multiqc
v1.33

General Stats
Strandedness checks
Strandedness: inference summary
Strandedness: read composition
FastQC (raw)
Sequence Counts
Sequence Quality Histograms
Per Sequence Quality Scores
Per Base Sequence Content
Per Sequence GC Content
Per Base N Content
Sequence Length Distribution
Sequence Duplication Levels
Overrepresented sequences by sample
Top overrepresented sequences
Adapter Content
Status Checks
Cutadapt
Filtered Reads
Trimmed Sequence Lengths (3')
FastQC (trimmed)
Sequence Counts
Sequence Quality Histograms
Per Sequence Quality Scores
Per Base Sequence Content
Per Sequence GC Content
Per Base N Content
Sequence Length Distribution
Sequence Duplication Levels
Overrepresented sequences by sample
Top overrepresented sequences
Adapter Content
Status Checks
BBTools
BBSplit alignment statistics
BBSplit Statistics
FastQC (filtered)
Sequence Counts
Sequence Quality Histograms
Per Sequence Quality Scores
Per Base Sequence Content
Per Sequence GC Content
Per Base N Content
Sequence Length Distribution
Sequence Duplication Levels
Overrepresented sequences by sample
Top overrepresented sequences
Adapter Content
Status Checks
DupRadar
Picard
QualiMap
Genomic origin of reads
Gene Coverage Profile
RSeQC
Read Distribution
Inner Distance
Read Duplication
Junction Annotation
Junction Saturation
Infer experiment
Bam Stat
Samtools
Percent mapped
Alignment stats
Flagstat
Flagstat: Percentage of total
Mapped reads per contig
STAR
Summary Statistics
Alignment Scores
RSEM
Mapped Reads
Multimapping rates
Salmon
Sample relationships
STAR_RSEM DESeq2 sample similarity
STAR_RSEM DESeq2 PCA plot
SALMON DESeq2 sample similarity
SALMON DESeq2 PCA plot
Biotype Counts
Software Versions
nf-core/rnaseq Methods Description
nf-core/rnaseq Workflow Summary

multiqc

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

This report has been generated by the nf-core/rnaseq analysis pipeline. For information about how to interpret these results, please see the documentation.

Report generated on 2026-08-25, 21:28 CDT based on data in: `/bulk/bioinfo/Nextflow_demos/work/26/93cae08e18dd0c459507e5df3332ca`

Summarize report

Welcome! Not sure where to start? Watch a tutorial video *(6:06)* don't show again

General Statistics

Copy table | Configure columns | Scatter plot | Violin plot | Export as CSV... Showing 5/5 rows and 29/53 columns. Summarize table

Sample Name	human Assigned	primary Assigned	dupInt	Duplication	5'-3' bias	M Aligned	Alignable	Proper Pairs	% Aligned	M Aligned	Li
RAP1_IAA_30M_REP1	0.0 M	0.1 M	0.0	12.7 %	0.88	0.0 M	85.7 %	74.4 %	81.0 %	0.0 M	IS
RAP1_UNINDUCED_REP1		0.0 M	0.0	75.3 %	0.94	0.0 M	88.9 %	0.0 %	81.0 %	0.0 M	SR
RAP1_UNINDUCED_REP2	0.0 M	0.1 M	0.0	81.7 %	0.90	0.1 M	88.7 %	0.0 %	80.8 %	0.1 M	SR
WT_REP1		0.2 M	0.0	21.1 %	1.07	0.1 M	83.4 %	65.8 %	79.9 %	0.1 M	IS
WT_REP2	0.0 M	0.1 M	0.0	13.3 %	1.17	0.0 M	84.4 %	72.8 %	79.8 %	0.0 M	IS

Strandedness checks

Strandedness declared in the samplesheet versus the values inferred by Salmon and/or RSeQC infer_experiment.

Strandedness: inference summary

One row per sample showing the strandedness declared in the samplesheet (auto means Salmon was asked to infer it) alongside the value each active inference method reported. The Salmon % and RSeQC % columns are a per-method certainty metric for forward / reverse calls: the inferred direction's share of the strand-classified fragment pool (the same quantity the configured stranded_threshold is applied to). Blank for unstranded or undetermined classifications, where different thresholds apply. Per-component Sense / Antisense / Unstranded percentages are hidden by default; toggle them on via the Configure Columns button above the table.

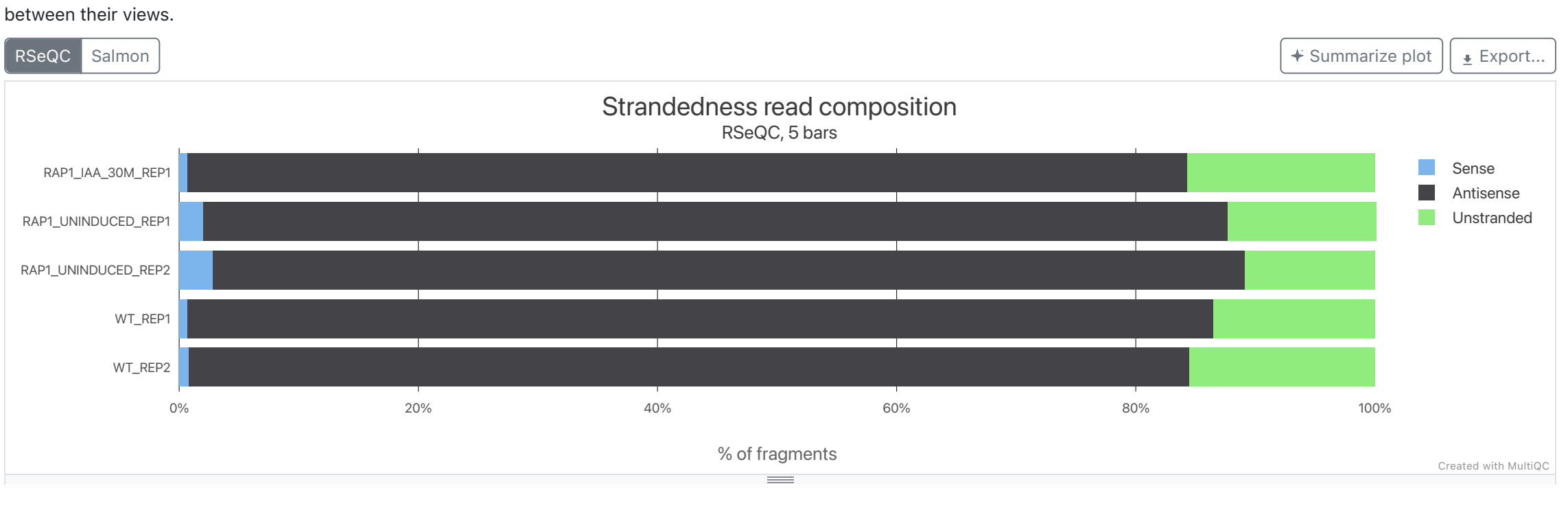
Copy table | Configure columns | Scatter plot | Violin plot | Export as CSV... Showing 5/5 rows and 6/12 columns. Summarize table

Sample Name	Provided	Salmon inferred	Salmon %	RSeQC inferred	RSeQC %	Status
RAP1_IAA_30M_REP1	reverse	-		reverse	99.2 %	pass
RAP1_UNINDUCED_REP1	reverse	-		reverse	97.8 %	pass
RAP1_UNINDUCED_REP2	reverse	-		reverse	96.8 %	pass
WT_REP1	auto	reverse	89.2 %	reverse	99.2 %	pass
WT_REP2	reverse	-		reverse	99.0 %	pass

Strandedness: read composition

Stacked percentage of sense, antisense and unstranded read fragments per sample. When both Salmon and RSeQC inferred strandedness, use the dataset switcher above the plot to flip between their views.

RSeQC | Salmon — Summarize plot | Export...



FastQC (raw)

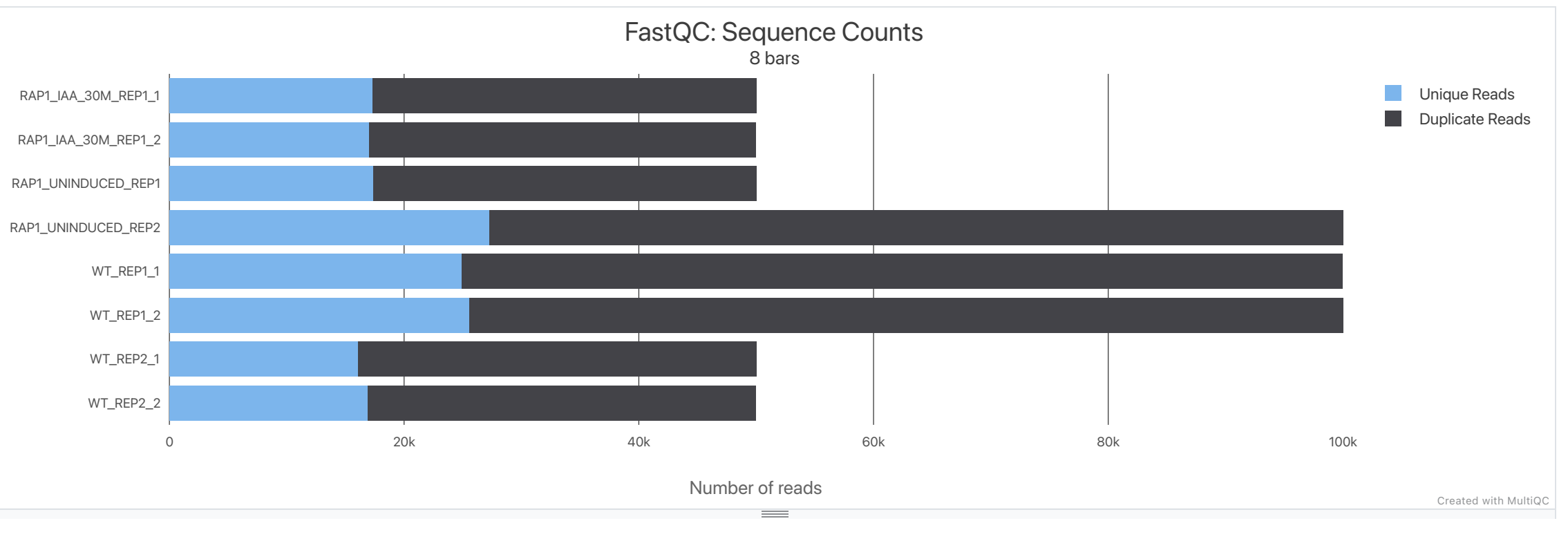
This section of the report shows FastQC results before adapter trimming. http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Sequence Counts

Help

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

Percentages — Summarize plot | Export...



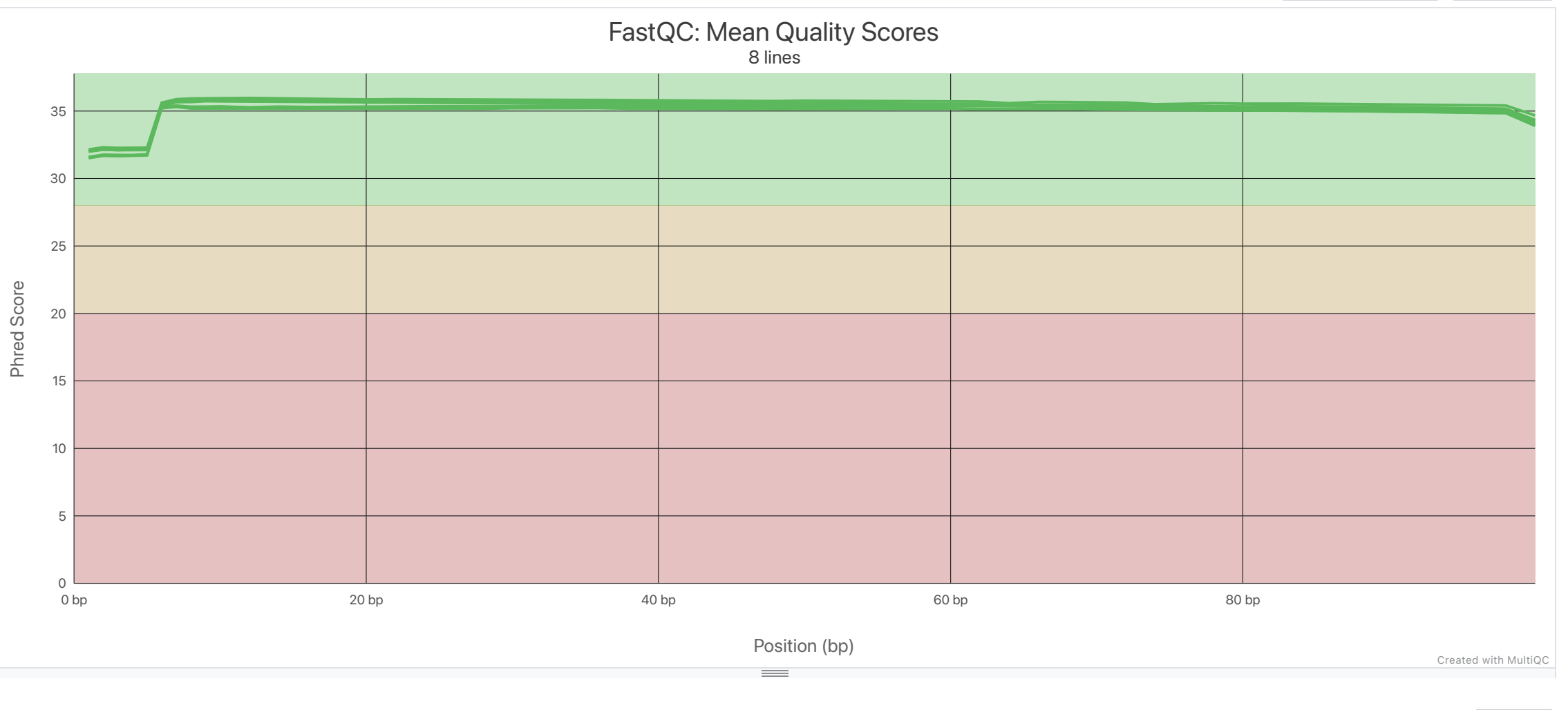
Sequence Quality Histograms

8

Help

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

Summarize plot | Export...



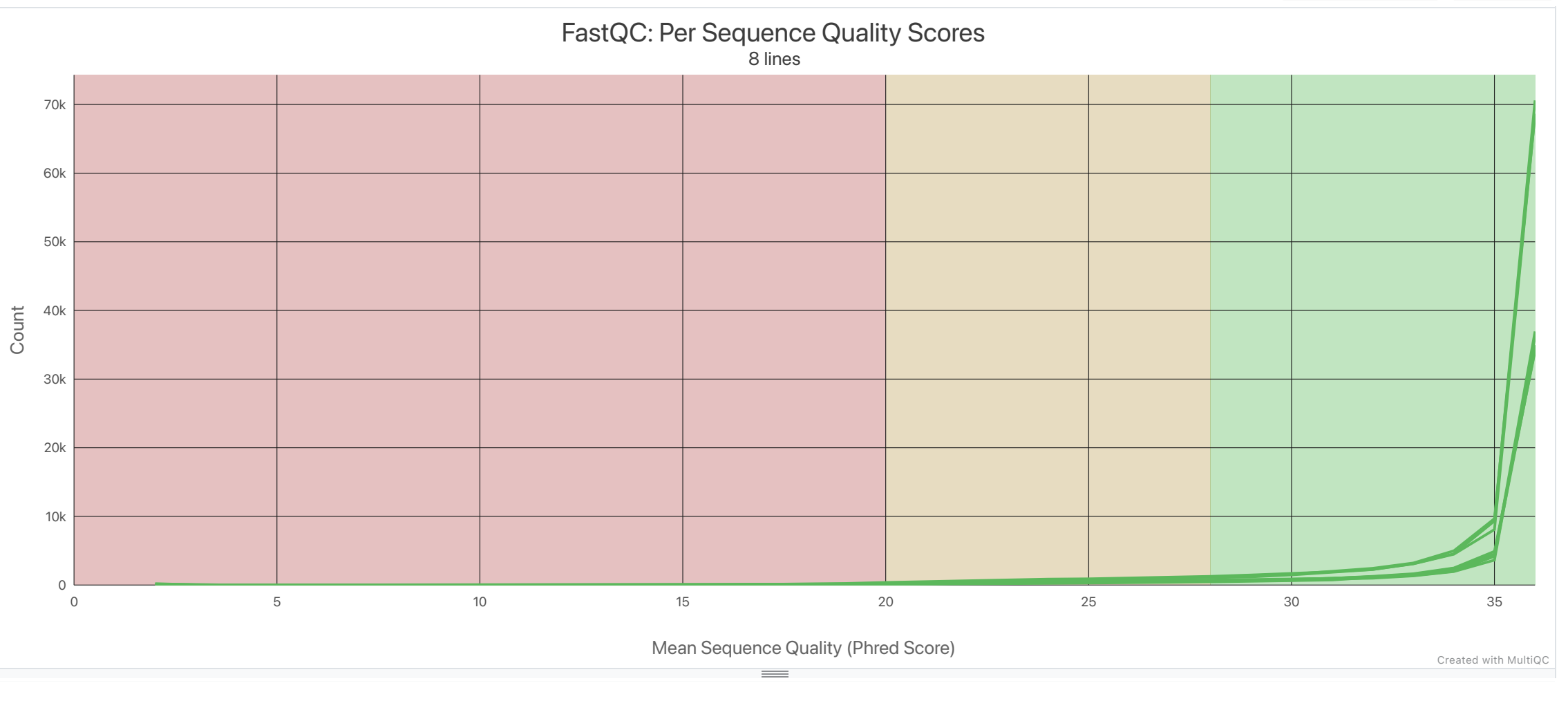
Per Sequence Quality Scores

8

Help

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

Summarize plot | Export...



Per Base Sequence Content

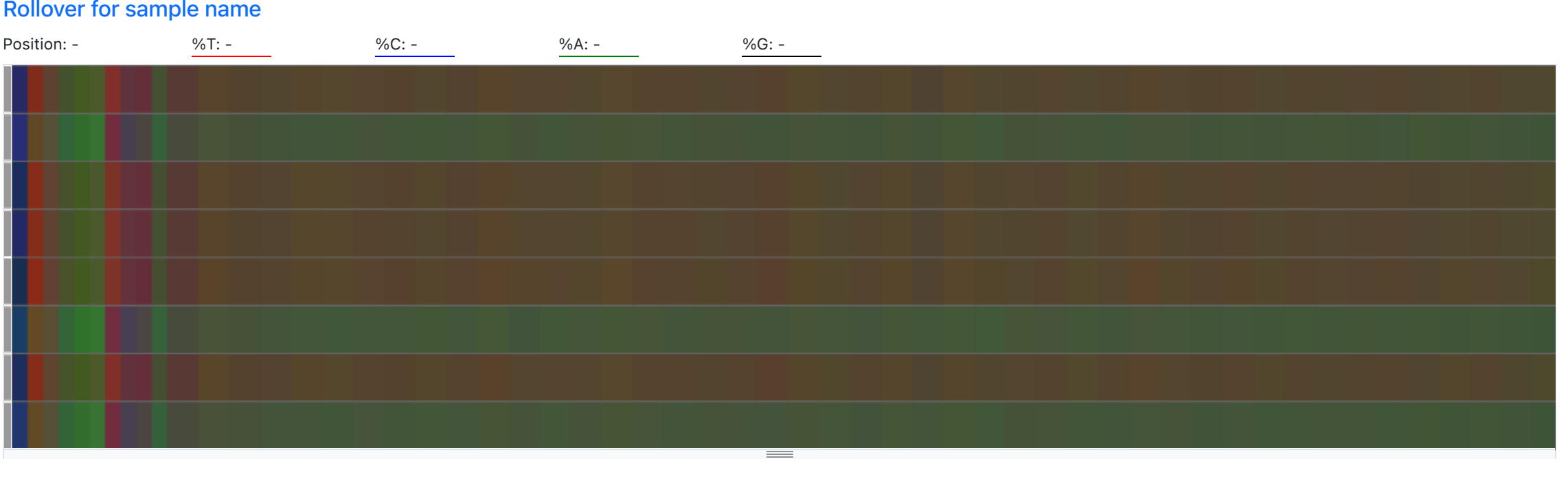
Help

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

Position: - %T: - %C: - %A: - %G: -



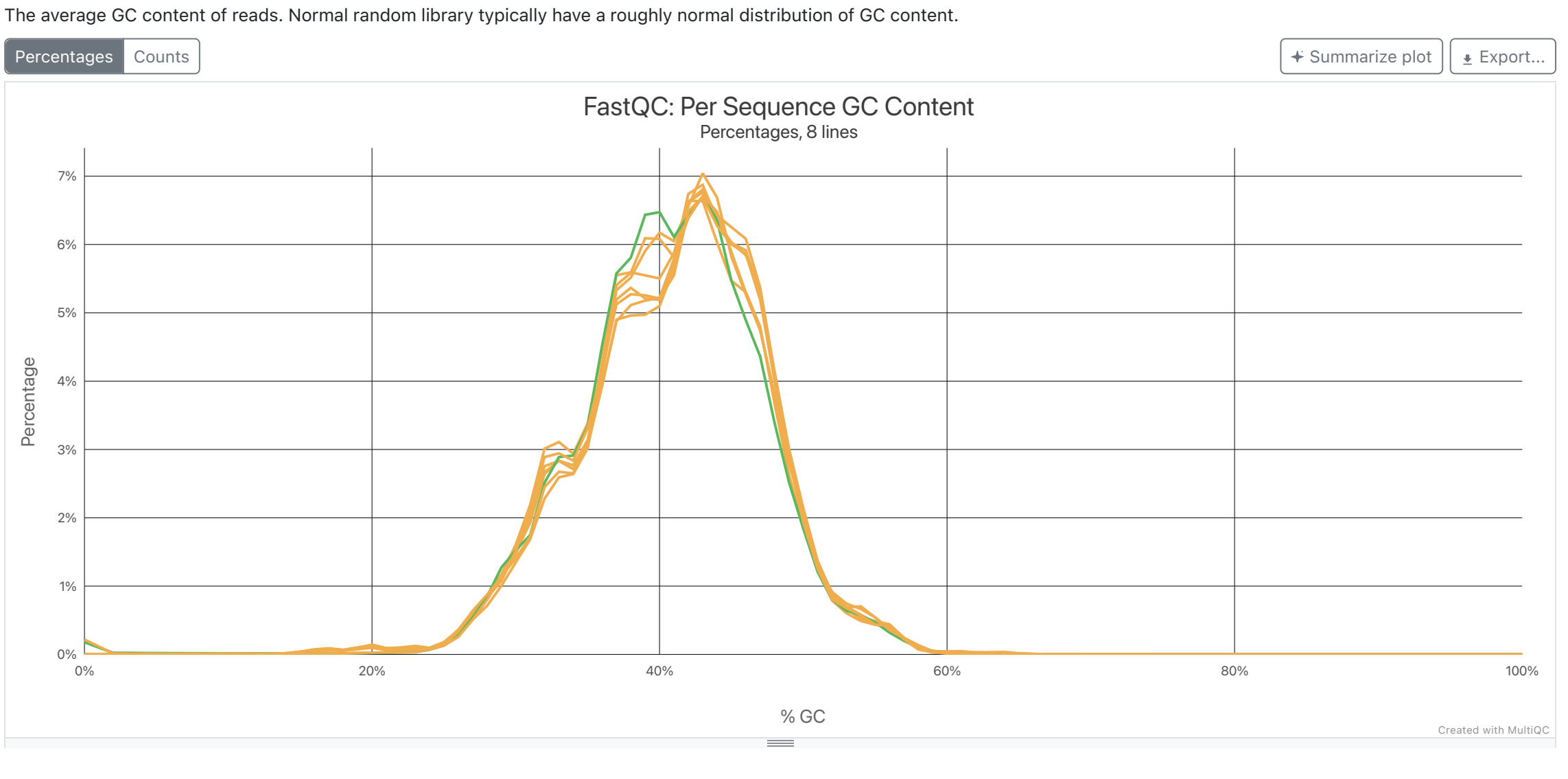
Per Sequence GC Content

1 7

Help

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

Percentages | Counts — Summarize plot | Export...



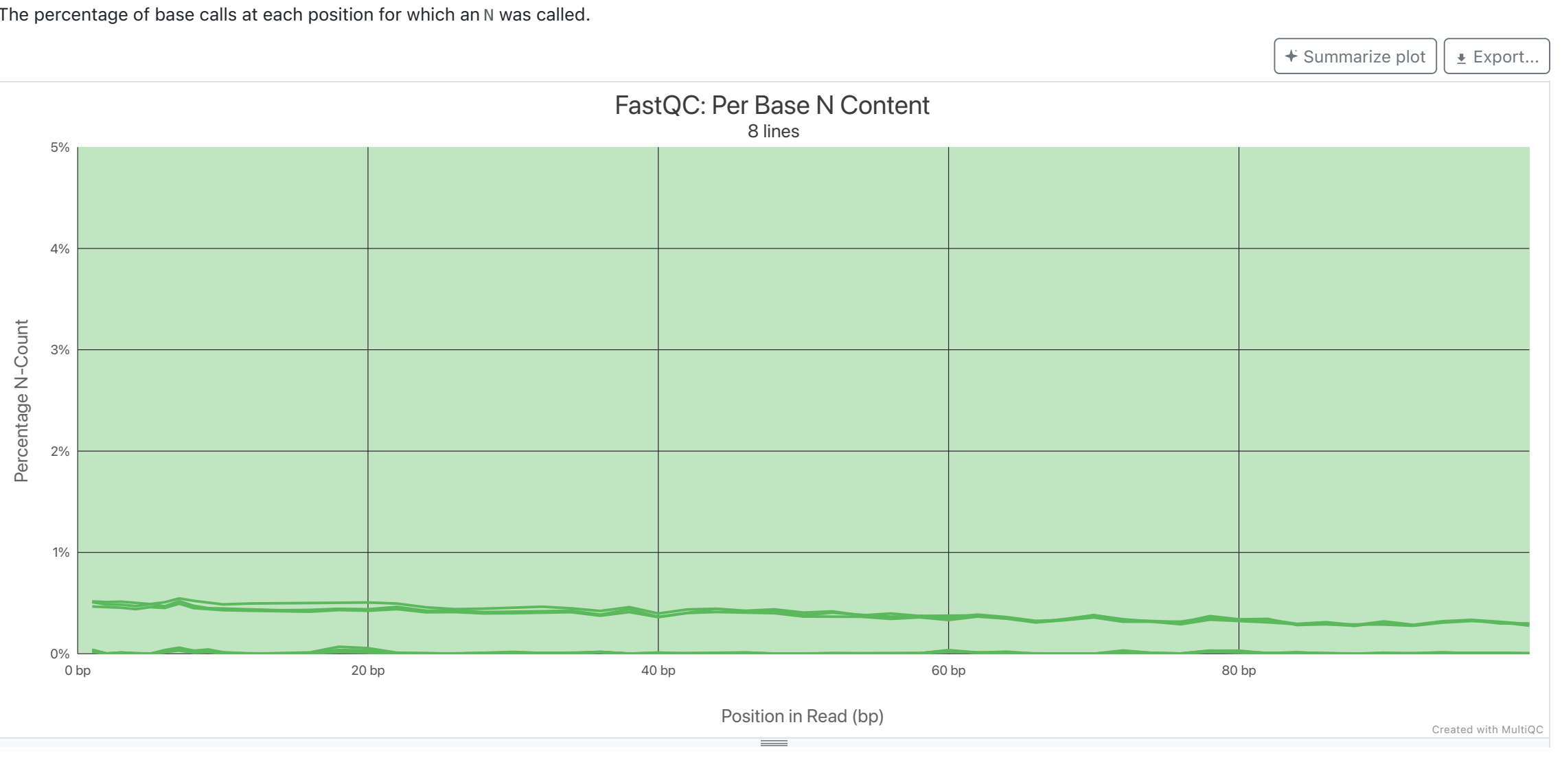
Per Base N Content

8

Help

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

Summarize plot | Export...



Sequence Length Distribution

8

All samples have sequences of a single length (101bp).

