



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

Report generated on 2024-02-23, 19:34 EST based on data in:
/athena/cayuga_0019/scratch/mij4011/homework/project/processed_fastqc_results

General Statistics

Copy table

Configure columns

Scatter plot

Violin plot

Showing 80/80 rows and 3/6 columns.

Sample Name	% Dups		% GC		M Seqs
SRR21190736_1_trimmed	53.6%	ed	44%		31.1 M
SRR21190736_2_trimmed	51.8%	red	44%		31.1 M
SRR21190737_1_trimmed	54.9%	ed	44%		32.0 M
SRR21190737_2_trimmed	52.7%	red	44%		32.0 M
SRR21190738_1_trimmed	57.2%	ed	43%		31.1 M
SRR21190738_2_trimmed	56.2%	ed	43%		31.1 M
SRR21190739_1_trimmed	58.7%	d	43%		29.8 M
SRR21190739_2_trimmed	56.9%	ed	43%		29.8 M
SRR21190740_1_trimmed	57.3%	ed	43%		33.7 M
SRR21190740_2_trimmed	56.0%	ed	43%		33.7 M
SRR21190741_1_trimmed	56.3%	ed	45%		37.0 M
SRR21190741_2_trimmed	54.9%	ed	45%		37.0 M
SRR21190742_1_trimmed	52.3%	red	44%		28.5 M
SRR21190742_2_trimmed	50.1%	ved	43%		28.5 M
SRR21190743_1_trimmed	51.7%	red	43%		29.9 M
SRR21190743_2_trimmed	49.7%	ved	43%		29.9 M
SRR21190744_1_trimmed	51.0%	red	44%		28.8 M
SRR21190744_2_trimmed	48.9%	ved	44%		28.8 M
SRR21190745_1_trimmed	51.9%	red	44%		29.8 M
SRR21190745_2_trimmed	49.8%	ved	44%		29.8 M

Sample Name	% Dups		% GC		M Seqs
SRR21190746_1_trimmed	47.5%	oved	44%		36.7 M
SRR21190746_2_trimmed	45.5%	oved	44%		36.7 M
SRR21190747_1_trimmed	43.6%	oved	43%		31.7 M
SRR21190747_2_trimmed	41.3%	noved	43%		31.7 M
SRR21190748_1_trimmed	52.6%	ed	44%		33.1 M
SRR21190748_2_trimmed	50.6%	ved	44%		33.1 M
SRR21190749_1_trimmed	43.8%	oved	44%		33.3 M
SRR21190749_2_trimmed	41.3%	noved	44%		33.3 M
SRR21190750_1_trimmed	48.0%	oved	44%		35.0 M
SRR21190750_2_trimmed	46.0%	oved	44%		35.0 M
SRR21190751_1_trimmed	43.6%	oved	42%		31.9 M
SRR21190751_2_trimmed	41.7%	noved	42%		31.9 M
SRR21190752_1_trimmed	38.6%	noved	43%		27.4 M
SRR21190752_2_trimmed	35.5%	removed	42%		27.4 M
SRR21190753_1_trimmed	41.8%	noved	43%		35.5 M
SRR21190753_2_trimmed	39.4%	noved	43%		35.5 M
SRR21190754_1_trimmed	42.4%	noved	42%		35.1 M
SRR21190754_2_trimmed	40.1%	noved	42%		35.1 M
SRR21190755_1_trimmed	46.4%	oved	42%		35.5 M
SRR21190755_2_trimmed	44.0%	oved	42%		35.5 M
SRR21190756_1_trimmed	51.5%	ved	46%		32.1 M
SRR21190756_2_trimmed	50.8%	ved	46%		32.1 M
SRR21190757_1_trimmed	46.8%	oved	45%		21.9 M
SRR21190757_2_trimmed	44.4%	oved	45%		21.9 M
SRR21190758_1_trimmed	49.4%	ved	45%		28.9 M
SRR21190758_2_trimmed	45.5%	oved	46%		28.9 M
SRR21190759_1_trimmed	48.3%	ved	45%		26.4 M
SRR21190759_2_trimmed	46.7%	oved	45%		26.4 M
SRR21190760_1_trimmed	49.8%	ved	45%		30.4 M
SRR21190760_2_trimmed	48.1%	ved	45%		30.4 M

Sample Name	% Dups		% GC	M Seqs
SRR21190761_1_trimmed	52.7%	red	42%	26.9 M
SRR21190761_2_trimmed	51.3%	red	42%	26.9 M
SRR21190762_1_trimmed	46.0%	oved	45%	24.8 M
SRR21190762_2_trimmed	44.9%	oved	45%	24.8 M
SRR21190763_1_trimmed	47.7%	oved	45%	25.5 M
SRR21190763_2_trimmed	45.9%	oved	45%	25.5 M
SRR21190764_1_trimmed	57.3%	ed	46%	33.3 M
SRR21190764_2_trimmed	56.4%	ed	46%	33.3 M
SRR21190765_1_trimmed	57.0%	ed	45%	33.6 M
SRR21190765_2_trimmed	57.0%	ed	45%	33.6 M
SRR21190766_1_trimmed	54.2%	ed	46%	35.8 M
SRR21190766_2_trimmed	52.3%	red	46%	35.8 M
SRR21190767_1_trimmed	52.2%	red	46%	32.8 M
SRR21190767_2_trimmed	51.1%	red	46%	32.8 M
SRR21190768_1_trimmed	54.1%	ed	46%	33.6 M
SRR21190768_2_trimmed	52.9%	ed	46%	33.6 M
SRR21190769_1_trimmed	53.2%	ed	46%	33.0 M
SRR21190769_2_trimmed	52.0%	red	46%	33.0 M
SRR21190770_1_trimmed	56.3%	ed	46%	27.1 M
SRR21190770_2_trimmed	53.4%	ed	46%	27.1 M
SRR21190771_1_trimmed	51.8%	red	45%	29.2 M
SRR21190771_2_trimmed	50.8%	red	45%	29.2 M
SRR21190772_1_trimmed	50.1%	ved	45%	28.7 M
SRR21190772_2_trimmed	48.7%	ved	45%	28.7 M
SRR21190773_1_trimmed	49.1%	ved	45%	26.0 M
SRR21190773_2_trimmed	47.3%	oved	45%	26.0 M
SRR21190774_1_trimmed	51.5%	red	45%	31.8 M
SRR21190774_2_trimmed	49.9%	ved	45%	31.8 M
SRR21190775_1_trimmed	51.2%	red	45%	33.2 M
SRR21190775_2_trimmed	49.1%	ved	45%	33.2 M

FastQC

Version: 0.12.1

FastQC is a quality control tool for high throughput sequence data, written by Simon Andrews at the Babraham Institute in Cambridge.

