



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

Report generated on 2024-02-20, 16:10 EST based on data in:
/athena/cayuga_0019/scratch/mij4011/homework/project/fastqc_results

General Statistics

Copy table

Configure Columns

Plot

Showing 80/80 rows and 3/6 columns.

Sample Name	% Dups	% GC	M Seqs
SRR21190736_1	53.1%	44%	32.0
SRR21190736_2	50.7%	44%	32.0
SRR21190737_1	54.6%	44%	33.0
SRR21190737_2	51.3%	44%	33.0
SRR21190738_1	56.8%	44%	31.9
SRR21190738_2	55.4%	44%	31.9
SRR21190739_1	58.4%	44%	30.6
SRR21190739_2	56.2%	44%	30.6
SRR21190740_1	56.8%	44%	34.5
SRR21190740_2	55.1%	44%	34.5
SRR21190741_1	55.9%	46%	38.0
SRR21190741_2	54.2%	46%	38.0
SRR21190742_1	52.0%	44%	29.2
SRR21190742_2	49.3%	44%	29.2
SRR21190743_1	51.6%	43%	30.7
SRR21190743_2	49.1%	43%	30.7
SRR21190744_1	50.8%	44%	29.6
SRR21190744_2	48.2%	44%	29.6
SRR21190745_1	51.7%	44%	30.6
SRR21190745_2	49.2%	44%	30.6

Sample Name	% Dups	% GC	M Seqs
SRR21190746_1	47.4%	44%	37.7
SRR21190746_2	45.0%	44%	37.7
SRR21190747_1	43.4%	43%	32.6
SRR21190747_2	40.7%	43%	32.6
SRR21190748_1	52.4%	44%	34.0
SRR21190748_2	49.6%	45%	34.0
SRR21190749_1	43.7%	44%	34.3
SRR21190749_2	40.7%	44%	34.3
SRR21190750_1	47.9%	44%	36.0
SRR21190750_2	45.5%	44%	36.0
SRR21190751_1	43.4%	42%	32.6
SRR21190751_2	41.3%	42%	32.6
SRR21190752_1	38.6%	43%	28.2
SRR21190752_2	35.3%	43%	28.2
SRR21190753_1	41.8%	43%	36.6
SRR21190753_2	39.2%	43%	36.6
SRR21190754_1	42.4%	43%	36.2
SRR21190754_2	39.8%	42%	36.2
SRR21190755_1	46.1%	42%	36.4
SRR21190755_2	43.6%	42%	36.4
SRR21190756_1	50.3%	46%	34.1
SRR21190756_2	48.5%	46%	34.1
SRR21190757_1	46.0%	45%	23.3
SRR21190757_2	42.5%	46%	23.3
SRR21190758_1	47.8%	46%	31.2
SRR21190758_2	42.5%	46%	31.2
SRR21190759_1	46.9%	45%	28.3
SRR21190759_2	44.2%	46%	28.3
SRR21190760_1	49.5%	46%	32.6
SRR21190760_2	46.6%	46%	32.6

Sample Name	% Dups	% GC	M Seqs
SRR21190761_1	51.8%	42%	28.5
SRR21190761_2	49.6%	42%	28.5
SRR21190762_1	45.0%	45%	26.7
SRR21190762_2	42.3%	45%	26.7
SRR21190763_1	47.0%	45%	27.3
SRR21190763_2	44.1%	45%	27.3
SRR21190764_1	56.9%	46%	34.4
SRR21190764_2	55.0%	46%	34.4
SRR21190765_1	56.6%	45%	34.6
SRR21190765_2	55.8%	45%	34.6
SRR21190766_1	53.9%	46%	37.1
SRR21190766_2	50.8%	46%	37.1
SRR21190767_1	51.9%	46%	34.0
SRR21190767_2	49.7%	46%	34.0
SRR21190768_1	53.7%	46%	34.7
SRR21190768_2	51.4%	46%	34.7
SRR21190769_1	52.9%	46%	34.1
SRR21190769_2	50.9%	46%	34.1
SRR21190770_1	56.1%	46%	28.0
SRR21190770_2	52.3%	46%	28.0
SRR21190771_1	51.4%	45%	30.2
SRR21190771_2	49.5%	45%	30.2
SRR21190772_1	49.7%	45%	29.6
SRR21190772_2	47.5%	45%	29.6
SRR21190773_1	48.9%	45%	27.0
SRR21190773_2	46.3%	45%	27.0
SRR21190774_1	51.1%	45%	32.7
SRR21190774_2	48.6%	45%	32.7
SRR21190775_1	50.9%	45%	34.4
SRR21190775_2	47.7%	45%	34.4

