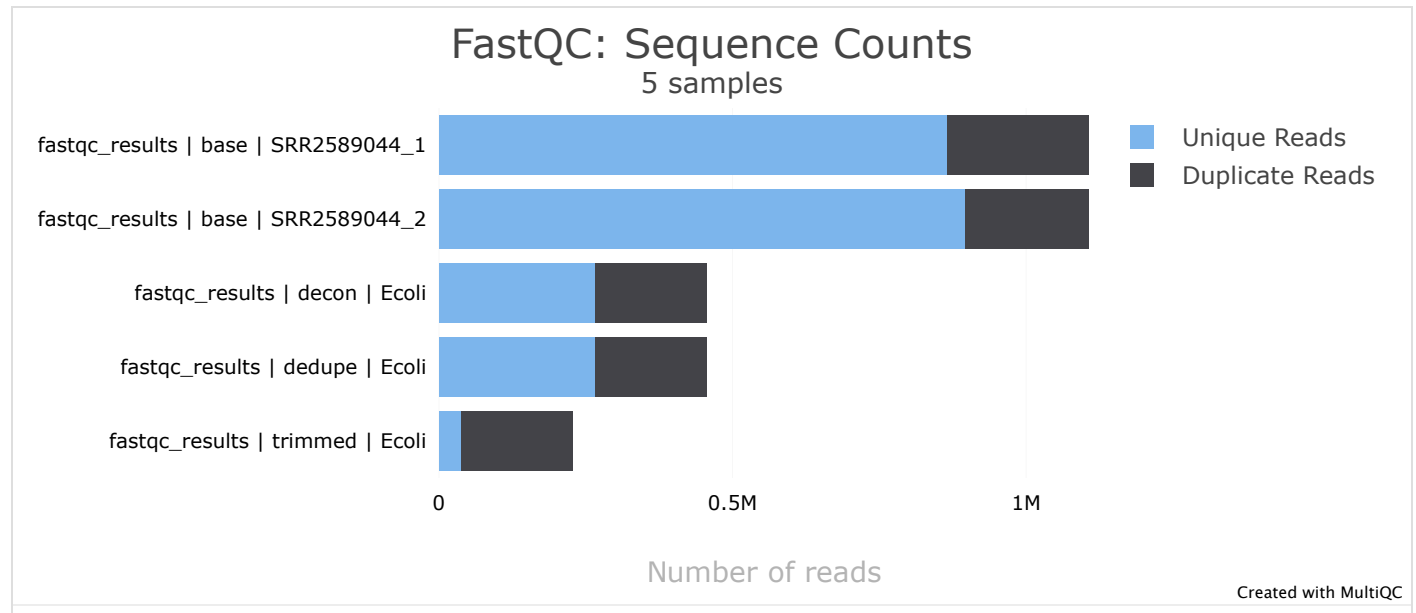
Version: 0.12.1

Quality control tool for high throughput sequencing data. URL: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>

## Sequence Counts

[? Help](#)

Sequence counts for each sample. Duplicate read counts are an estimate only.

[Percentages](#)[✦ Summarize plot](#)[↓ Export...](#)

# Sequence Quality Histograms

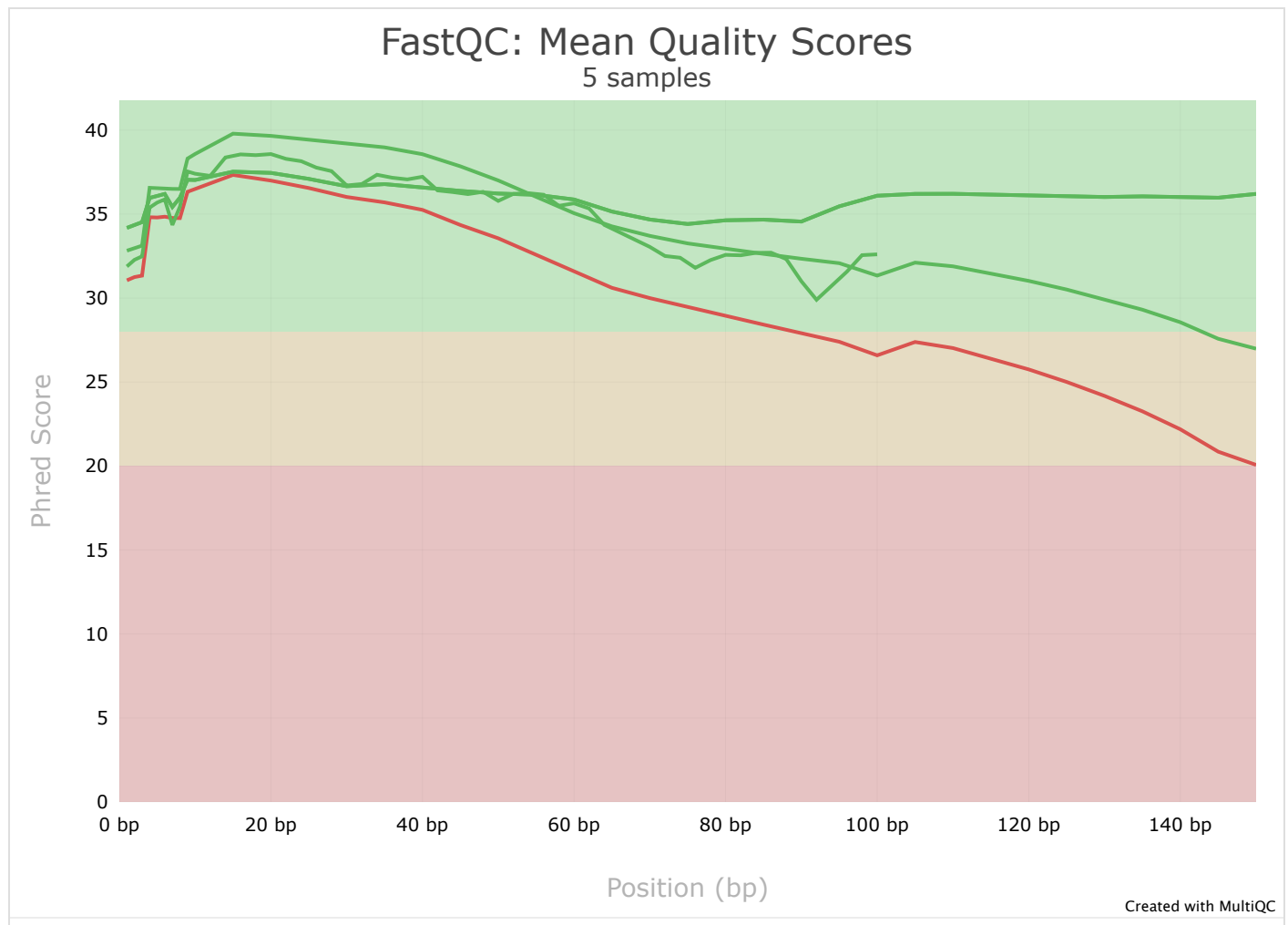
[? Help](#)