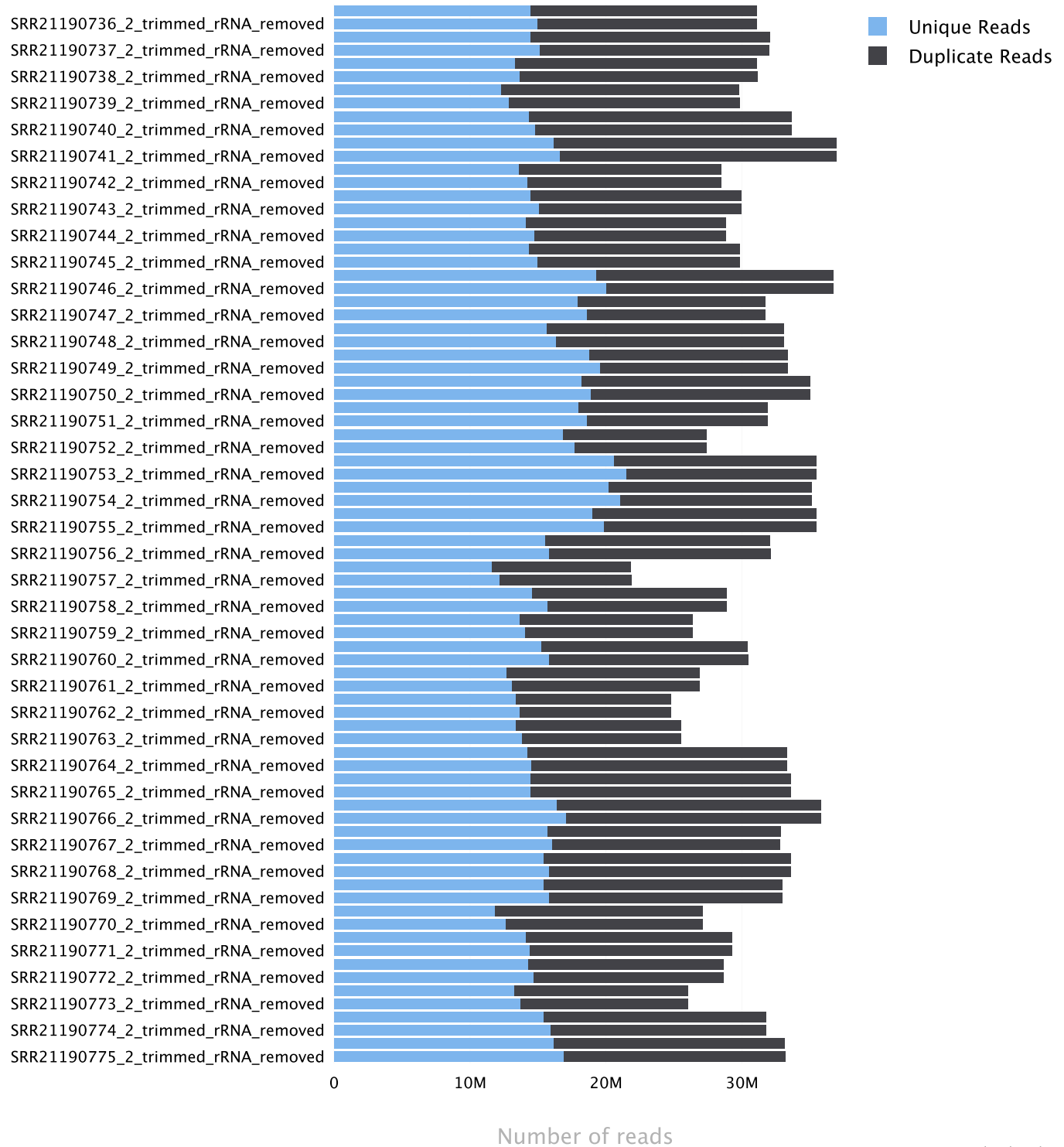
Sequence Counts

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

Percentages

Export Plot

FastQC: Sequence Counts

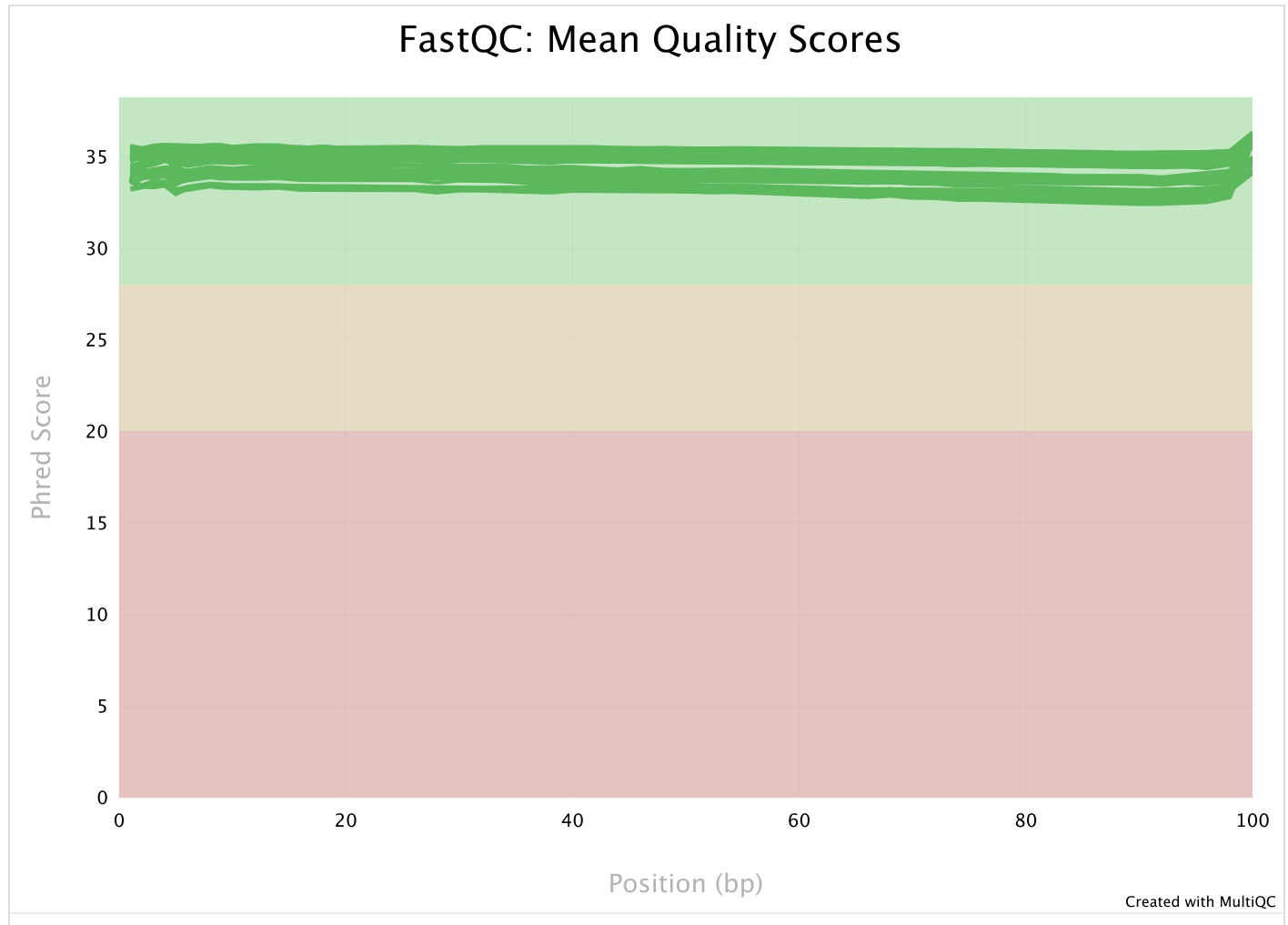


Created with MultiQC

Sequence Quality Histograms

80

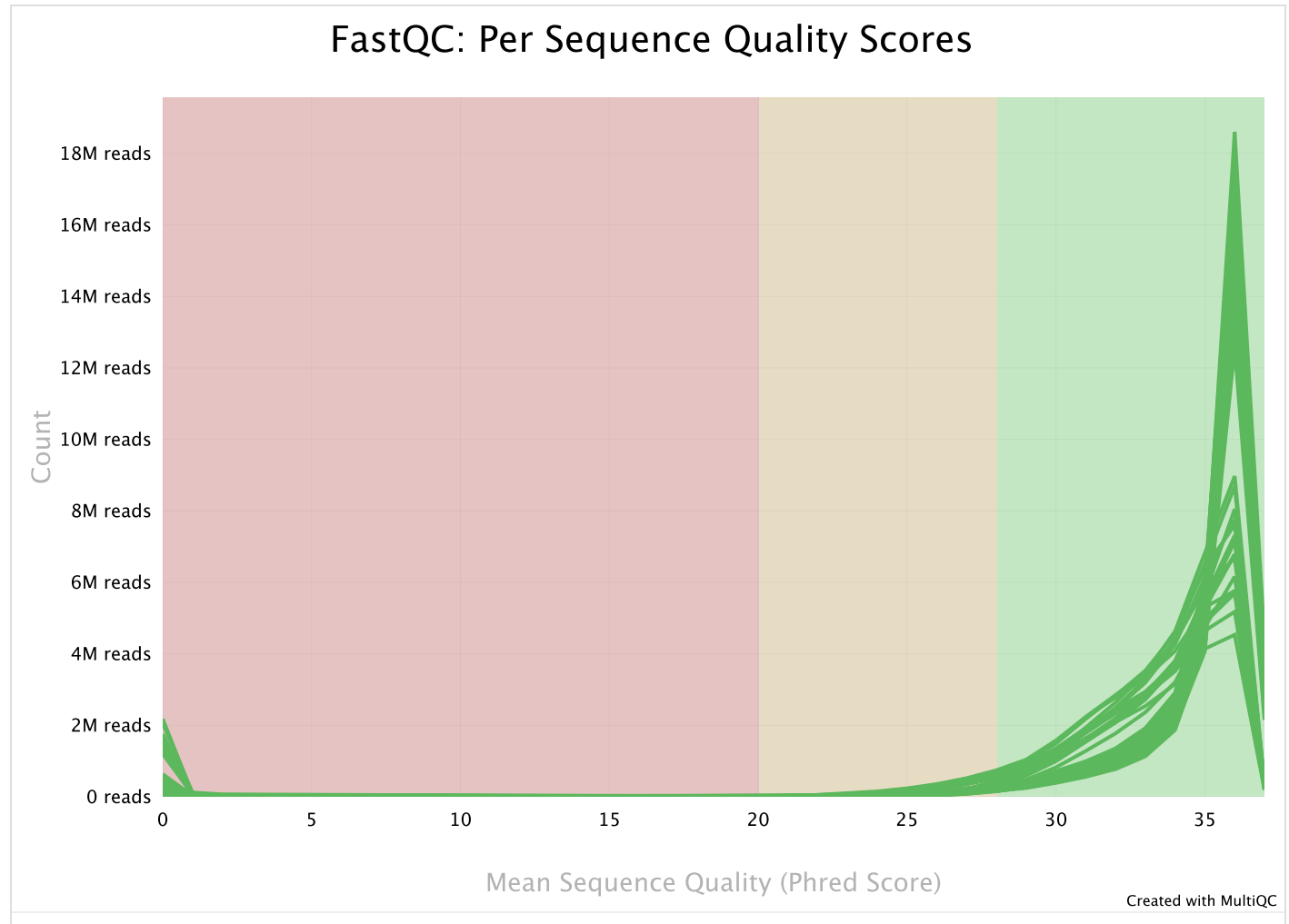
The mean quality value across each base position in the read.

[Export Plot](#)

Per Sequence Quality Scores

80

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

[Export Plot](#)

Per Base Sequence Content

39

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.

