



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 2024-11-15, 16:28 UTC based on data in: /media/jochum00/Aagaard_Raid3/jochum00/k_pennington/LIF_project/BaseSpace/edwards-8184-434104856/work/2a/8c5685237f521eeccac9f96833c44

General Statistics

	Copy table	Configure columns	Scatter plot	Violin plot	Showing 0/75 rows and 19/34 columns.										Export as CSV	
Sampled up	Duplication	M	Proper	Error	Non-primary	Reads %	%	Total	Reads mapped	Aligned	Uniq	GC	Seqs	Trimmed	GC	Seqs bases
Name	bias	AlignedPairs	AlignedPairs	rate	narily	napped	Proper	seqs	mapped	aligned	aligned	pairs				
CONTROL_0h	90.2%	1.44	20.2 M	5.7%	0.36%	23.2 M	40.4 M	100.0%	100.0%	40.4 M	63.6 M	91.3%	48.8%			
CONTROL_0hr_1_1												59.0%	23.9 M	15.2%	58.0%	22.2 M
CONTROL_0hr_1_2												59.0%	23.9 M	11.0%	58.0%	22.2 M
CONTROL_0hr_1h	96.3%	1.65	23.8 M	1.4%	0.38%	50.0 M	47.7 M	100.0%	100.0%	47.7 M	97.7 M	92.0%	26.0%			
CONTROL_0hr_2_1												62.0%	27.2 M	12.6%	62.0%	25.9 M
CONTROL_0hr_2_2												63.0%	27.2 M	10.1%	63.0%	25.9 M
CONTROL_0hr_3_1	92.4%	1.55	28.5 M	3.5%	0.39%	52.1 M	57.0 M	100.0%	100.0%	57.0 M	109.1 M	91.0%	32.3%			
CONTROL_0hr_3_2												62.0%	32.2 M	10.6%	62.0%	31.3 M
CONTROL_0hr_4_1	83.0%	1.51	26.0 M	11.0%	0.32%	21.6 M	52.0 M	100.0%	100.0%	52.0 M	73.5 M	89.0%	59.2%			
CONTROL_0hr_4_2												56.0%	31.4 M	14.9%	54.0%	29.2 M
CONTROL_0hr_5_1	86.4%	1.43	32.6 M	7.0%	0.24%	48.8 M	65.1 M	100.0%	100.0%	65.1 M	113.9 M	93.0%	41.9%			
CONTROL_0hr_5_2												56.0%	35.5 M	9.1%	56.0%	35.0 M
CONTROL_1h	74.4%	1.46	32.0 M	17.5%	0.37%	24.2 M	63.9 M	100.0%	100.0%	63.9 M	88.1 M	90.9%	65.5%			
CONTROL_1hr_1_1												56.0%	35.9 M	8.9%	56.0%	35.2 M
CONTROL_1hr_1_2												56.0%	35.9 M	7.7%	56.0%	35.2 M
CONTROL_1h	88.7%	1.45	20.0 M	5.4%	0.33%	30.7 M	40.0 M	100.0%	100.0%	40.0 M	70.6 M	92.9%	42.7%			
CONTROL_1hr_2_1												57.0%	22.8 M	14.0%	57.0%	21.5 M
CONTROL_1hr_2_2												58.0%	22.8 M	10.4%	57.0%	21.5 M
CONTROL_1h	44.0%	1.43	39.6 M	47.5%	0.30%	11.4 M	79.2 M	100.0%	100.0%	79.2 M	90.6 M	93.4%	84.0%			
CONTROL_1hr_4_1												51.0%	42.5 M	6.1%	51.0%	42.4 M
CONTROL_1hr_4_2												51.0%	42.5 M	6.1%	51.0%	42.4 M
CONTROL_1h	76.2%	1.45	34.9 M	15.4%	0.36%	30.7 M	69.8 M	100.0%	100.0%	69.8 M	100.5 M	90.0%	59.5%			
CONTROL_1hr_5_1												56.0%	39.2 M	8.2%	56.0%	38.8 M
CONTROL_1hr_5_2												57.0%	39.2 M	7.5%	57.0%	38.8 M
CONTROL_24	90.3%	1.36	27.1 M	5.3%	0.23%	33.7 M	54.2 M	100.0%	100.0%	54.2 M	87.9 M	94.0%	49.2%			
CONTROL_24hr_1_1												57.0%	31.9 M	17.6%	56.0%	28.8 M
CONTROL_24hr_1_2												58.0%	31.9 M	11.5%	56.0%	28.8 M
CONTROL_24	88.4%	1.40	31.5 M	5.4%	0.31%	54.2 M	63.0 M	100.0%	100.0%	63.0 M	117.2 M	92.4%	41.9%			
CONTROL_24hr_2_1												57.0%	35.2 M	11.4%	57.0%	34.1 M

Sampled	Duplication	M	Proper Error	Non-primary	Reads %	% primary	Total	Reads	Aligned	GC	Seqs	Trimmed	GC	Seqs	
Name	bias	AlignedPairs	rate	mapped	mapped	proper seqs	mapped	proper	aligned	aligned	bases	bases	bases	bases	
CONTROL_24hr_2_2											58.0%	35.2 M	9.0%	58.0%	34.1 M
CONTROL_24hr_2_2	89.1%	1.54	31.4 M	5.8%	0.24%	39.5 M	62.8 M	100.0%	100.0%	62.8 M	102.3 M	93.4%	50.0%		
CONTROL_24hr_3_1											55.0%	34.9 M	11.5%	54.0%	33.6 M
CONTROL_24hr_3_2											55.0%	34.9 M	8.9%	54.0%	33.6 M
CONTROL_24hr_4_1	76.0%	1.47	39.9 M	13.5%	0.29%	49.9 M	79.9 M	100.0%	100.0%	79.9 M	129.8 M	93.1%	56.1%		
CONTROL_24hr_4_1											55.0%	43.5 M	10.0%	55.0%	42.9 M
CONTROL_24hr_4_2											56.0%	43.5 M	8.9%	56.0%	42.9 M
CONTROL_24hr_4_2	90.9%	1.41	24.2 M	4.3%	0.22%	40.1 M	48.4 M	100.0%	100.0%	48.4 M	88.6 M	96.2%	39.9%		
CONTROL_24hr_5_1											56.0%	26.1 M	13.1%	55.0%	25.2 M
CONTROL_24hr_5_2											56.0%	26.1 M	10.8%	55.0%	25.2 M
LIF_1hr_0_1	79.4%	1.38	34.6 M	13.4%	0.44%	31.1 M	69.3 M	100.0%	100.0%	69.3 M	100.4 M	87.8%	60.5%		
LIF_1hr_0_1											58.0%	40.2 M	9.6%	58.0%	39.5 M
LIF_1hr_1_2											59.0%	40.2 M	8.6%	59.0%	39.5 M
LIF_1hr_2_2	94.5%	1.40	24.4 M	2.6%	0.32%	43.5 M	48.8 M	100.0%	100.0%	48.8 M	92.3 M	91.0%	35.7%		
LIF_1hr_2_2											62.0%	30.6 M	20.6%	60.0%	26.8 M
LIF_1hr_2_1											62.0%	30.6 M	11.4%	61.0%	26.8 M
LIF_1hr_3_2											60.0%	34.9 M	8.3%	60.0%	34.2 M
LIF_1hr_3_3	84.8%	1.40	31.0 M	8.2%	0.39%	41.3 M	61.9 M	100.0%	100.0%	61.9 M	103.2 M	90.4%	46.8%		
LIF_1hr_3_3											59.0%	34.9 M	9.8%	59.0%	34.2 M
LIF_1hr_3_1											60.0%	34.9 M	8.3%	60.0%	34.2 M
LIF_1hr_4_2											57.0%	38.6 M	8.4%	57.0%	37.9 M
LIF_1hr_4_1											56.0%	38.6 M	9.5%	56.0%	37.9 M
LIF_1hr_4_2											57.0%	38.6 M	8.4%	57.0%	37.9 M
LIF_1hr_4_3	89.4%	1.45	47.1 M	5.8%	0.26%	59.2 M	94.3 M	100.0%	100.0%	94.3 M	153.5 M	92.3%	55.1%		
LIF_1hr_4_3											54.0%	52.1 M	10.1%	54.0%	51.0 M
LIF_1hr_5_1											55.0%	52.1 M	8.7%	54.0%	51.0 M
LIF_1hr_5_2											56.0%	32.7 M	11.1%	56.0%	31.6 M
LIF_24hr_1_1	86.8%	1.42	29.6 M	7.6%	0.29%	34.2 M	59.3 M	100.0%	100.0%	59.3 M	93.5 M	93.8%	51.5%		
LIF_24hr_1_1											56.0%	32.7 M	9.1%	56.0%	31.6 M
LIF_24hr_1_2											57.0%	32.7 M	10.0%	60.0%	31.5 M
LIF_24hr_1_2	86.9%	1.40	24.6 M	6.9%	0.31%	33.8 M	49.2 M	100.0%	100.0%	49.2 M	83.0 M	93.5%	43.5%		
LIF_24hr_1_2											57.0%	27.3 M	11.9%	57.0%	26.3 M
LIF_24hr_2_1											57.0%	27.3 M	9.2%	57.0%	26.3 M
LIF_24hr_2_2											57.0%	27.3 M	11.4%	57.0%	26.3 M
LIF_24hr_3_3	94.1%	1.66	28.4 M	2.6%	0.34%	52.4 M	56.8 M	100.0%	100.0%	56.8 M	109.2 M	90.3%	39.9%		
LIF_24hr_3_3											60.0%	33.6 M	14.7%	60.0%	31.5 M
LIF_24hr_3_1											61.0%	33.6 M	10.0%	60.0%	31.5 M
LIF_24hr_3_2											61.0%	33.6 M	11.4%	60.0%	31.5 M
LIF_24hr_4_4	53.0%	1.55	42.3 M	36.5%	0.25%	20.6 M	84.5 M	100.0%	100.0%	84.5 M	105.2 M	96.8%	81.0%		
LIF_24hr_4_4											50.0%	43.9 M	7.8%	50.0%	43.7 M
LIF_24hr_4_2											51.0%	43.9 M	7.5%	50.0%	43.7 M
LIF_24hr_5_5	92.3%	1.58	20.0 M	4.0%	0.23%	27.1 M	40.1 M	100.0%	100.0%	40.1 M	67.2 M	92.3%	44.6%		
LIF_24hr_5_5											56.0%	23.4 M	16.4%	55.0%	21.7 M
LIF_24hr_5_1											57.0%	23.4 M	11.4%	55.0%	21.7 M
LIF_24hr_5_2											57.0%	23.4 M	11.4%	55.0%	21.7 M
control_0.6hr_3	77.5%	1.38	29.7 M	14.1%	0.38%	29.0 M	59.5 M	100.0%	100.0%	59.5 M	88.5 M	87.0%	52.8%		
control_1hr_3_1											59.0%	35.0 M	10.1%	59.0%	34.2 M
control_1hr_3_2											59.0%	35.0 M	8.3%	59.0%	34.2 M

Sample status checks

Reports on sample strandedness status, and any failures in trimming or mapping.