FastQC

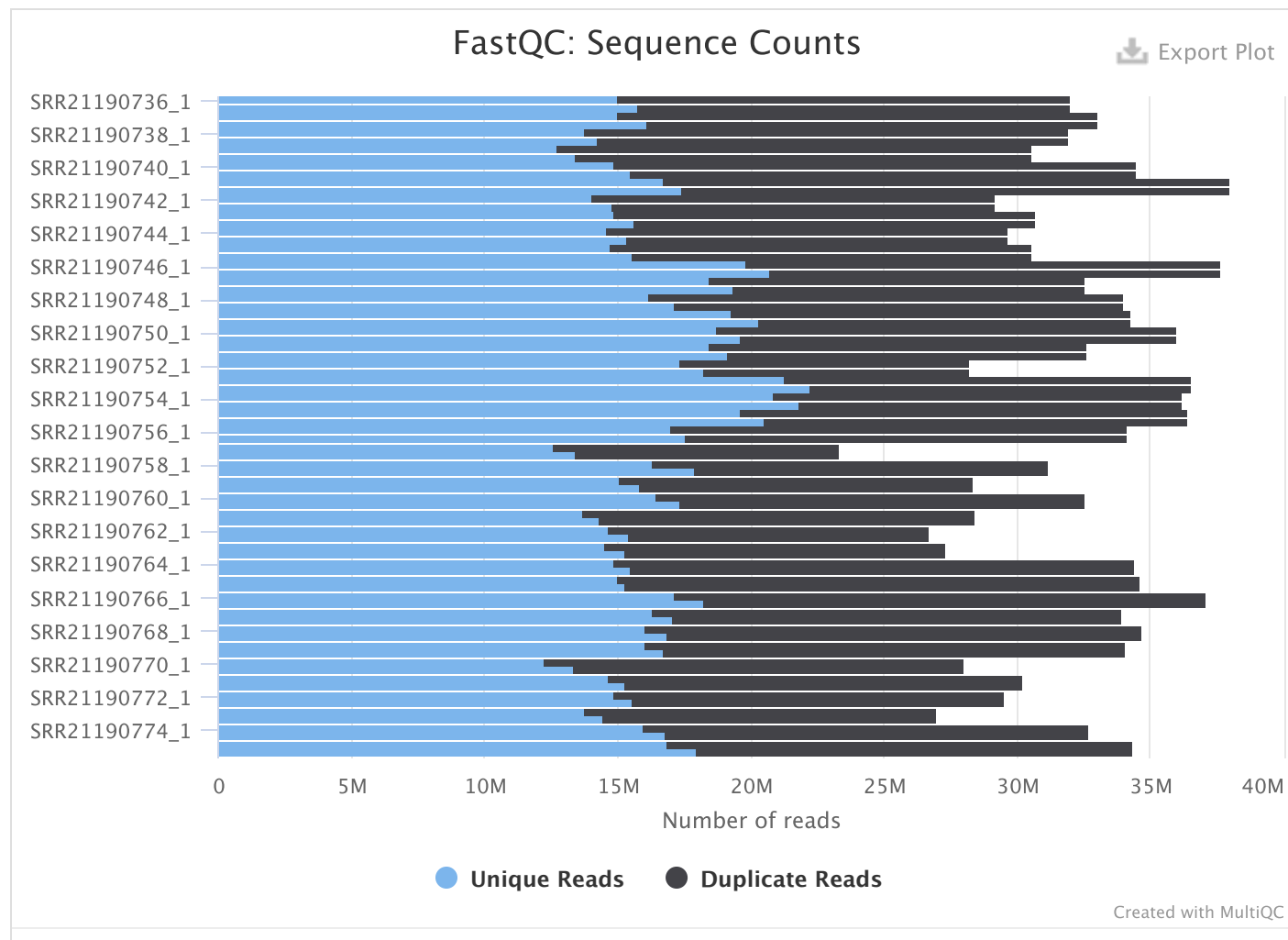
Version: 0.12.1

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.