i Rollover for sample name

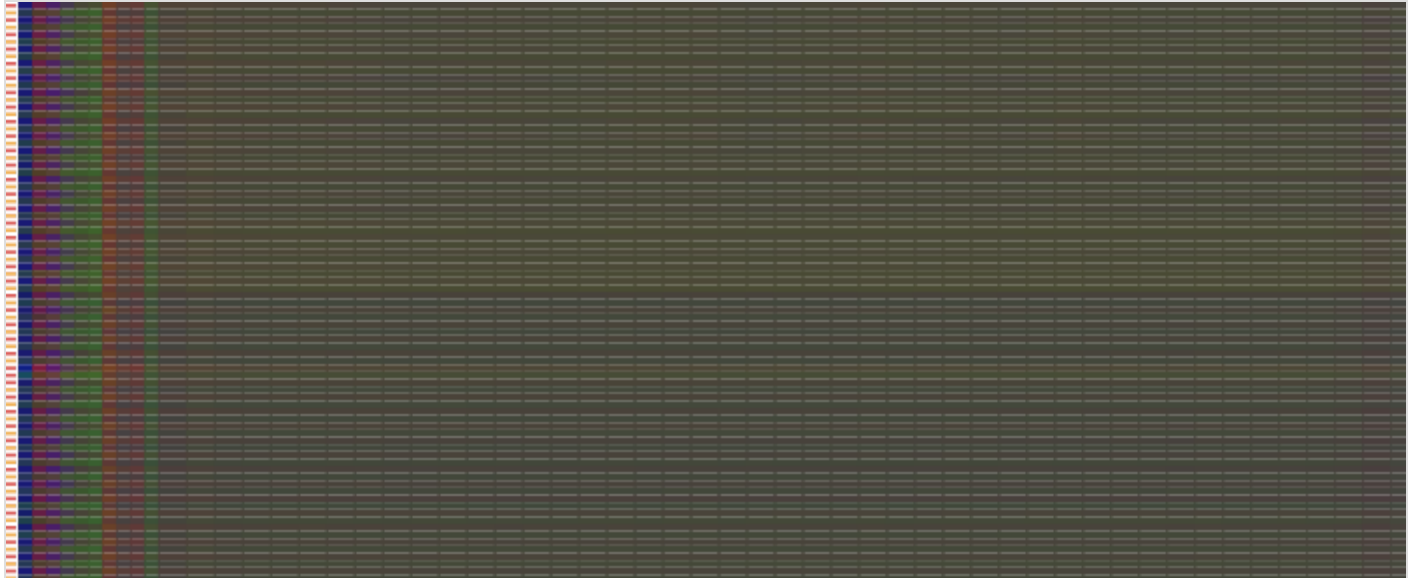
Position: -

%T: -

%C: -

%A: -

%G: -



Per Sequence GC Content

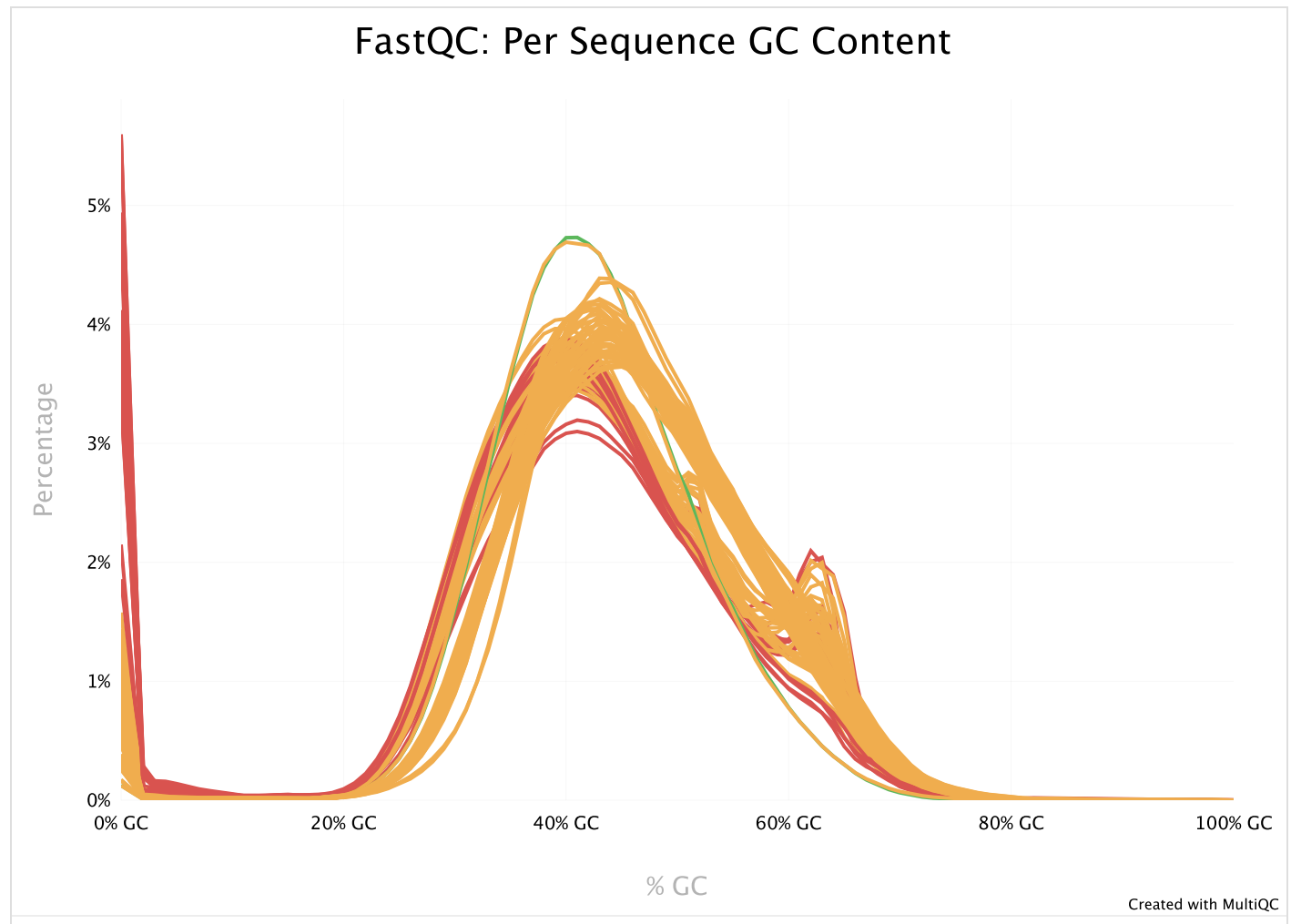
65

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