The mean quality value across each base position in the read.

To enable multiple samples to be plotted on the same graph, only the mean quality scores are plotted (unlike the box plots seen in FastQC reports).

Taken from the FastQC help:

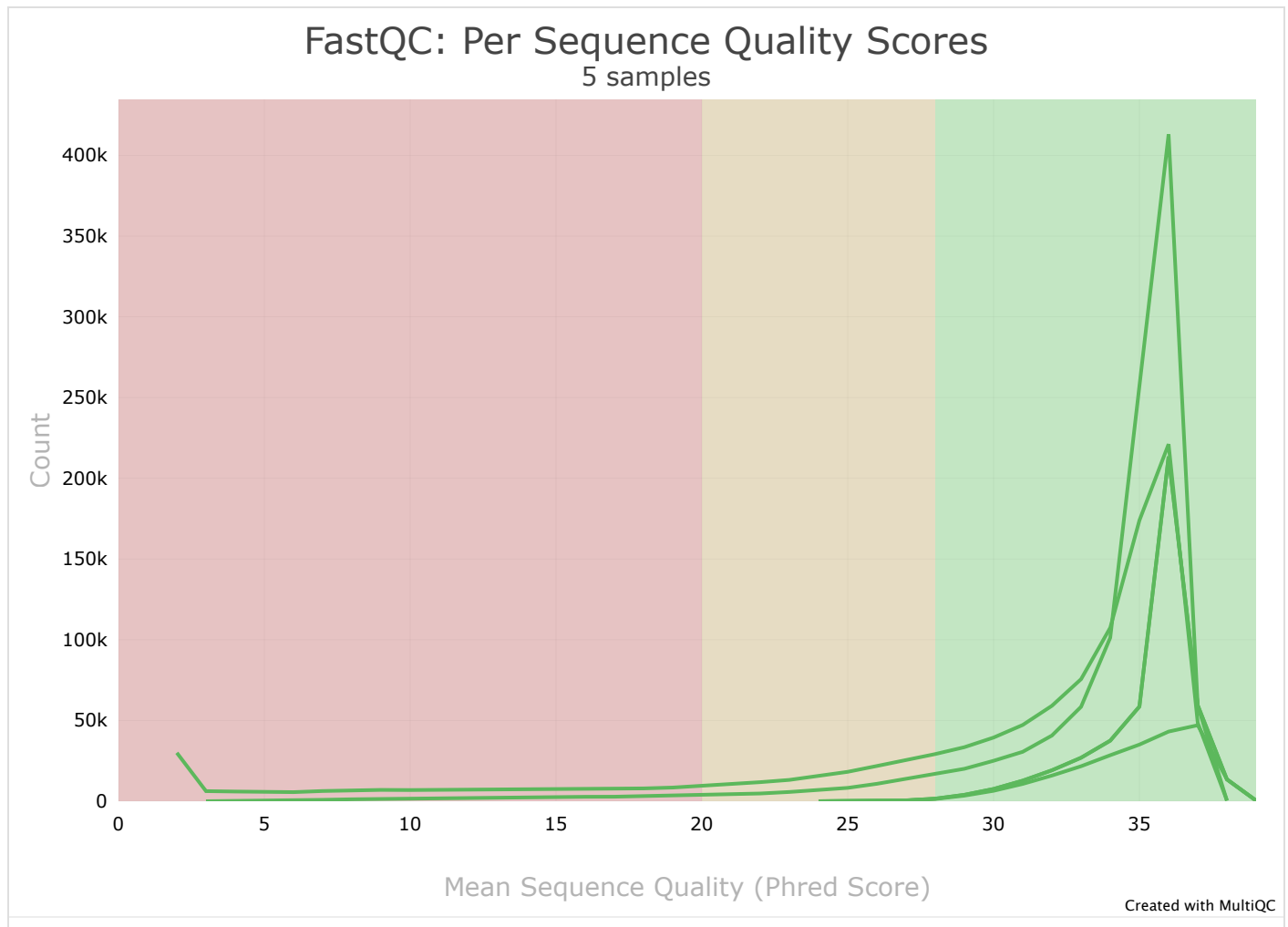
*The y-axis on the graph shows the quality scores. The higher the score, the better the base call. The background of the graph divides the y axis into very good quality calls (green), calls of reasonable quality (orange), and calls of poor quality (red). The quality of calls on most platforms will degrade as the run progresses, so it is common to see base calls falling into the orange area towards the end of a read.*

[✦ Summarize plot](#)[↓ Export...](#)

# Per Sequence Quality Scores

[? Help](#)

The number of reads with average quality scores. Shows if a subset of reads has poor quality.