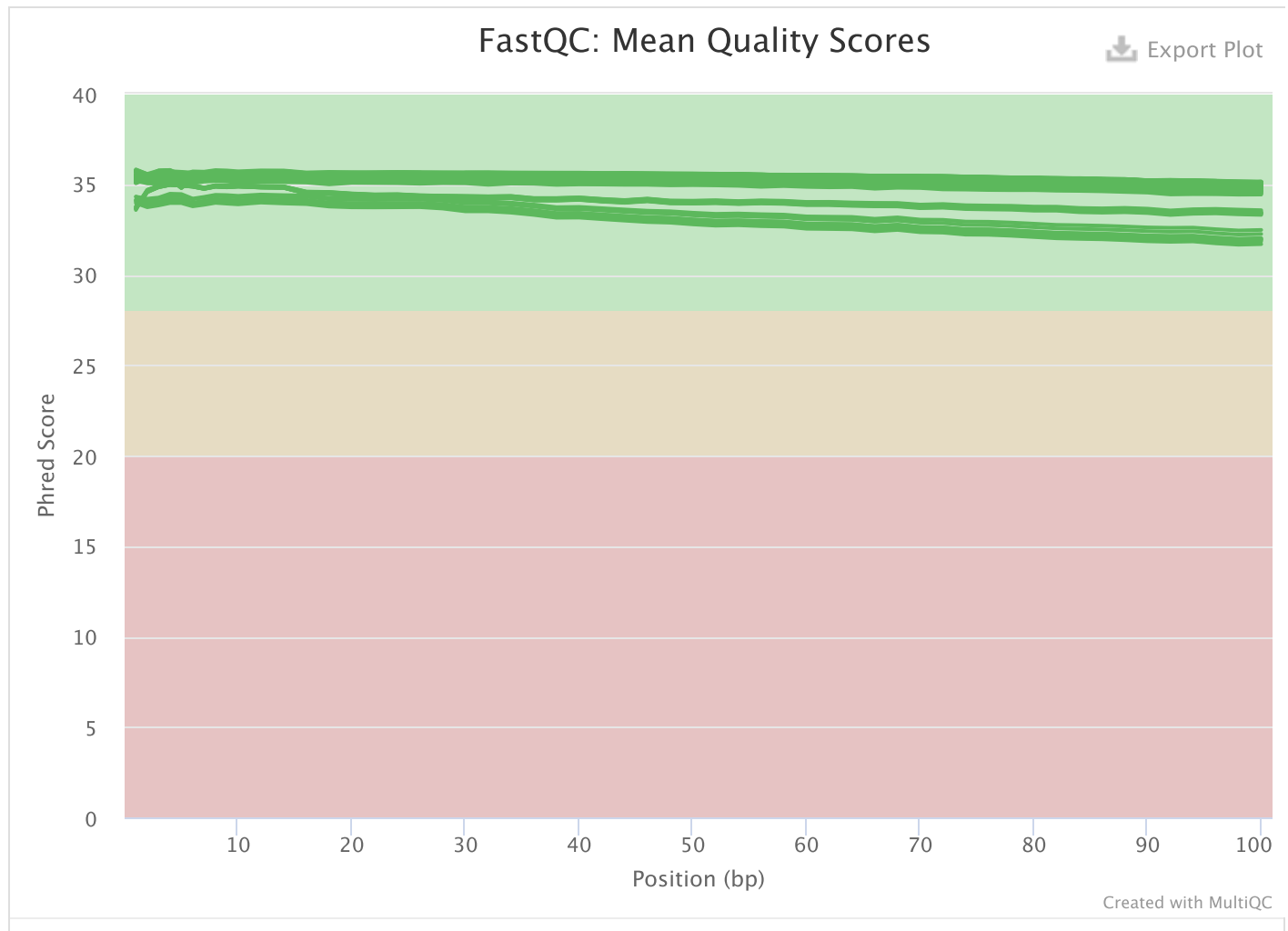
Number of reads Percentages



Sequence Quality Histograms

80

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



Per Sequence Quality Scores

80

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