Percentages

Counts

Export Plot

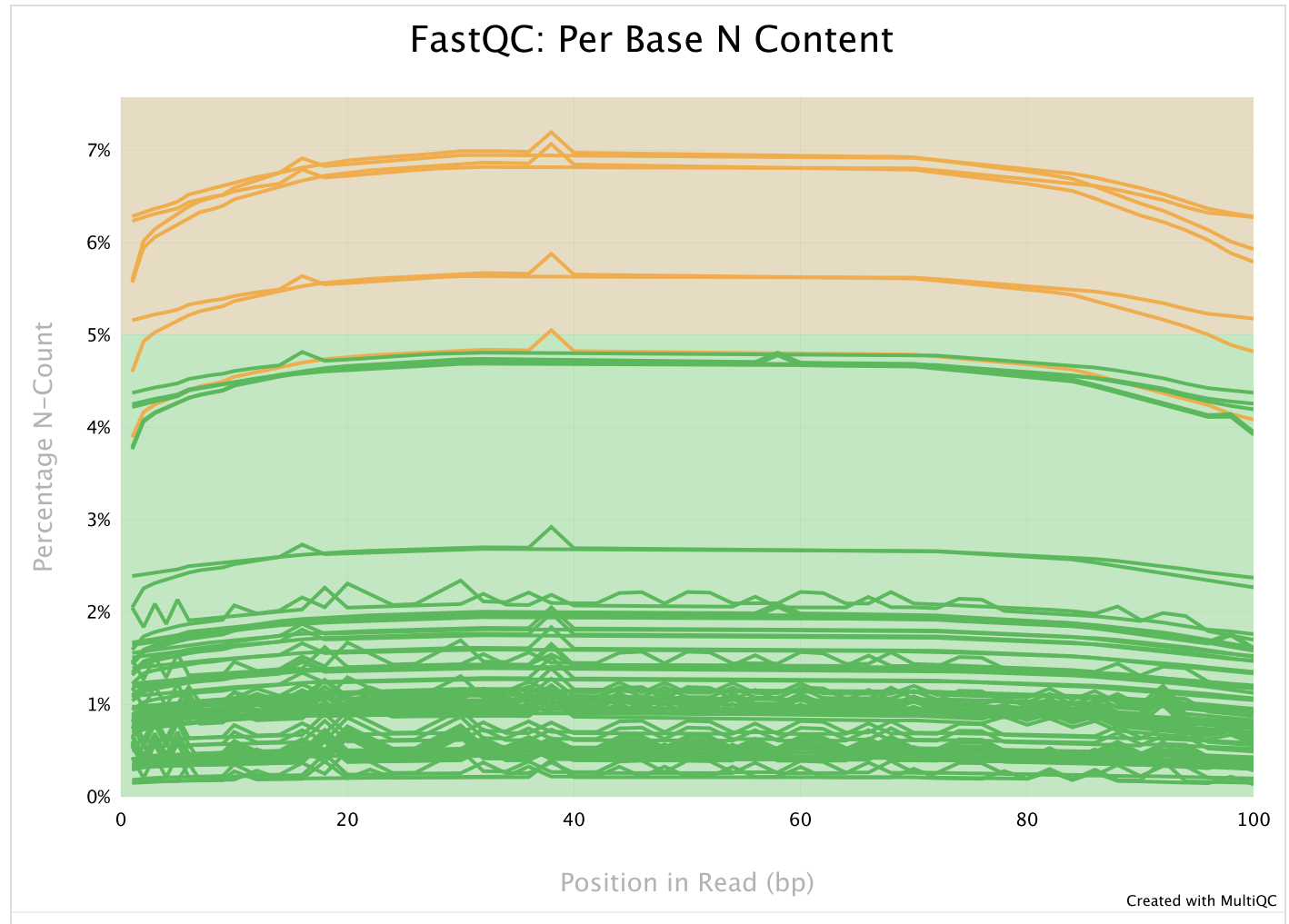


Per Base N Content

73

7

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

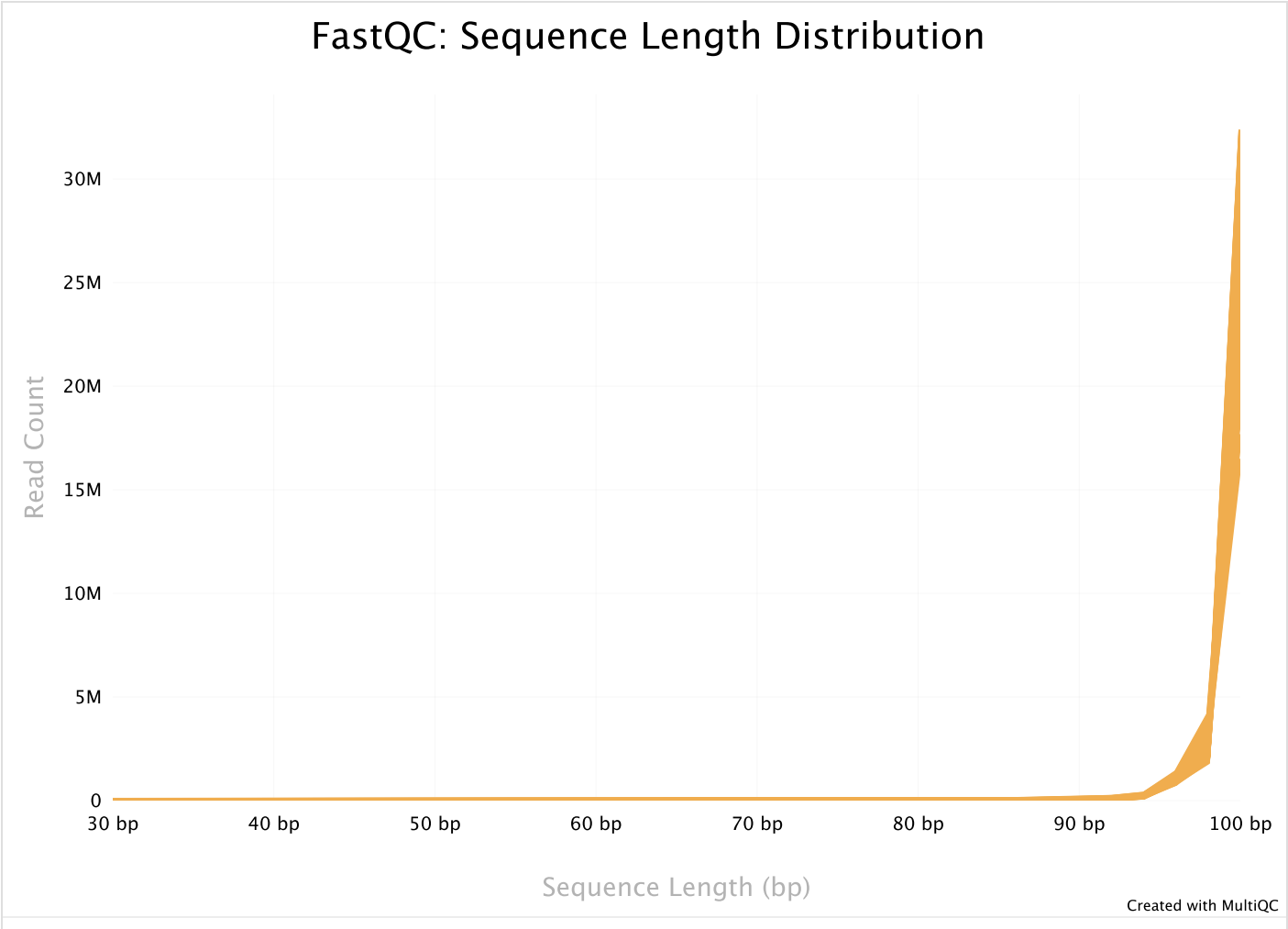
[Export Plot](#)