Strandedness Checks

Copy table	Configure columns	Scatter plot	Violin plot	Showing 0/50 rows and 7/7 columns.				Export as CSV
Sample	Status	Strand inference method	Provided strandedness	Inferred strandedness	Sense (%)	Antisense (%)	Unstranded (%)	
CONTROL_Ohr_1	pass	RSeQC	auto	reverse	1.5	92.5	6.1	
CONTROL_Ohr_1	pass	Salmon	auto	reverse	0.3	99.7	0.0	
CONTROL_Ohr_2	pass	RSeQC	auto	reverse	18.8	78.5	2.7	
CONTROL_Ohr_2	pass	Salmon	auto	reverse	0.2	99.8	0.0	
CONTROL_Ohr_3	pass	RSeQC	auto	reverse	5.8	89.6	4.5	
CONTROL_Ohr_3	pass	Salmon	auto	reverse	0.2	99.8	0.0	
CONTROL_Ohr_4	pass	RSeQC	auto	reverse	1.2	86.3	12.5	
CONTROL_Ohr_4	pass	Salmon	auto	reverse	0.3	99.7	0.0	
CONTROL_Ohr_5	pass	RSeQC	auto	reverse	0.9	89.8	9.4	
CONTROL_Ohr_5	pass	Salmon	auto	reverse	0.3	99.7	0.0	
CONTROL_1hr_1	pass	RSeQC	auto	reverse	0.7	98.4	0.9	
CONTROL_1hr_1	pass	Salmon	auto	reverse	0.4	99.6	0.0	
CONTROL_1hr_2	pass	RSeQC	auto	reverse	1.3	93.8	4.9	
CONTROL_1hr_2	pass	Salmon	auto	reverse	0.3	99.7	0.0	
CONTROL_1hr_4	pass	RSeQC	auto	reverse	0.4	99.4	0.2	
CONTROL_1hr_4	pass	Salmon	auto	reverse	0.4	99.6	0.0	
CONTROL_1hr_5	pass	RSeQC	auto	reverse	0.8	98.2	1.0	
CONTROL_1hr_5	pass	Salmon	auto	reverse	0.3	99.7	0.0	
CONTROL_24hr_1	pass	RSeQC	auto	reverse	0.7	92.4	6.9	
CONTROL_24hr_1	pass	Salmon	auto	reverse	0.3	99.7	0.0	
CONTROL_24hr_2	pass	RSeQC	auto	reverse	0.8	91.1	8.1	
CONTROL_24hr_2	pass	Salmon	auto	reverse	0.2	99.8	0.0	
CONTROL_24hr_3	pass	RSeQC	auto	reverse	0.6	91.1	8.3	
CONTROL_24hr_3	pass	Salmon	auto	reverse	0.2	99.8	0.0	
CONTROL_24hr_4	pass	RSeQC	auto	reverse	0.8	98.2	1.0	
CONTROL_24hr_4	pass	Salmon	auto	reverse	0.3	99.7	0.0	
CONTROL_24hr_5	pass	RSeQC	auto	reverse	6.8	88.0	5.2	
CONTROL_24hr_5	pass	Salmon	auto	reverse	0.3	99.7	0.0	
LIF_1hr_1	pass	RSeQC	auto	reverse	0.9	97.6	1.5	
LIF_1hr_1	pass	Salmon	auto	reverse	0.4	99.6	0.0	
LIF_1hr_2	pass	RSeQC	auto	reverse	5.2	90.8	4.0	
LIF_1hr_2	pass	Salmon	auto	reverse	0.2	99.8	0.0	
LIF_1hr_3	pass	RSeQC	auto	reverse	0.9	89.9	9.2	
LIF_1hr_3	pass	Salmon	auto	reverse	0.2	99.8	0.0	
LIF_1hr_4	pass	RSeQC	auto	reverse	0.8	98.0	1.2	
LIF_1hr_4	pass	Salmon	auto	reverse	0.3	99.7	0.0	
LIF_1hr_5	pass	RSeQC	auto	reverse	0.8	94.0	5.2	
LIF_1hr_5	pass	Salmon	auto	reverse	0.3	99.7	0.0	
LIF_24hr_1	pass	RSeQC	auto	reverse	0.6	90.3	9.1	
LIF_24hr_1	pass	Salmon	auto	reverse	0.3	99.7	0.0	
LIF_24hr_2	pass	RSeQC	auto	reverse	0.7	91.6	7.6	
LIF_24hr_2	pass	Salmon	auto	reverse	0.3	99.7	0.0	

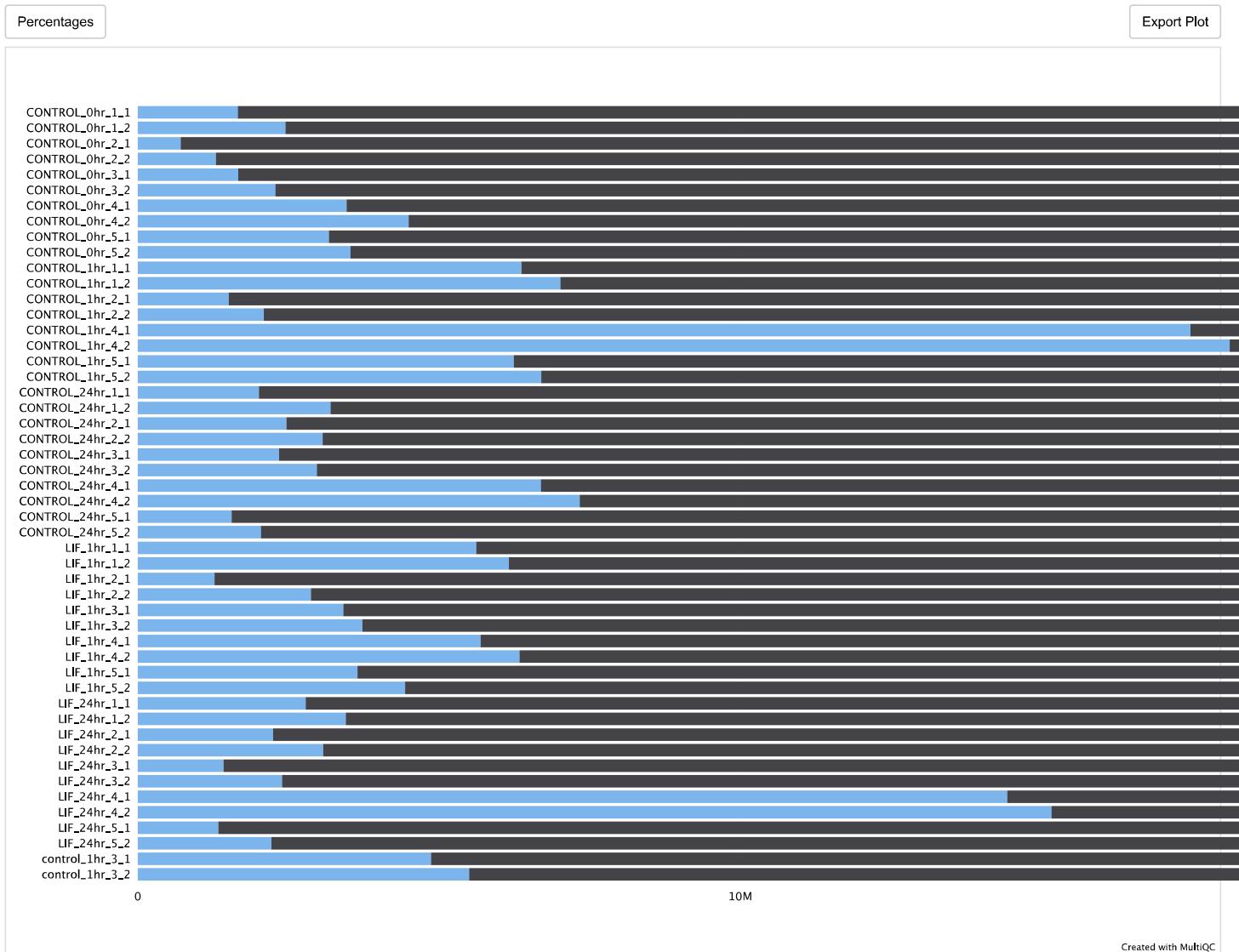
Sample	Status	Strand inference method	Provided strandedness	Inferred strandedness	Sense (%)	Antisense (%)	Unstranded (%)
LIF_24hr_3	pass	RSeQC	auto	reverse	0.9	94.4	4.7
LIF_24hr_3	pass	Salmon	auto	reverse	0.3	99.7	0.0
LIF_24hr_4	pass	RSeQC	auto	reverse	0.5	99.4	0.0
LIF_24hr_4	pass	Salmon	auto	reverse	0.4	99.6	0.0
LIF_24hr_5	pass	RSeQC	auto	reverse	3.6	92.0	4.4
LIF_24hr_5	pass	Salmon	auto	reverse	0.3	99.7	0.0
control_1hr_3	pass	RSeQC	auto	reverse	0.9	97.1	2.1
control_1hr_3	pass	Salmon	auto	reverse	0.3	99.7	0.0