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

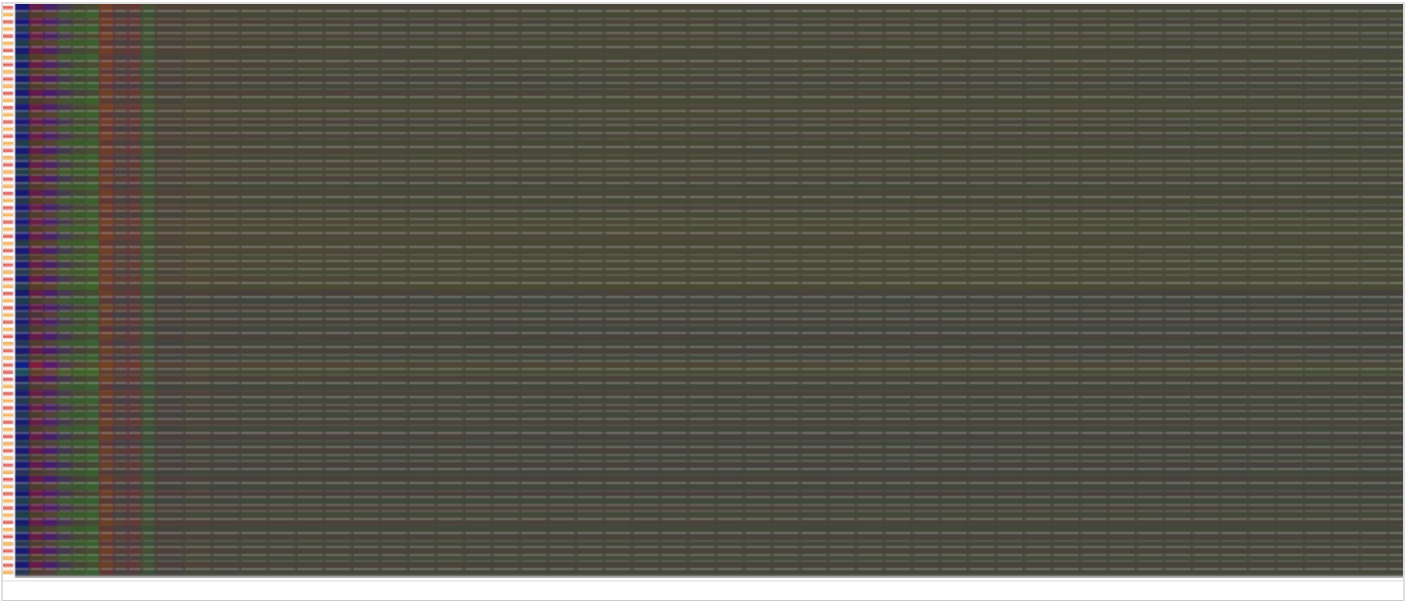
Position: -
%G: -

%T: -

%C: -

%A: -

⬇️ Export Plot



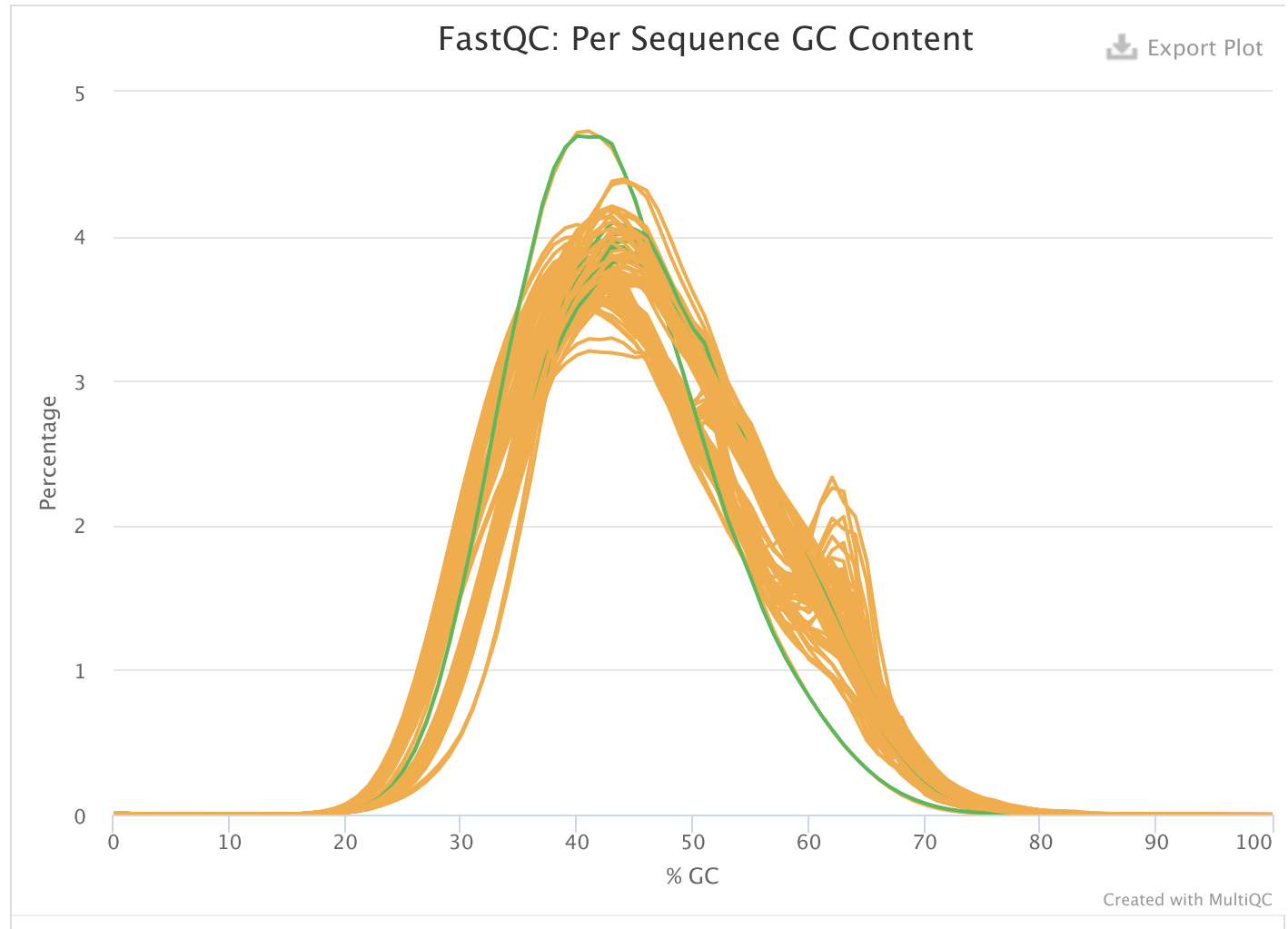
Per Sequence GC Content

7 73

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