Sequence Length Distribution

80

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

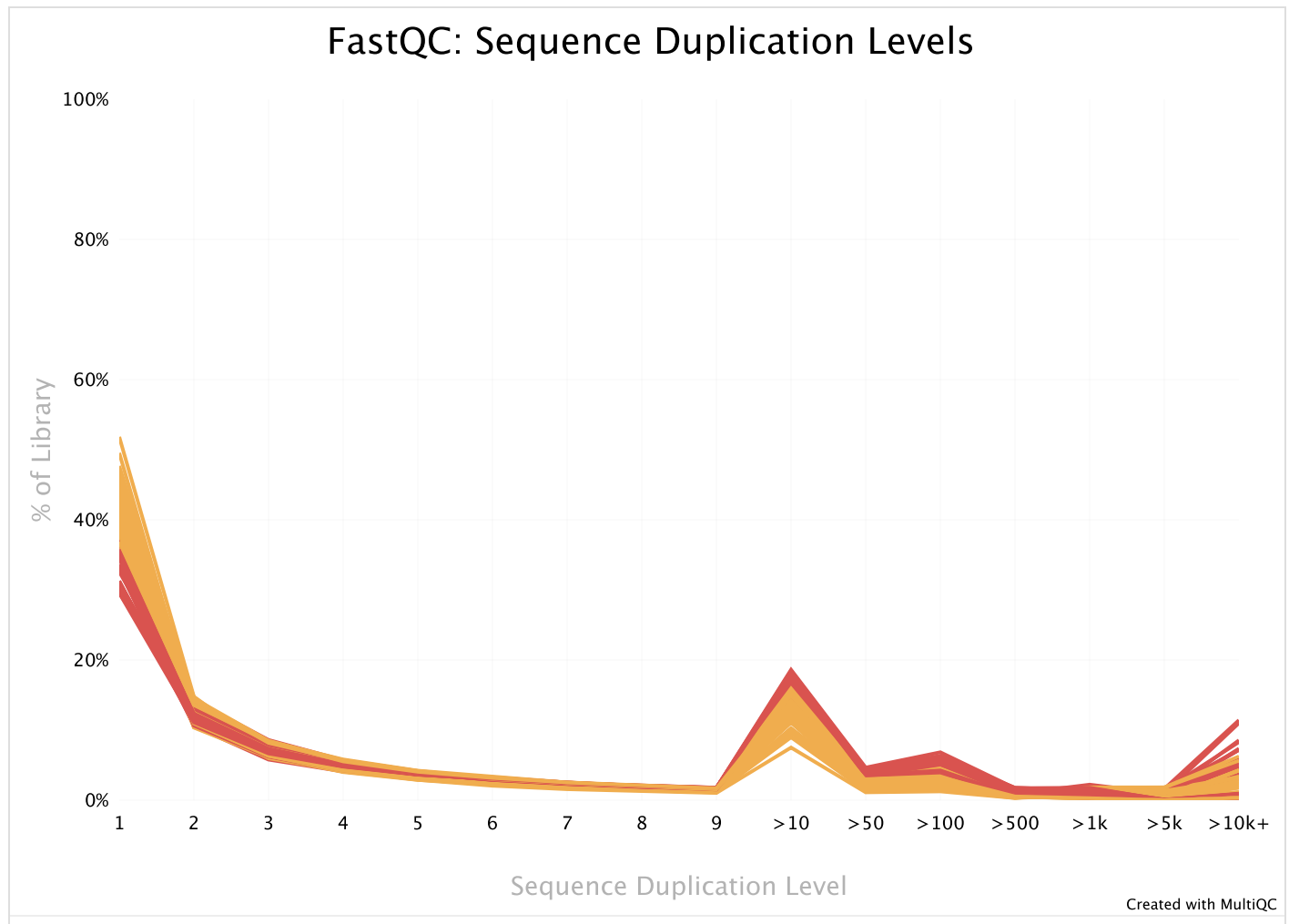
Export Plot



Sequence Duplication Levels

38

The relative level of duplication found for every sequence.