FastQC (raw)

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

Sequence Counts

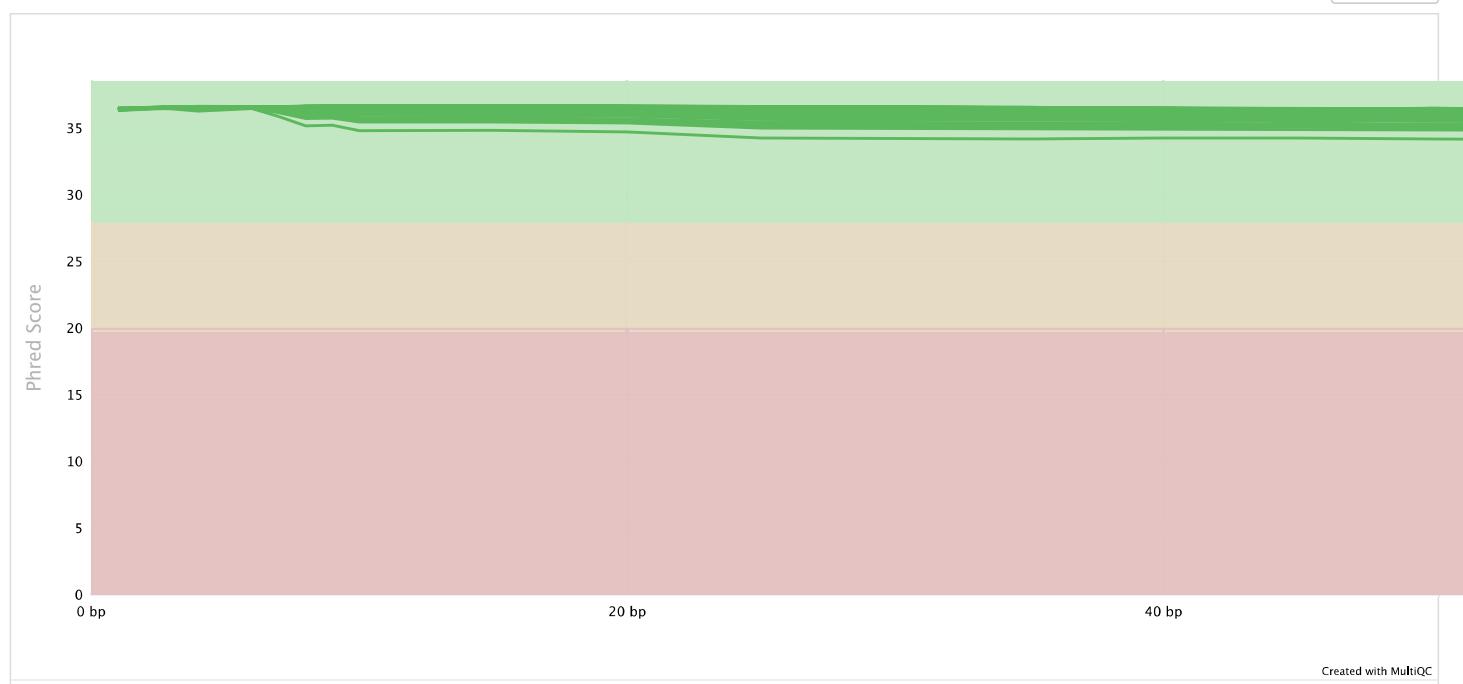
Sequence counts for each sample. Duplicate read counts are an estimate only.



Sequence Quality Histograms

50

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

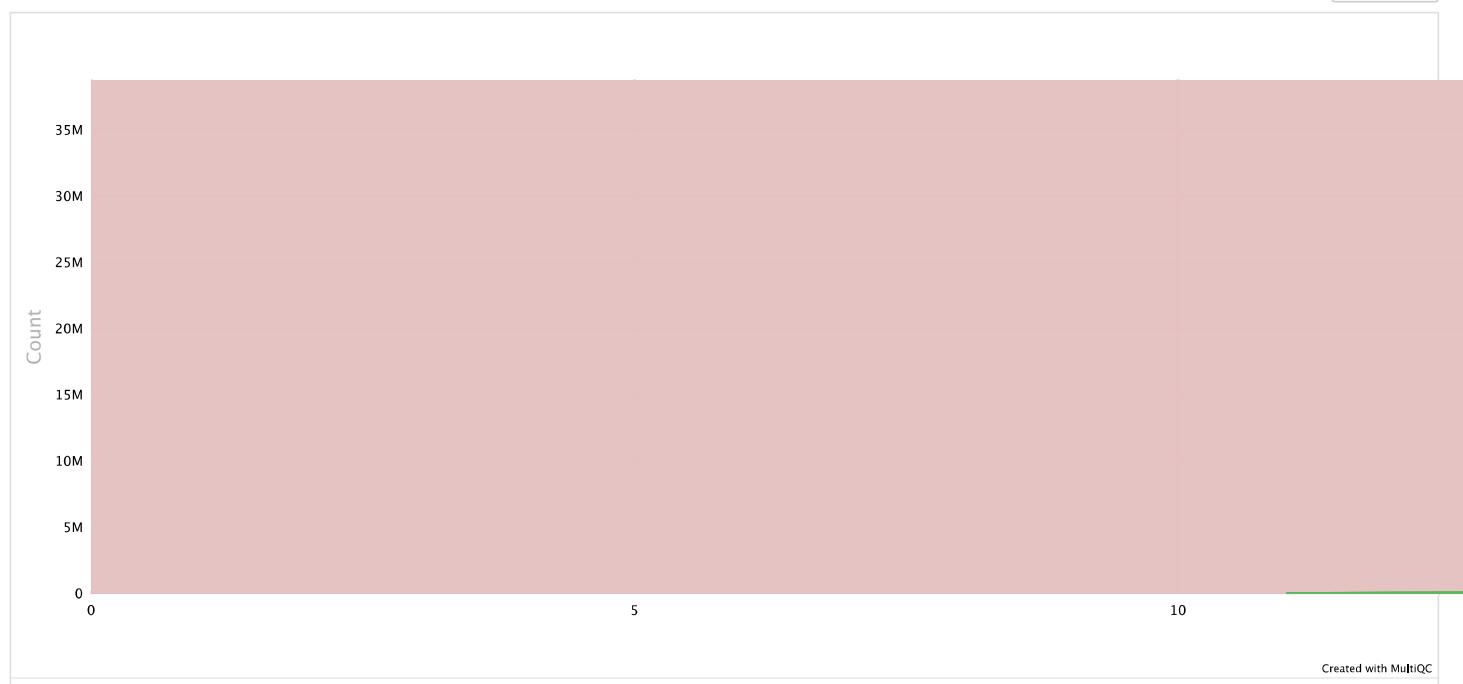
[Export Plot](#)

Created with MultiQC

Per Sequence Quality Scores

50

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

[Export Plot](#)

Created with MultiQC

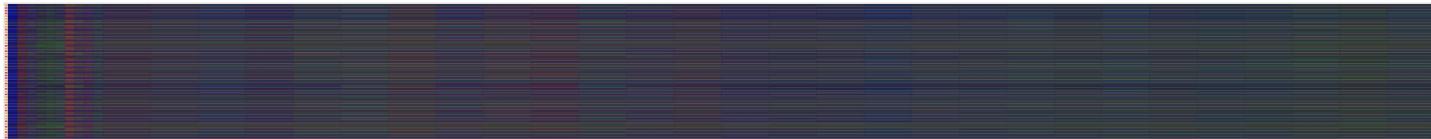
Per Base Sequence Content

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: -



29

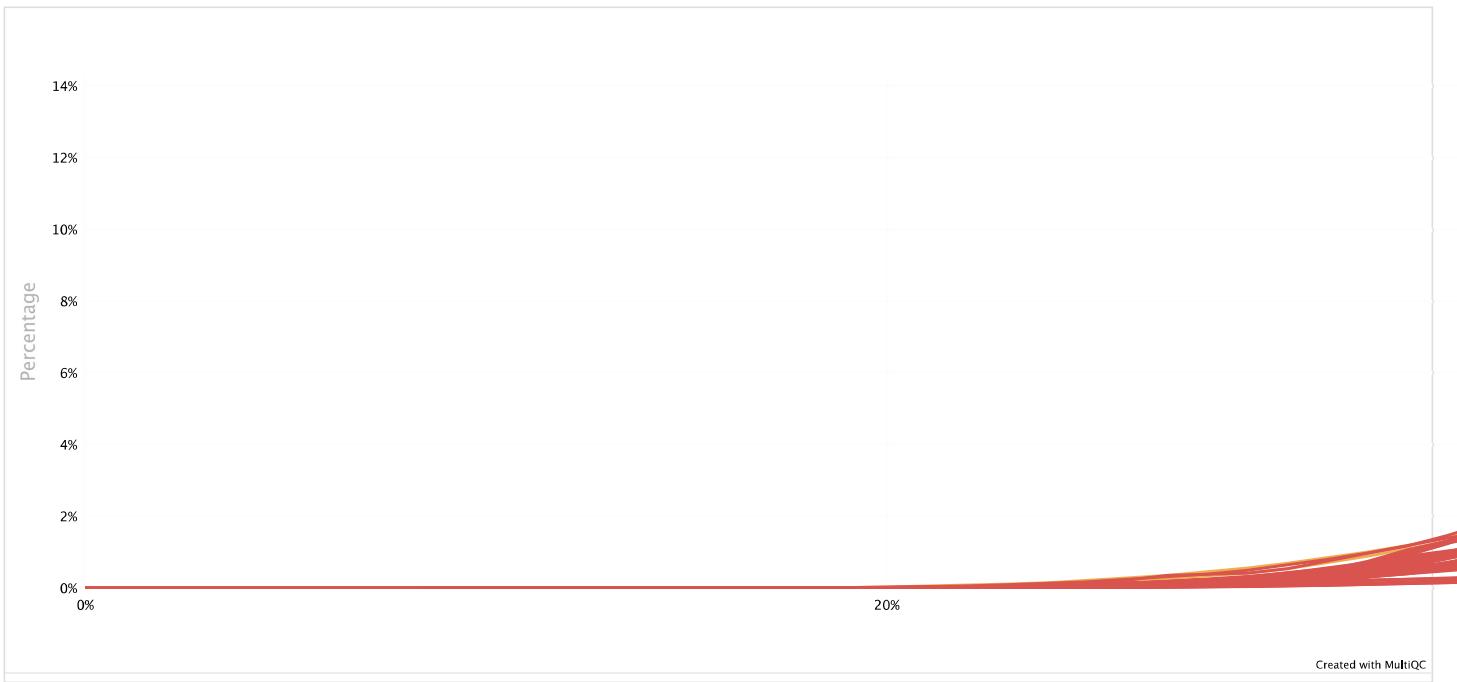
Per Sequence GC Content



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

Percentages Counts

Export Plot



Created with MultiQC

Per Base N Content

50

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

[Export Plot](#)