Percentages

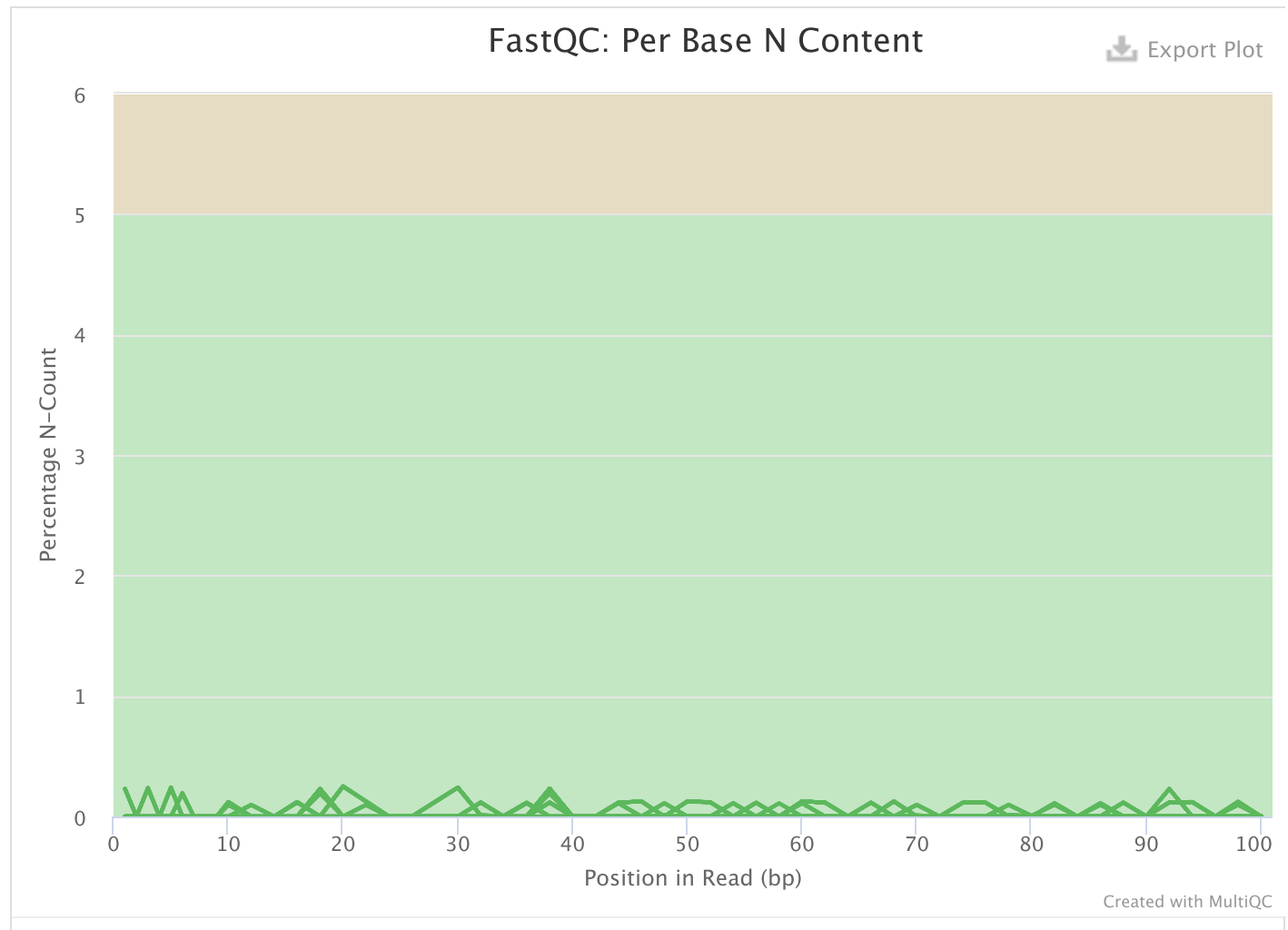
Counts



Per Base N Content

80

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



Sequence Length Distribution

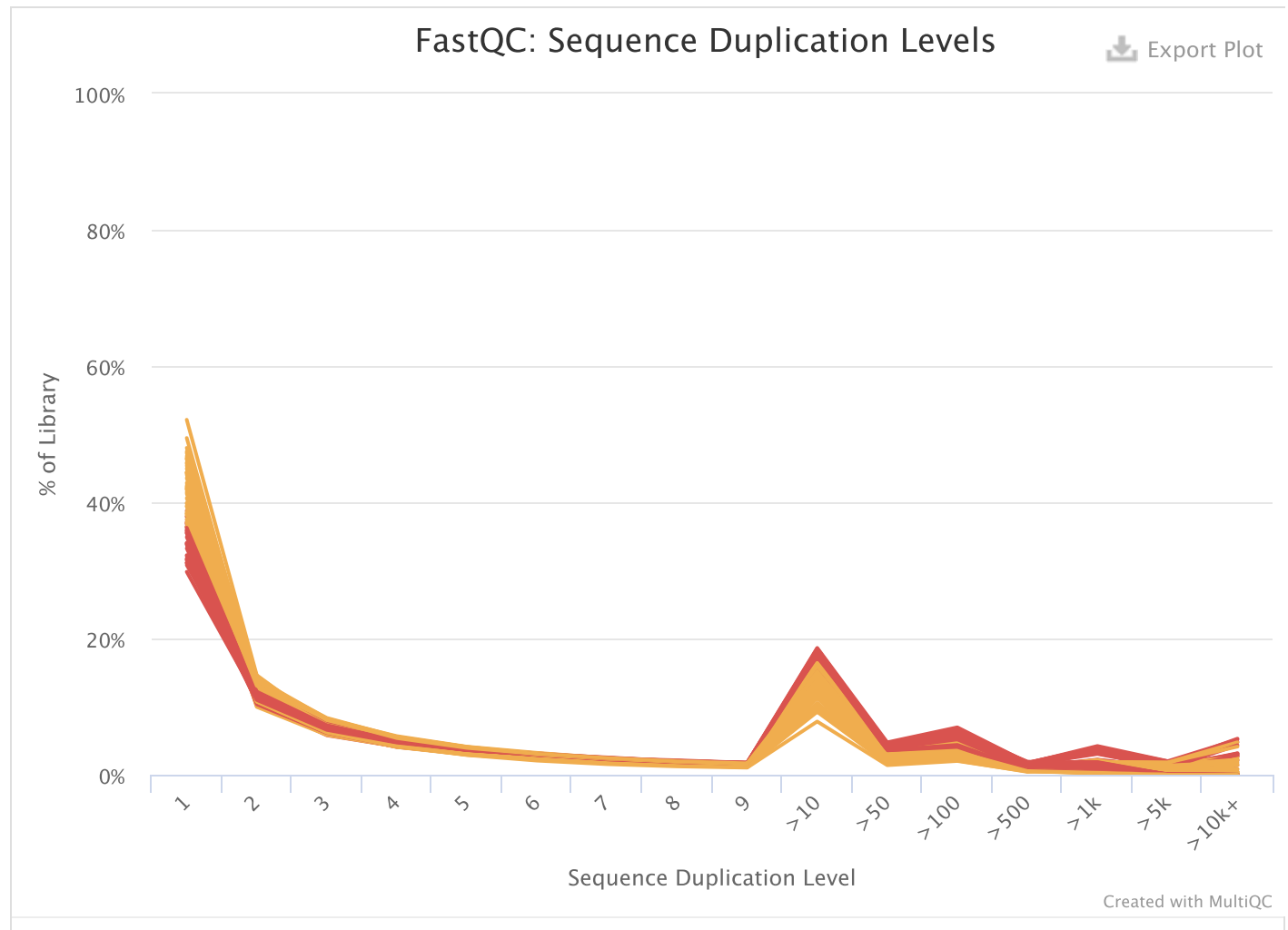
80

All samples have sequences of a single length (100bp).