[✚ Summarize plot](#)[↓ Export...](#)


5

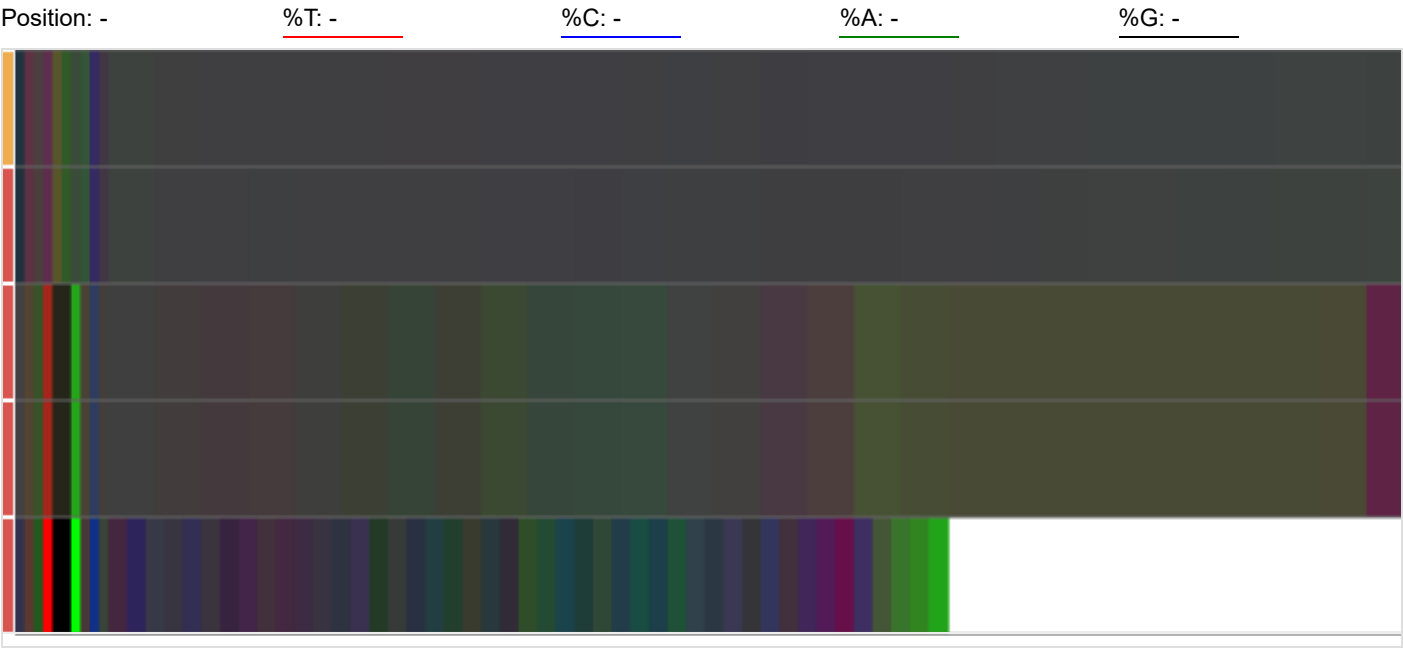
# Per Base Sequence Content

 Help

The proportion of each base position for which each of the four normal DNA bases has been called.

 Click a sample row to see a line plot for that dataset.