Sequence Length Distribution

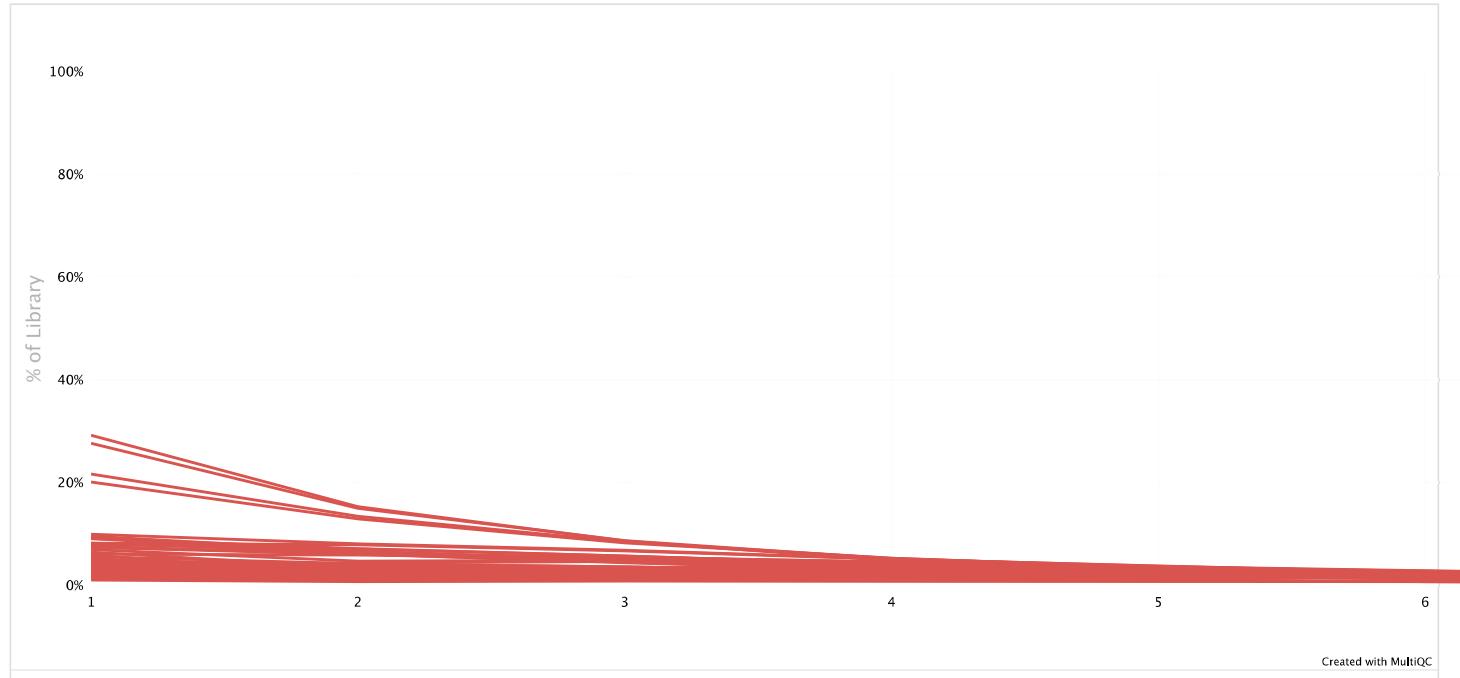
50

All samples have sequences of a single length (151bp)

[Export Plot](#)

Sequence Duplication Levels

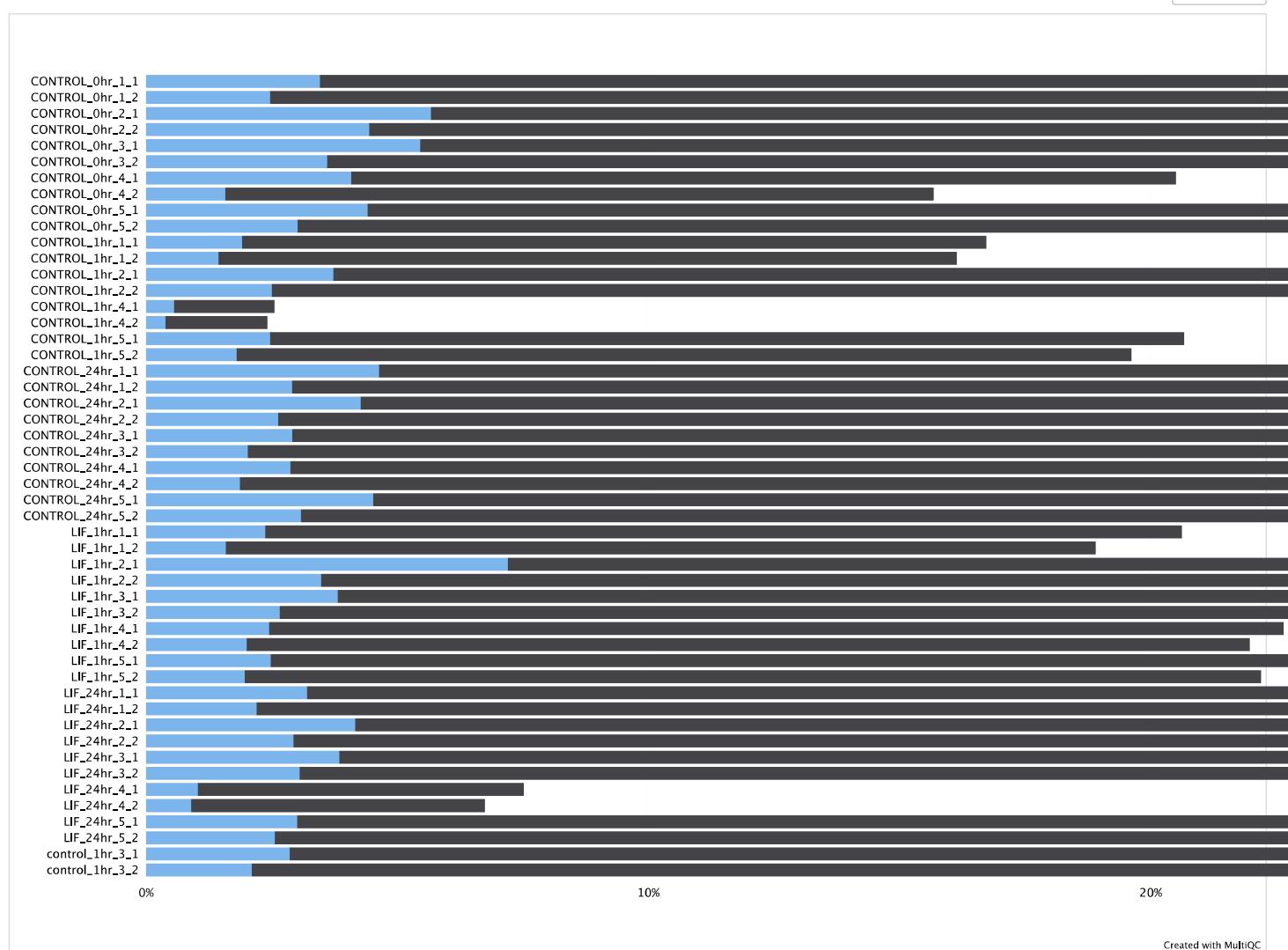
The relative level of duplication found for every sequence.

[Export Plot](#)

Overrepresented sequences by sample



The total amount of overrepresented sequences found in each library.