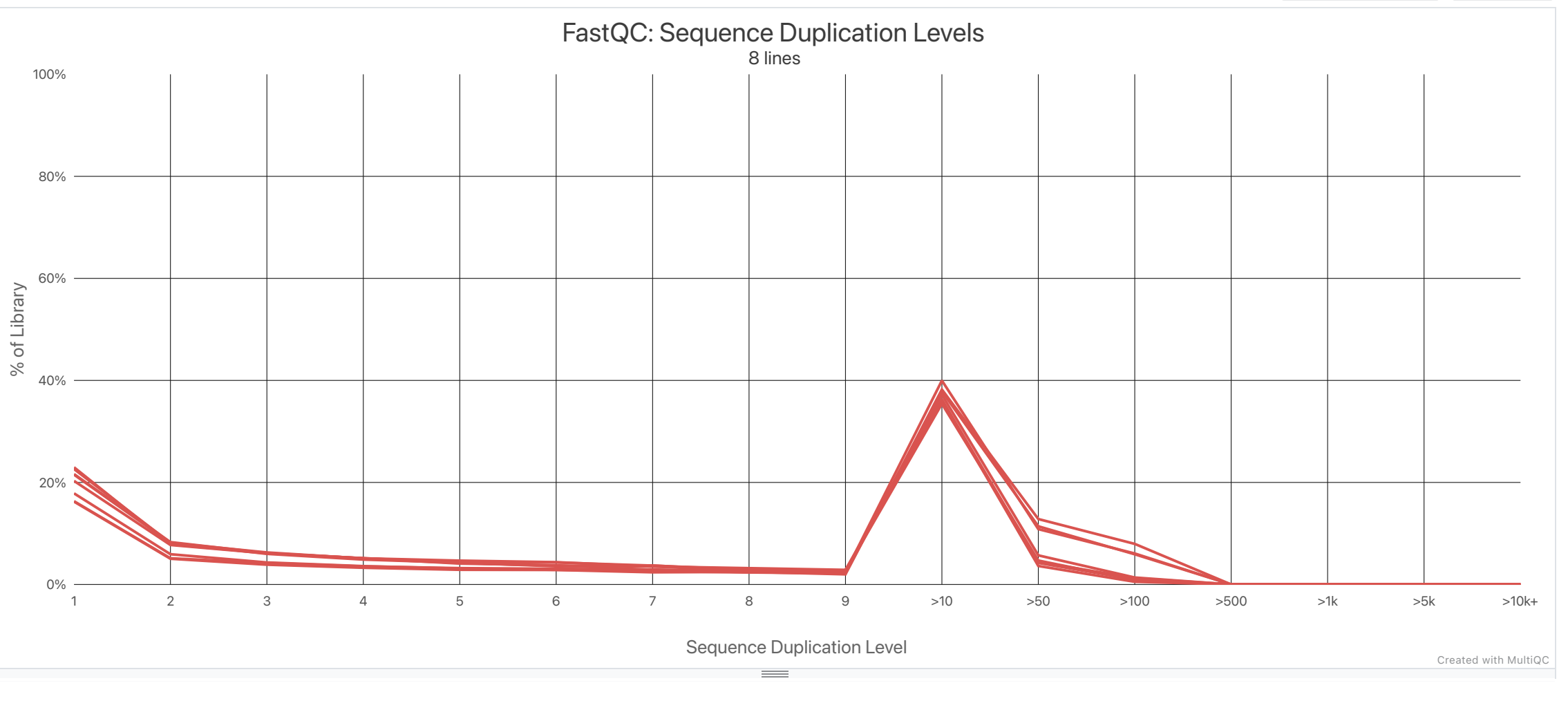
Sequence Duplication Levels

8

Help

The relative level of duplication found for every sequence.

Summarize plot | Export...

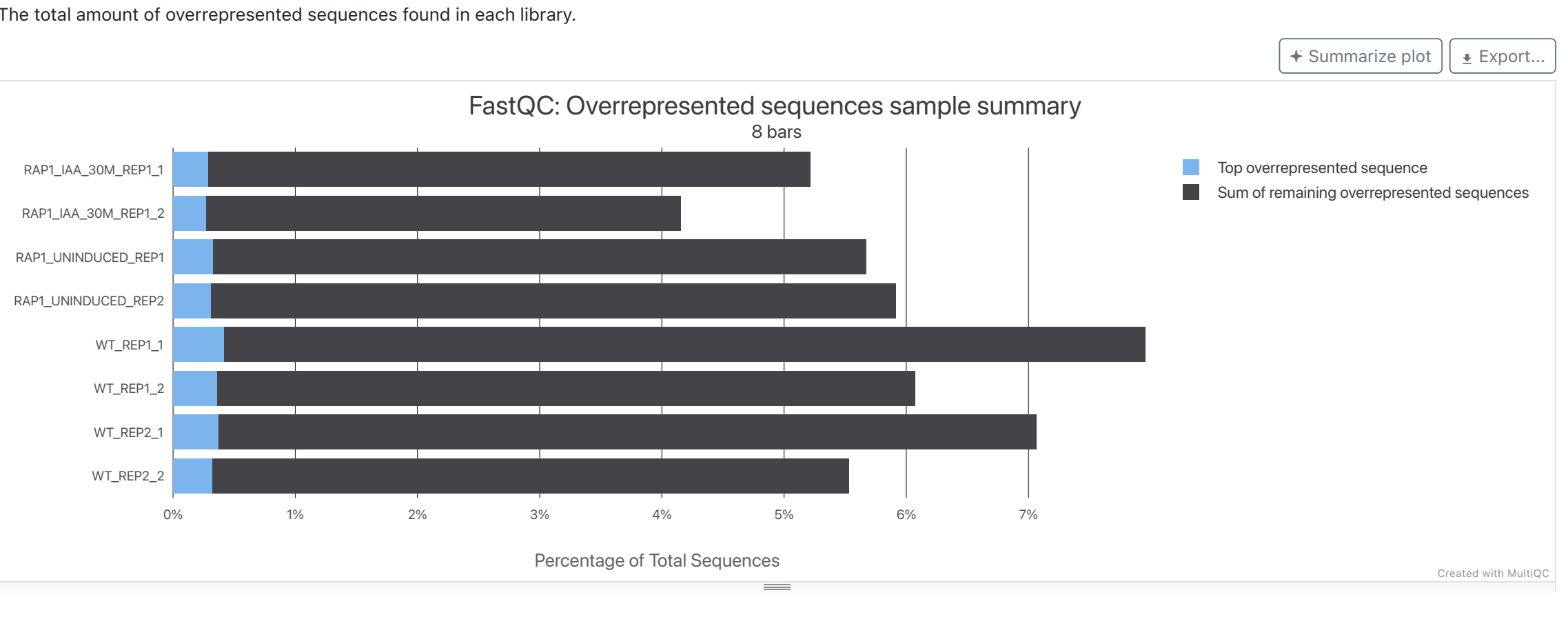


Overrepresented sequences by sample

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
CTTTGACCCAGGTAGGTAGGAATTGAGAGGGAGCAAATATTTGAAAAGTG	5	1 217	0.2213 %
GTTAGGAGTGTGCAATACCTTTATTGATGTTTTGGTGTTGTCCTGGAAGT	5	1 037	0.1885 %
CTTCGATCTCGGAGGTTCTAAACTACGCATATTCTTAGTATTCCATGTGT	5	826	0.1502 %
CGAGATTCGAGTTTTGTAAAAATTTGATGTATGTTTTAACCCAATTTGGA	5	784	0.1425 %
CCTCATTCCGTTTTATATGTTTATATTCATTGATCCTATTACATTATCAA	5	736	0.1338 %
CTTTAAATGGGTCAGGGAATTCGGTAGGAGATTCTGGAGGTAGATCAGGG	5	691	0.1256 %
CTGGAAGTGAAATTGTAGGTCACCAATAGCGTTAATTGGTATATTTCTTT	5	633	0.1151 %
GTTTGATTGATTTTACTGCTTTTACAGCTGCAATCAGGTGAATTCGTTTC	5	624	0.1135 %
GTCAACTTCTCTGGGTGCGCGAATATTGGTTTTAGAAACACGTTTCGAAT	5	613	0.1115 %
CTAAAACTTGCCTAGAGGGGAATAGTATATGTTGAGCAAGTGTGTTGATG	5	592	0.1076 %
AGAGAATCTAGCAGCTCTTGGGCATCTGGTAACCAAGATGATTGGACAGT	5	578	0.1051 %
CTCATTTTCACTGTACGAGTCTGAGTGTCTGAAATCTTTAGATTTACTGG	5	560	0.1018 %
CAATGGTCTACCTGGGTACAATTCTTCGGACTTTCTGGCGTTGTCAATGA	5	551	0.1002 %
GTTCAGTCTGTCGATAATGTTTGTGACTTTTGTTTCAAATGCATCTGCAG	5	535	0.0973 %
GTGATATCTACTGCAGCCAATTCATTCAAACTGAGTAAGTCATAGGCTAT	5	526	0.0956 %

Expand table

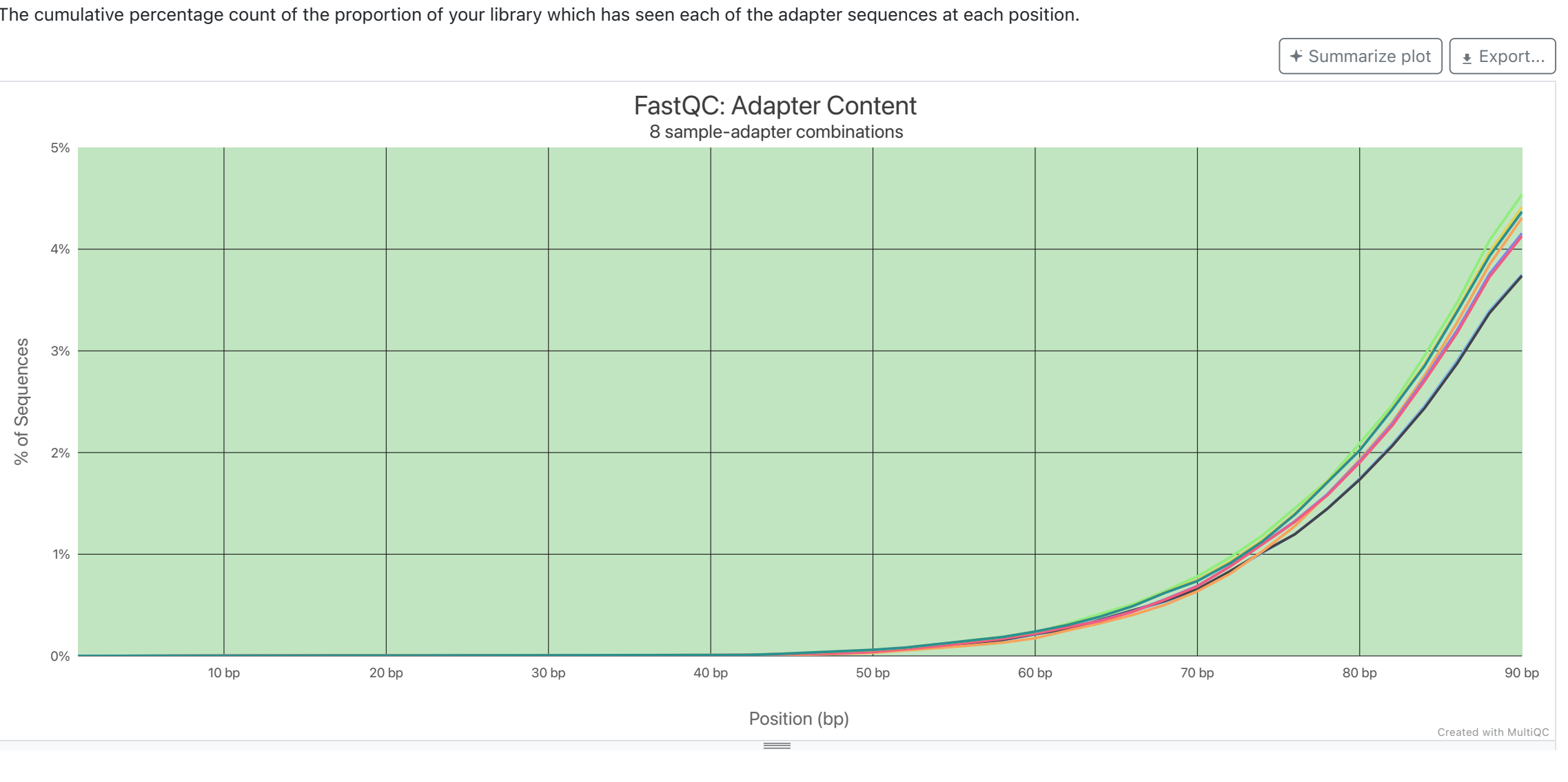
Adapter Content

8

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...

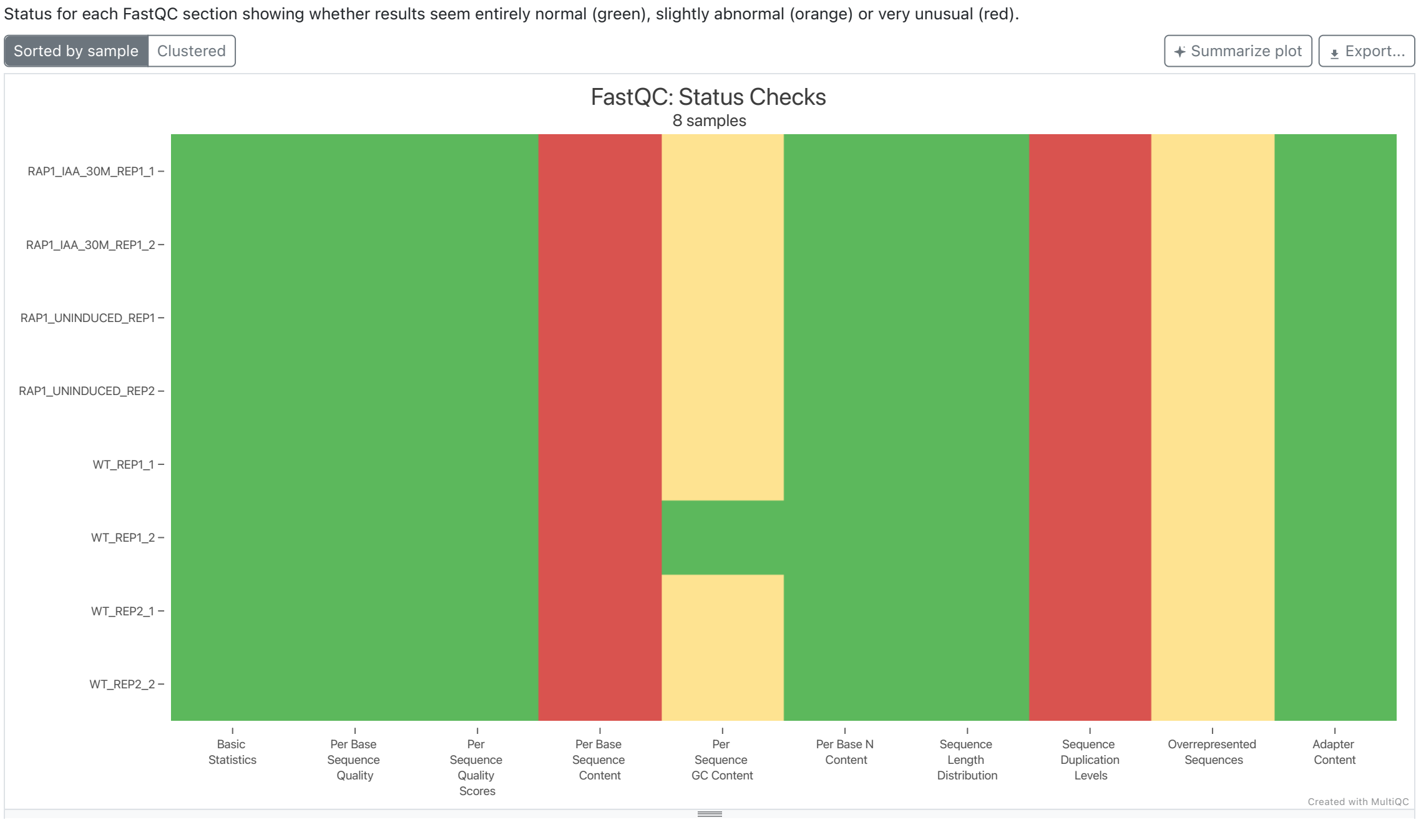


Status Checks

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...



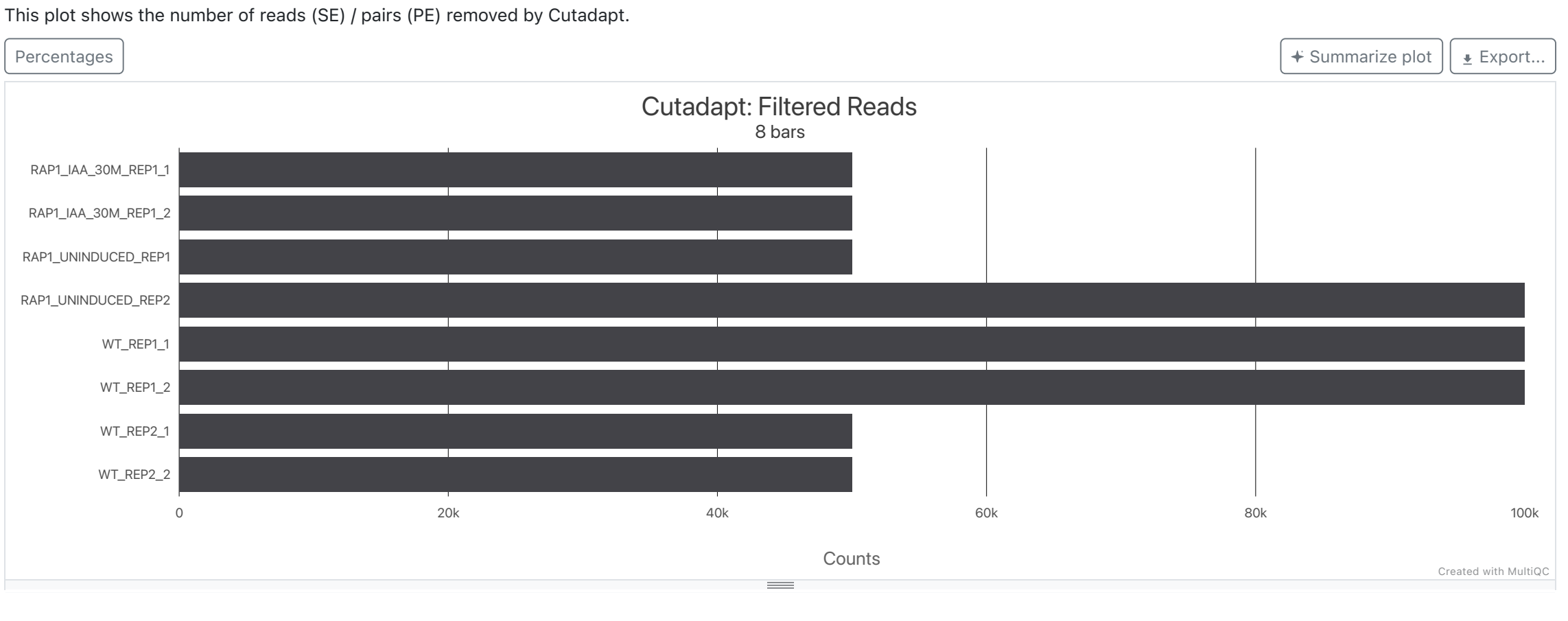
Cutadapt

Finds and removes adapter sequences, primers, poly-A tails, and other types of unwanted sequences. https://cutadapt.readthedocs.io DOI: 10.14806/ej.17.1.200

Filtered Reads

This plot shows the number of reads (SE) / pairs (PE) removed by Cutadapt.

Percentages — Summarize plot | Export...

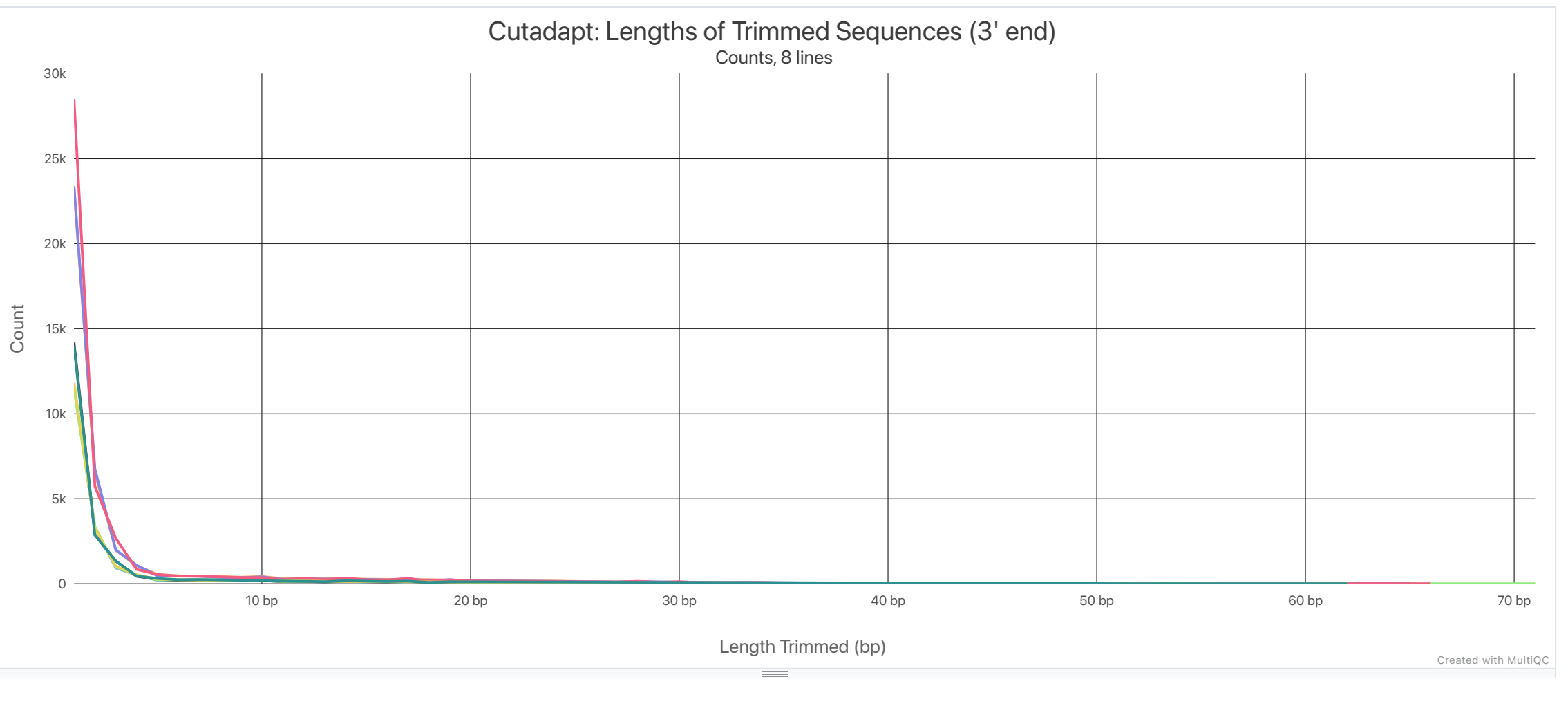


Trimmed Sequence Lengths (3')

Help

This plot shows the number of reads with certain lengths of adapter trimmed for the 3' end.

Counts | Obs/Exp — Summarize plot | Export...



FastQC (trimmed)

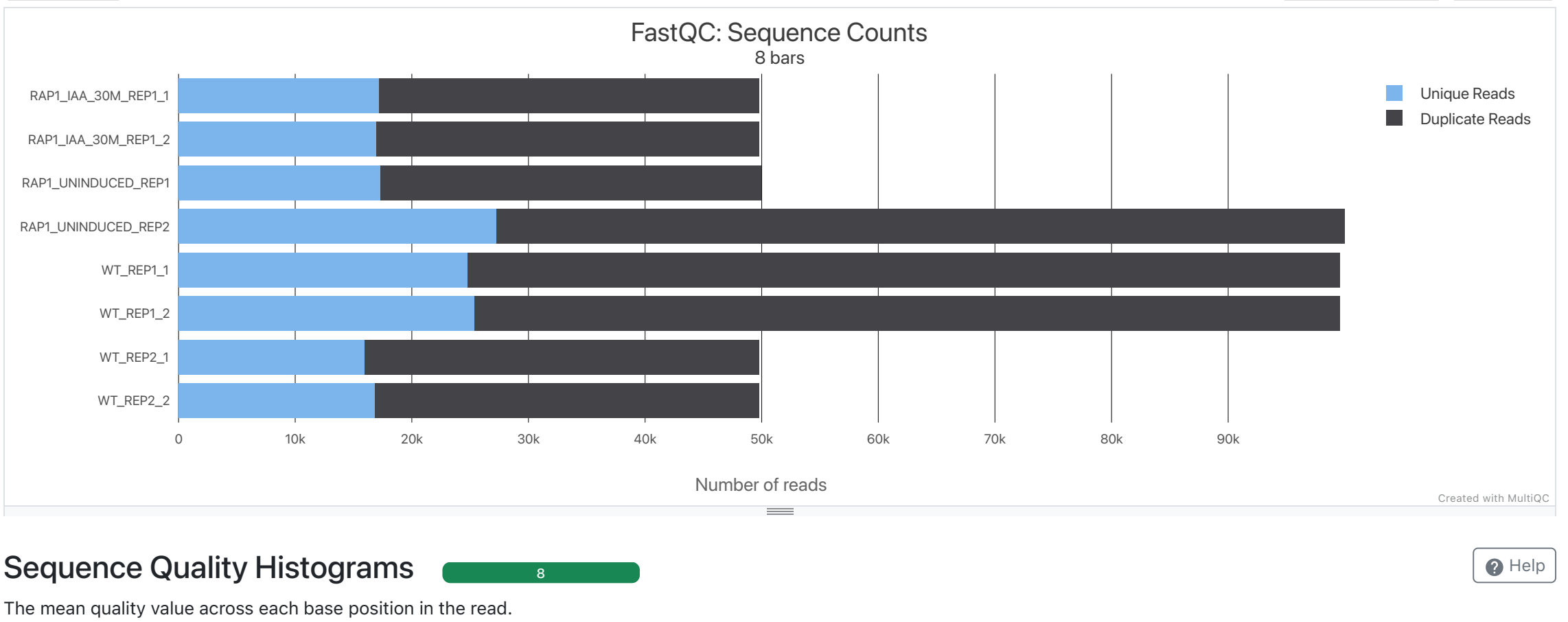
This section of the report shows FastQC results after adapter trimming. http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Sequence Counts

Help

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

Percentages — Summarize plot | Export...



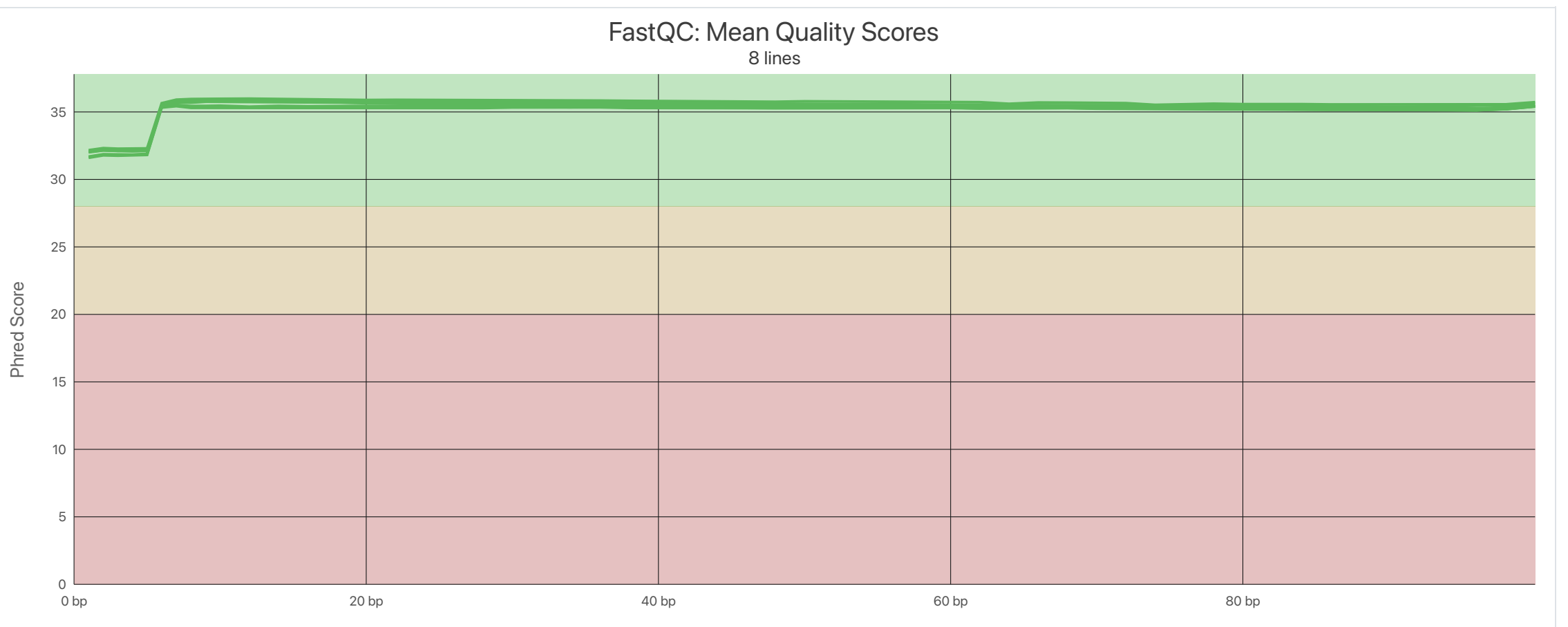
Sequence Quality Histograms

8

Help

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

Summarize plot | Export...