 Rollover for sample name



1

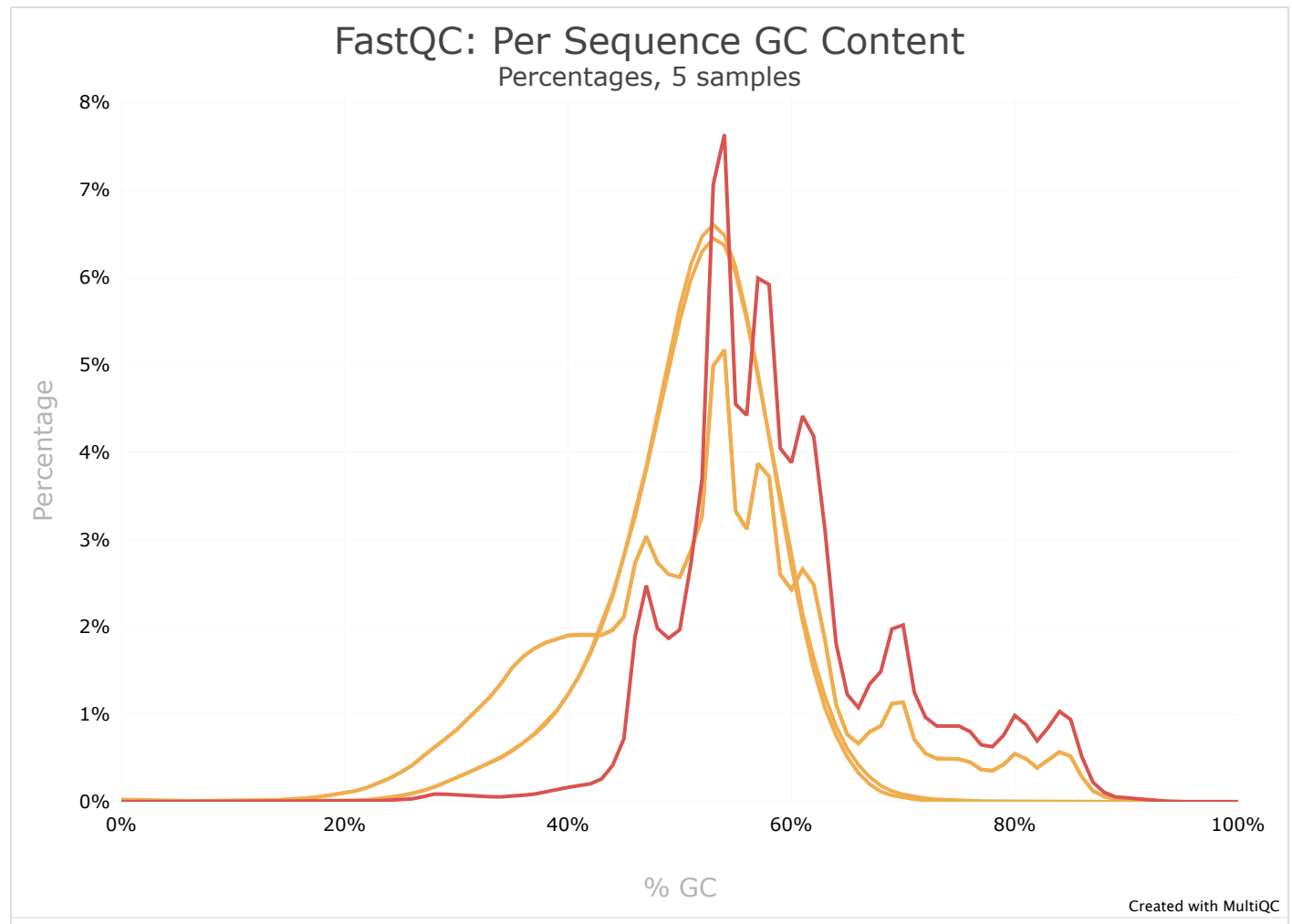
# Per Sequence GC Content

[? Help](#)

The average GC content of reads. Normal random library typically have a roughly normal distribution of GC content.

Percentages

Counts

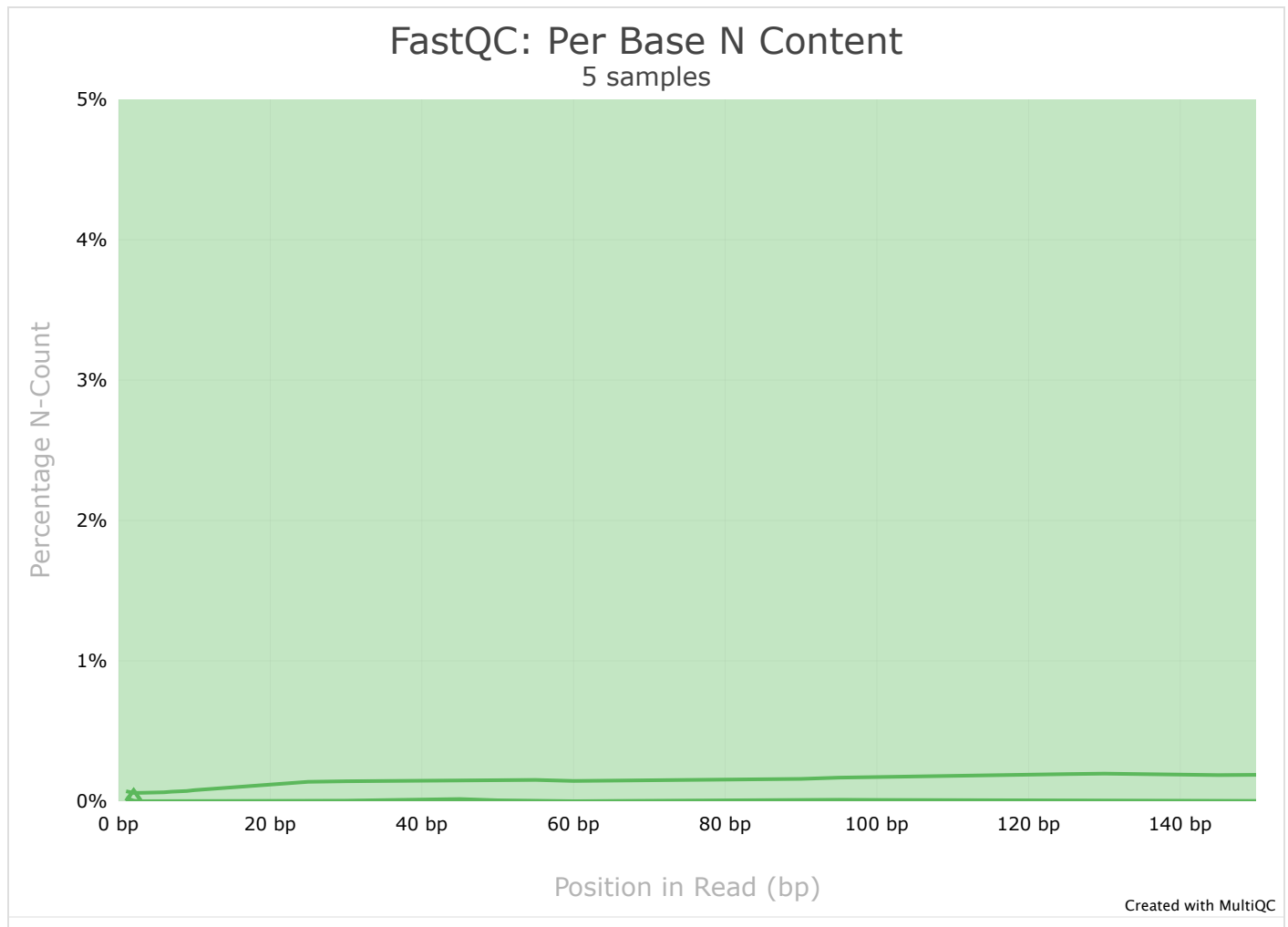
[✚ Summarize plot](#)[↓ Export...](#)

4

# Per Base N Content

[? Help](#)

The percentage of base calls at each position for which an N was called.