[Export Plot](#)

Created with MultiQC

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 0/20 rows and 3/3 columns.
[Export as CSV](#)

Overrepresented sequence	Reports	Occurrences	% of all reads
CGGTGGCGCACGCCCTGAGTCCCA	25	19 060 141	1.1154 %
CTTGAGTCAGGAGTTCTGGGCTG	25	14 246 169	0.8337 %
CCCAGCTACTCGGGAGGCTGAGAC	25	10 321 454	0.6040 %
CGCTTGAGTCCAGGAGTTCTGGGC	25	11 263 948	0.6592 %
CTGGGCTGTAGTGCCTATGCCGA	25	11 240 910	0.6578 %
CGCACGCCTGTAGTCCCAGCTACTC	25	9 323 839	0.5456 %
CAGGAGGATCGCTTGAGTCCAGGA	25	8 614 624	0.5041 %
GGCGCACGCCTGTAGTCCCAGCTA	25	9 239 622	0.5407 %
CAGCTACTCGGGAGGCTGAGACAG	25	6 088 185	0.3563 %
GGTGGCGCACGCCCTGAGTCCCAG	25	6 012 206	0.3518 %
GGAGGGATCGCTTGAGTCCAGGAGT	25	5 887 814	0.3446 %
GCGCACGCCTGTAGTCCCAGCTAC	25	4 270 854	0.2499 %
CCAGGCTGGAGTGCAGTGGCTATTG	25	26 714 366	1.5633 %
CCTTAGGCAACCTGGTGGCCCCCG	25	12 527 267	0.7331 %
CCCTCCTTAGGCAACCTGGTGGTC	25	11 989 416	0.7016 %
CCCTCCTTAGGCAACCTGGTGGT	25	11 096 651	0.6494 %
CTCGCTATGTTGCCAGGCTGGAGT	25	12 818 163	0.7501 %
CAGGCTGGAGTGCAGTGGCTATTG	25	9 767 193	0.5716 %
CTTAGGCAACCTGGTGGCCCCCG	25	8 571 197	0.5016 %
CTCCTTAGGCAACCTGGTGGTCCC	25	8 811 293	0.5156 %

Adapter Content

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



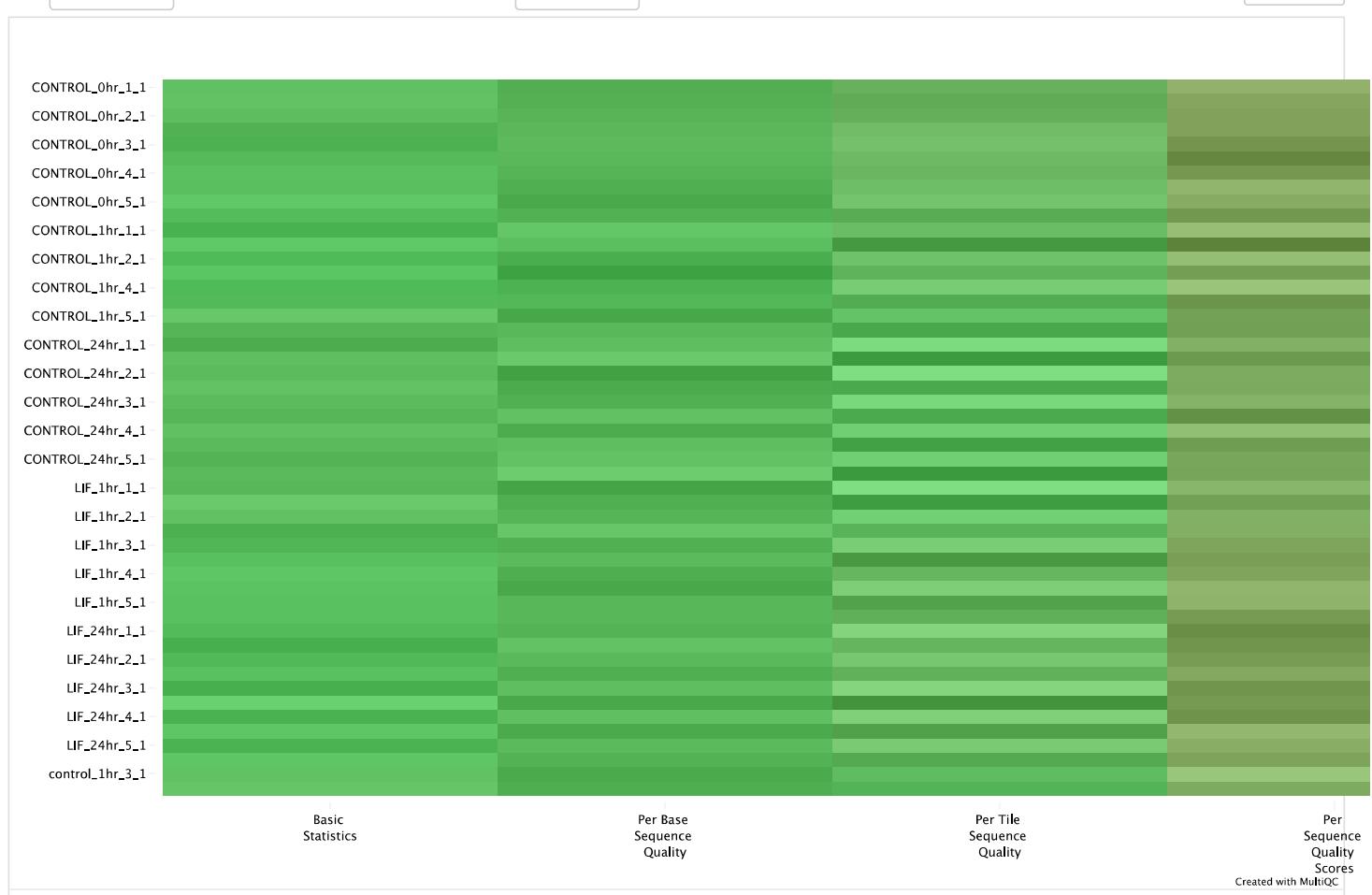
Status Checks

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

Min: 0.0



Max: 1.0

[Export Plot](#)

Cutadapt

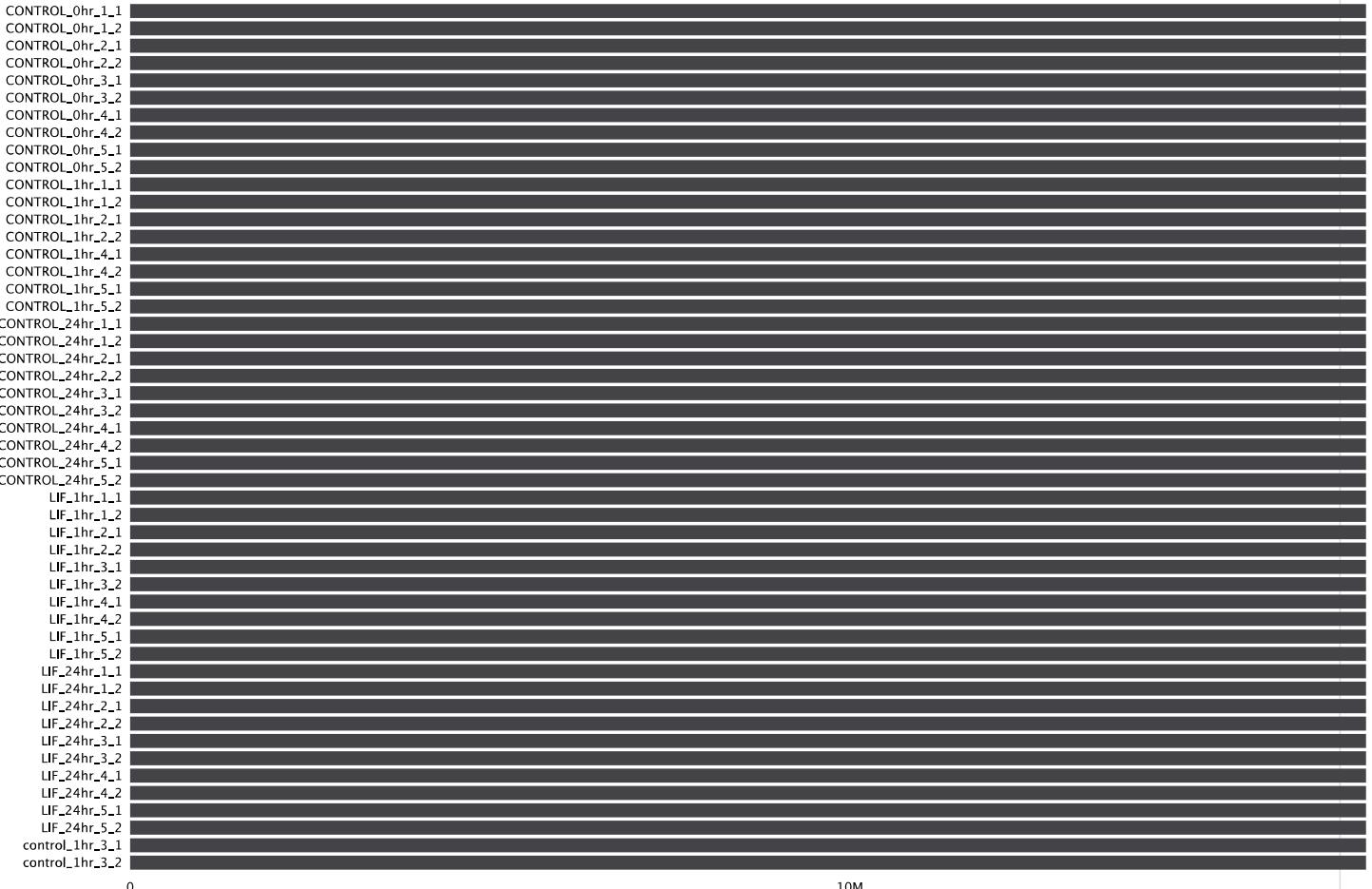
Finds and removes adapter sequences, primers, poly-A tails, and other types of unwanted sequences. URL: <https://cutadapt.readthedocs.io> DOI: 10.14806/ej.17.1.200