Sequence Duplication Levels

45

The relative level of duplication found for every sequence.

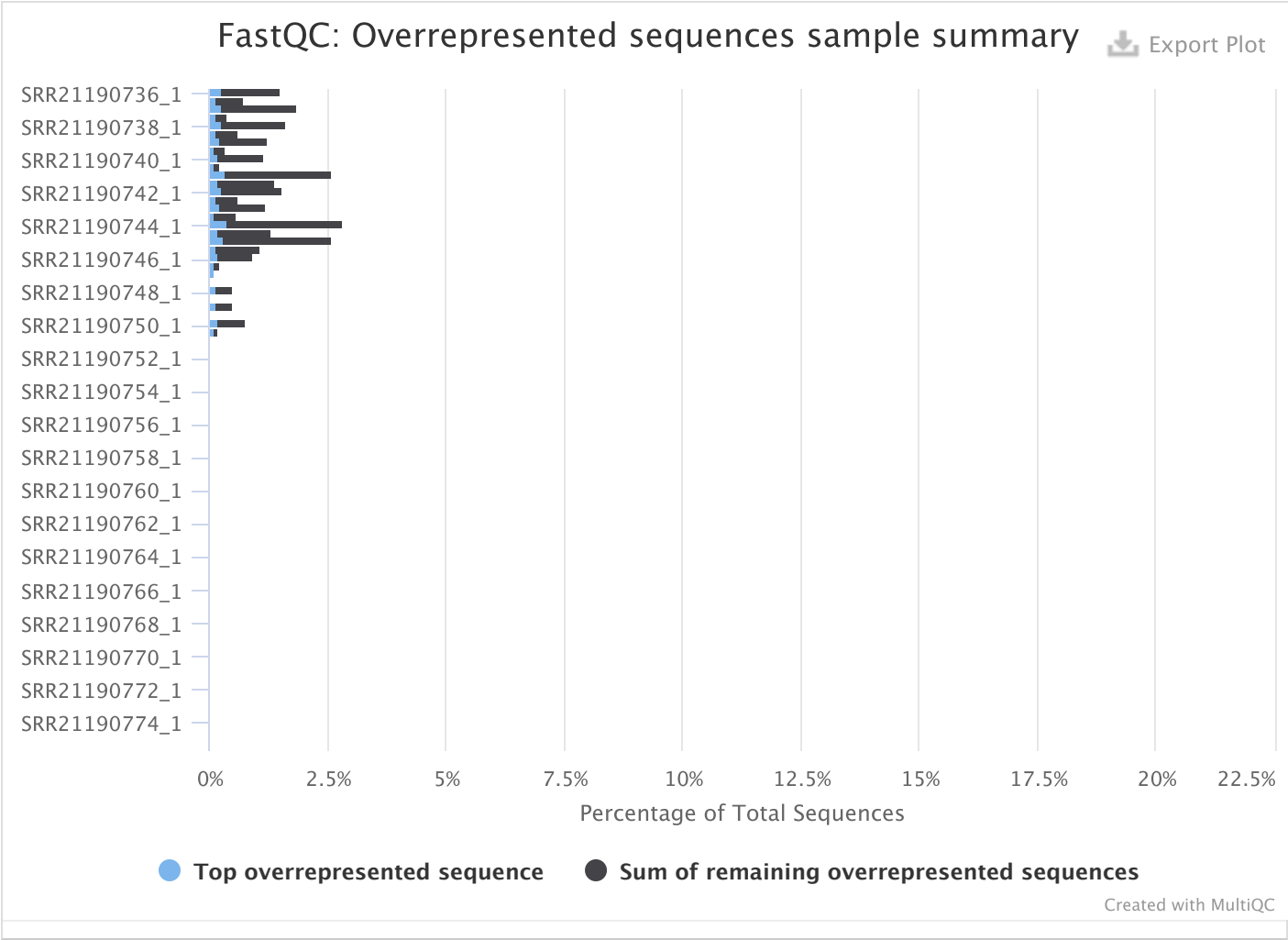


Overrepresented sequences by sample

53

27

The total amount of overrepresented sequences found in each library.