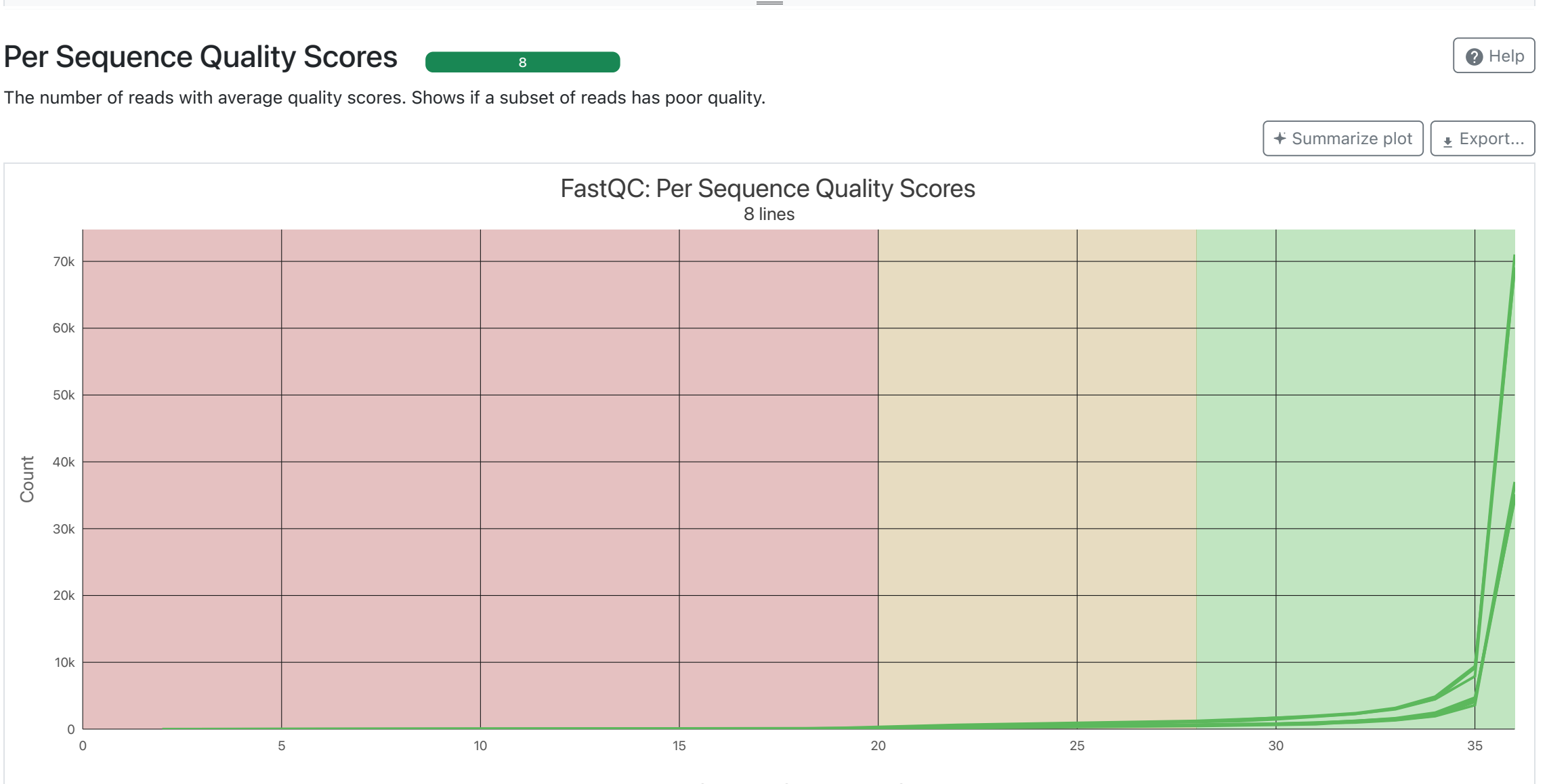
Per Sequence Quality Scores

8

Help

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

Summarize plot | Export...



Per Base Sequence Content

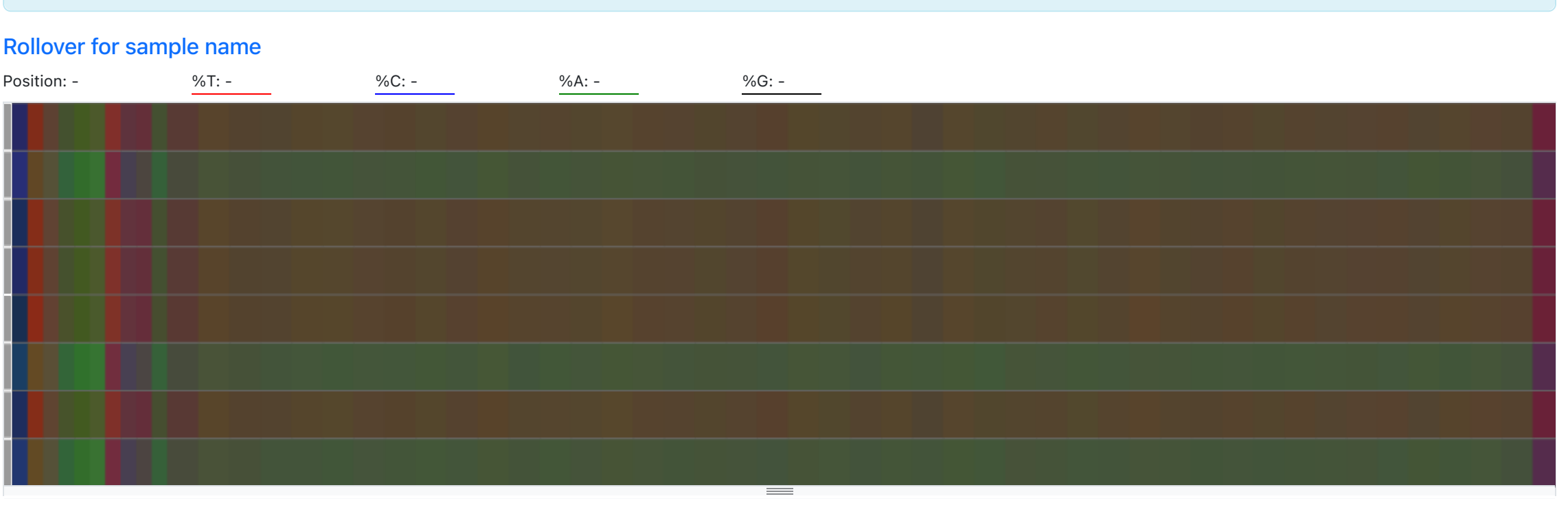
Help

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

Position: - %T: - %C: - %A: - %G: -



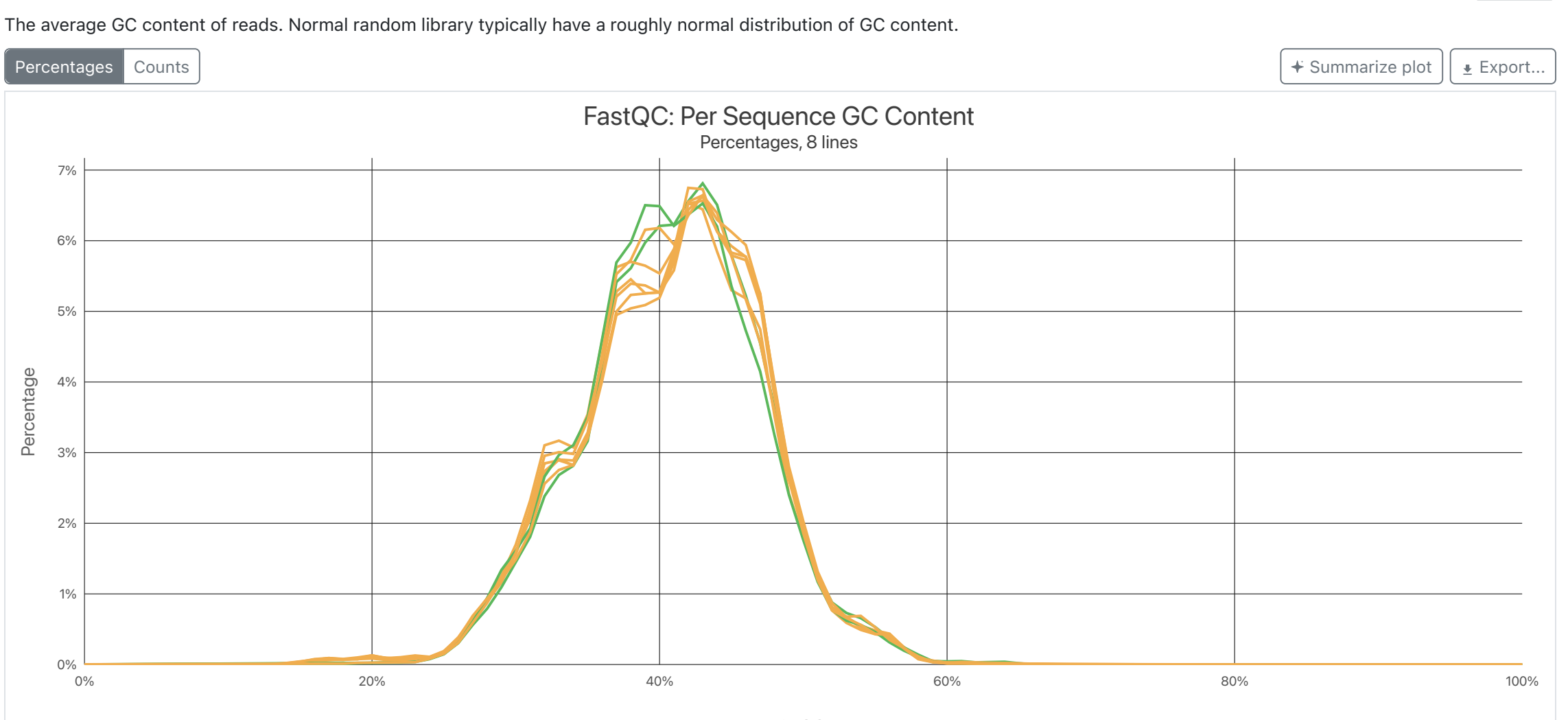
Per Sequence GC Content

2 6

Help

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

Percentages | Counts — Summarize plot | Export...



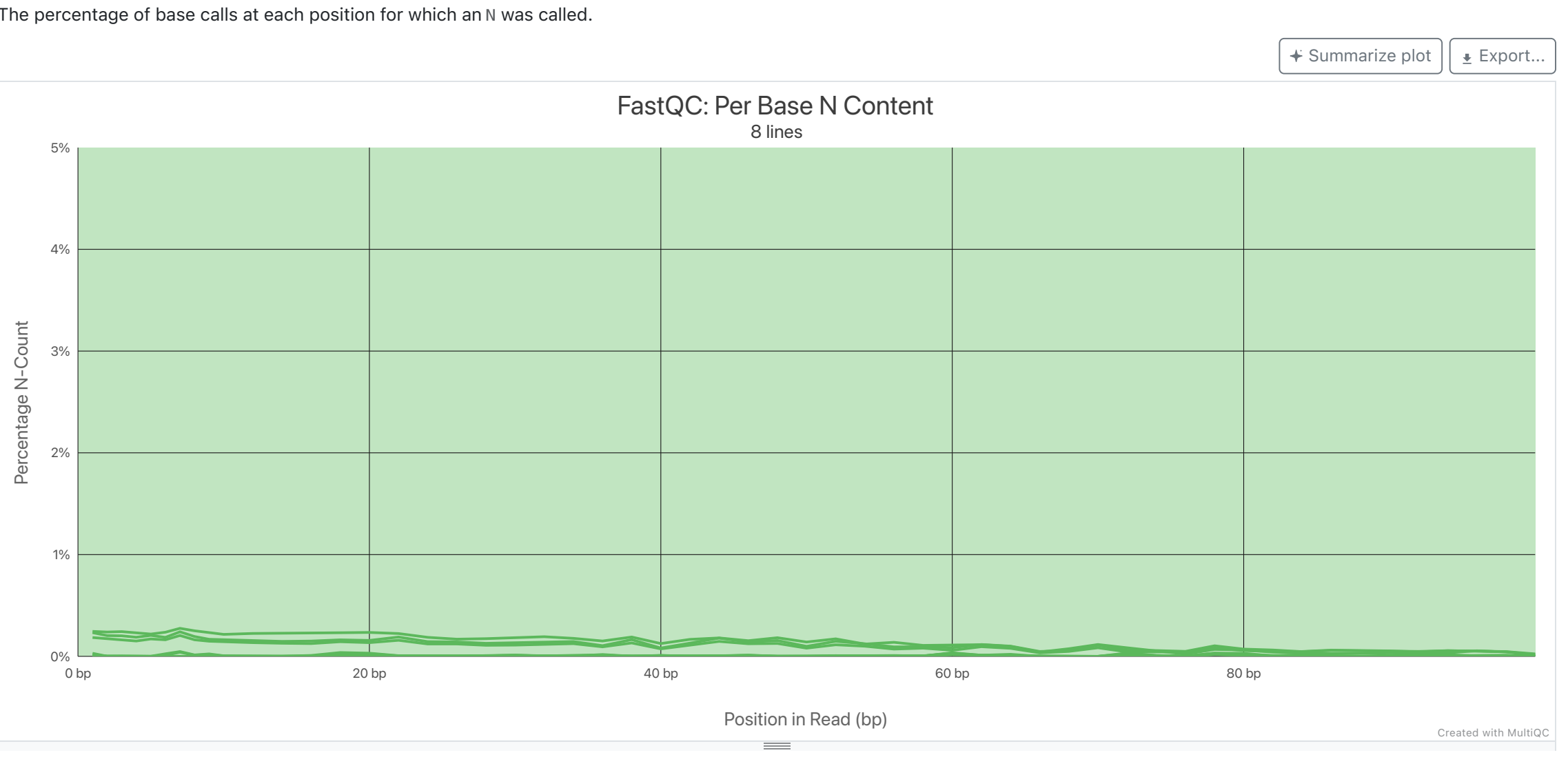
Per Base N Content

8

Help

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

Summarize plot | Export...

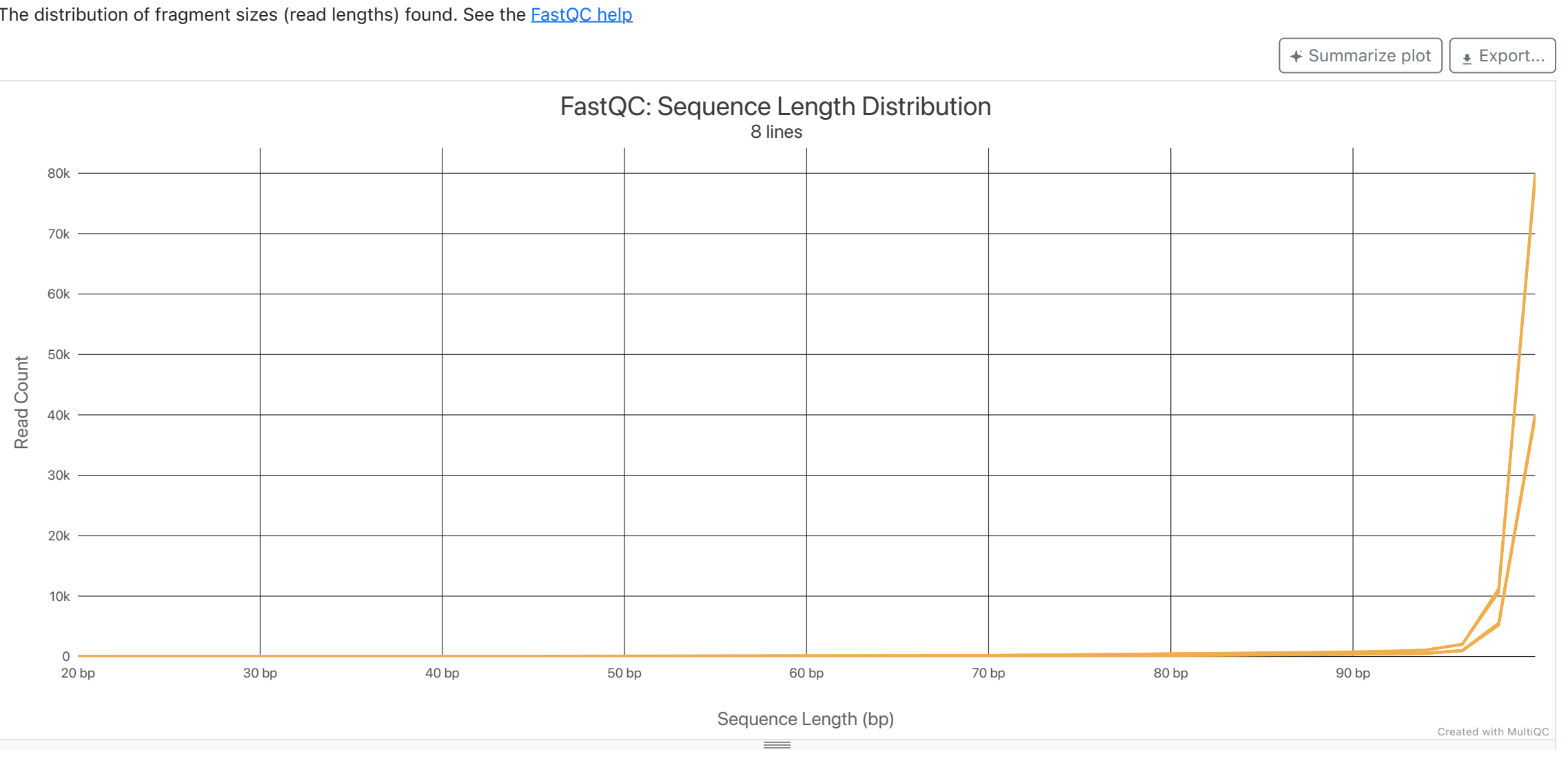


Sequence Length Distribution

8

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

Summarize plot | Export...



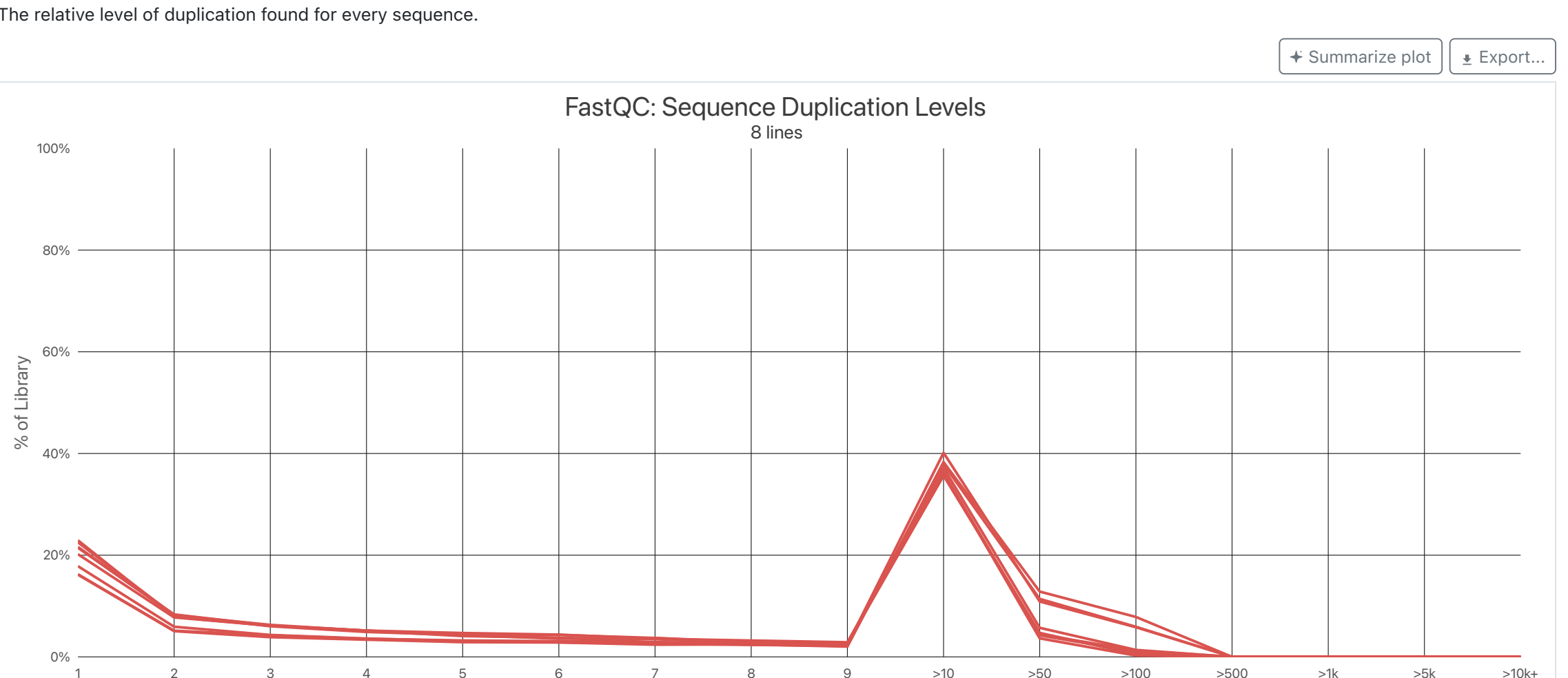
Sequence Duplication Levels

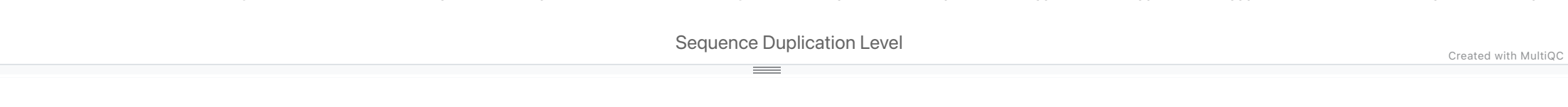
8

Help

The relative level of duplication found for every sequence.

Summarize plot | Export...



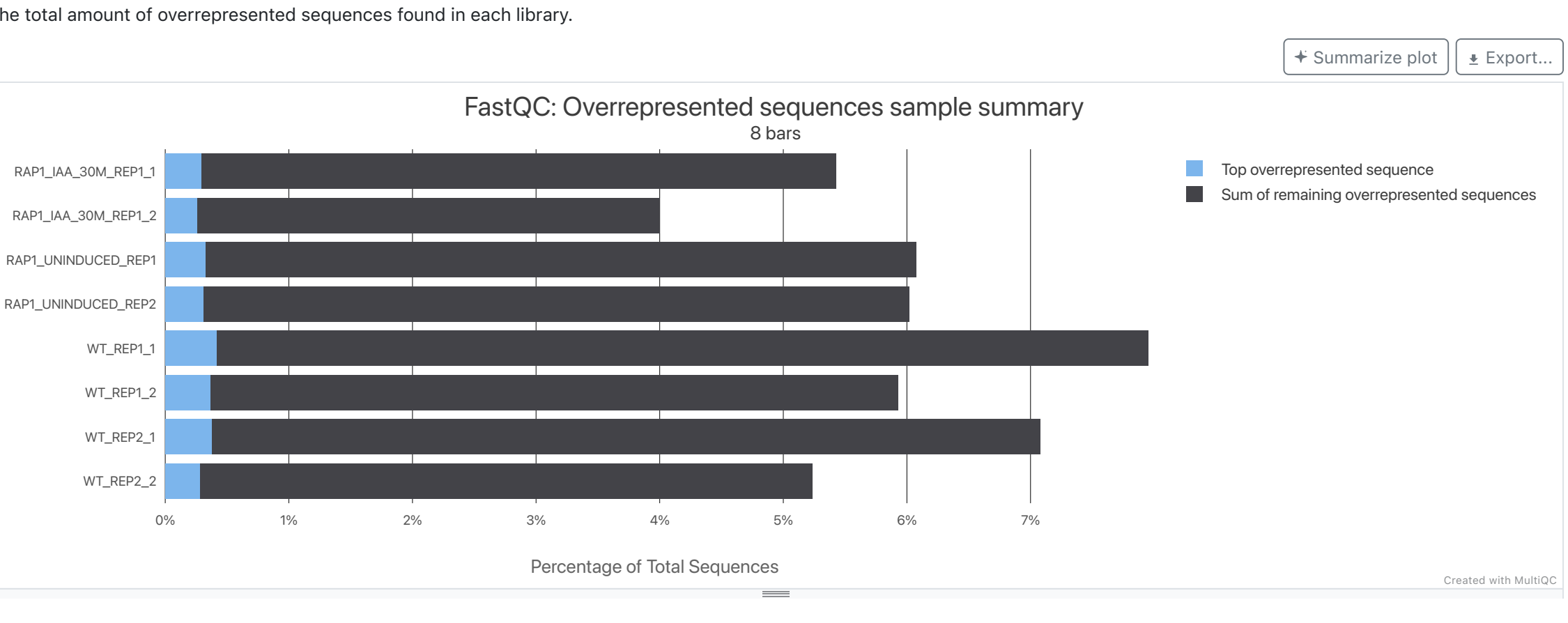


Overrepresented sequences by sample

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
CTTTGACCCAGGTAGGTAGGAATTGAGAGGGAGCAAATATTTGAAAAGTG	5	1214	0.2215 %
GTTAGGAGTGTGCAATACCTTTATTGATGTTTTGGTGTTGTCCTGGAAGT	5	1031	0.1881 %
CTTCGATCTCGGAGGTTCTAAACTACGCATATTCTTAGTATTCCATGTGT	5	823	0.1501 %
CGAGATTCGAGTTTTGTAAAAATTTGATGTATGTTTTAACCCAATTTGGA	5	783	0.1428 %
CCTCATTCCGTTTTATATGTTTATATTCATTGATCCTATTACATTATCAA	5	736	0.1343 %
CTTTAAATGGGTCAGGGAATTCGGTAGGAGATTCTGGAGGTAGATCAGGG	5	691	0.1261 %
CTGGAAGTGAAATTGTAGGTCACCAATAGCGTTAATTGGTATATTTCTTT	5	633	0.1155 %
GTTTGATTGATTTTACTGCTTTTACAGCTGCAATCAGGTGAATTCGTTTC	5	620	0.1131 %
GTCAACTTCTCTGGGTGCGCGAATATTGGTTTTAGAAACACGTTTCGAAT	5	613	0.1118 %
CTAAAACTTGCCTAGAGGGGAATAGTATATGTTGAGCAAGTGTGTTGATG	5	590	0.1076 %
AGAGAATCTAGCAGCTCTTGGGCATCTGGTAACCAAGATGATTGGACAGT	5	578	0.1054 %
CTCATTTTCACTGTACGAGTCTGAGTGTCTGAAATCTTTAGATTTACTGG	5	560	0.1022 %
CAATGGTCTACCTGGGTACAATTCTTCGGACTTTCTGGCGTTGTCAATGA	5	551	0.1005 %
GTTCAGTCTGTCGATAATGTTTGTGACTTTTGTTTCAAATGCATCTGCAG	5	534	0.0974 %
GTGATATCTACTGCAGCCAATTCATTCAAACTGAGTAAGTCATAGGCTAT	5	525	0.0958 %

Expand table

Adapter Content

8

Help

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

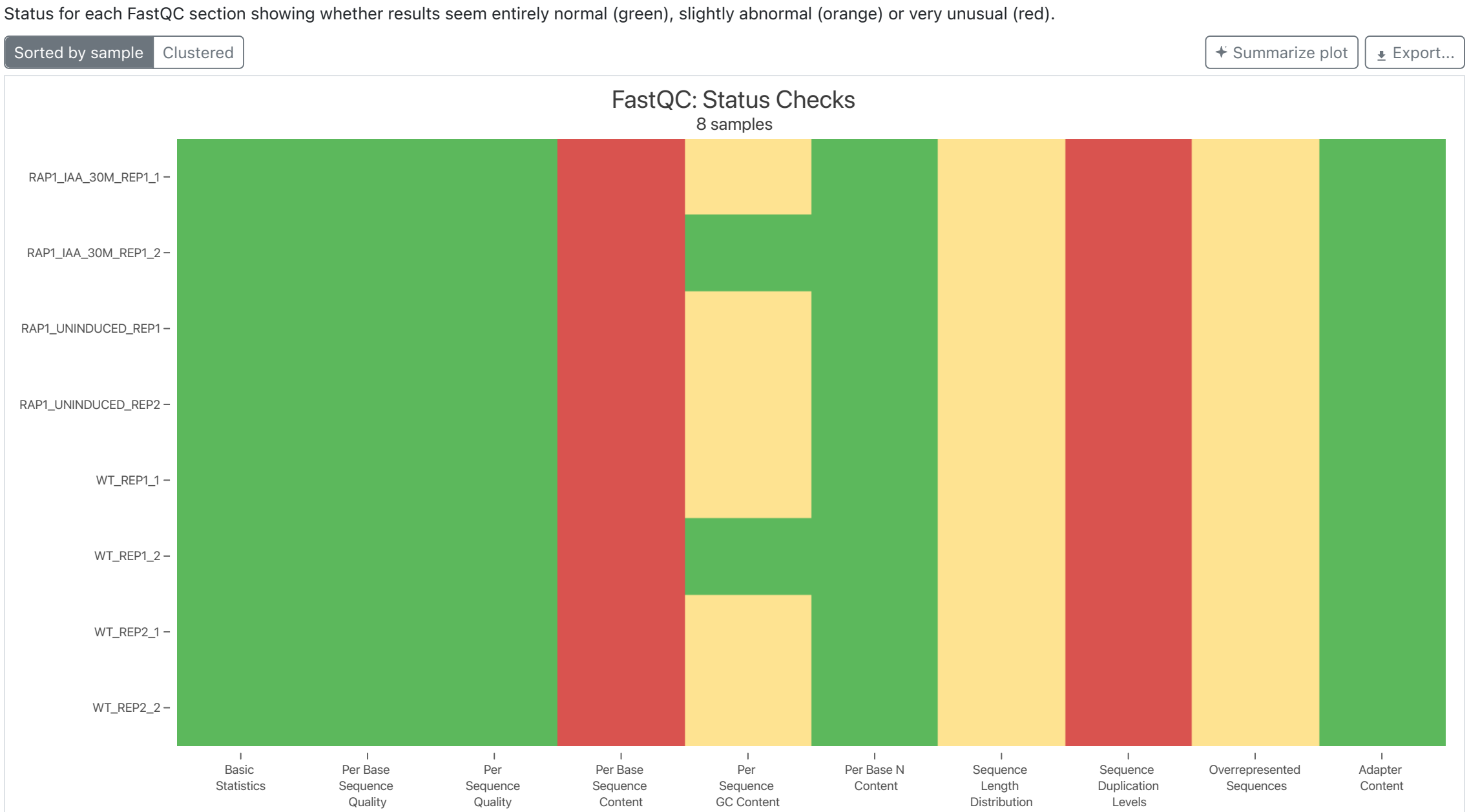
No samples found with any adapter contamination > 0.1%

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...