Filtered Reads

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

Percentages

Export Plot



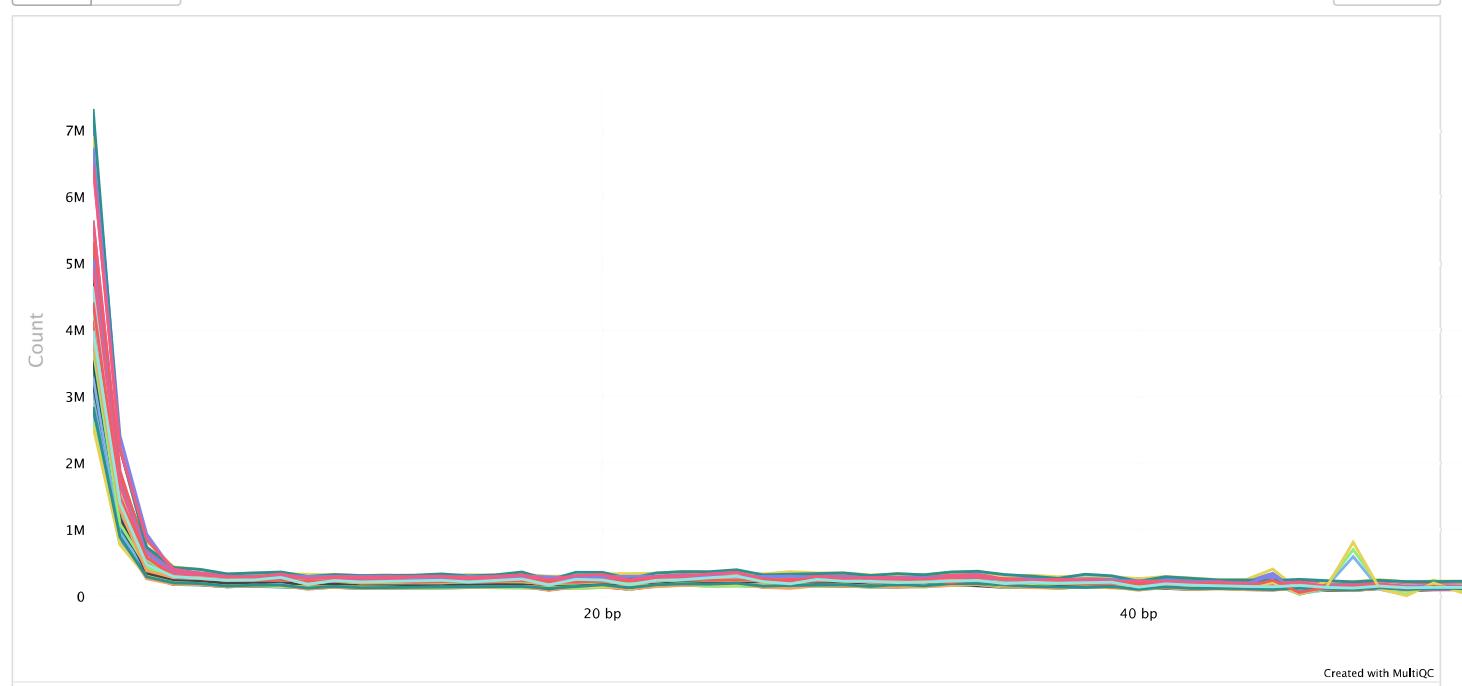
Created with MultiQC

Trimmed Sequence Lengths (3')

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

Counts Obs/Exp

Export Plot

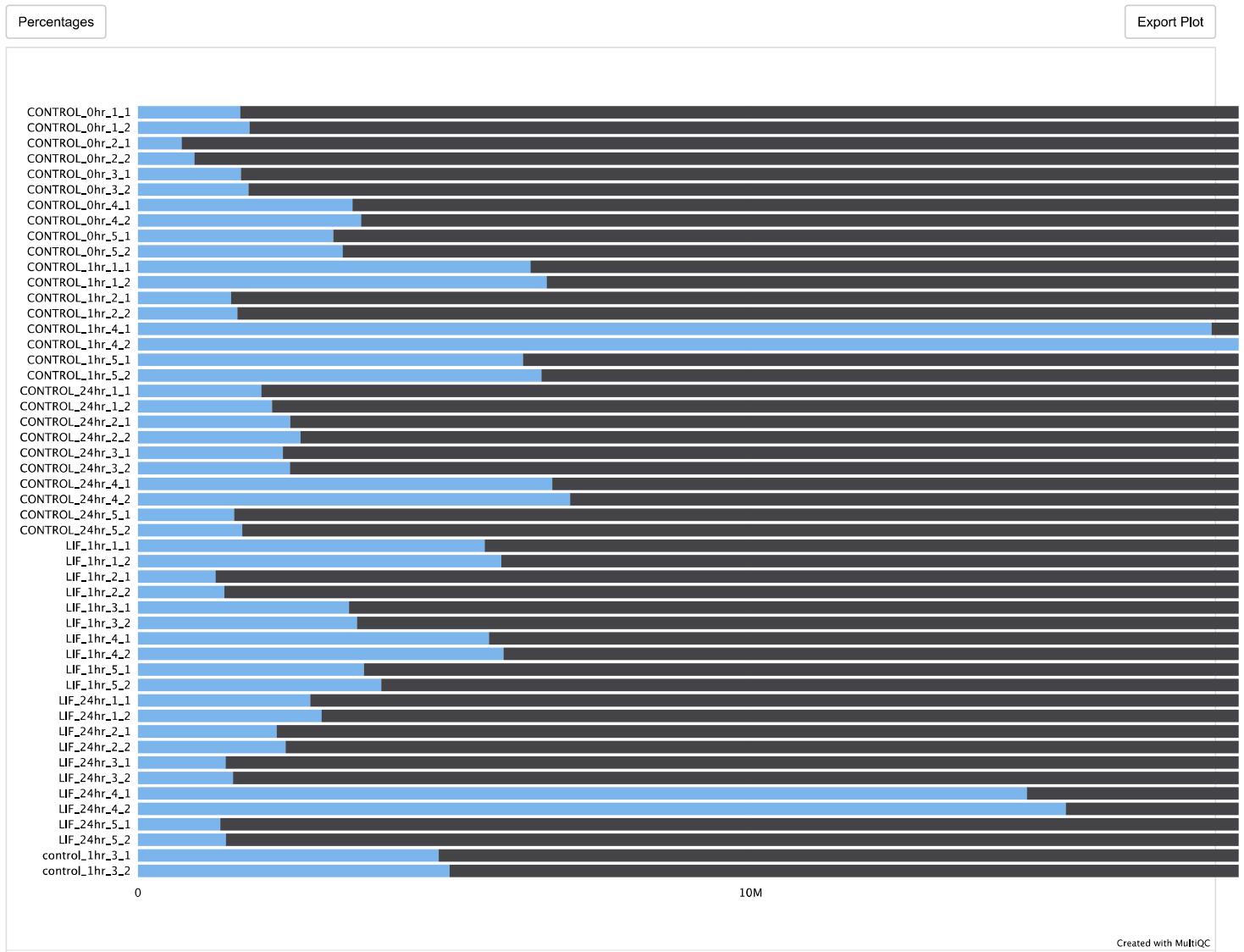


FastQC (trimmed)

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

Sequence Counts

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



Sequence Quality Histograms

50

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

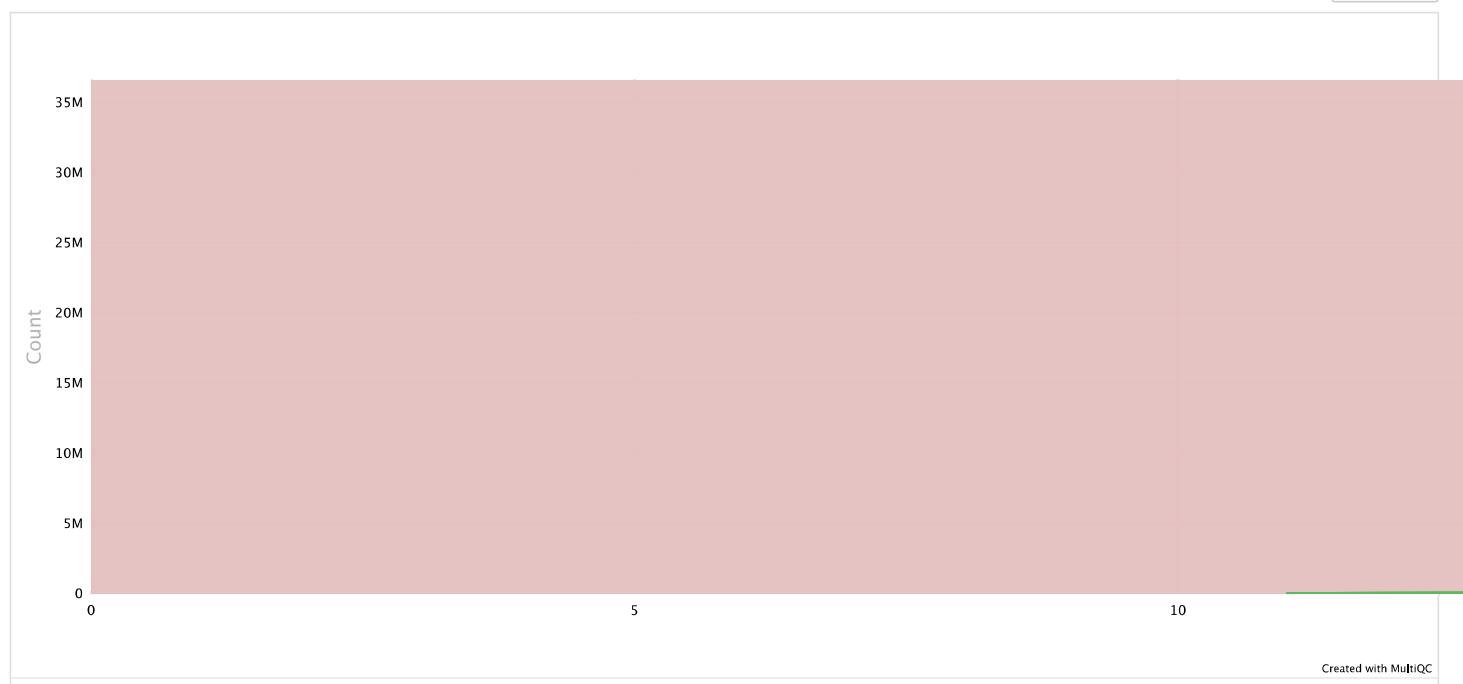
[Export Plot](#)

Created with MultiQC

Per Sequence Quality Scores

50

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

[Export Plot](#)