[+ Summarize plot](#)[↓ Export...](#)

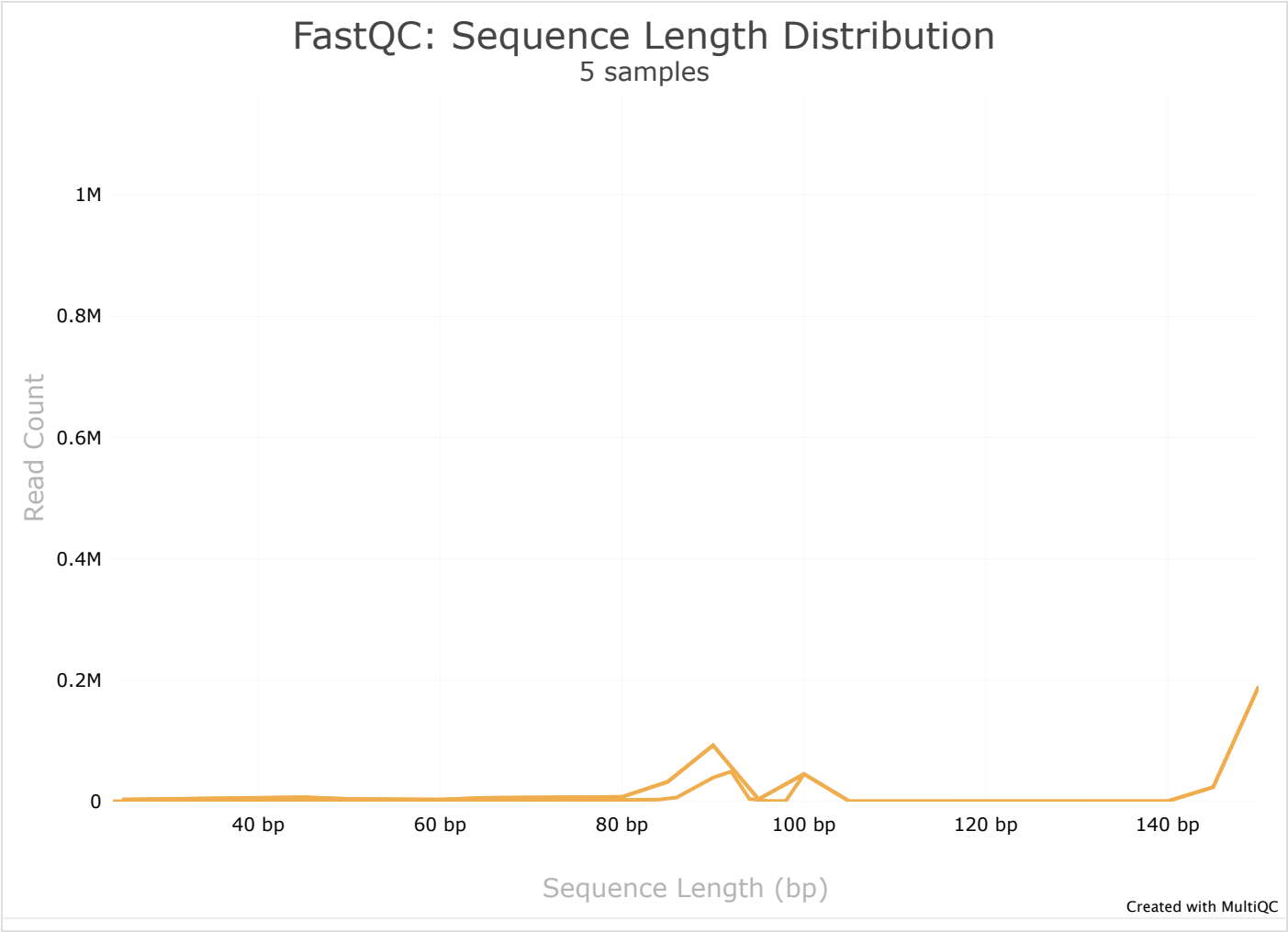
5

# Sequence Length Distribution

The distribution of fragment sizes (read lengths) found. See the FastQC help

✦ Summarize plot

⬇ Export...