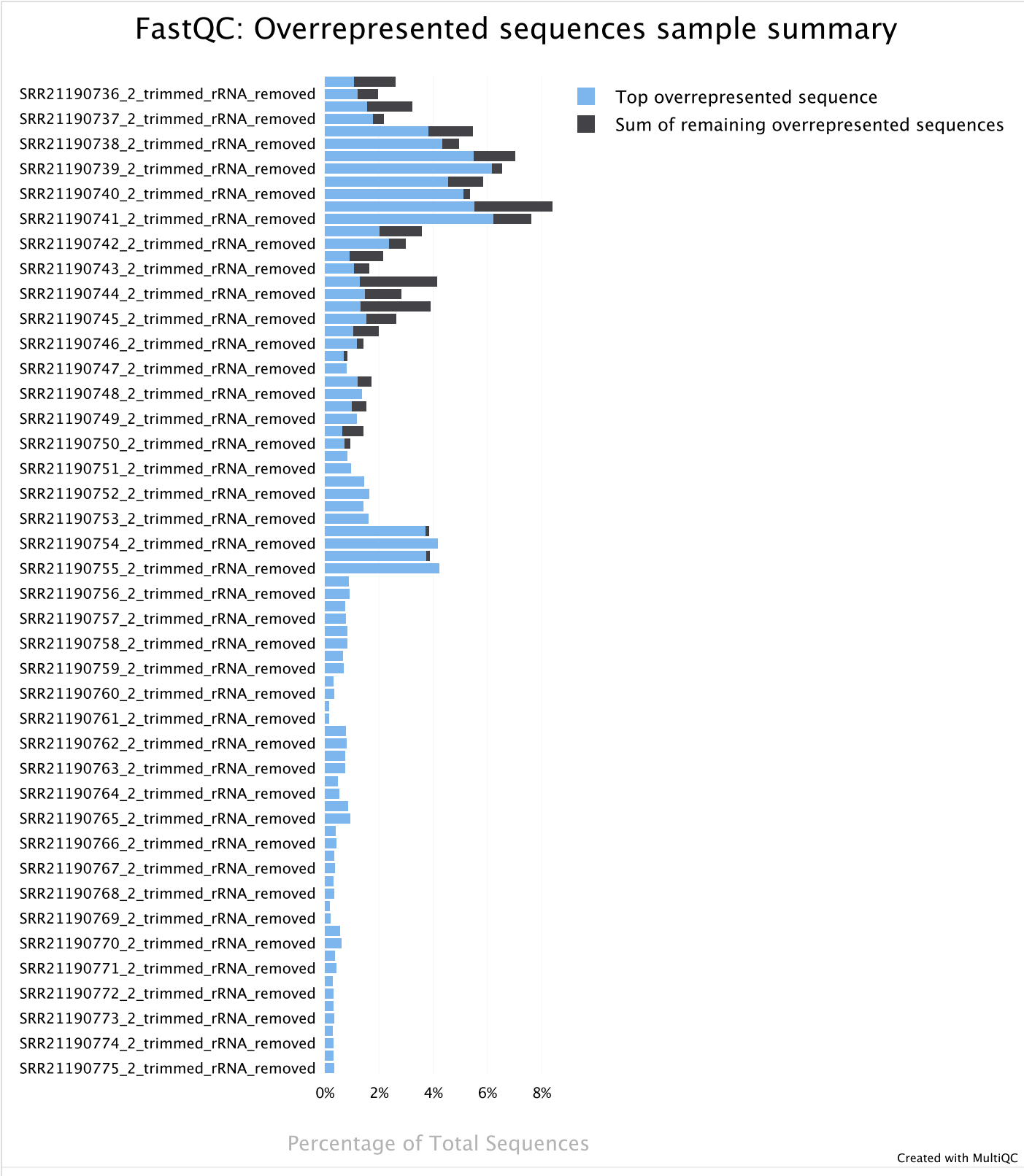
[Export Plot](#)

Overrepresented sequences by sample

47


The total amount of overrepresented sequences found in each library.

Export Plot



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.

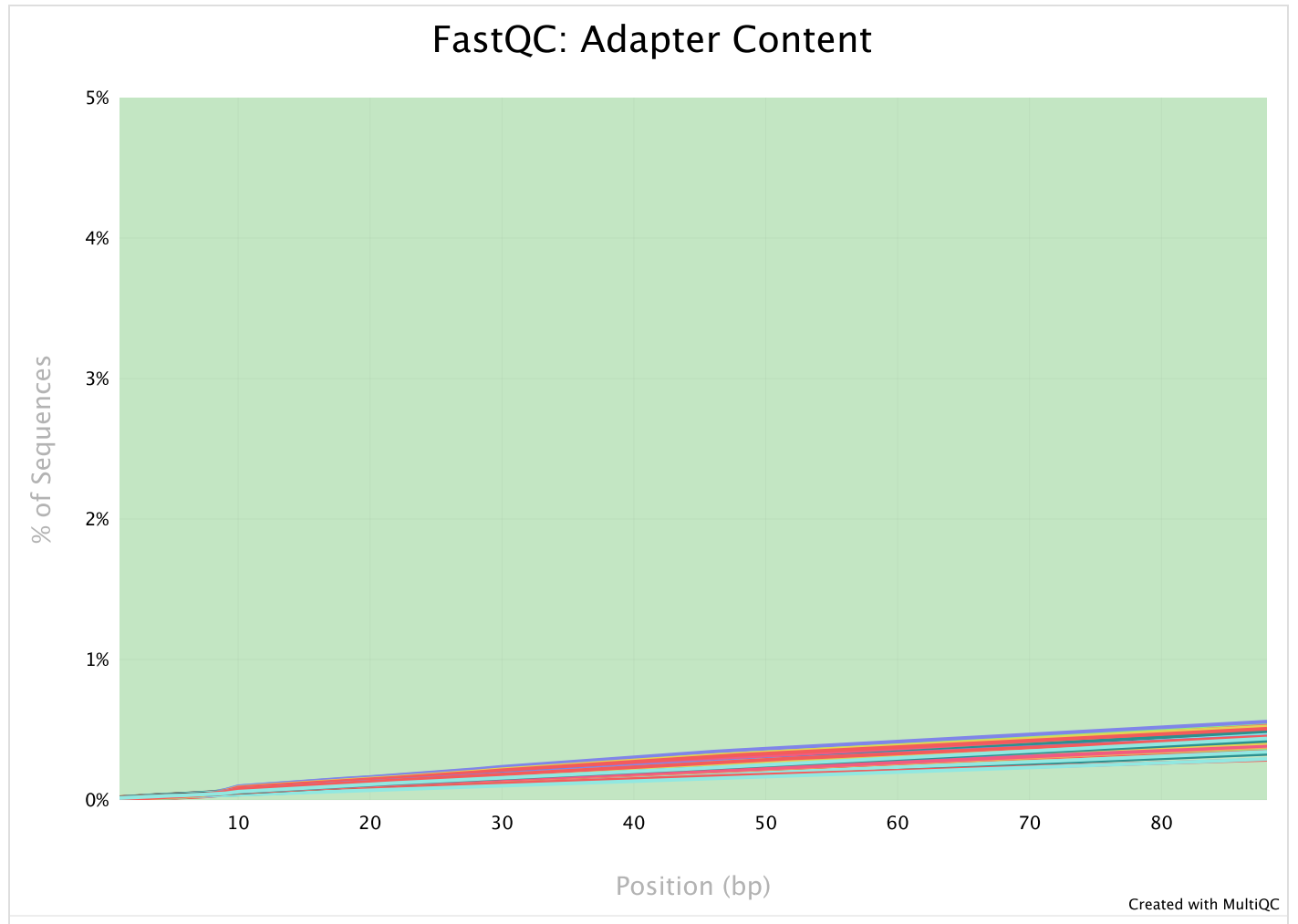
 Copy table
  Configure columns
  Scatter plot
  Violin plot
 Showing ²⁰/₂₀ rows and ³/₃ columns.