BBTools

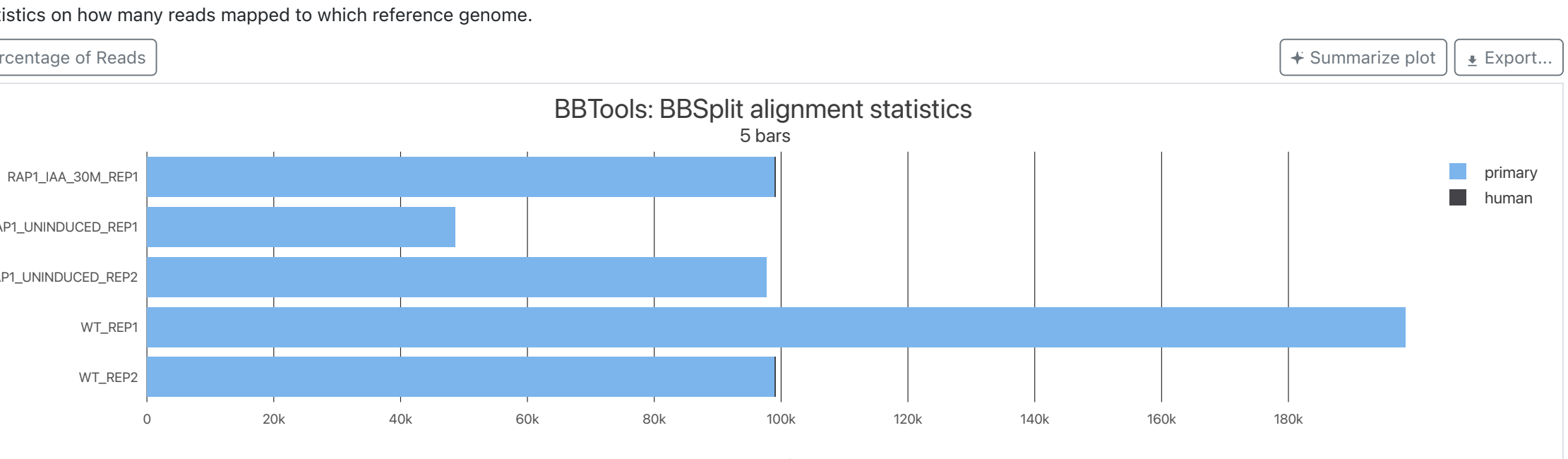
Pre-processing, assembly, alignment, and statistics tools for DNA/RNA sequencing reads. http://jgi.doe.gov/data-and-tools/bbtools

BBSplit alignment statistics

Help

Statistics on how many reads mapped to which reference genome.

Percentage of Reads
+ Summarize plot
Export...



BBSplit Statistics

Statistics on how many reads mapped to which reference genome.

Copy table
Configure columns
Scatter plot
Violin plot
Export as CSV...
Showing $^{5}/_{5}$ rows and $^{4}/_{10}$ columns.
+ Summarize table

Sample Name	human % Unambig	human % Ambig	primary % Unambig	primary % Ambig
RAP1_IAA_30M_REP1	0.0 %	0.0 %	99.6 %	0.0 %
RAP1_UNINDUCED_REP1			97.2 %	0.0 %
RAP1_UNINDUCED_REP2	0.0 %	0.0 %	97.7 %	0.0 %
WT_REP1			99.7 %	0.0 %
WT_REP2	0.0 %	0.0 %	99.6 %	0.0 %

FastQC (filtered)

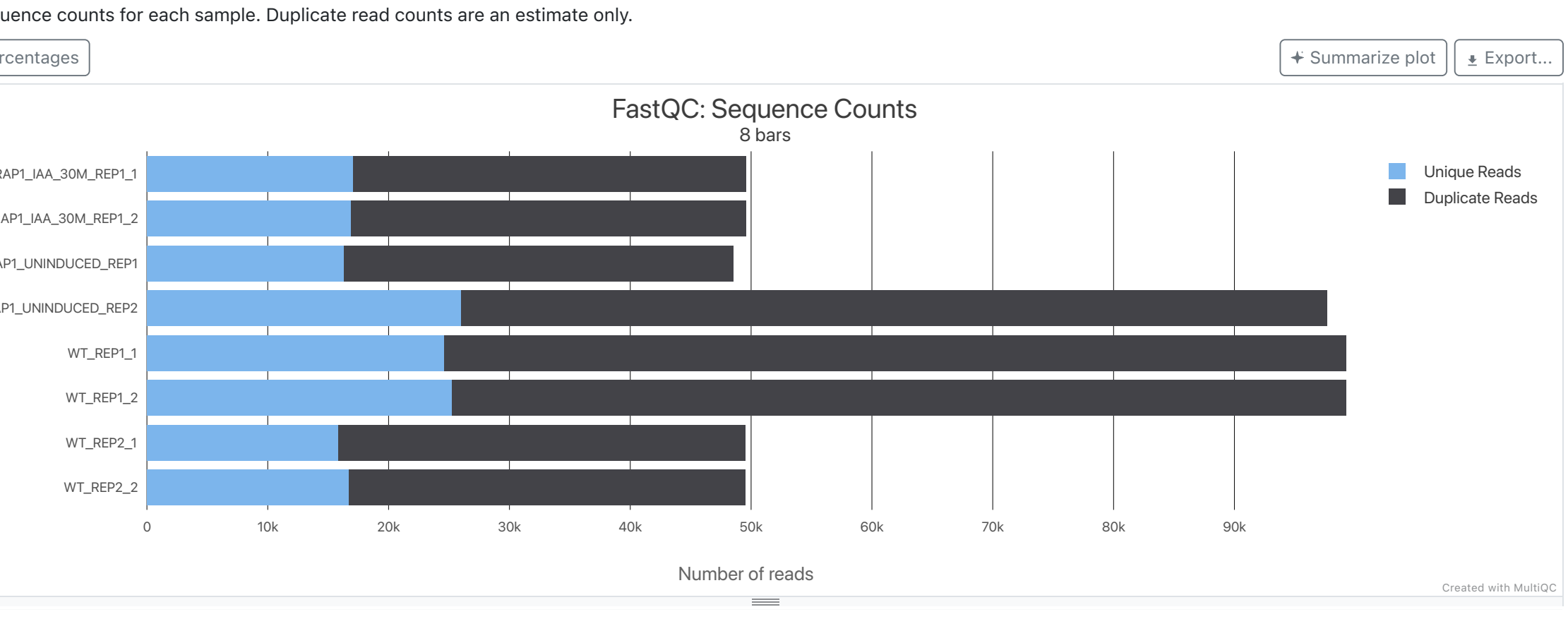
This section of the report shows FastQC results after contaminant removal (BBSplit) and/or ribosomal RNA removal. http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Sequence Counts

Help

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

Percentages
+ Summarize plot
Export...



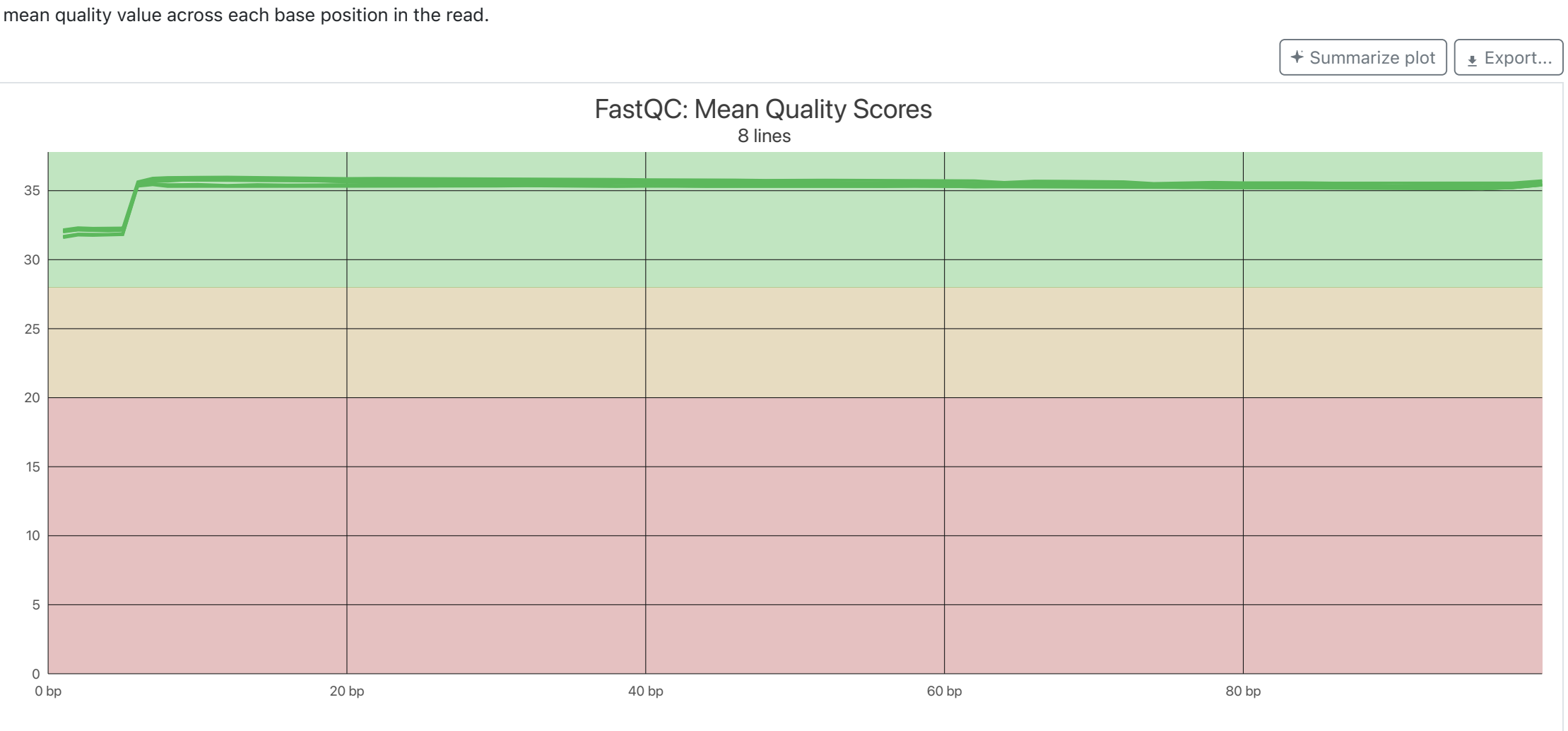
Sequence Quality Histograms

8

Help

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

+ Summarize plot
Export...



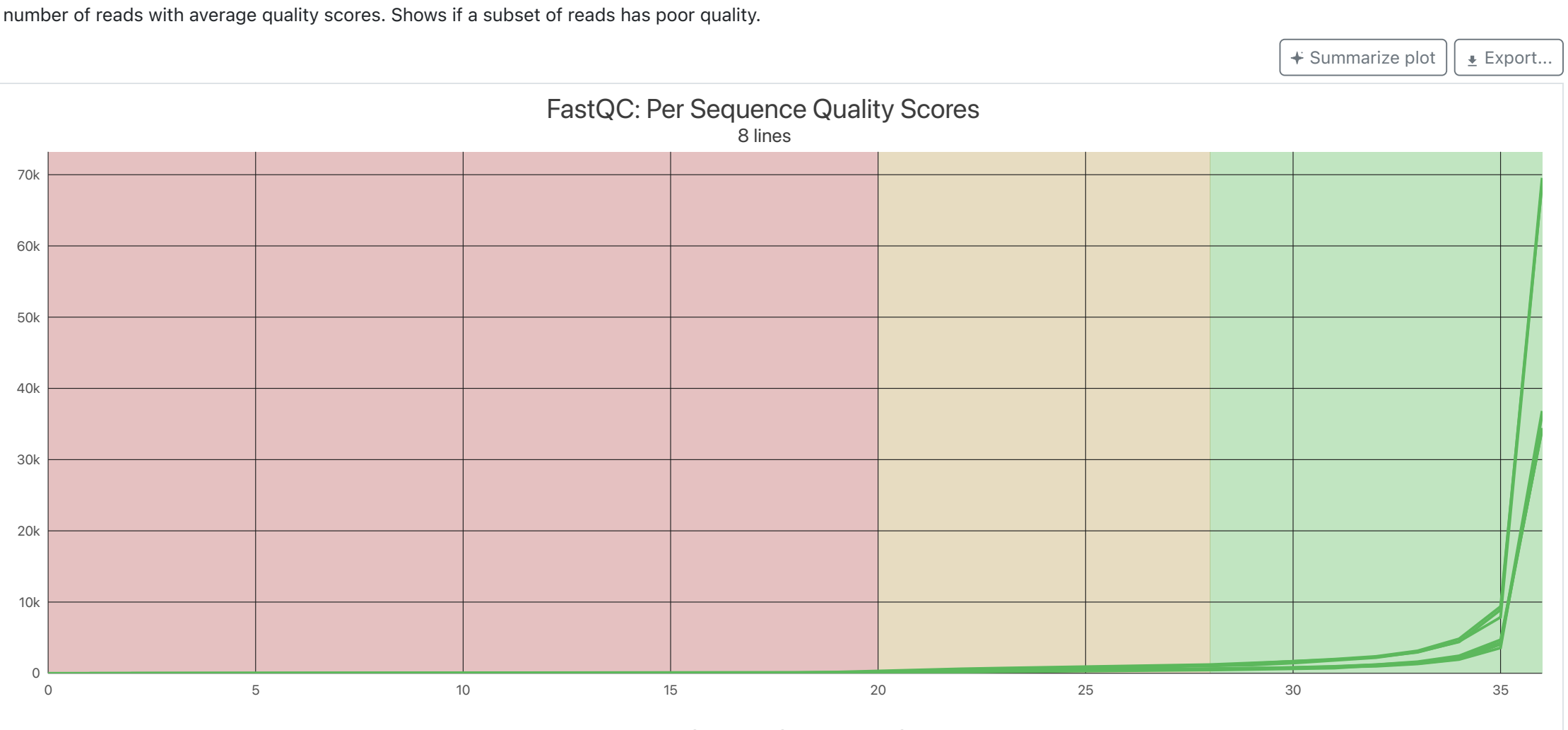
Per Sequence Quality Scores

8

Help

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

+ Summarize plot
Export...

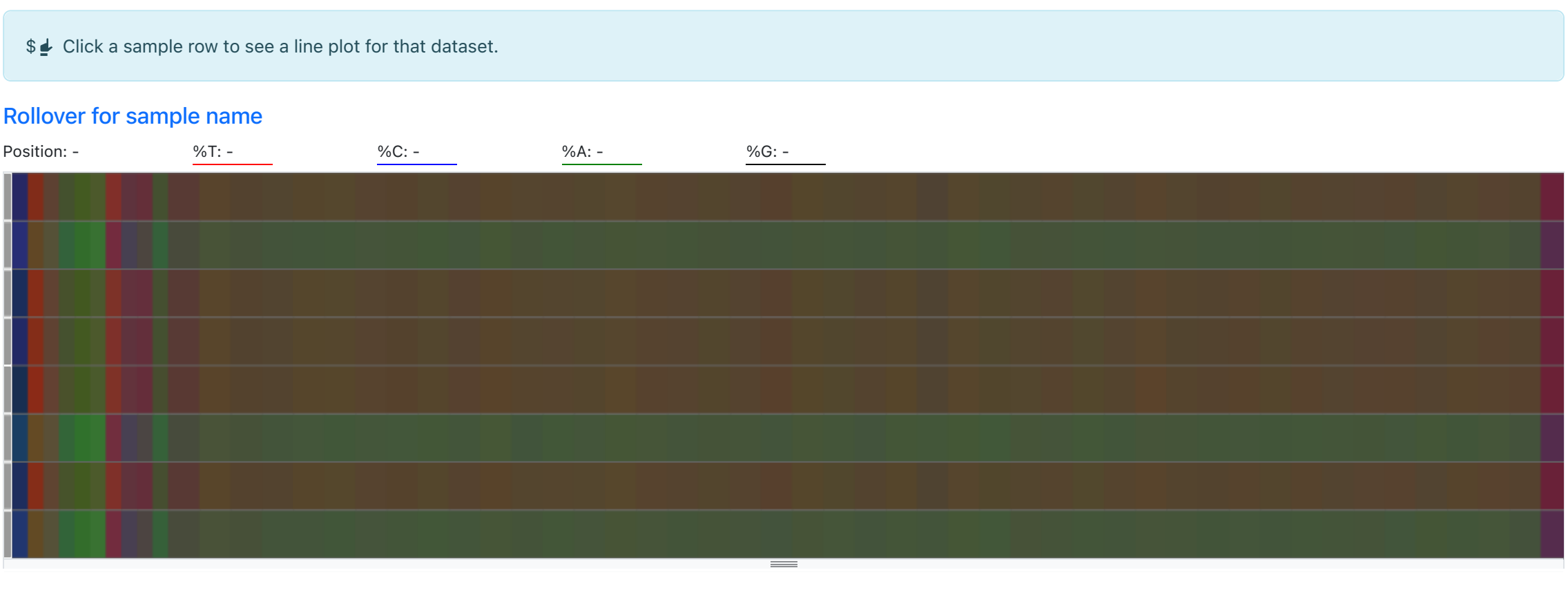


Per Base Sequence Content

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.



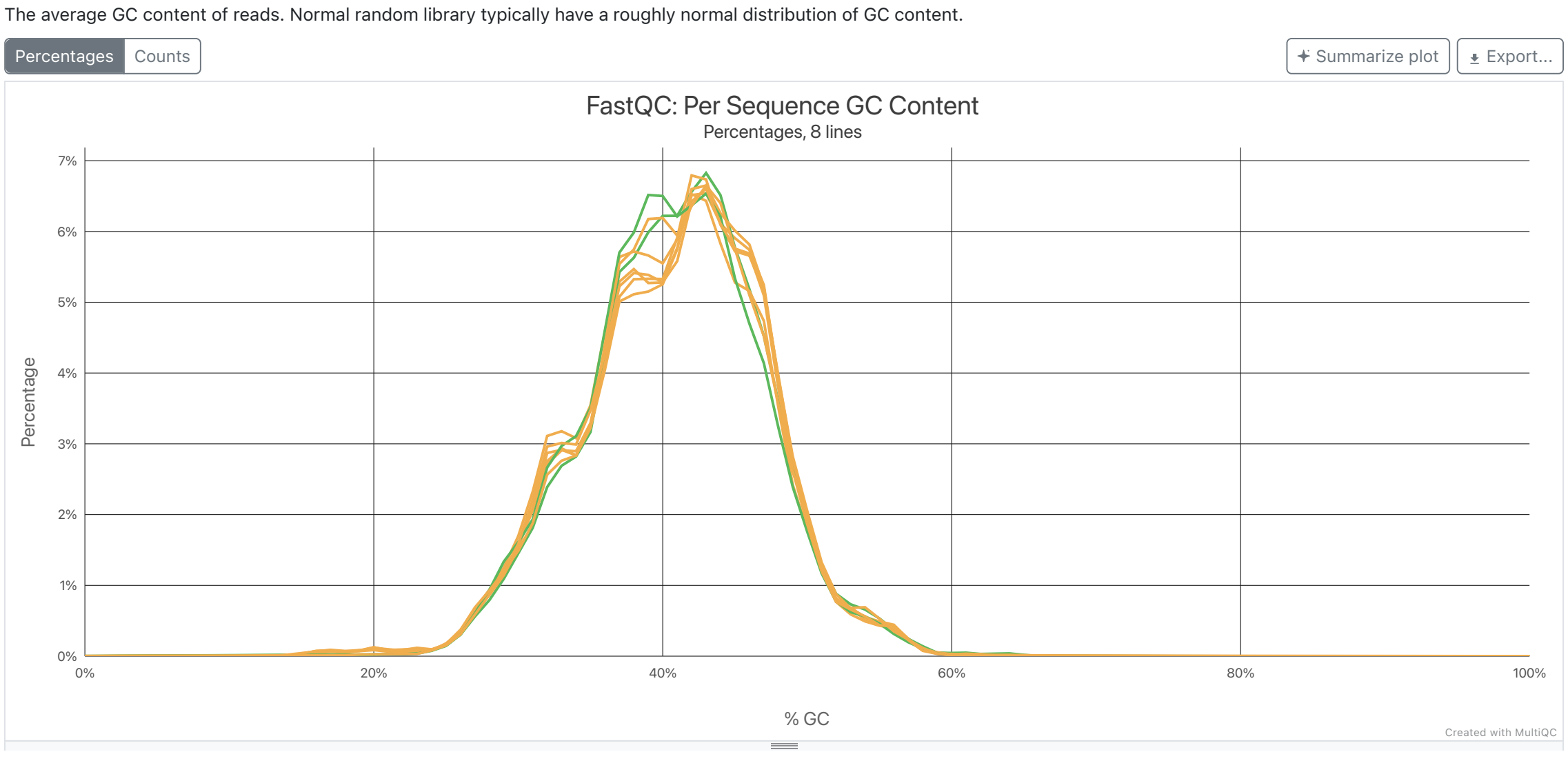
Per Sequence GC Content

2
6

Help

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

Percentages
Counts
+ Summarize plot
Export...



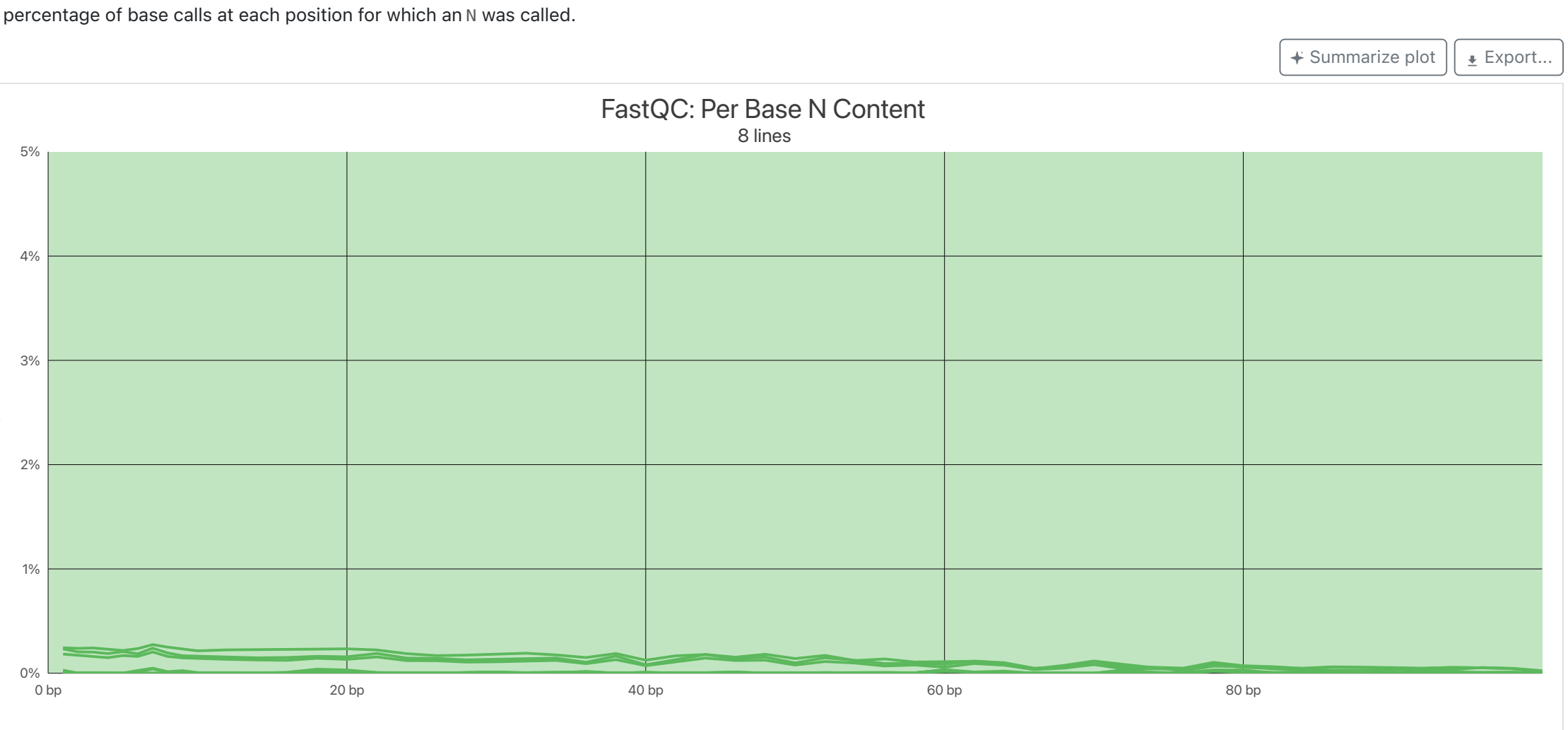
Per Base N Content

8

Help

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

+ Summarize plot
Export...