2

3



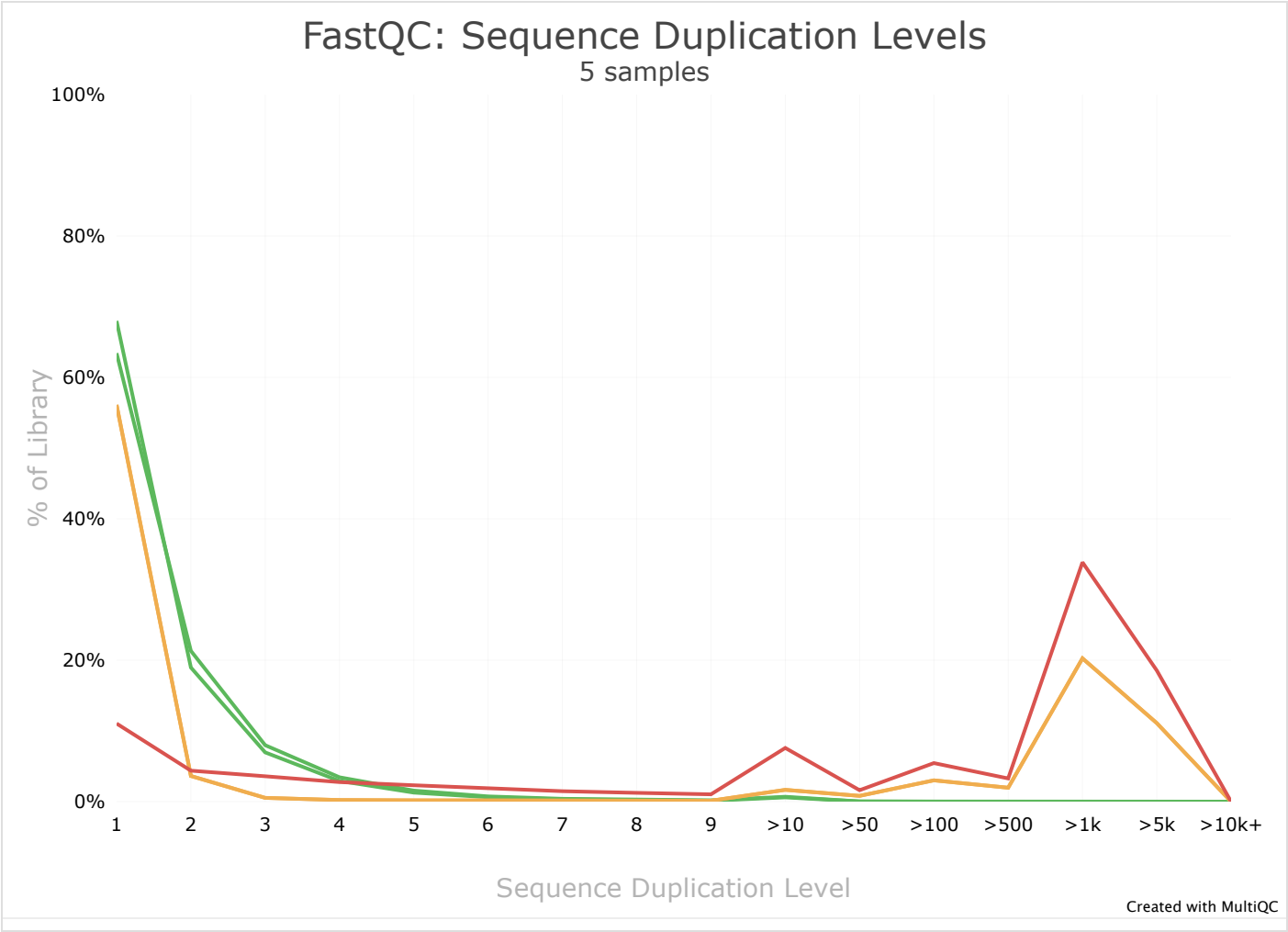
# Sequence Duplication Levels

 Help

The relative level of duplication found for every sequence.

 Summarize plot

 Export...



2 2

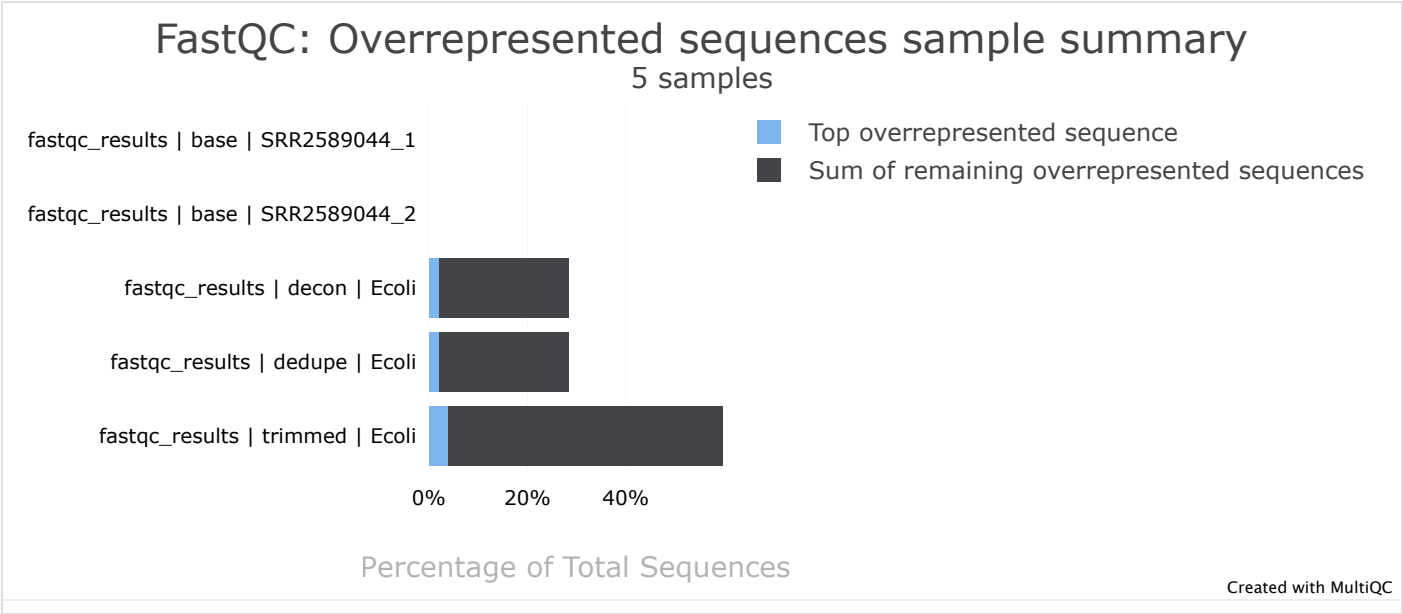
# Overrepresented sequences by sample

[? Help](#)

The total amount of overrepresented sequences found in each library.