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](#)
[Plot](#)

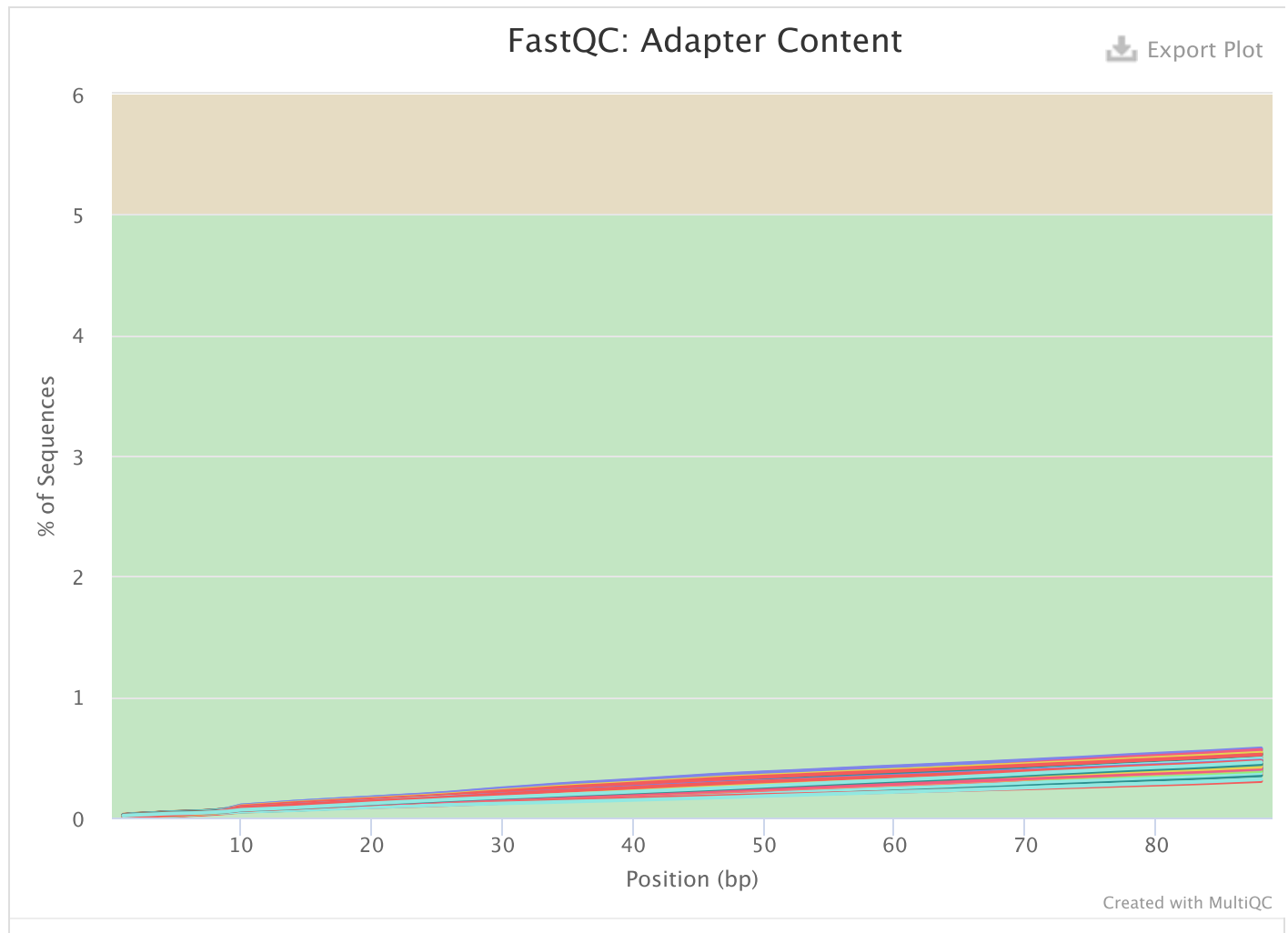
Showing 20/20 rows and 3/3 columns.