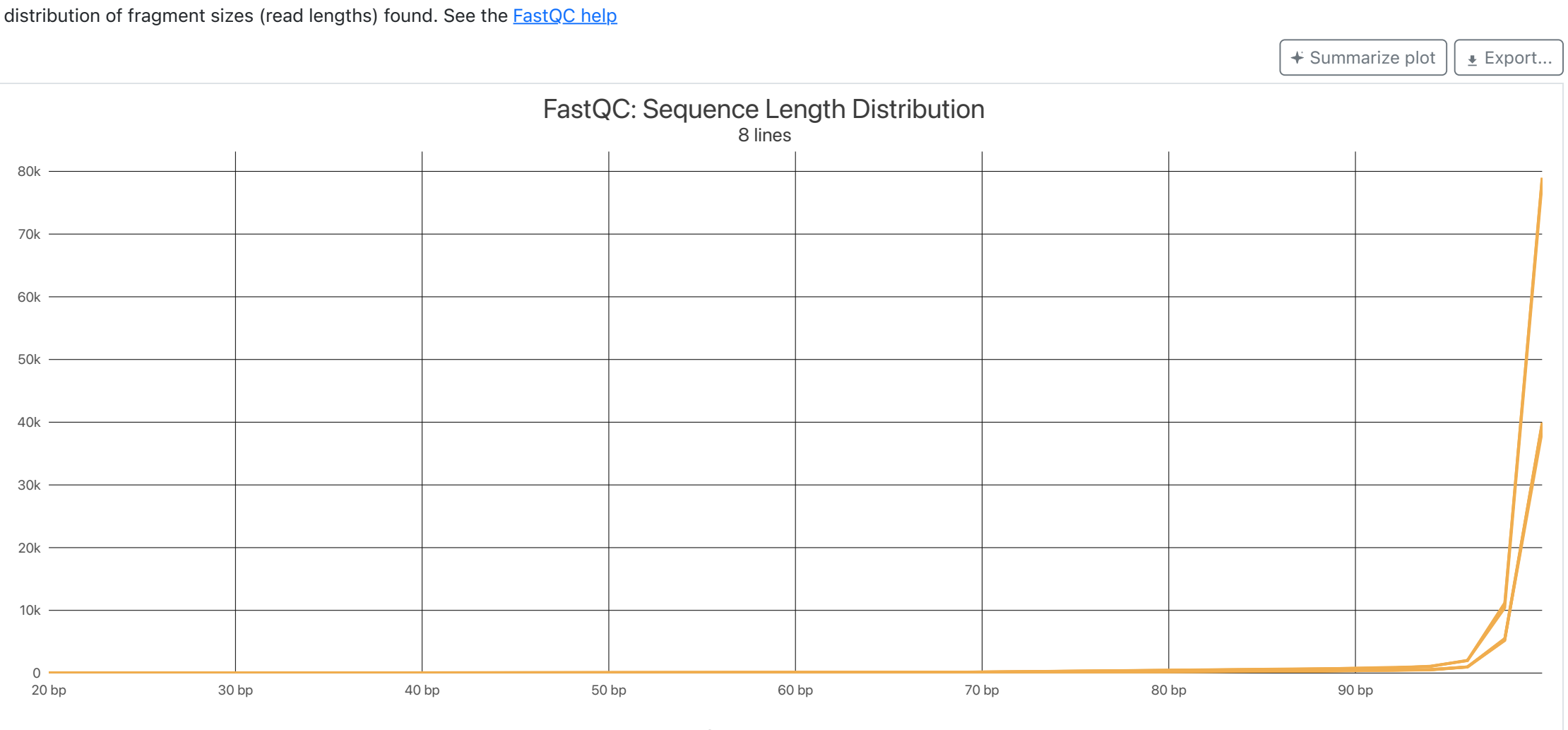
Sequence Length Distribution

8

Help

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

+ Summarize plot
Export...



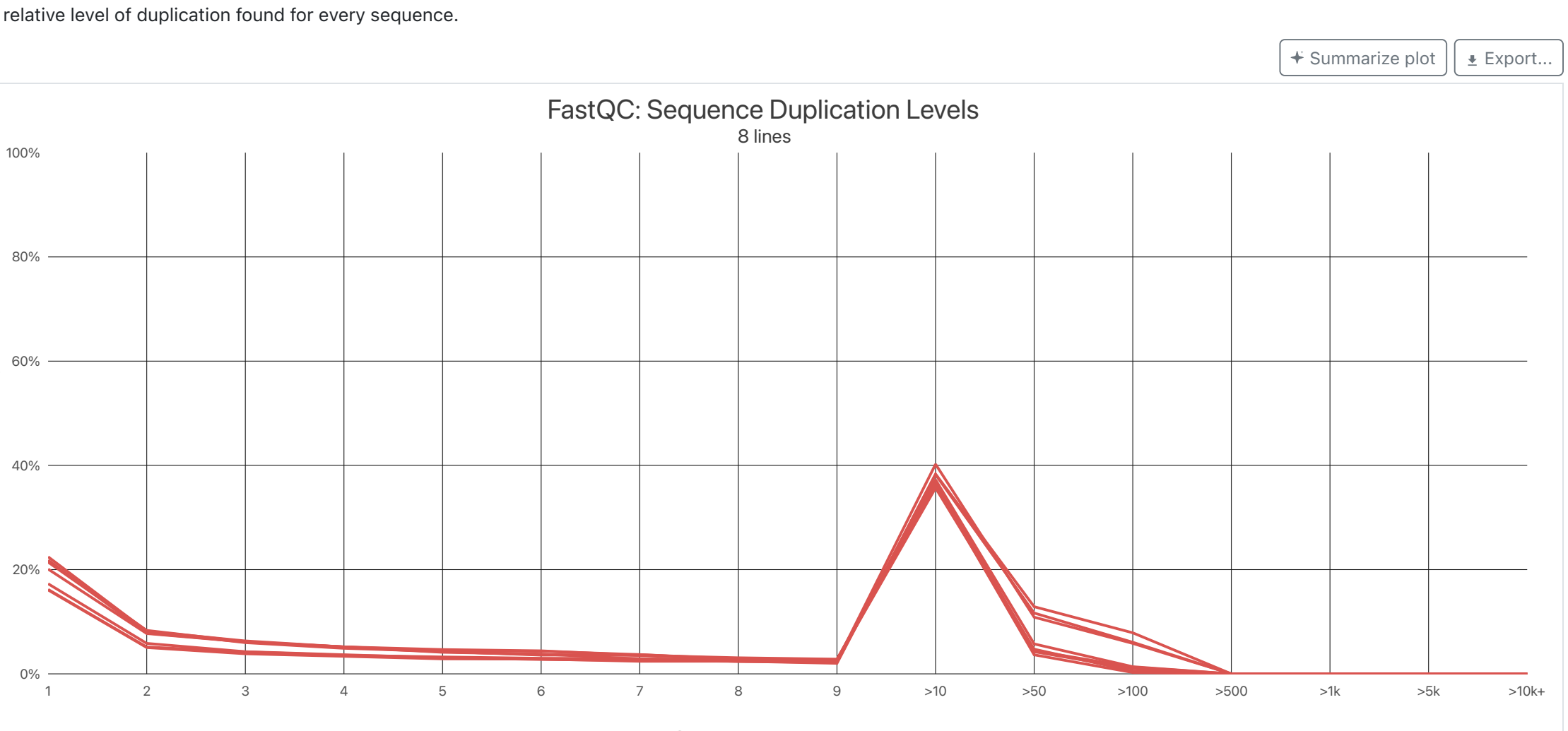
Sequence Duplication Levels

8

Help

The relative level of duplication found for every sequence.

+ Summarize plot
Export...

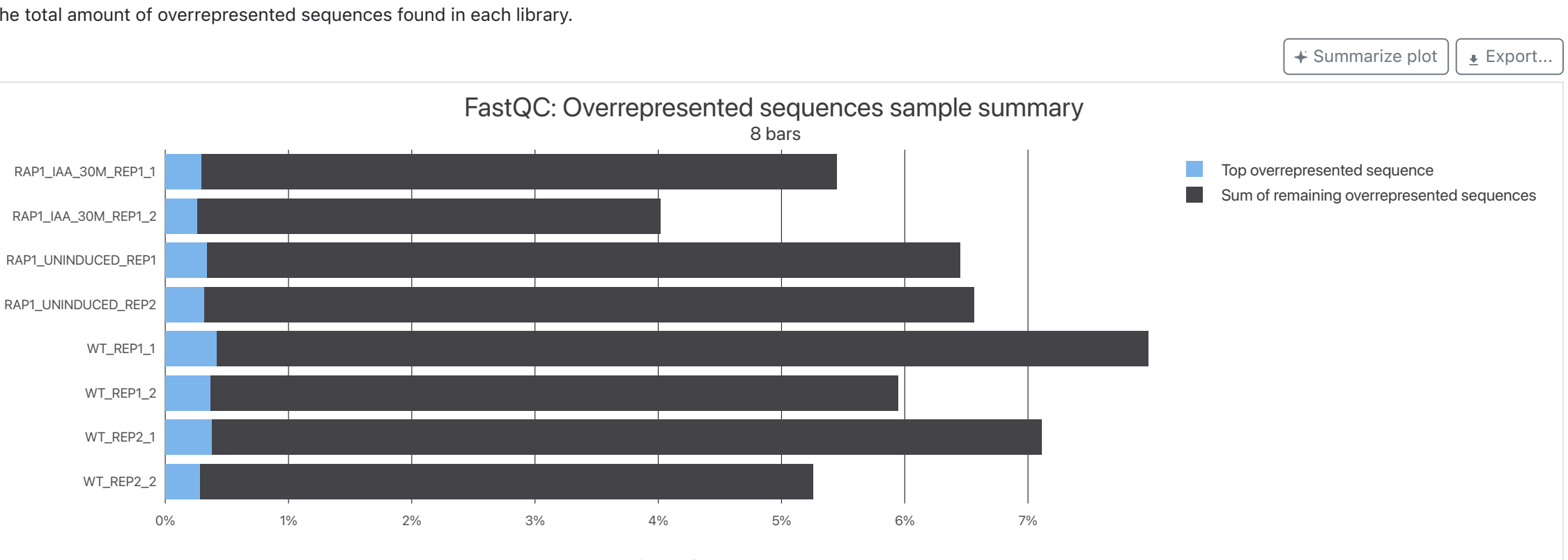


Overrepresented sequences by sample

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
CTTTGACCCAGGTAGGTAGGAATTGAGAGGGAGCAAATATTTGAAAAGTG	5	1214	0.2236 %
GTTAGGAGTGTGCAATACCTTTATTGATGTTTTGGTGTTGTCCTGGAAGT	5	1030	0.1897 %
CTTCGATCTCGGAGGTTCTAAACTACGCATATTCTTAGTATTCCATGTGT	5	823	0.1516 %
CGAGATTCGAGTTTTGTAAAAATTTGATGTATGTTTTAACCCAATTTGGA	5	783	0.1442 %
CCTCATTCCGTTTTATATGTTTATATTCATTGATCCTATTACATTATCAA	5	736	0.1356 %
CTTTAAATGGGTCAGGGAATTCGGTAGGAGATTCTGGAGGTAGATCAGGG	5	691	0.1273 %
CTGGAAGTGAAATTGTAGGTCACCAATAGCGTTAATTGGTATATTTCTTT	5	633	0.1166 %
GTTTGATTGATTTTACTGCTTTTACAGCTGCAATCAGGTGAATTCGTTTC	5	620	0.1142 %
GTCAACTTCTCTGGGTGCGCGAATATTGGTTTTAGAAACACGTTTCGAAT	5	613	0.1129 %
CTAAAACTTGCCTAGAGGGGAATAGTATATGTTGAGCAAGTGTGTTGATG	5	590	0.1087 %
AGAGAATCTAGCAGCTCTTGGGCATCTGGTAACCAAGATGATTGGACAGT	5	578	0.1065 %
CTCATTTTCACTGTACGAGTCTGAGTGTCTGAAATCTTTAGATTTACTGG	5	560	0.1031 %
CAATGGTCTACCTGGGTACAATTCTTCGGACTTTCTGGCGTTGTCAATGA	5	551	0.1015 %
GTTCAGTCTGTCGATAATGTTTGTGACTTTTGTTTCAAATGCATCTGCAG	5	534	0.0983 %
GTGATATCTACTGCAGCCAATTCATTCAAACTGAGTAAGTCATAGGCTAT	5	525	0.0967 %

Expand table

Adapter Content

8

Help

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

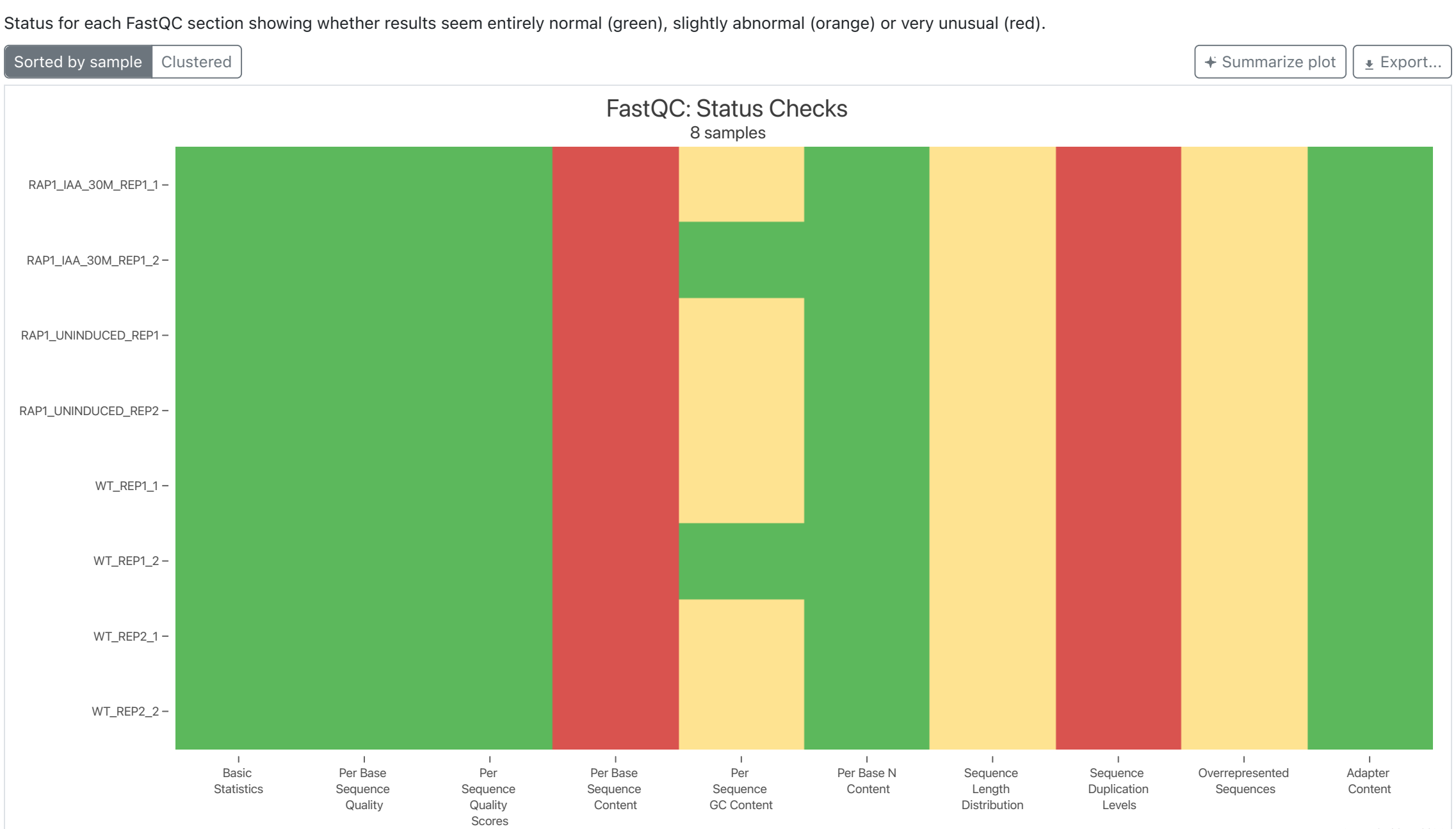
No samples found with any adapter contamination > 0.1%

Status Checks

Help

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

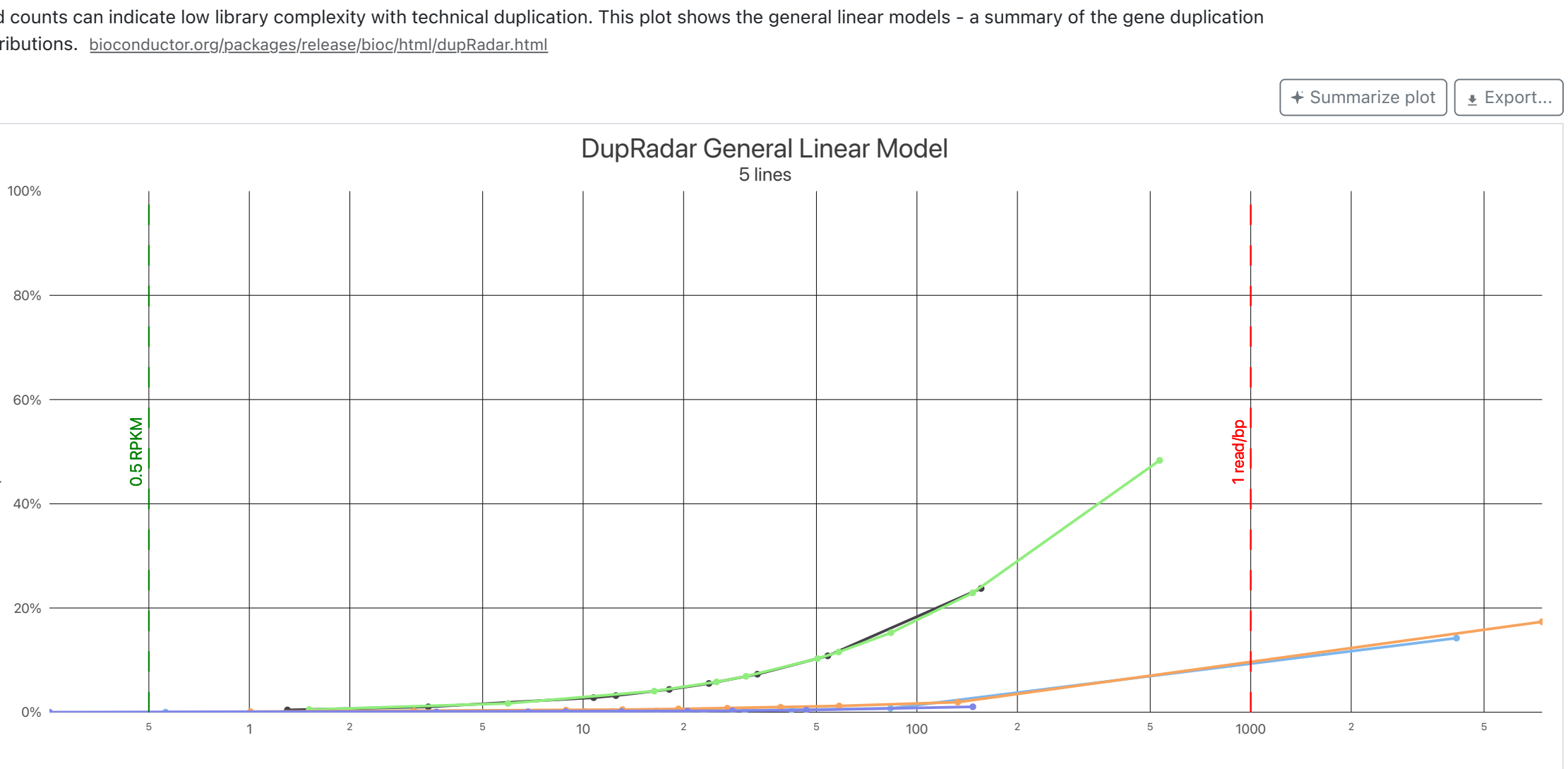
Sorted by sample
Clustered
+ Summarize plot
Export...



DupRadar

DupRadar provides duplication rate quality control for RNA-Seq datasets. Highly expressed genes can be expected to have a lot of duplicate reads, but high numbers of duplicates at low read counts can indicate low library complexity with technical duplication. This plot shows the general linear models - a summary of the gene duplication distributions. bioconductor.org/packages/release/bioc/html/dupRadar.html

+ Summarize plot
Export...



Picard

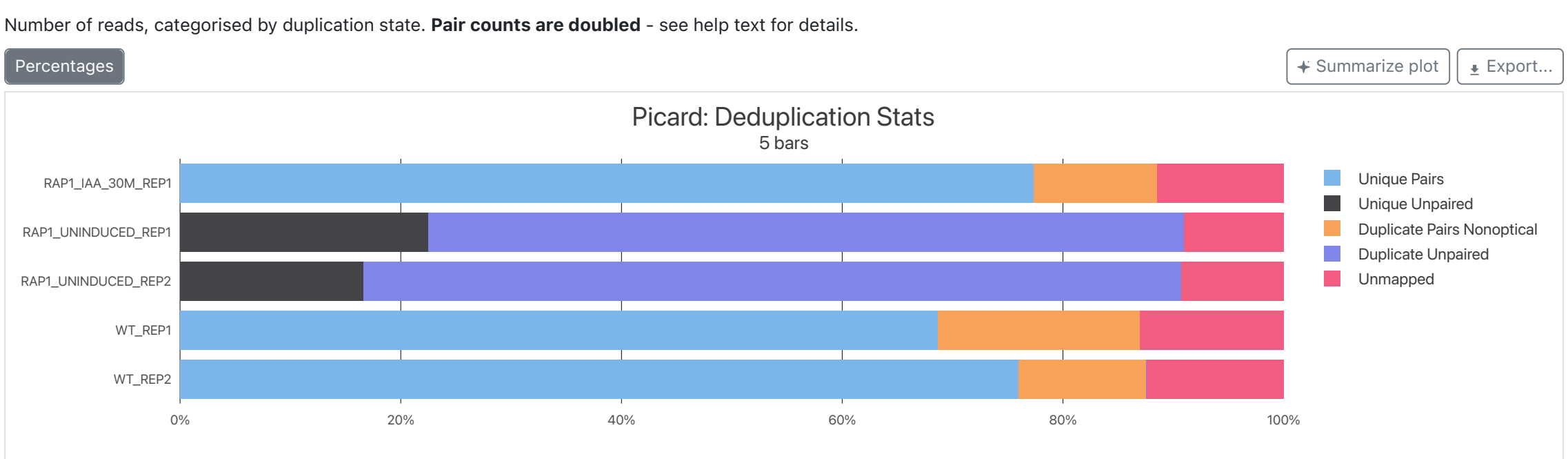
Tools for manipulating high-throughput sequencing data. http://broadinstitute.github.io/picard

Mark Duplicates

Help

Number of reads, categorised by duplication state. **Pair counts are doubled** - see help text for details.

Percentages
+ Summarize plot
Export...



QualiMap

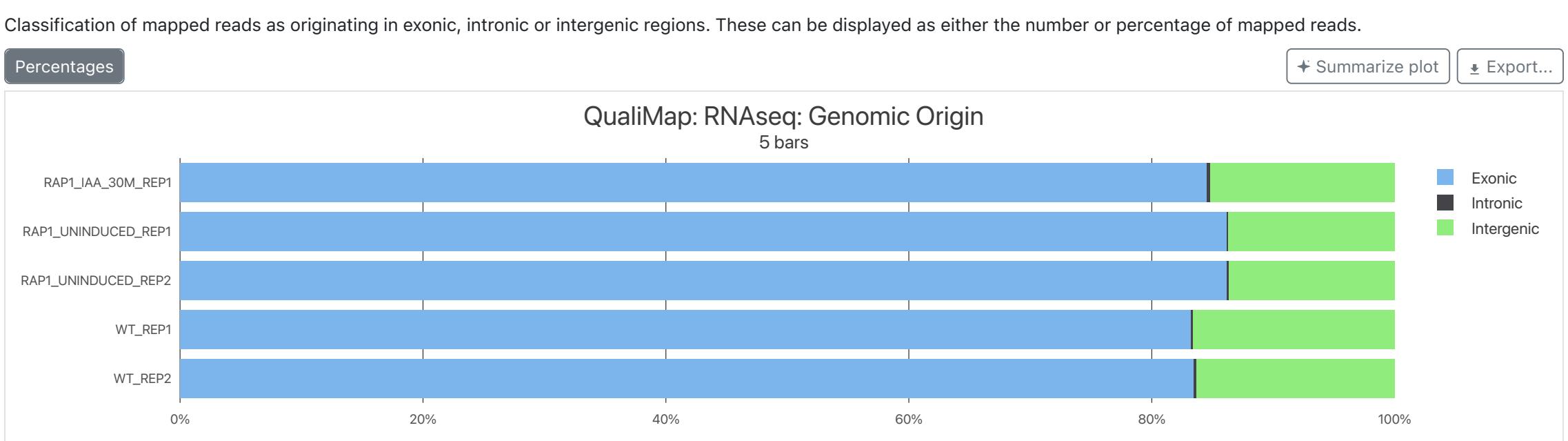
Quality control of alignment data and its derivatives like feature counts. http://qualimap.bioinfo.cipf.es DOI: 10.1093/bioinformatics/btv566; 10.1093/bioinformatics/bts503

Genomic origin of reads

Help

Classification of mapped reads as originating in exonic, intronic or intergenic regions. These can be displayed as either the number or percentage of mapped reads.

Percentages
+ Summarize plot
Export...

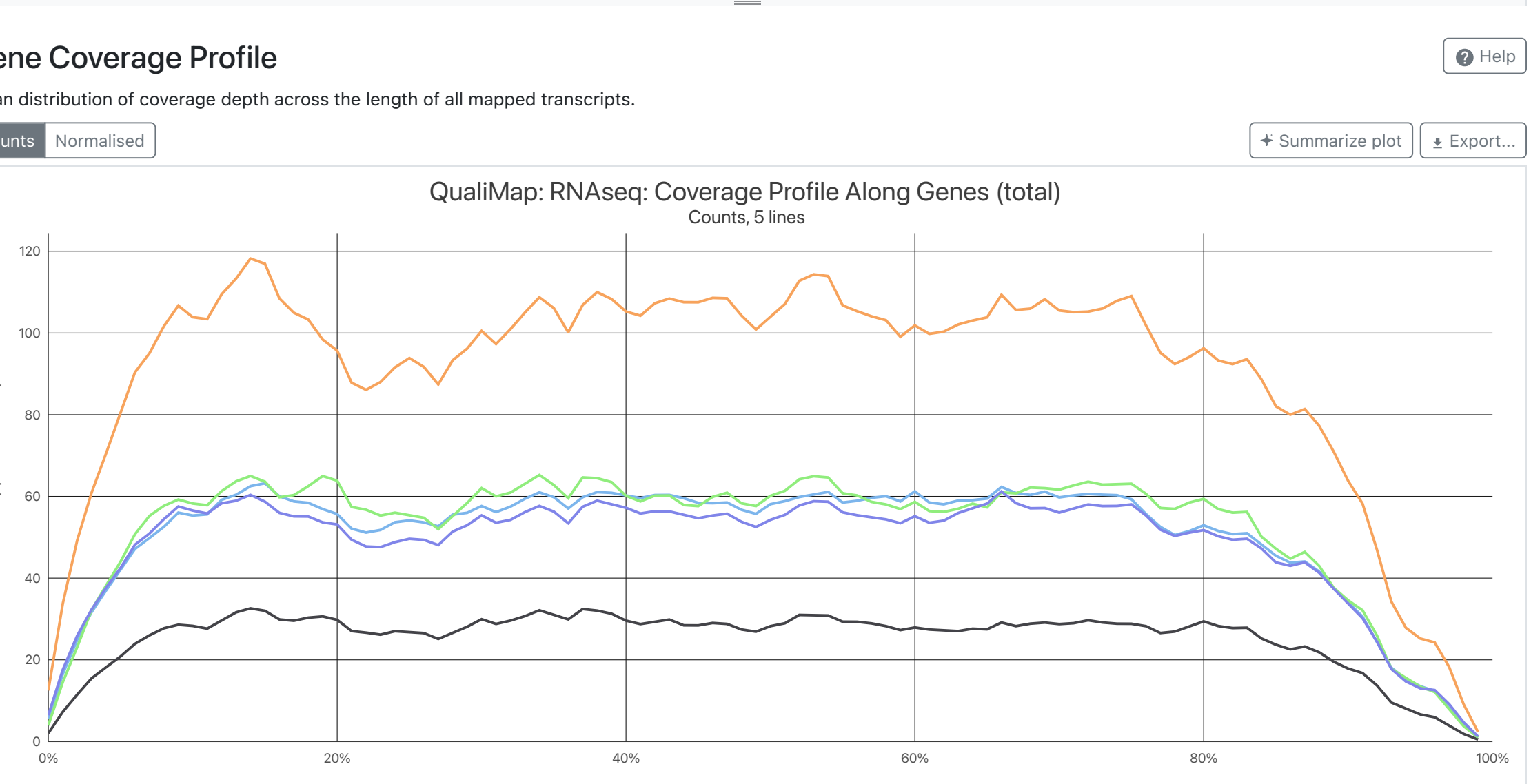


Gene Coverage Profile

Help

Mean distribution of coverage depth across the length of all mapped transcripts.