[+ Summarize plot](#)

[↓ Export...](#)



# Top overrepresented sequences

Top overrepresented sequences across all samples. The table shows 20 most overrepresented sequences across all samples, ranked by the number of samples they occur in.

Copy table

Configure columns

Scatter plot

Violin plot

Export as CSV...

Showing 20/20 rows and 3/3 columns.

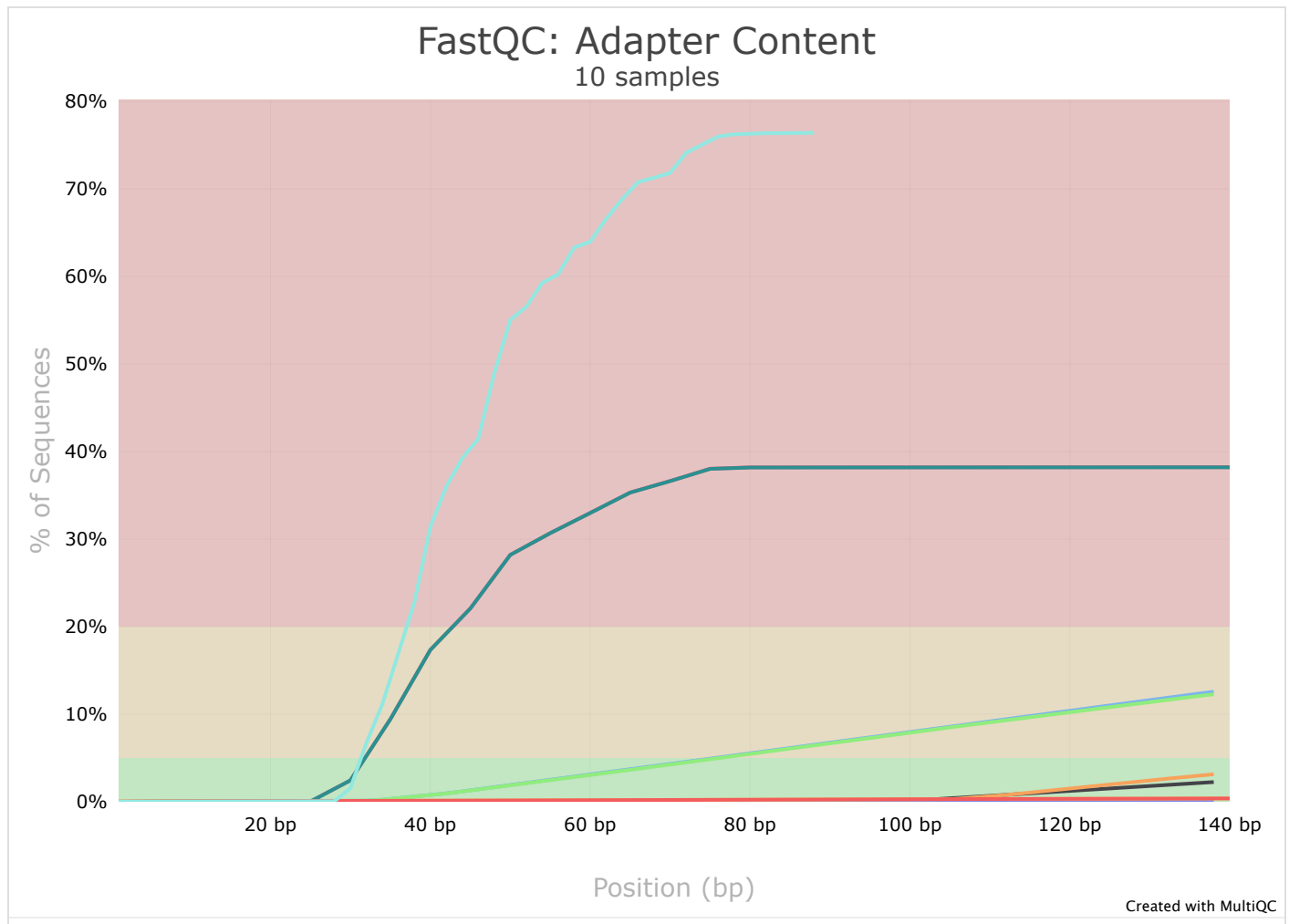
Summarize table

Overrepresented sequence	Reports	Occurrences	% of all reads
CTGTGGAGCGGTGACT	3	27 366	0.2161 %
CTATGGAGCATTACCG	3	24 536	0.17317 %
TGGTGGACAAGCTTGT	3	20 923	0.16240 %
GTGTGGAAATCGCCCC	3	20 619	0.16049 %
GCATGGAAGTGTGCT	3	16 776	0.135003 %
CTATGGATACCGCCGC	3	16 548	0.12935 %
GAATGGAGCACATCCT	3	14 965	0.114463 %
GCCTGGAGGTGCACAC	3	14 961	0.114462 %
TTTTGGAGCTAGTCCC	3	13 304	0.09968 %
TCTTGGATGAGGGTGG	3	12 823	0.09824 %
AAATGGATCCCCAGTG	3	11 853	0.08535 %
AGATGGATTATTGAGG	3	10 614	0.08165 %
GCTTGGAGGGGGGTC	3	10 599	0.08161 %