Overrepresented sequence	Samples	Occurrences	% of all reads
GTCTGGAGTCTTGGAAGC	15	1013033	0.0393%
CTCCGTTTCCGACCTGG	14	1088496	0.0423%
CCCCTCCTTAGGCAACCT	14	916574	0.0356%
CTGGAGTCTTGGAAGCTT	14	762143	0.0296%
CCCTCCTTAGGCAACCTC	12	654533	0.0254%
GCTCCGTTTCCGACCTGG	11	634853	0.0247%
CCTCCTTAGGCAACCTGG	10	453868	0.0176%
CGGTGGCGCGTGCCTGT	10	464796	0.0180%
GGTGGCGCGTGCCTGTAT	9	418434	0.0162%
GTTCTGGGCTGTAGTGCC	9	378328	0.0147%
CATCTGTCACCCCATTGA	9	346110	0.0134%
CCTTAGGCAACCTGGTGC	7	309024	0.0120%
GGGCGATCTGGCTGCGA	7	264468	0.0103%
CTCAGGCTGGAGTGCAG	5	225016	0.0087%
GCTCAGGCTGGAGTGCA	5	219166	0.0085%
GTGGCTATTCACAGGCGC	5	193551	0.0075%
CCAGGCTGGAGTGCAGT	5	176601	0.0069%
CGGTGGCGCGTGCCTGT	5	218872	0.0085%
CCCAGCTACTCGGGAGG	5	180549	0.0070%
CTCCTTAGGCAACCTGGT	4	152845	0.0059%

Adapter Content

80

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



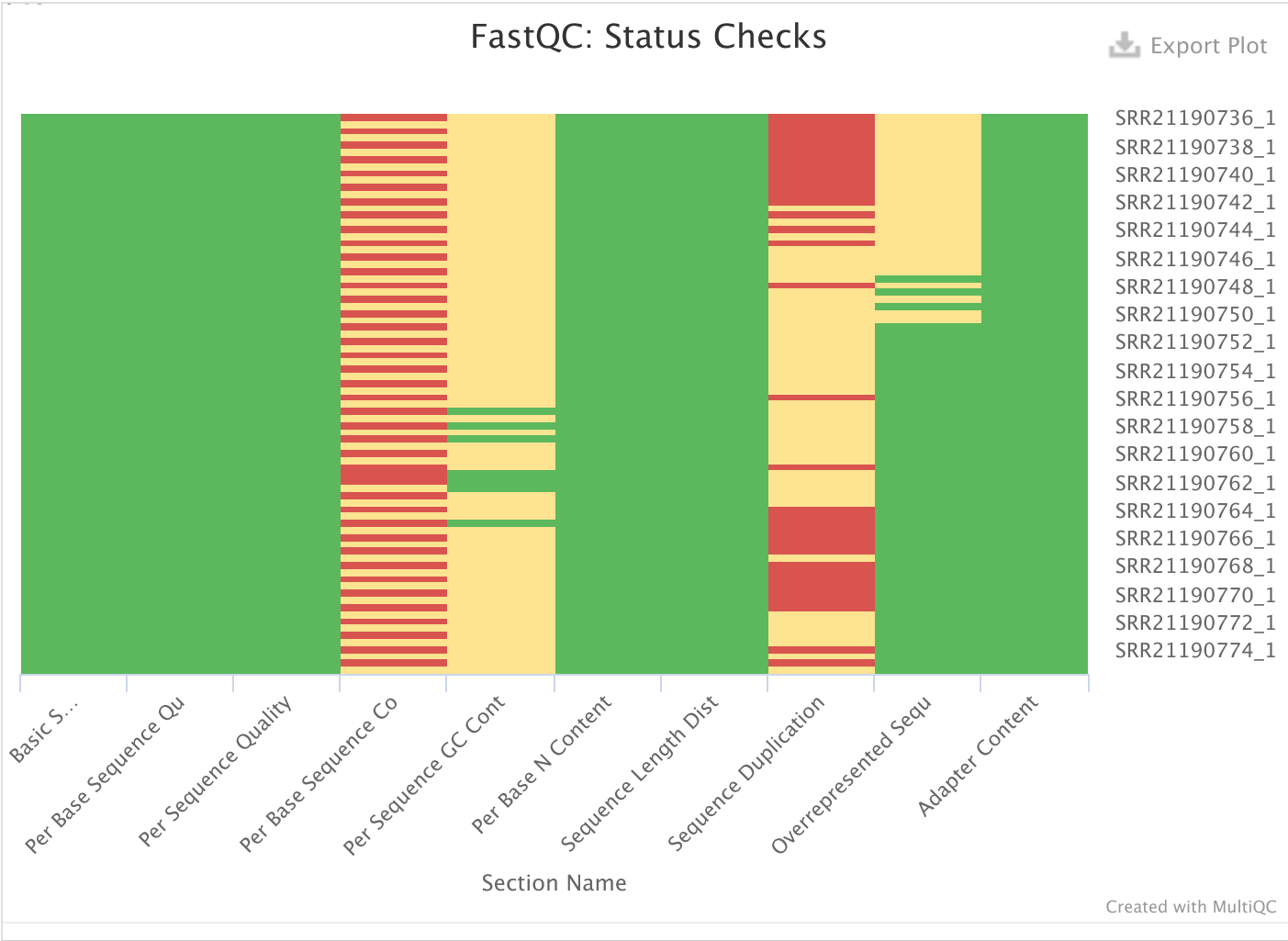
Status Checks

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

Sort by highlight

Min: 0

Max: 1



Software Versions

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

 Copy table

Software	Version
FastQC	0.12.1

MultiQC v1.17 - Written by [Phil Ewels](#), available on [GitHub](#).

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