Counts
Normalised
+ Summarize plot
Export...



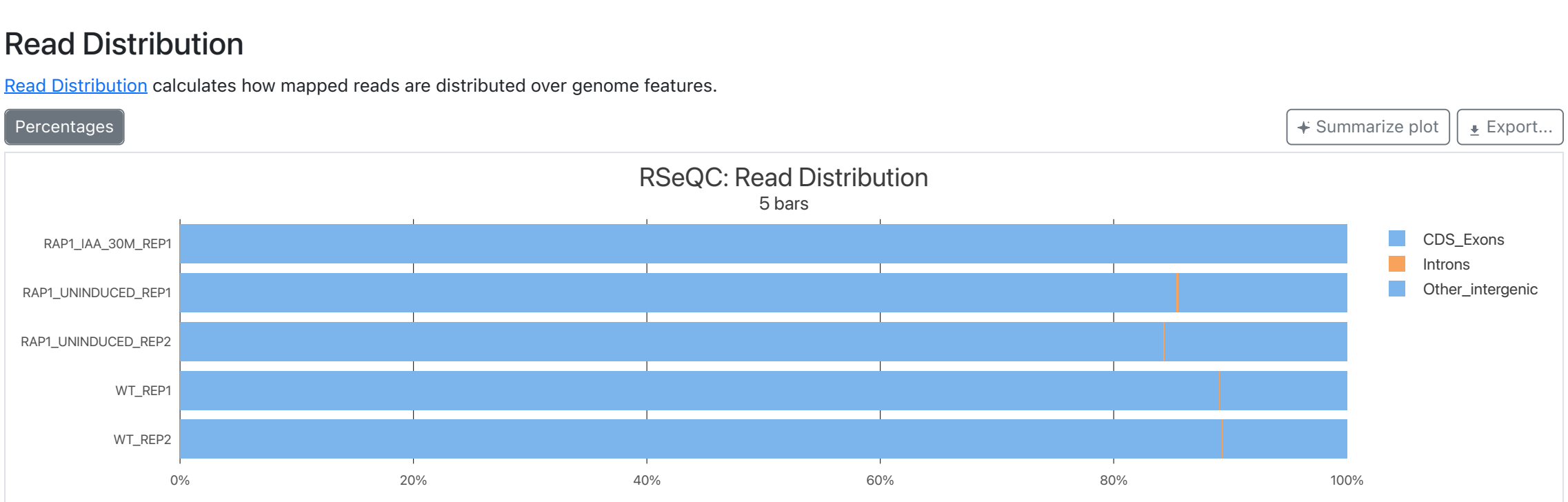
RSeQC

Evaluates high throughput RNA-seq data. http://rseqc.sourceforge.net DOI: 10.1093/bioinformatics/bts356

Read Distribution

Read Distribution calculates how mapped reads are distributed over genome features.

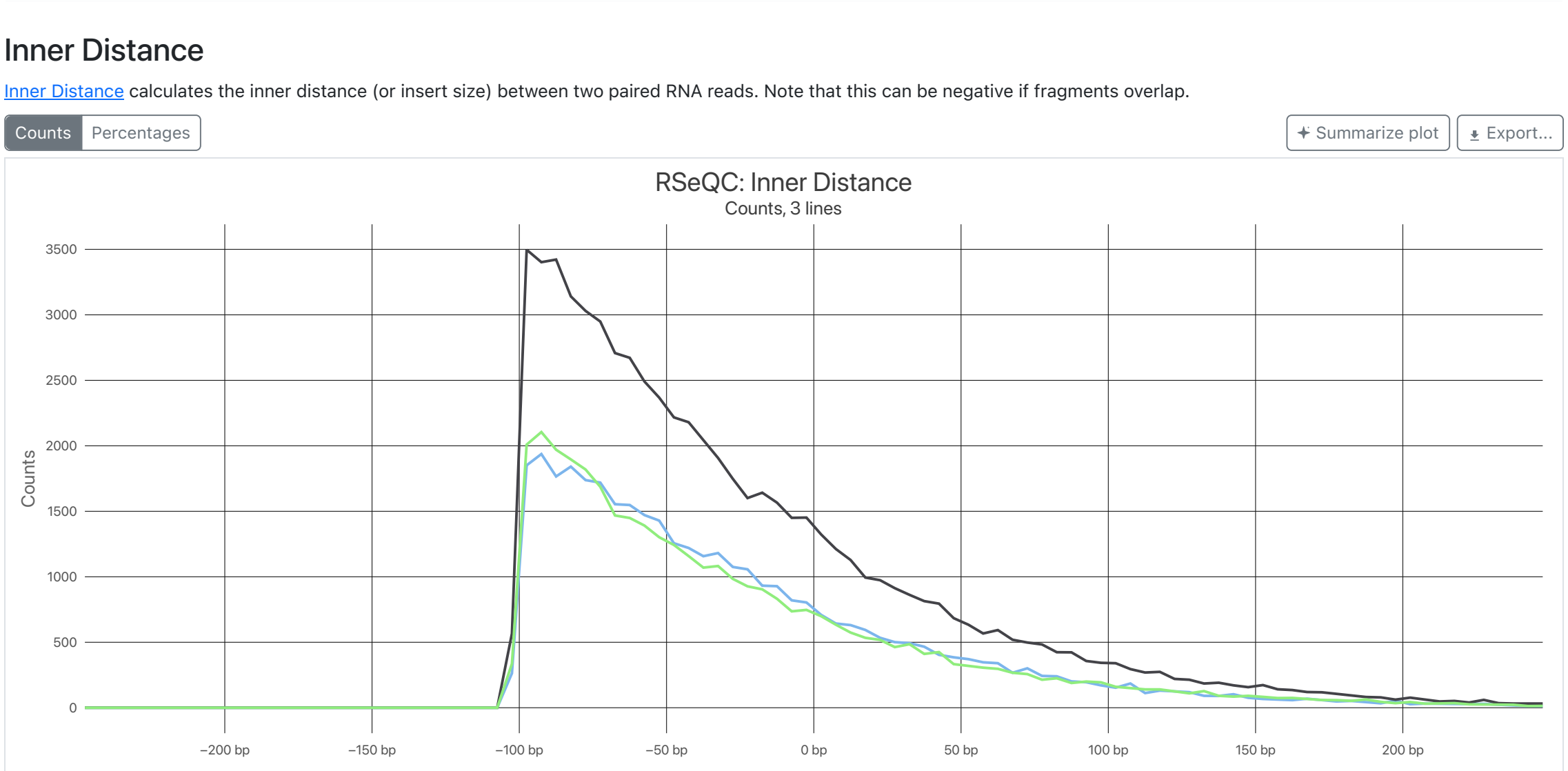
Percentages
+ Summarize plot
Export...



Inner Distance

Inner Distance calculates the inner distance (or insert size) between two paired RNA reads. Note that this can be negative if fragments overlap.

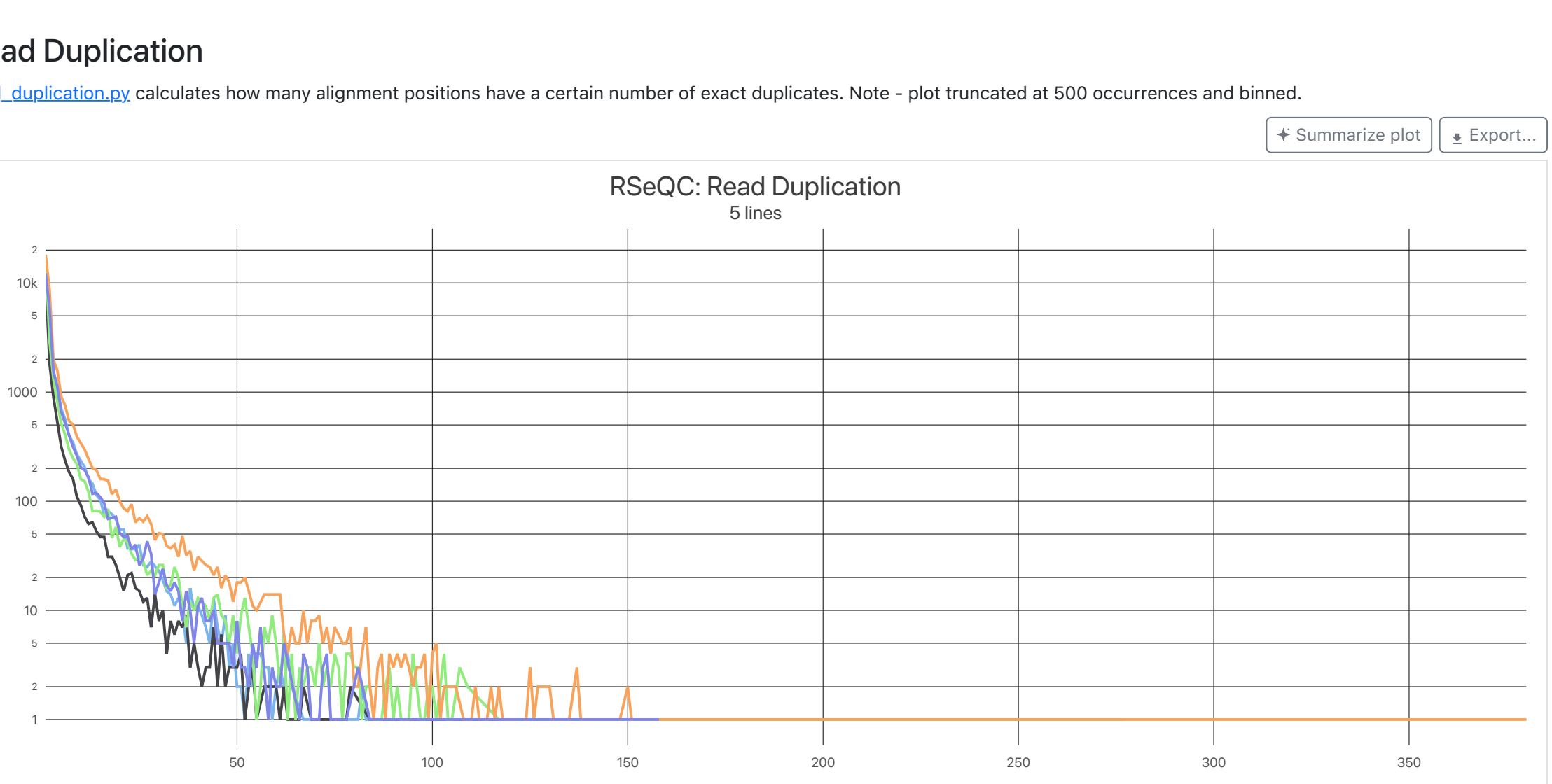
Counts
Percentages
+ Summarize plot
Export...



Read Duplication

read_duplication.py calculates how many alignment positions have a certain number of exact duplicates. Note - plot truncated at 500 occurrences and binned.

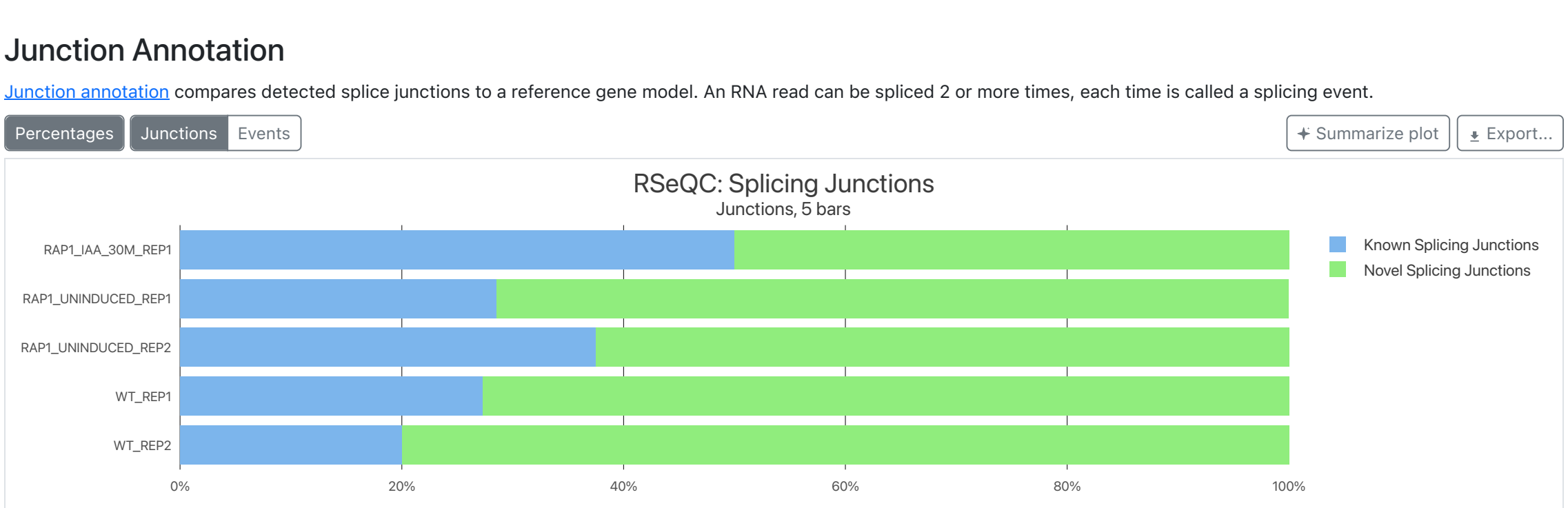
+ Summarize plot
Export...



Junction Annotation

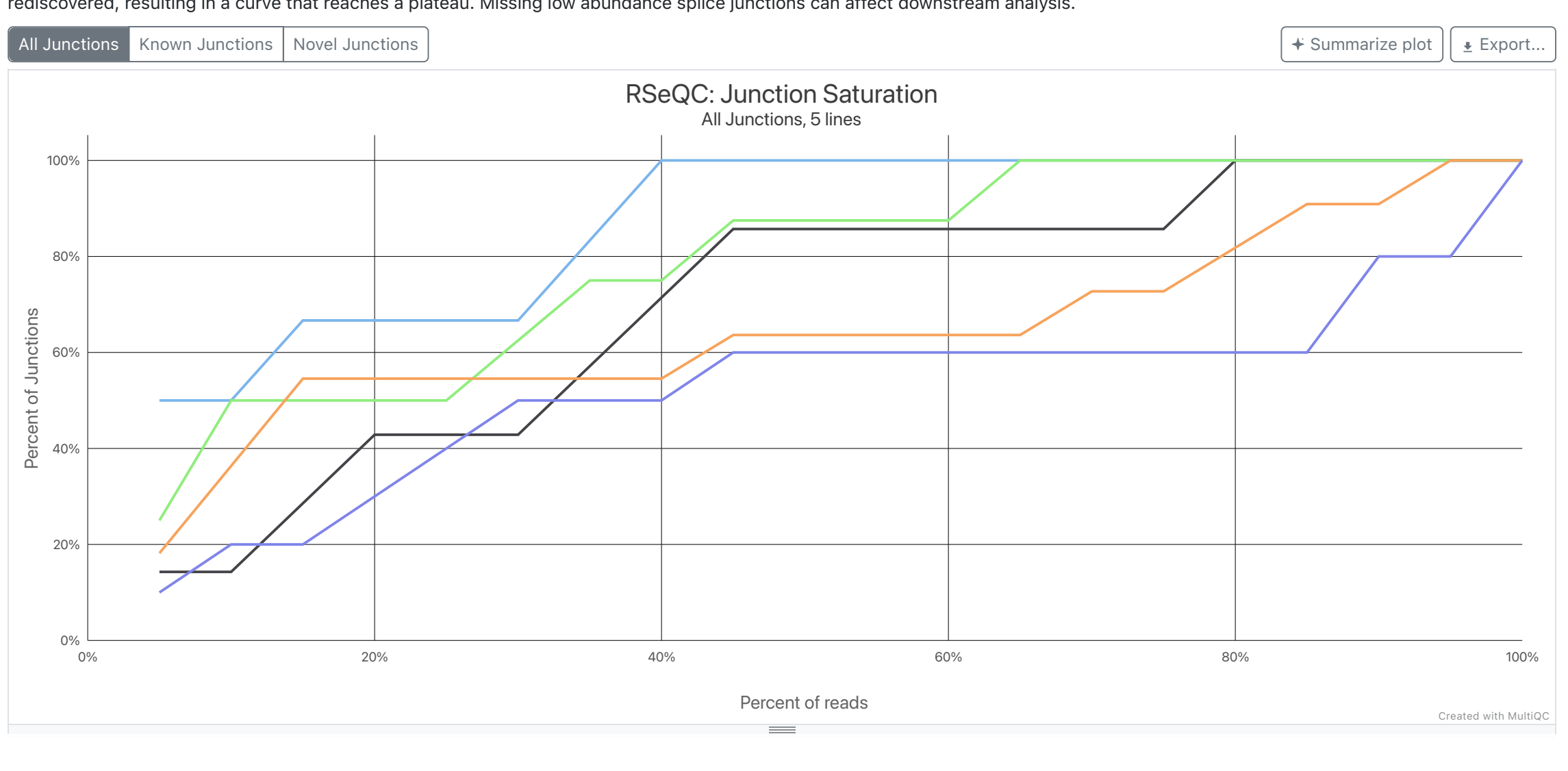
Junction annotation compares detected splice junctions to a reference gene model. An RNA read can be spliced 2 or more times, each time is called a splicing event.

Percentages
Junctions
Events
+ Summarize plot
Export...



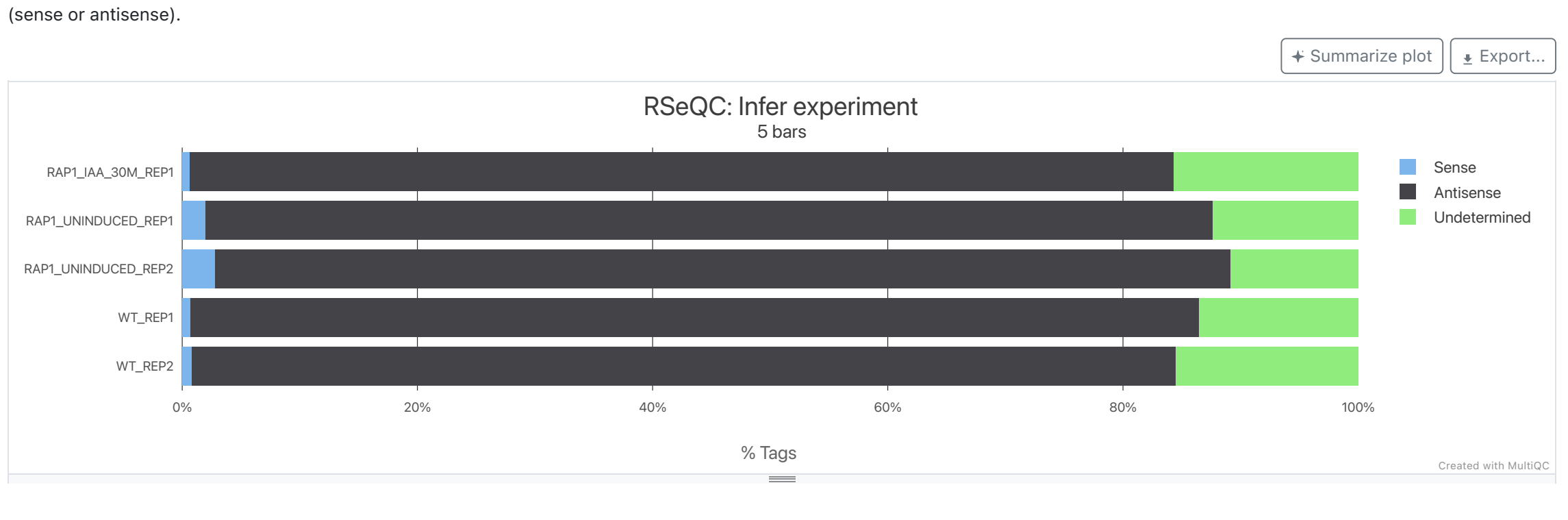
Junction Saturation

Junction Saturation counts the number of known splicing junctions that are observed in each dataset. If sequencing depth is sufficient, all (annotated) splice junctions should be rediscovered, resulting in a curve that reaches a plateau. Missing low abundance splice junctions can affect downstream analysis.



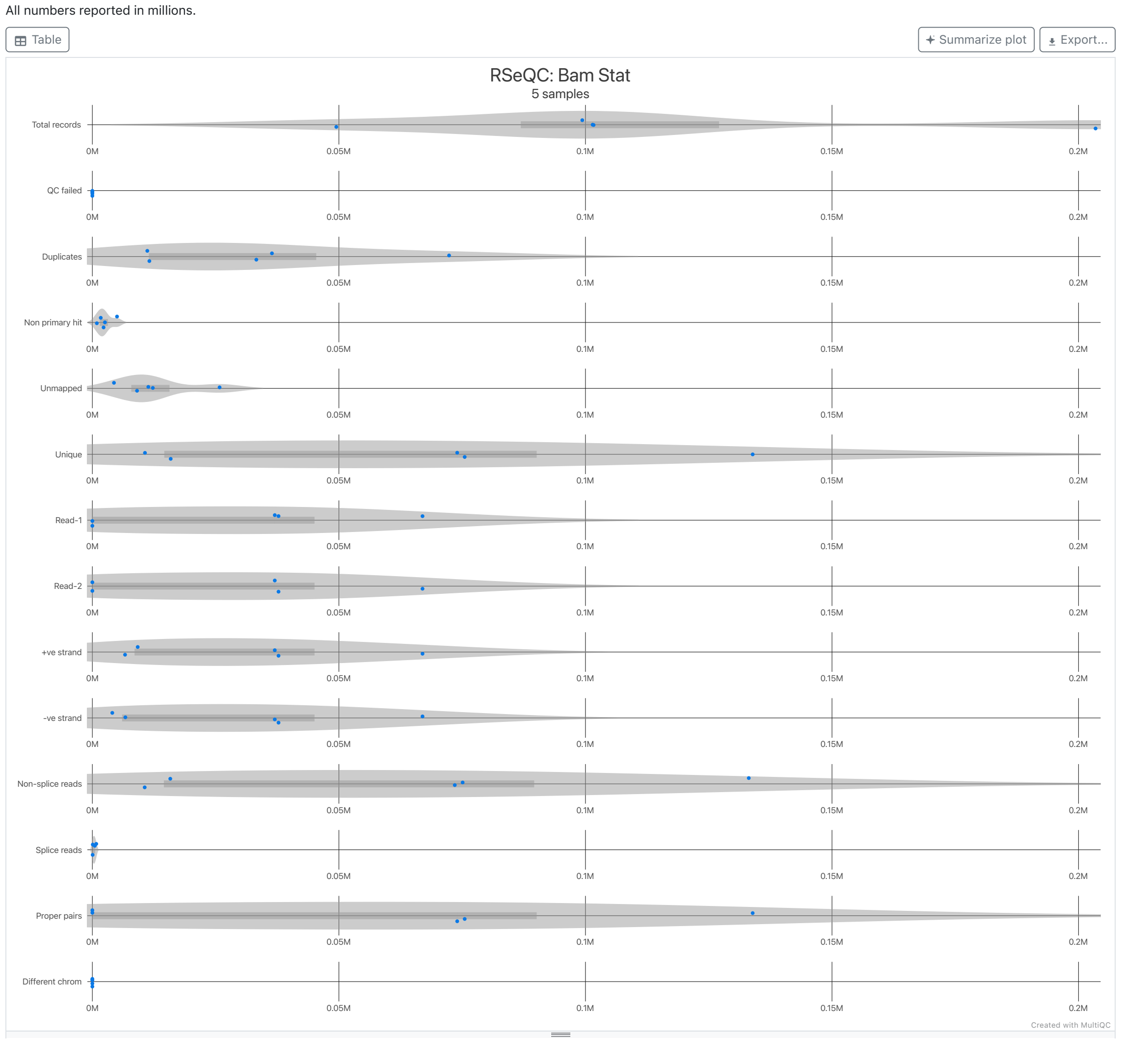
Infer experiment

Infer experiment counts the percentage of reads and read pairs that match the strandedness of overlapping transcripts. It can be used to infer whether RNA-seq library preps are stranded (sense or antisense).



Bam Stat

All numbers reported in millions.

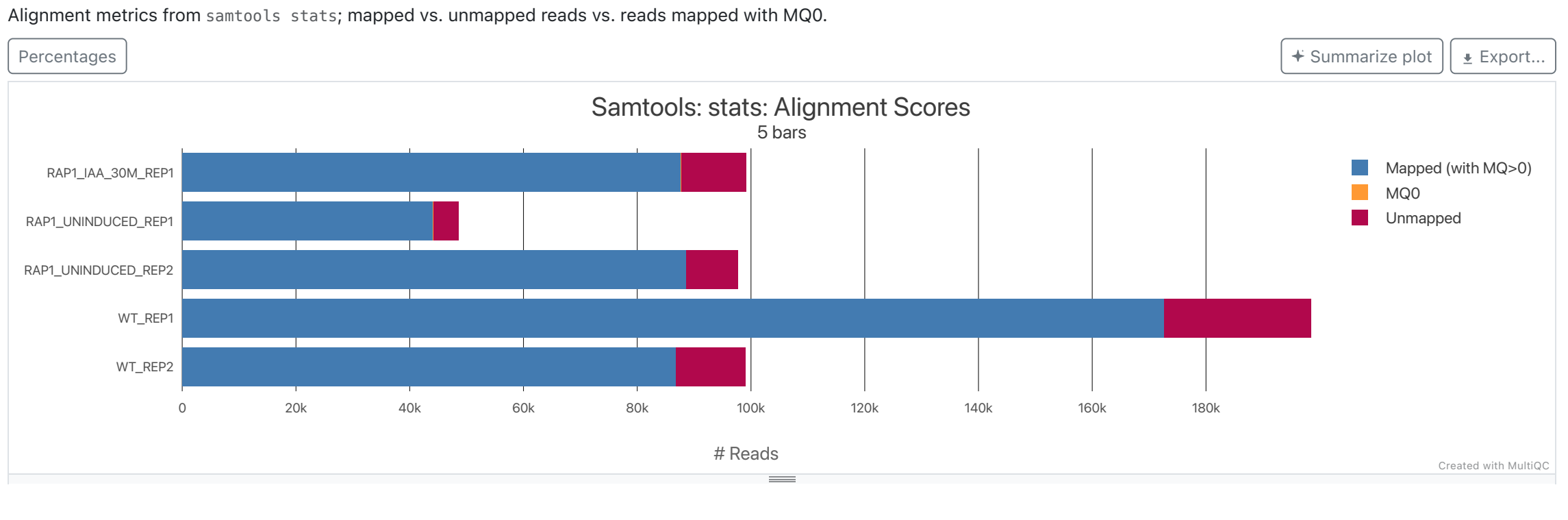


Samtools

Toolkit for interacting with BAM/CRAM files. http://www.htslib.org DOI: 10.1093/bioinformatics/btp352

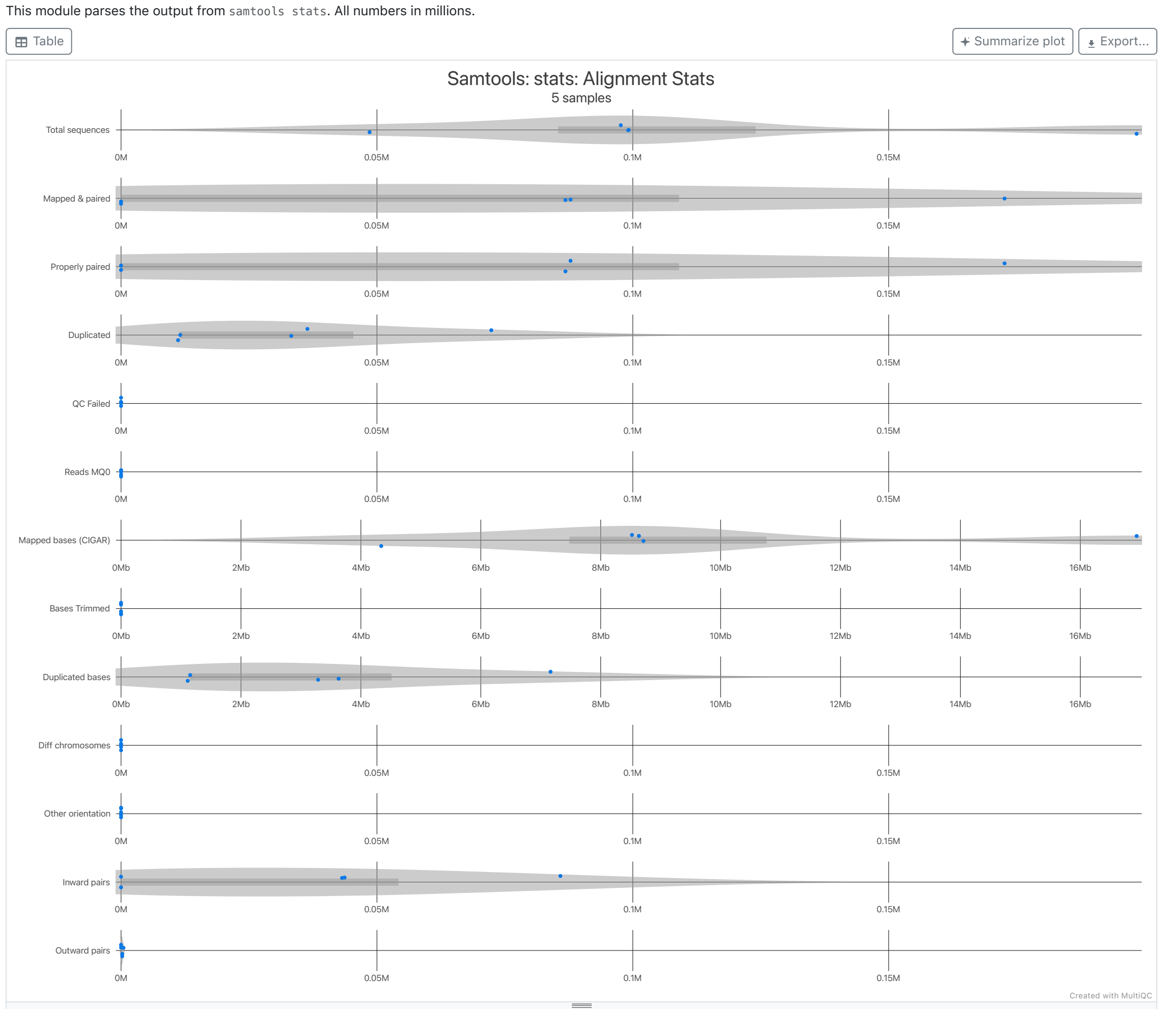
Percent mapped

Alignment metrics from `samtools stats`; mapped vs. unmapped reads vs. reads mapped with MQ0.