GGGTGGAGGAGCCTG	3	10 350	0.08087 %
CTGTGGAACAAGGGA	3	9 261	0.0762 %
CGGTGGACACAGCGC	3	8 619	0.06570 %
GTGTGGATCCCCTGTG	3	8 358	0.06493 %
ATCTGGAAGGCTGCAA	3	7 381	0.05201 %
GAGTGGAAGAGTGGC	3	6 498	0.04938 %
AGGTGGAGGTGTGGA	3	6 447	0.04923 %

# Adapter Content

[? Help](#)

The cumulative percentage count of the proportion of your library which has seen each of the adapter sequences at each position.

[✚ Summarize plot](#)[↓ Export...](#)

# Status Checks

[? Help](#)

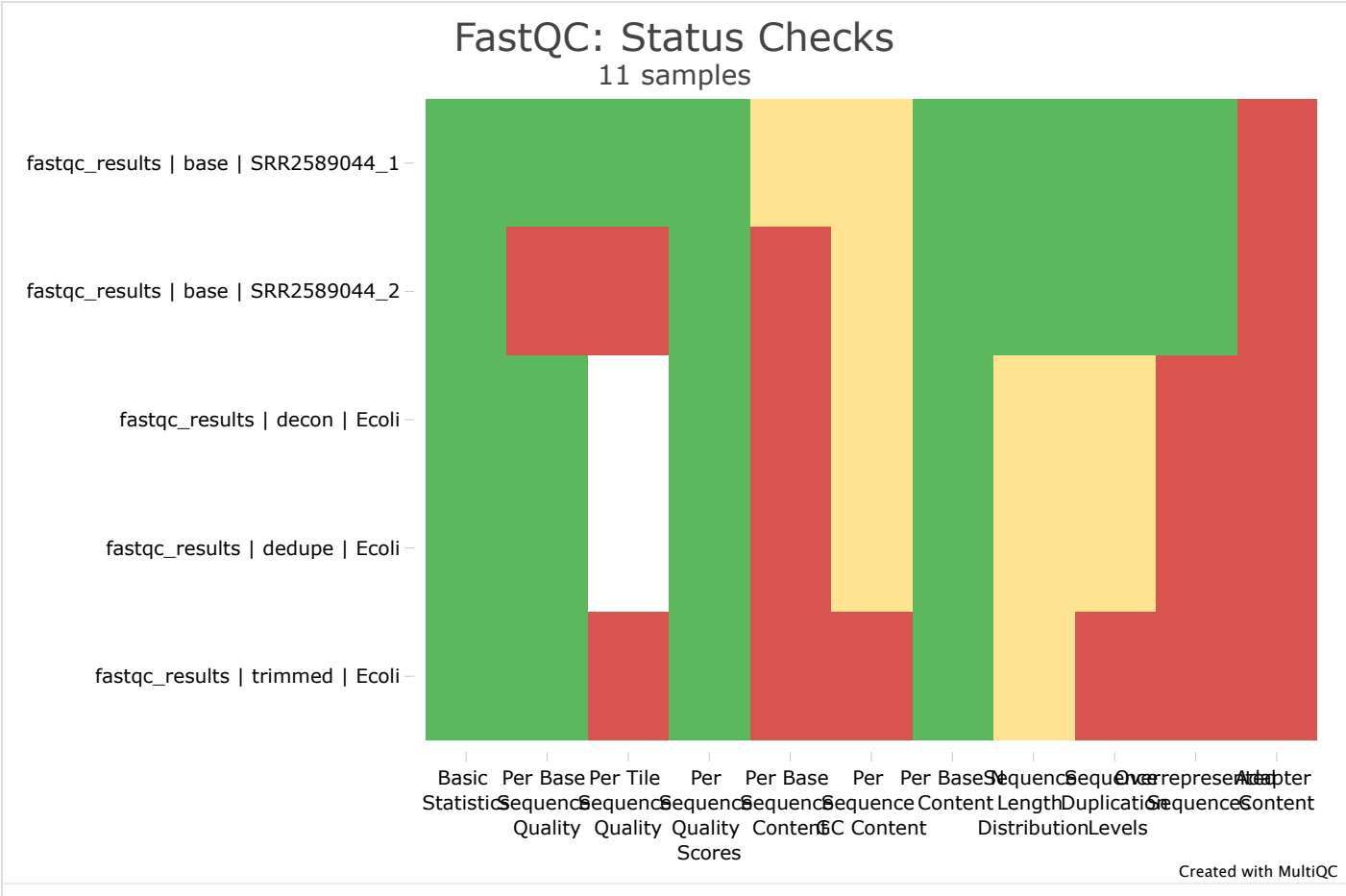
Status for each FastQC section showing whether results seem entirely normal (green), slightly abnormal (orange) or very unusual (red).

Sorted by sample

Clustered

✦ Summarize plot

⬇ Export...



# Software Versions

Software Versions lists versions of software tools extracted from file contents.

[📄 Copy table](#)

Software	Version
FastQC	0.12.1

This report uses [Plotly](#), [jQuery](#), [jQuery UI](#), [Bootstrap](#) and [FileSaver.js](#).