Created with MultiQC

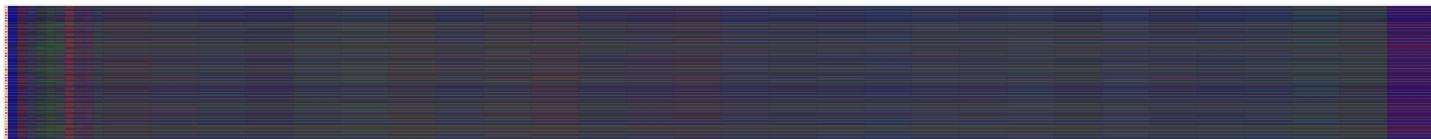
Per Base Sequence Content

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: -



24

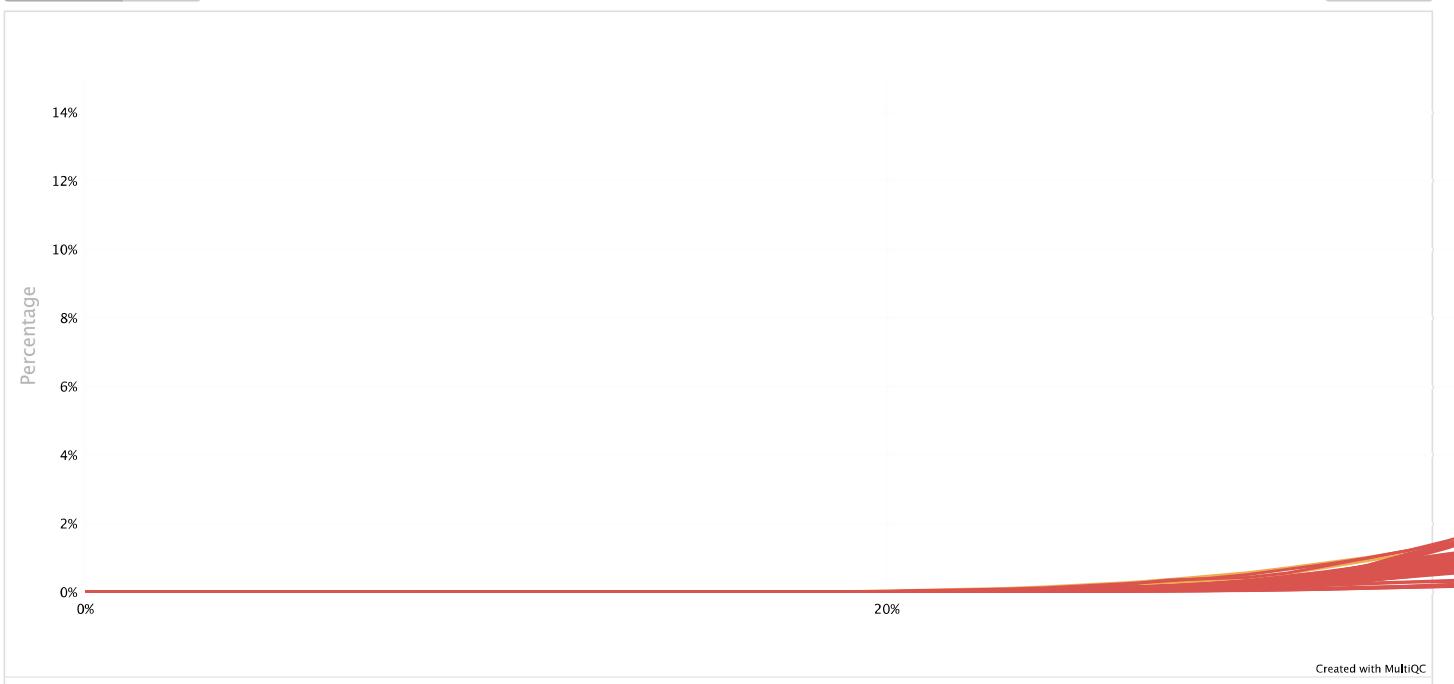
Per Sequence GC Content



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

50

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

[Export Plot](#)

Sequence Length Distribution

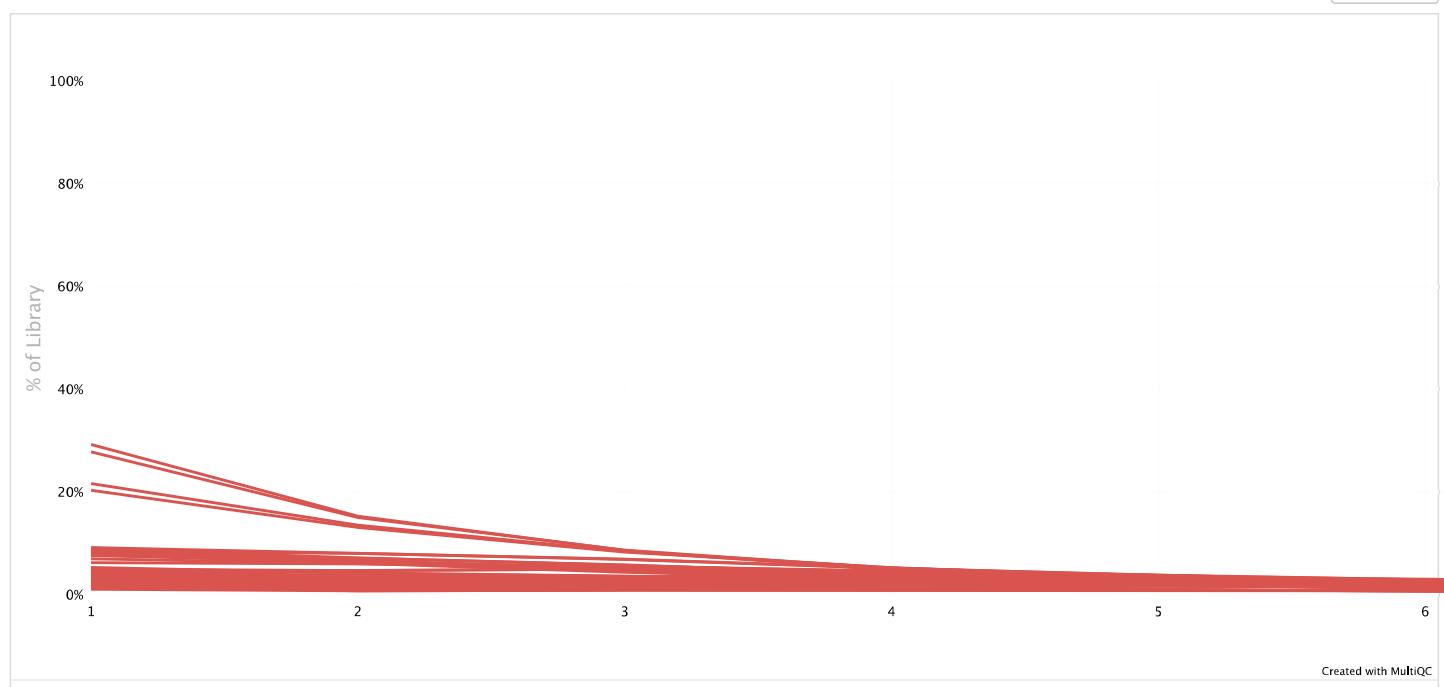
50

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

[Export Plot](#)

Sequence Duplication Levels

The relative level of duplication found for every sequence.

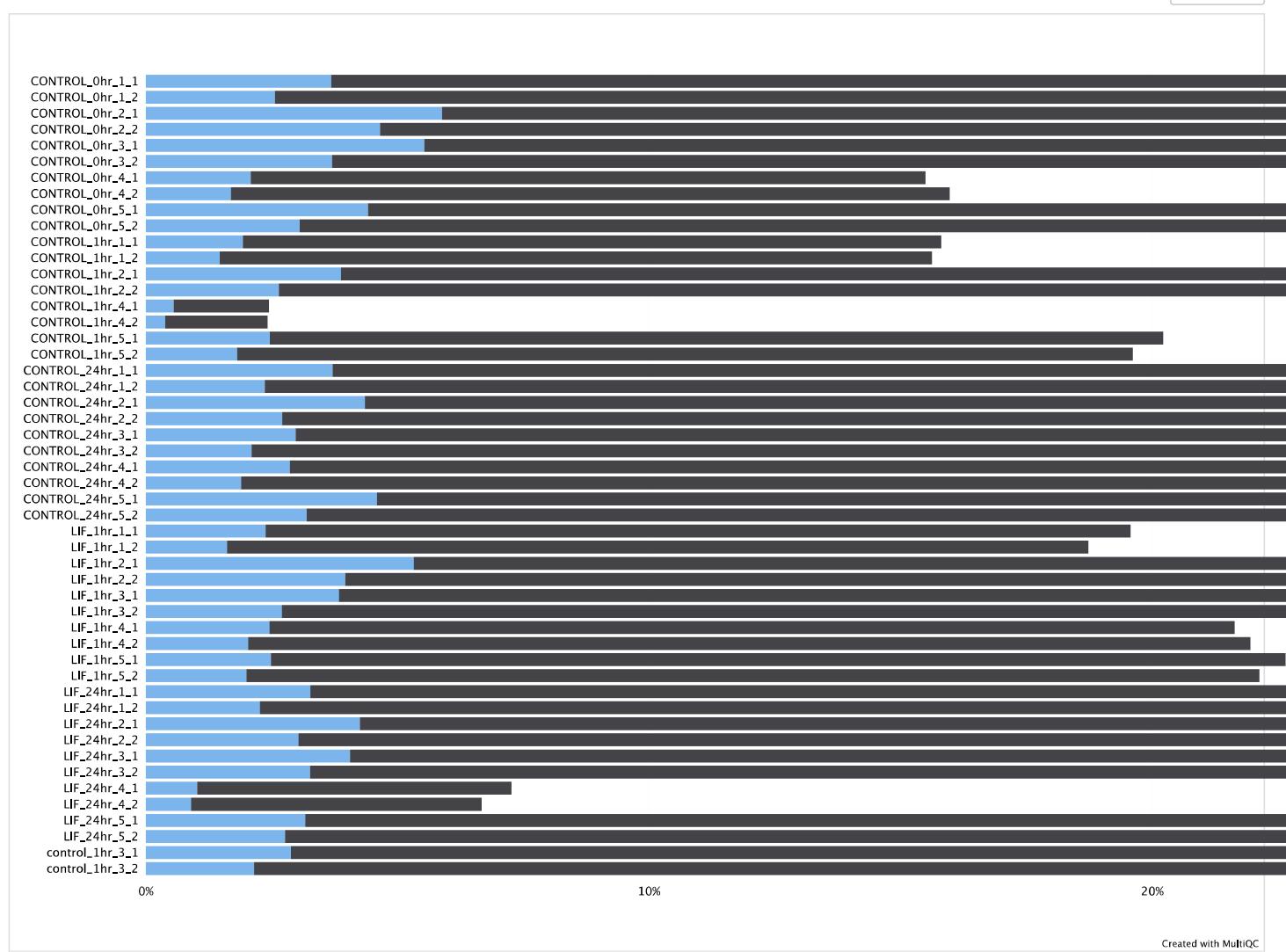
[Export Plot](#)

Created with MultiQC

Overrepresented sequences by sample

3

The total amount of overrepresented sequences found in each library.