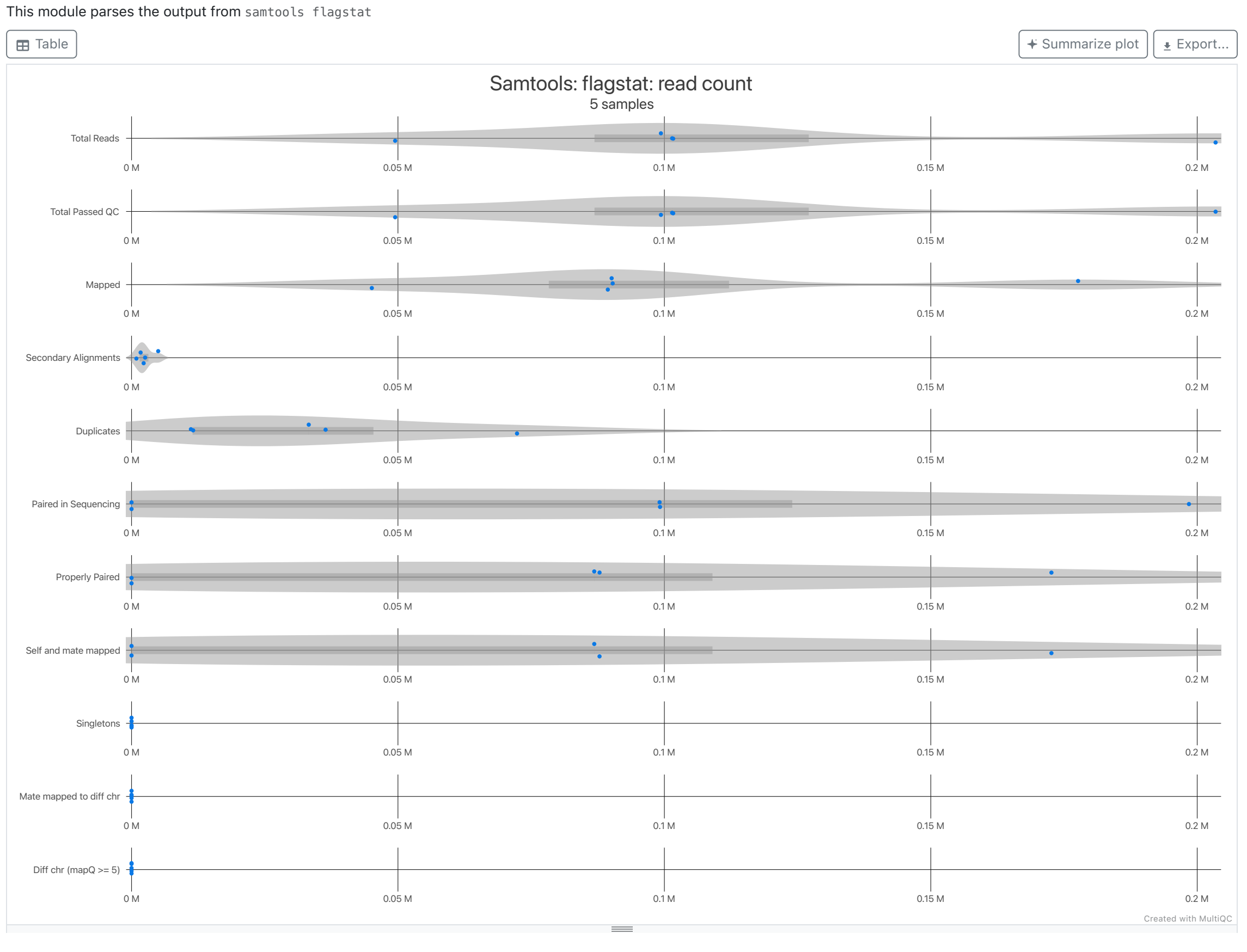
Alignment stats

This module parses the output from `samtools stats`. All numbers in millions.



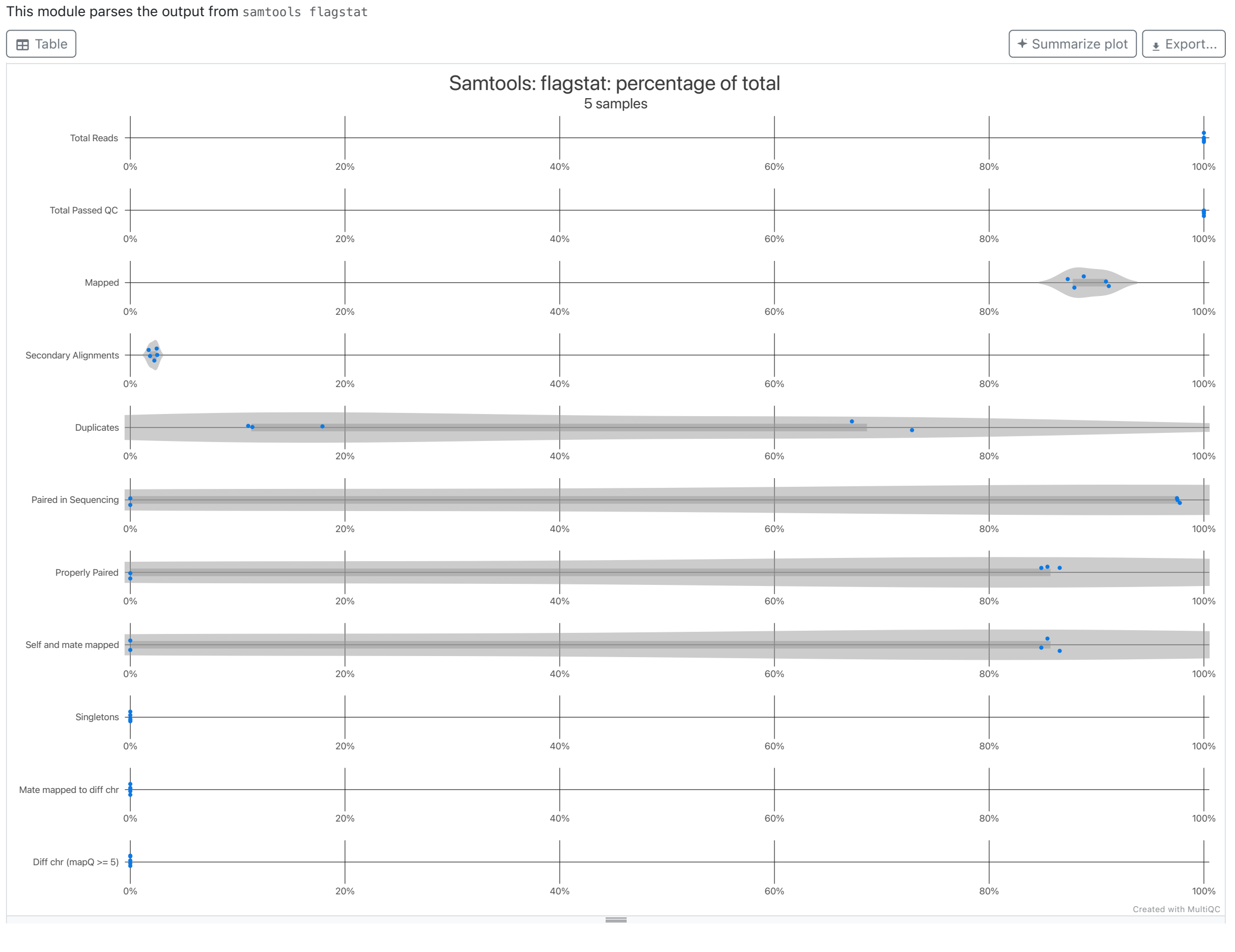
Flagstat

This module parses the output from `samtools flagstat`



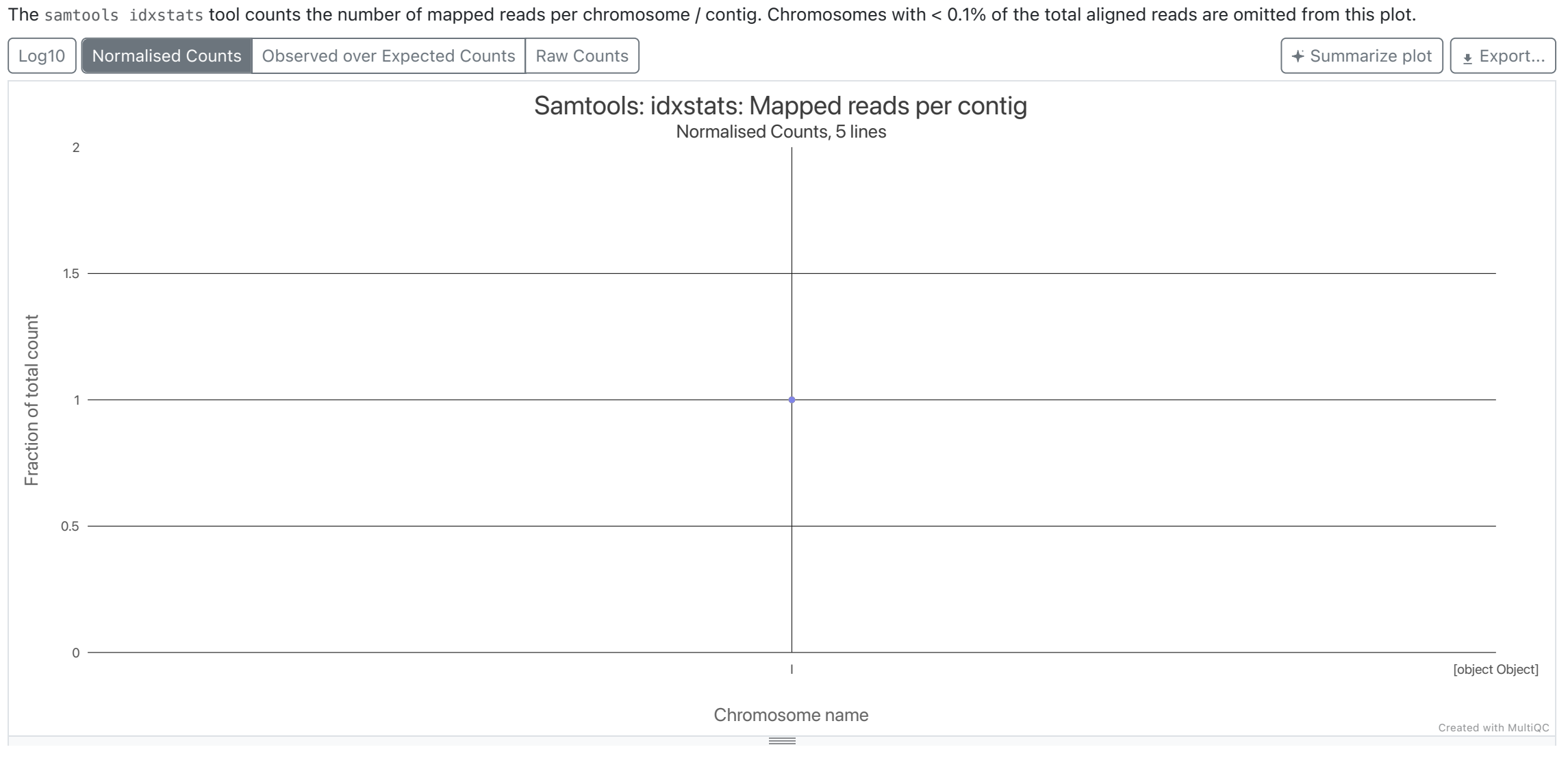
Flagstat: Percentage of total

This module parses the output from `samtools flagstat`



Mapped reads per contig

The `samtools idxstats` tool counts the number of mapped reads per chromosome / contig. Chromosomes with < 0.1% of the total aligned reads are omitted from this plot.



STAR

Universal RNA-seq aligner. https://github.com/alexdobin/STAR DOI: 10.1093/bioinformatics/bts635

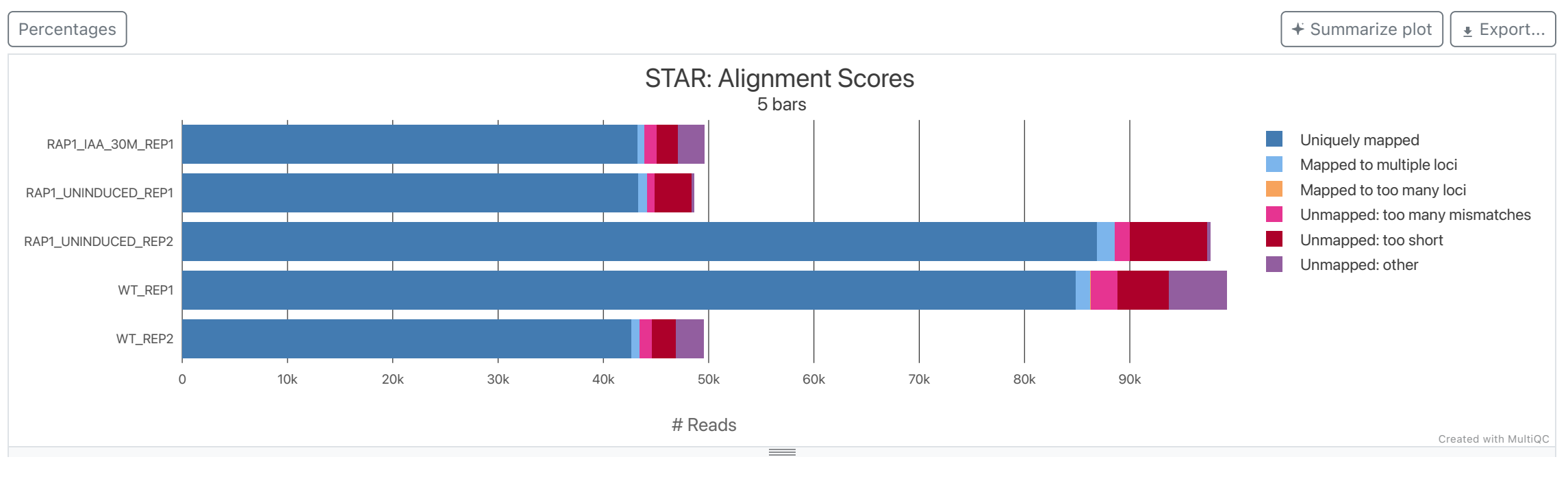
Summary Statistics

Summary statistics from the STAR alignment

Copy table | Configure columns | Scatter plot | Violin plot | Export as CSV... | Showing $^5/_5$ rows and $^{10}/_{19}$ columns. | Summarize table

Sample Name	Total reads	Aligned	Uniq aligned	Avg. mapped len	Annotated splices	Mismatch rate	Del rate	Del len	Ins rate	Ins len
RAP1_IAA_30M_REP1	0.0 M	88.5 %	87.2 %	196.8 bp	0.0 M	0.6 %	0.0 %	1.1 bp	0.0 %	1.1 bp
RAP1_UNINDUCED_REP1	0.0 M	91.0 %	89.2 %	98.3 bp	0.0 M	0.5 %	0.0 %	1.2 bp	0.0 %	1.1 bp
RAP1_UNINDUCED_REP2	0.1 M	90.7 %	89.0 %	98.4 bp	0.0 M	0.5 %	0.0 %	1.2 bp	0.0 %	1.1 bp
WT_REP1	0.1 M	87.0 %	85.5 %	196.3 bp	0.0 M	0.7 %	0.0 %	1.1 bp	0.0 %	1.1 bp
WT_REP2	0.0 M	87.6 %	86.2 %	196.4 bp	0.0 M	0.6 %	0.0 %	1.1 bp	0.0 %	1.1 bp

Alignment Scores

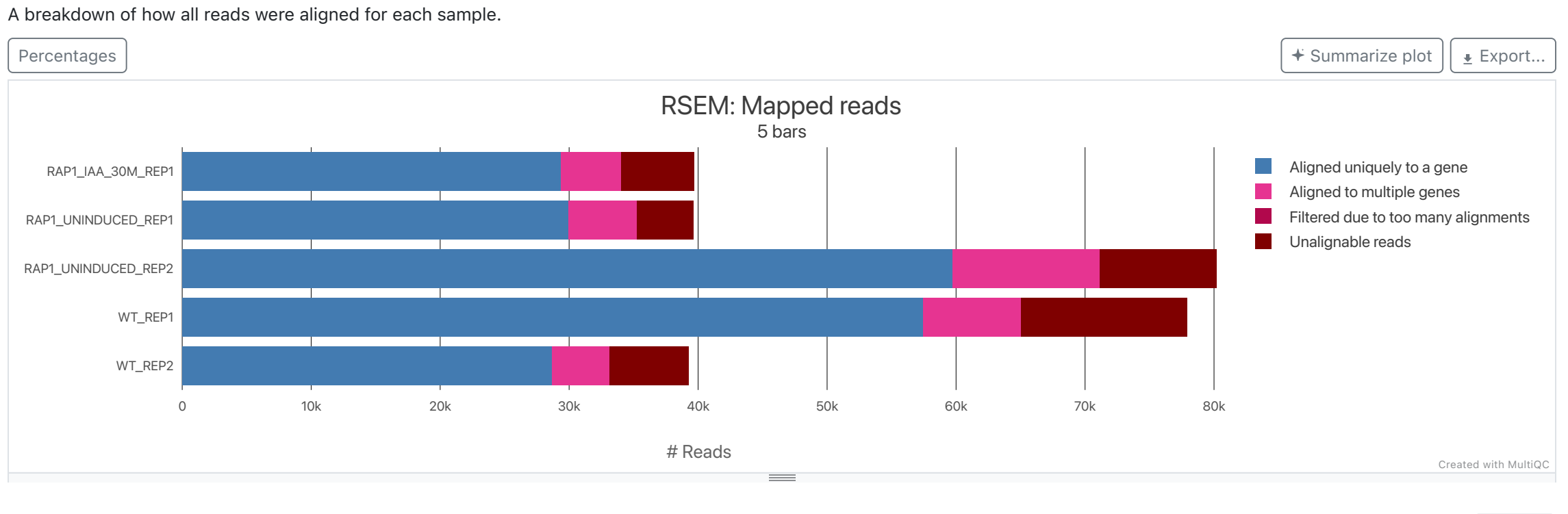


RSEM

Estimates gene and isoform expression levels from RNA-Seq data. https://deweylab.github.io/RSEM DOI: 10.1186/1471-2105-12-323

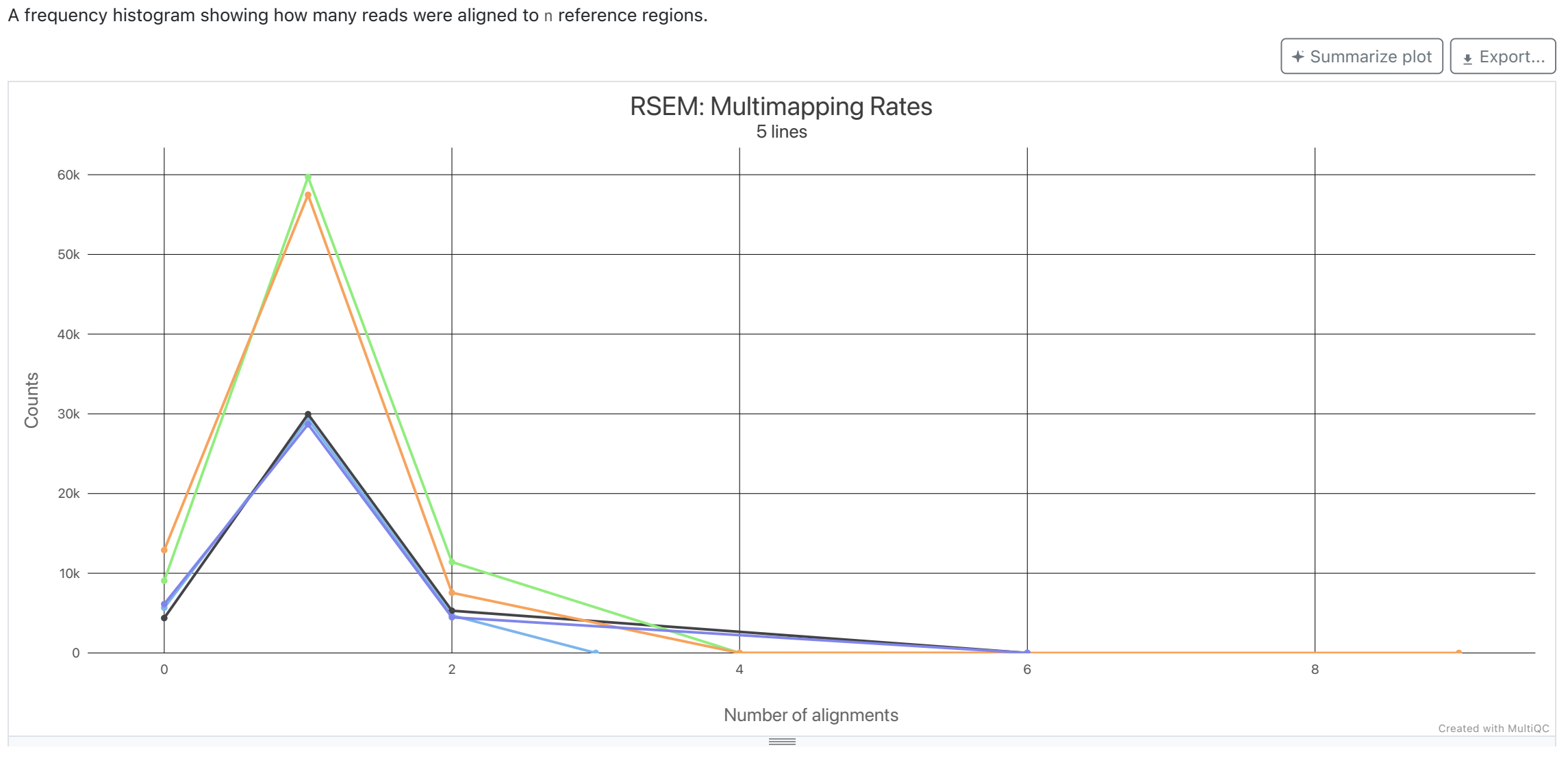
Mapped Reads

A breakdown of how all reads were aligned for each sample.



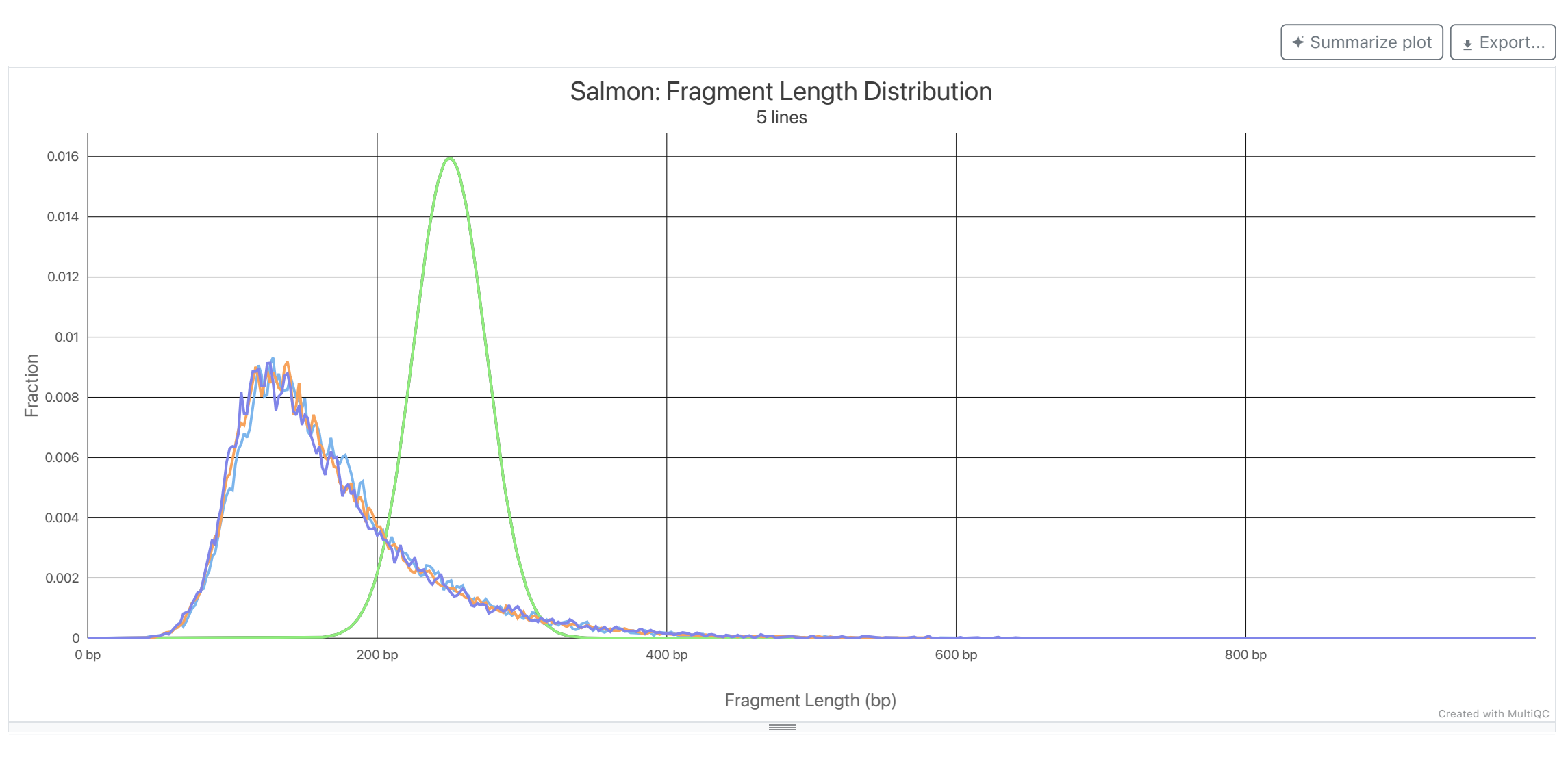
Multimapping rates

A frequency histogram showing how many reads were aligned to `n` reference regions.