Overrepresented sequence	Samples	Occurrences	% of all reads
NNNNNNNNNNNNNNNNNNNN	80	36 208 925	1.4578%
GTCTGGAGTCTTGAAGCT	15	ACTACCCTACGTTCTCTTACAAATGGACC	0.0401%
CTCCGTTTCCGACCTGG	14	GGTTCACCCCTCCTTGGCAACCTGGTGGT	0.0424%
CCCCTCCTTAGGCAACC	14	GGTCCCCCGCTCCCCGAGGTCACCATAT	0.0360%
CTGGAGTCTTGAAGCT	14	TACCCTACGTTCTCTTACAAATGGACCTT	0.0302%
CCCTCCTTAGGCAACCT	12	GGTCCCCCGCTCCCCGAGGTCACCATATT	0.0256%
GCTCCGTTTCCGACCTG	11	CGGTTACCCCTCCTTACGCAACCTGGTGG	0.0246%
CCTCCTTAGGCAACCTG	10	GTCCCCCGCTCCCCGAGGTCACCATATTG	0.0176%
CGGTGGCGCGTGCCTGT	10	TCCCAGCTACTCGGGAGGCTGAGGCTGGAGG	0.0181%
CATCTGTCACCCCATG	9	GCCAGGGTTGATTCCGGTCTGCTGGCTGGC	0.0137%
GGTGGCGCGTGCCTGT	9	CCCAGCTACTCGGGAGGCTGAGGCTGGAGGA	0.0165%
GTTCTGGGCTGTAGTGCC	9	TATGCCGATCGGGTGTCTGCACTAAGTTCGG	0.0149%
GGGCGATCTGGCTGCGA	7	TCTGTCACCCCATGATGCGCAGGGTTGATT	0.0106%
CNNNNNNNNNNNNNNNNNN	6	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	0.0102%
CCTTAGGCAACCTGGTG	6	CCCCCGCTCCCGGGAGGTCACCATATTGATG	0.0107%
CTCAGGCTGGAGTGCA	5	GGCTATTCACAGGCGCGATCTCCACTACTGATC	0.0088%
GTGGCTATTCACAGGCGC	5	TATCCCACTACTGATCAGGACCGGAGTTTGA	0.0075%
CGGTGGCGCGTGCCTGT	5	GTCCCAGCTACTCGGGAGGCTGAGGCTGGGAGG	0.0085%
CCCAGCTACTCGGGAGG	5	TGAGGCTGGAGGATCGCTTCCAGGAGTT	0.0071%
GCTCAGGCTGGAGTGCA	4	TGGCTATTCACAGGCGCGATGCTCCACTACTGAT	0.0073%

Adapter Content

80

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

[Export Plot](#)

Status Checks

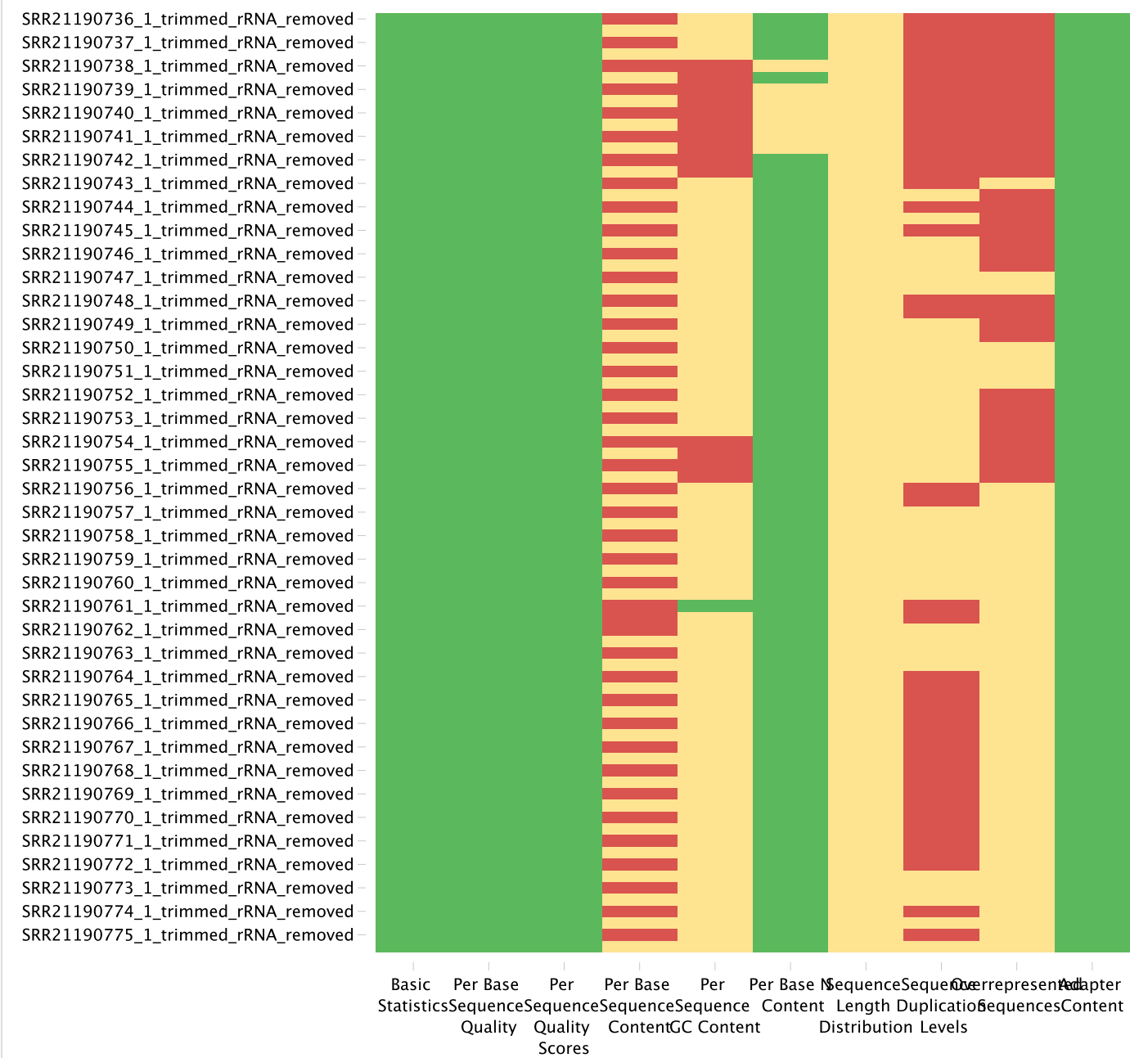
Status for each FastQC section showing whether results seem entirely normal (green), slightly abnormal (orange) or very unusual (red).

Export Plot

Min:

Max:

FastQC: Status Checks



Created with MultiQC

Software Versions

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

 Copy table

Software	Version
FastQC	0.12.1

MultiQC v1.20 - Written by [Phil Ewels](#), available on [GitHub](#).

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