Salmon

Quantifies expression of transcripts using RNA-seq data. https://combine-lab.github.io/salmon DOI: 10.1038/nmeth.4197

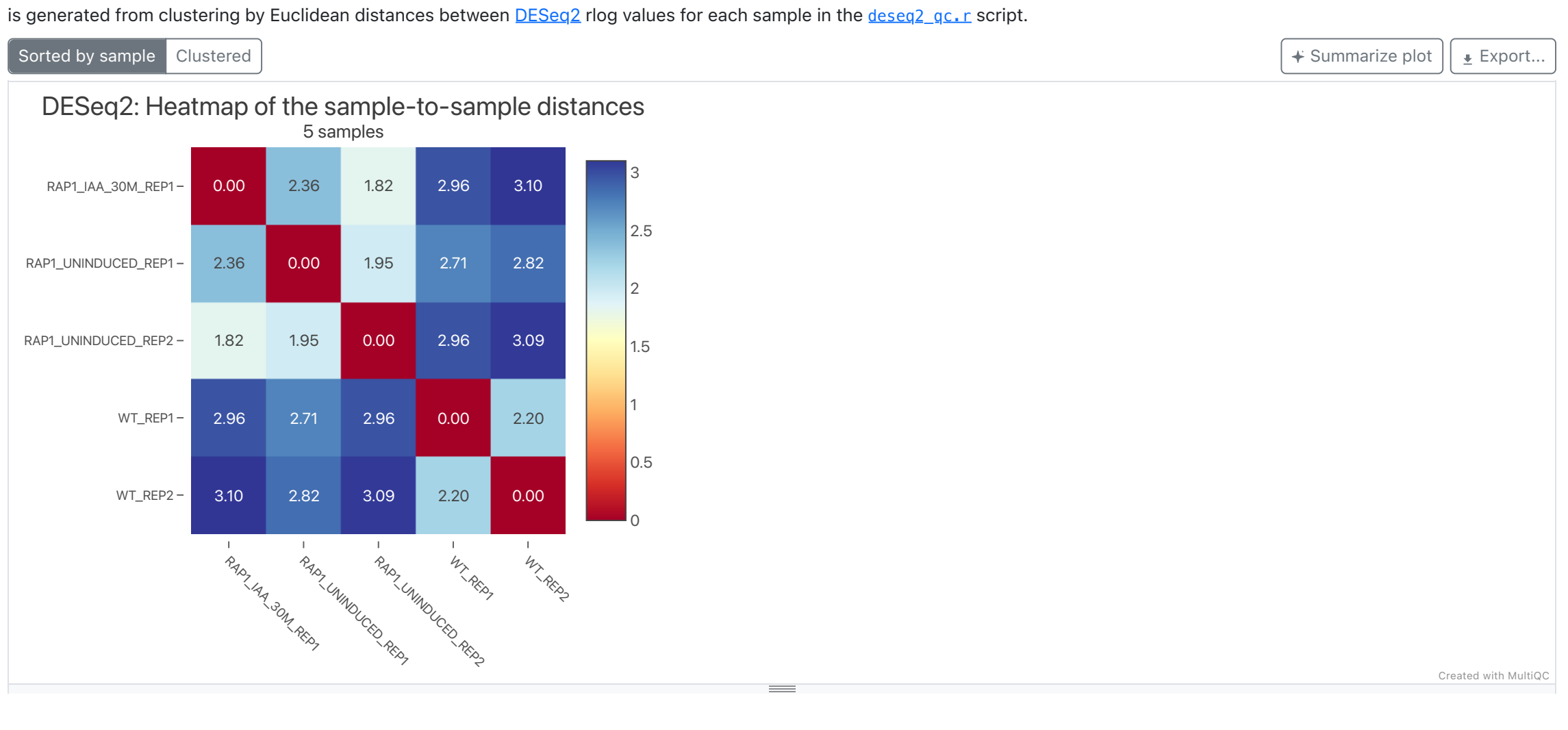


Sample relationships

Plots interrogating sample relationships, based on final count matrices.

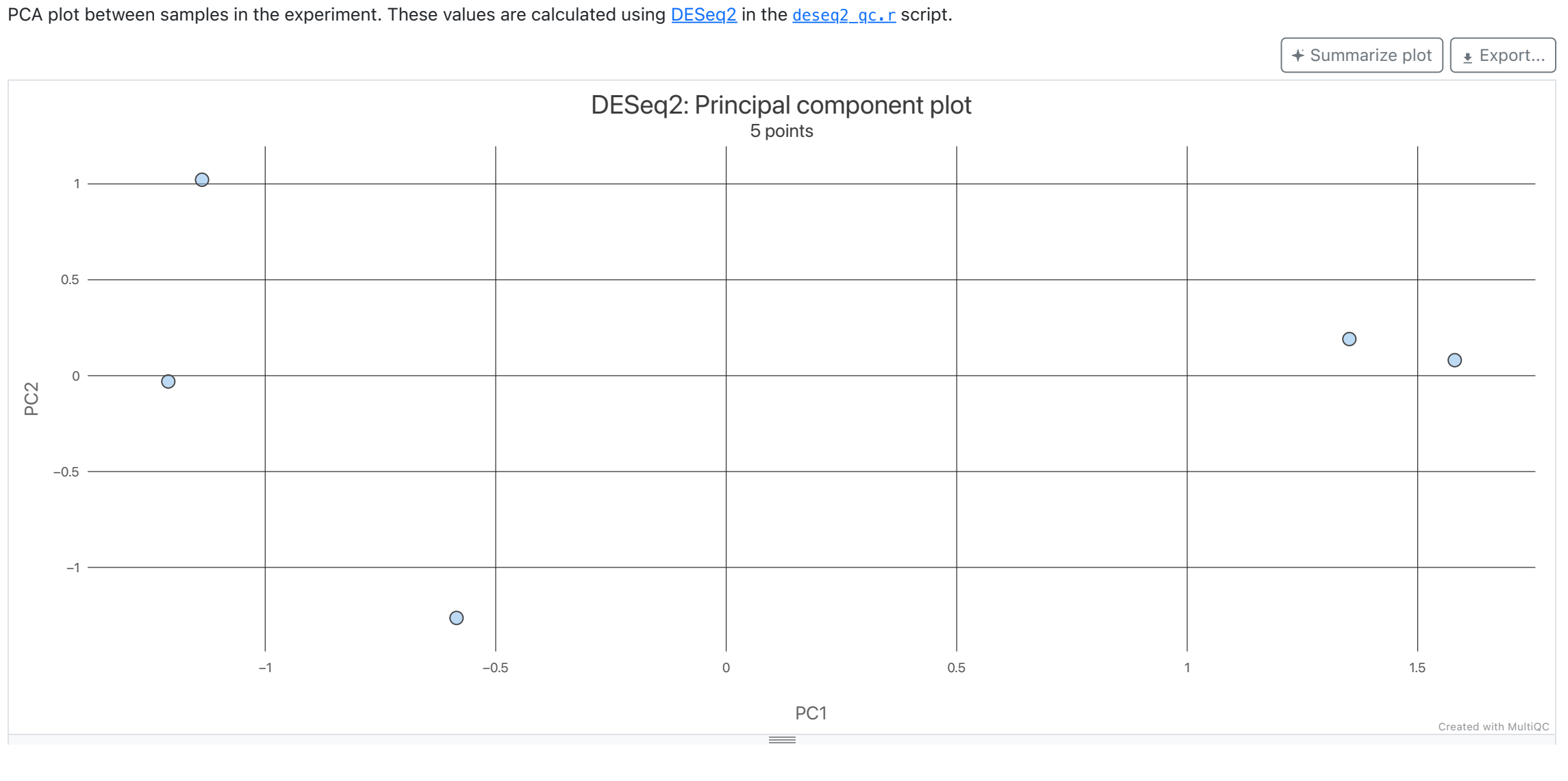
STAR_RSEM DESeq2 sample similarity

is generated from clustering by Euclidean distances between DESeq2 rlog values for each sample in the `deseq2_qc.r` script.



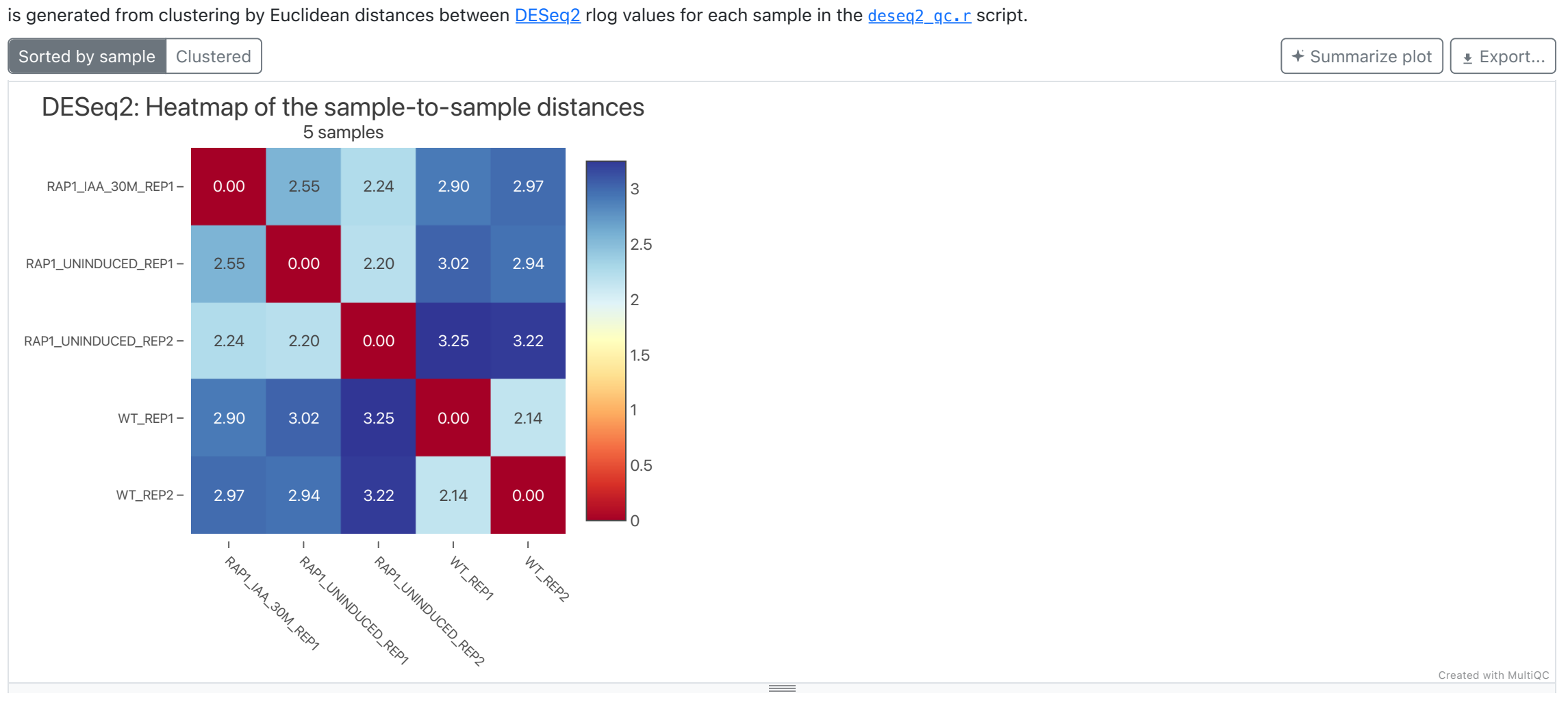
STAR_RSEM DESeq2 PCA plot

PCA plot between samples in the experiment. These values are calculated using DESeq2 in the `deseq2_qc.r` script.



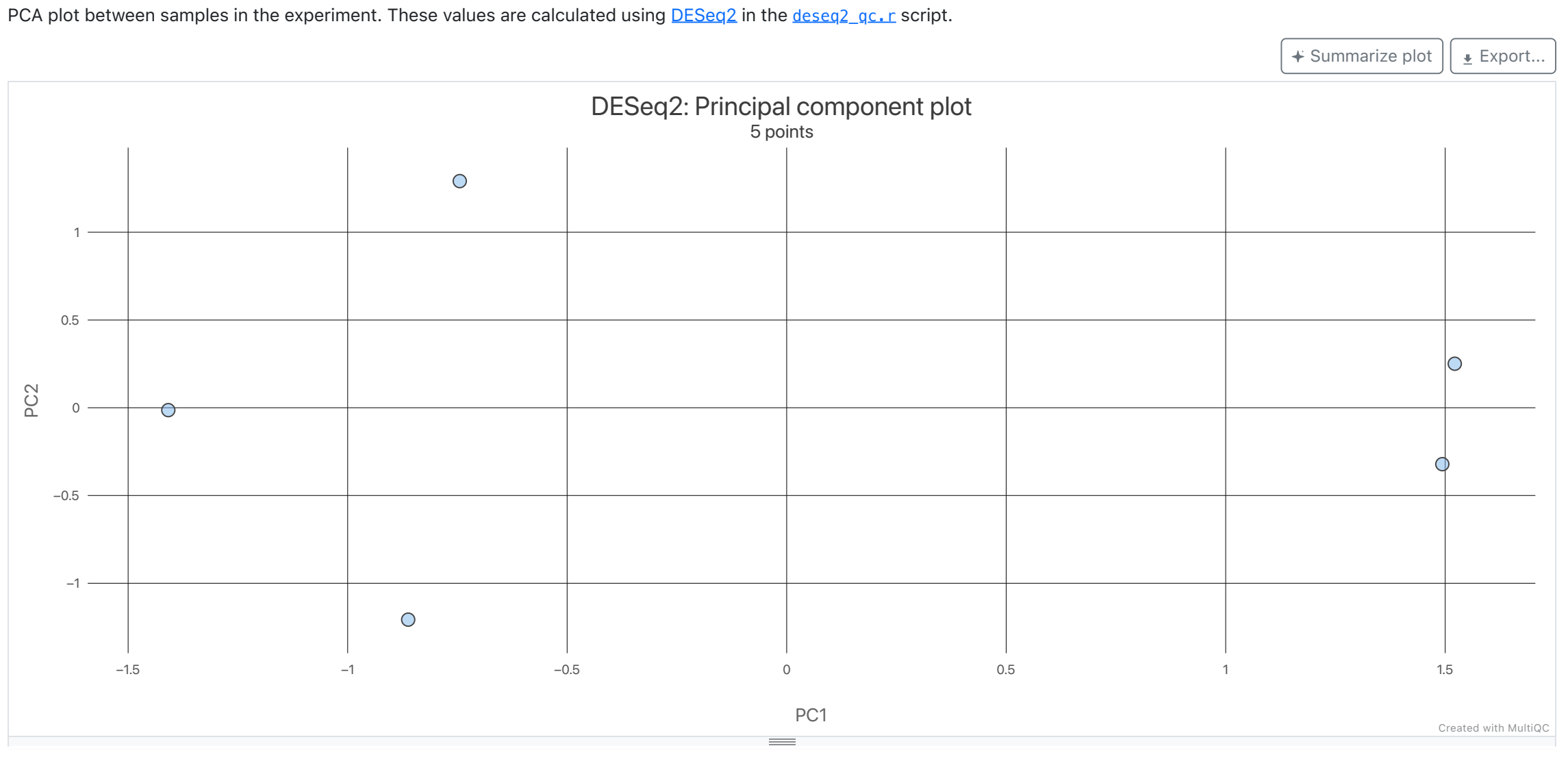
SALMON DESeq2 sample similarity

is generated from clustering by Euclidean distances between DESeq2 rlog values for each sample in the `deseq2_qc.r` script.



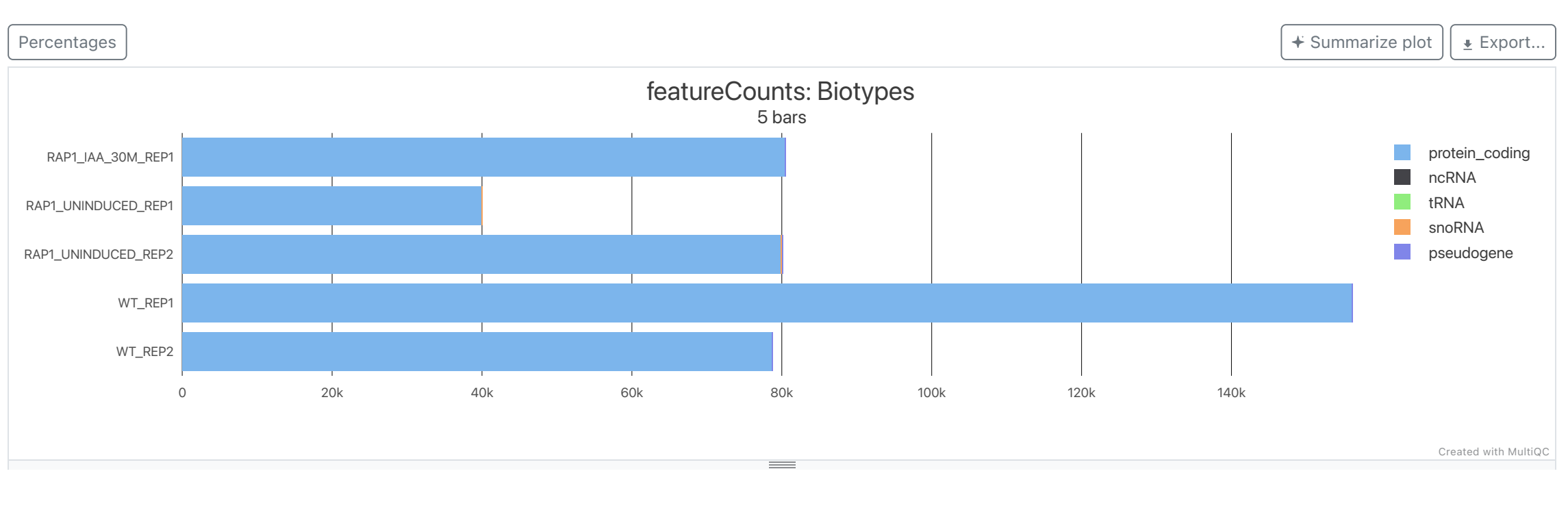
SALMON DESeq2 PCA plot

PCA plot between samples in the experiment. These values are calculated using DESeq2 in the `deseq2_qc.r` script.



Biotype Counts

Biotype Counts shows reads overlapping genomic features of different biotypes, counted by featureCounts.



Software Versions

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

Copy table

Group	Software	Version
BBMAP_BBSPLIT	bbmap	39.18
BEDTOOLS_GENOMECOV_COMBINED	bedtools	2.31.1
BEDTOOLS_GENOMECOV_FW	bedtools	2.31.1
BEDTOOLS_GENOMECOV_REV	bedtools	2.31.1
CAT_FASTQ	cat	9.5
CUSTOM_CATADDITIONALFASTA	python	3.12.2
CUSTOM_GTFFILTER	python	3.9.5
CUSTOM_MULTIQCCUSTOMBIOTYPE	python	3.12.12
CUSTOM_RSEMMERGECOUNTS	sed	4.7
CUSTOM_TX2GENE	python	3.10.4
DESEQ2_QC_PSEUDO	bioconductor-deseq2	1.46.0
	r-base	4.4.2
DESEQ2_QC_RSEM	bioconductor-deseq2	1.46.0
	r-base	4.4.2
DupRadar	bioconductor-dupradar	1.38.0
EAUTILS_GTF2BED	perl	5.26.2
FASTQC	fastqc	0.12.1
FASTQC_FILTERED	fastqc	0.12.1
FQ_LINT	fq	0.12.0
FQ_LINT_AFTER_BBSPLIT	fq	0.12.0
FQ_LINT_AFTER_TRIMMING	fq	0.12.0
FQ_SUBSAMPLE	fq	0.12.0
GUNZIP_ADDITIONAL_FASTA	gunzip	1.13
GUNZIP_GTF	gunzip	1.13
PICARD_MARKDUPLICATES	picard	3.4.0
QUALIMAP_RNASEQ	qualimap	2.3
RSEM_CALCULATEEXPRESSION	rsem	1.3.1
RSEM_PREPAREREFERENCE_GENOME	rsem	1.3.1
RSEQC_BAMSTAT	rseqc	5.0.4
RSEQC_INFEREXPERIMENT	rseqc	5.0.4
RSEQC_INNERDISTANCE	rseqc	5.0.4
RSEQC_JUNCTIONANNOTATION	rseqc	5.0.4
RSEQC_JUNCTIONSATURATION	rseqc	5.0.4
RSEQC_READDISTRIBUTION	rseqc	5.0.4
RSEQC_READDUPLICATION	rseqc	5.0.4
SALMON_QUANT	salmon	1.10.3
SAMTOOLS_FAIDX	samtools	1.23.1
SAMTOOLS_FLAGSTAT	samtools	1.23.1
SAMTOOLS_IDXSTATS	samtools	1.23.1
SAMTOOLS_INDEX	samtools	1.23.1
SAMTOOLS_SORT	samtools	1.23.1
SAMTOOLS_SORT_QUALIMAP	samtools	1.23.1
SAMTOOLS_STATS	samtools	1.23.1
SE_GENE_UNIFIED	bioconductor-summarizedexperiment	1.32.0
SE_TRANSCRIPT_UNIFIED	bioconductor-summarizedexperiment	1.32.0
STAR_ALIGN	gawk	5.1.0
	samtools	1.21
	star	2.7.11b
STAR_GENOMEGENERATE	gawk	5.1.0
	samtools	1.21
	star	2.7.11b
STRINGTIE_STRINGTIE	stringtie	2.2.3
SUBREAD_FEATURECOUNTS	subread	2.0.6
TRIMGALORE	trimgalore	2.1.0
TXIMETA_TXIMPORT	bioconductor-tximeta	1.20.1
UCSC_BEDCLIP	ucsc	377
UCSC_BEDGRAPHTOBIGWIG	ucsc	469
UNTAR_SALMON_INDEX	untar	1.34
Workflow	Nextflow	26.04.2
	nf-core/rnaseq	v3.26.0-ge7ca462

nf-core/rnaseq Methods Description

Suggested text and references to use when describing pipeline usage within the methods section of a publication. https://github.com/nf-core/rnaseq

Methods

Data was processed using nf-core/rnaseq v3.26.0 (doi: 10.5281/zenodo.1400710) of the nf-core collection of workflows (Ewels *et al.*, 2020), utilising reproducible software environments from the Bioconda (Grüning *et al.*, 2018) and Biocontainers (da Veiga Leprevost *et al.*, 2017) projects.

The pipeline was executed with Nextflow v26.04.2 (Di Tommaso *et al.*, 2017) with the following command:

```
nextflow run nf-core/rnaseq -r 3.26.0 -profile test,singularity -c beocat_rnaseq.config --outdir rnaseq_star_rsem_test_output --aligner star_rsem
```

References

- Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., & Notredame, C. (2017). Nextflow enables reproducible computational workflows. Nature Biotechnology, 35(4), 316-319. doi: 10.1038/nbt.3820

- Ewels, P. A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M. U., Di Tommaso, P., & Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. Nature Biotechnology, 38(3), 276-278. doi: 10.1038/s41587-020-0439-x
- Grüning, B., Dale, R., Sjödin, A., Chapman, B. A., Rowe, J., Tomkins-Tinch, C. H., Valieris, R., Köster, J., & Bioconda Team. (2018). Bioconda: sustainable and comprehensive software distribution for the life sciences. Nature Methods, 15(7), 475–476. doi: 10.1038/s41592-018-0046-7
- da Veiga Leprevost, F., Grüning, B. A., Alves Aflitos, S., Röst, H. L., Uszkoreit, J., Barsnes, H., Vaudel, M., Moreno, P., Gatto, L., Weber, J., Bai, M., Jimenez, R. C., Sachsenberg, T., Pfeuffer, J., Vera Alvarez, R., Griss, J., Nesvizhskii, A. I., & Perez-Riverol, Y. (2017). BioContainers: an open-source and community-driven framework for software standardization. Bioinformatics (Oxford, England), 33(16), 2580–2582. doi: 10.1093/bioinformatics/btx192

Notes:

- The command above does not include parameters contained in any configs or profiles that may have been used. Ensure the config file is also uploaded with your publication!
- You should also cite all software used within this run. Check the "Software Versions" of this report to get version information.

nf-core/rnaseq Workflow Summary

- this information is collected when the pipeline is started. https://github.com/nf-core/rnaseq

Input/output options

input
`https://raw.githubusercontent.com/nf-core/test-datasets/626c8fab639062eade4b10747e919341cbf9b41a/samplesheet/v3.10/samplesheet_test.csv`

outdir
`rnaseq_star_rsem_test_output`

Reference genome options

additional_fasta
`https://raw.githubusercontent.com/nf-core/test-datasets/626c8fab639062eade4b10747e919341cbf9b41a/reference/gfp.fa.gz`

fasta
`https://raw.githubusercontent.com/nf-core/test-datasets/626c8fab639062eade4b10747e919341cbf9b41a/reference/genome.fasta`

gff
`https://raw.githubusercontent.com/nf-core/test-datasets/626c8fab639062eade4b10747e919341cbf9b41a/reference/genes.gff.gz`

gtf
`https://raw.githubusercontent.com/nf-core/test-datasets/626c8fab639062eade4b10747e919341cbf9b41a/reference/genes_with_empty_tid.gtf.gz`

hisat2_index
`https://raw.githubusercontent.com/nf-core/test-datasets/626c8fab639062eade4b10747e919341cbf9b41a/reference/hisat2.tar.gz`

salmon_index
`https://raw.githubusercontent.com/nf-core/test-datasets/626c8fab639062eade4b10747e919341cbf9b41a/reference/salmon.tar.gz`

transcript_fasta
`https://raw.githubusercontent.com/nf-core/test-datasets/626c8fab639062eade4b10747e919341cbf9b41a/reference/transcriptome.fasta`

Read filtering options

bbsplit_fasta_list
`https://raw.githubusercontent.com/nf-core/test-datasets/626c8fab639062eade4b10747e919341cbf9b41a/reference/bbsplit_fasta_list.txt`

UMI options

umi_discard_read
`0`

umitools_bc_pattern
`NNNN`

Alignment options

aligner
`star_rsem`

pseudo_aligner
`salmon`

Quality Control

kraken_db
`https://raw.githubusercontent.com/nf-core/test-datasets/eb0cbf73c3f103f8aeda9878ba200e92b4d045d8/data/genomics/sarscov2/genome/db/kraken2.tar.gz`

Process skipping options

skip_bbsplit
`false`

Institutional config options

config_profile_contact
`/Teresa Shippy (tshippy@ksu.edu)`

config_profile_description
`Beocat at Kansas State University cluster profile provided by nf-core/configs.`

config_profile_name
`Test profile`

Generic options

trace_report_suffix
`2026-08-25_21-15-59`

Core Nextflow options

configFiles
`/homes/tshippy/.nextflow/assets/.repos/nf-core/rnaseq/clones/e7ca46272c8f9d5ceee3f71759f4ba551d3217a4/nextflow.config, /bulk/bioinfo/Nextflow_demos/beocat_rnaseq.config`

container
`[.*PARABRICKS_STARGENOMEGENERATE:community.wave.seqera.io/library/htslib_samtools_star_gawk:4de2f983041d42e6]`

containerEngine
`singularity`

launchDir
`/bulk/bioinfo/Nextflow_demos`

profile
`test,singularity`

projectDir
`/homes/tshippy/.nextflow/assets/.repos/nf-core/rnaseq/clones/e7ca46272c8f9d5ceee3f71759f4ba551d3217a4`

revision
`master`

runName
`friendly_fermat`

userName
`tshippy`

workDir
`/bulk/bioinfo/Nextflow_demos/work`

MultiQC v1.33 - Written by Phil Ewels, available on GitHub.
This report uses Plotly, jQuery, jQuery UI, Bootstrap and FileSaver.js.