[Export Plot](#)

Created with MultiQC

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.

Overrepresented sequence	Reports	Occurrences	% of all reads
CGGTGGCGCACGCCCTGAGTCCCACGTC	25	19 047 960	1.1550 %
CTTGAGTCCAGGAGTTCTGGGCTG	25	14 229 624	0.8628 %
CGCTTGAGTCCAGGAGTTCTGGGC	25	11 254 429	0.6824 %
GGCGCACGCCCTGAGTCCCAGCTA	25	9 235 964	0.5600 %
CCCAGCTACTCGGGAGGCTGAGAC	25	10 303 652	0.6248 %
CTGGGCTGTAGTGCCTATGCCGA	25	11 220 609	0.6804 %
CGCACGCCCTGAGTCCCAGCTACTC	25	9 319 387	0.5651 %
CAGGAGGATCGCTTGAGTCCAGGA	25	8 604 974	0.5218 %
GGTGGCGCACGCCCTGAGTCCCAG	25	6 008 363	0.3643 %
CAGCTACTCGGGAGGCTGAGACAG	25	6 078 488	0.3686 %
GGAGGATCGCTTGAGTCCAGGAGT	25	5 882 512	0.3567 %
GCGCACGCCCTGAGTCCCAGCTAC	25	4 269 404	0.2589 %
CCAGGCTGGAGTCAGTGGCTATTG	25	26 332 527	1.5967 %
CTCGCTATGTTGCCAGGCTGGAGT	25	12 758 021	0.7736 %
CCCTCCTTAGGCAACCTGGTGGTC	25	11 965 414	0.7255 %
CCTTAGGCAACCTGGTGGCCCCC	25	12 490 774	0.7574 %
CCCCTCCTTAGGCAACCTGGTGGTC	25	11 071 041	0.6713 %
CAGGCTGGAGTCAGTGGCTATTG	25	9 632 049	0.5841 %
CTCCTTAGGCAACCTGGTGGCCCC	25	8 785 042	0.5327 %
CTCCGTTCCGACCTGGGCCGGTT	25	8 602 973	0.5217 %

Adapter Content

50

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%

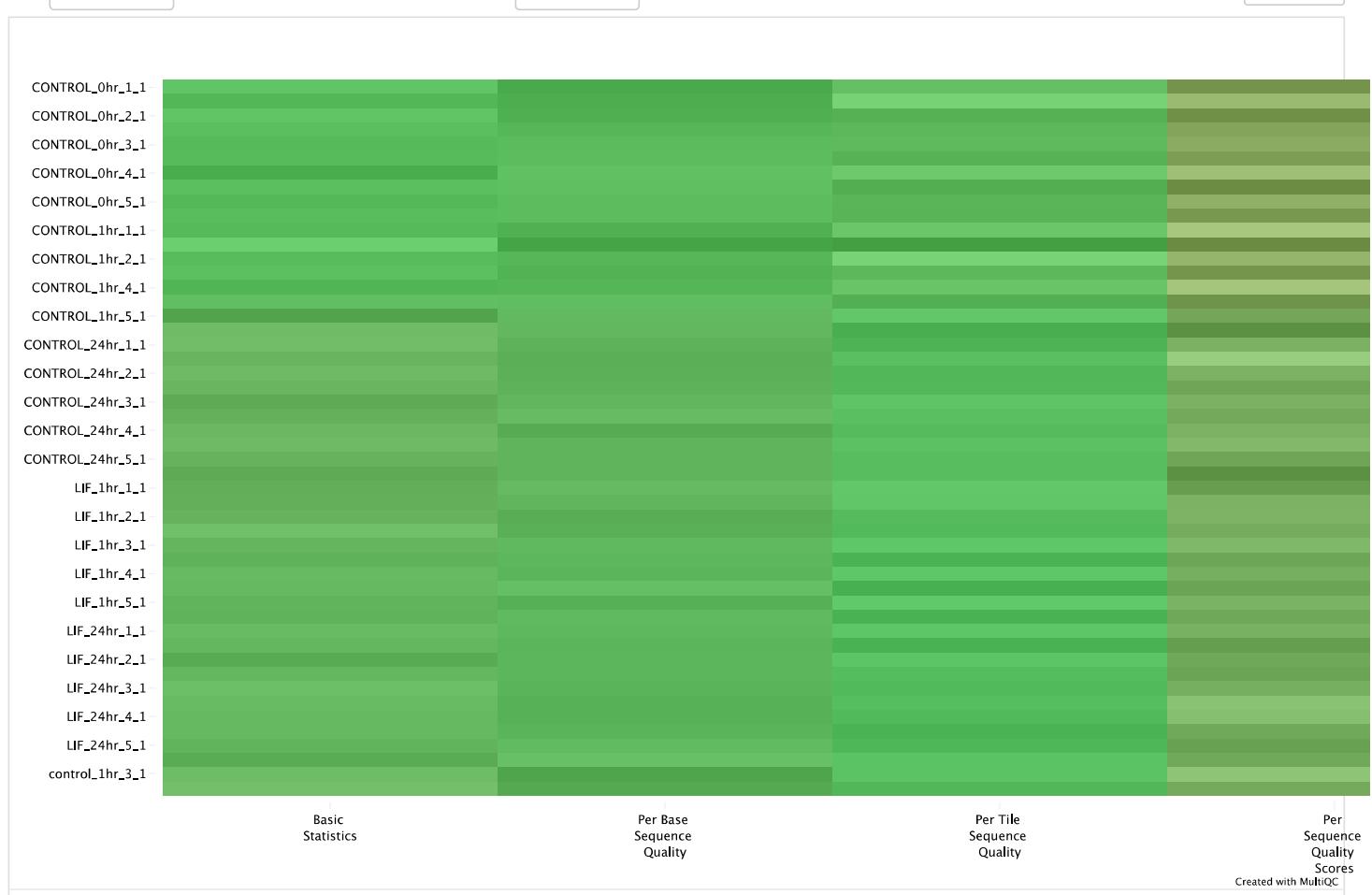
Status Checks

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

Min: 0.0

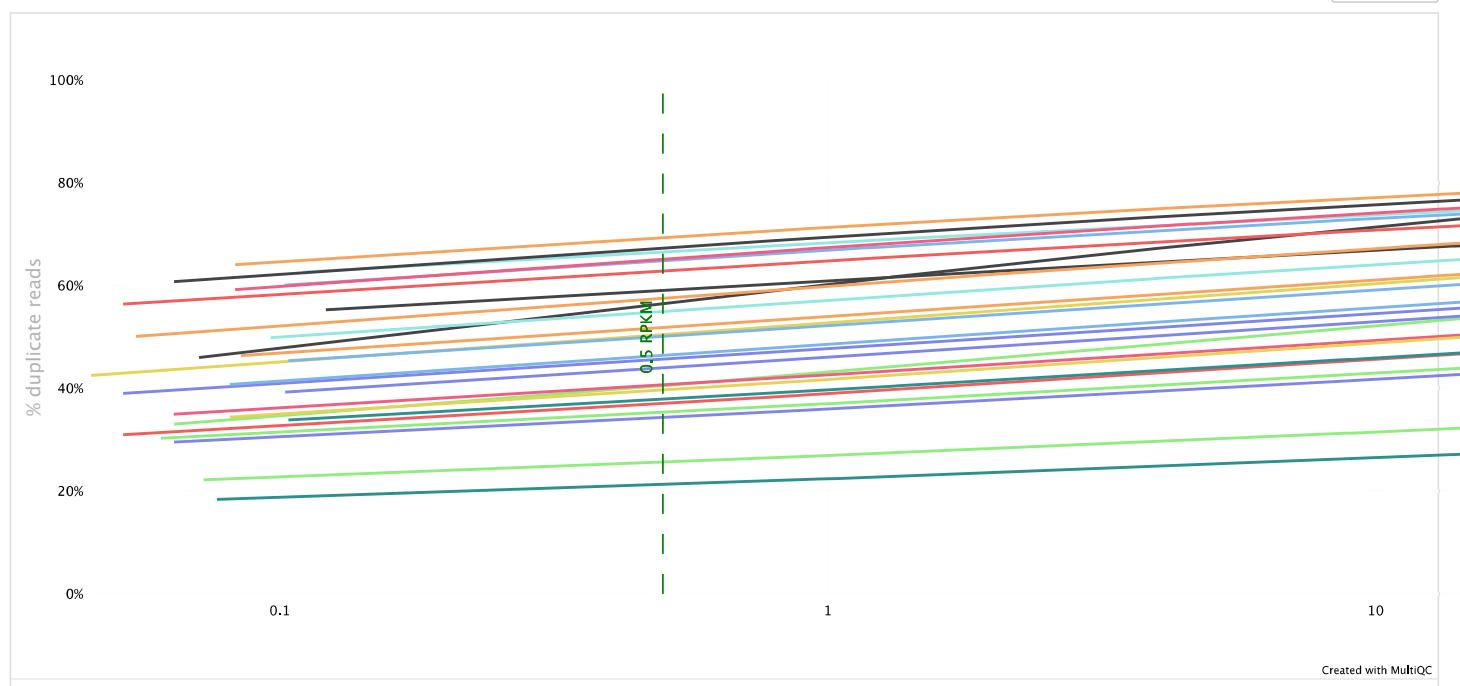


Max: 1.0

[Export Plot](#)

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. URL: bioconductor.org/packages/release/bioc/html/dupRadar.html

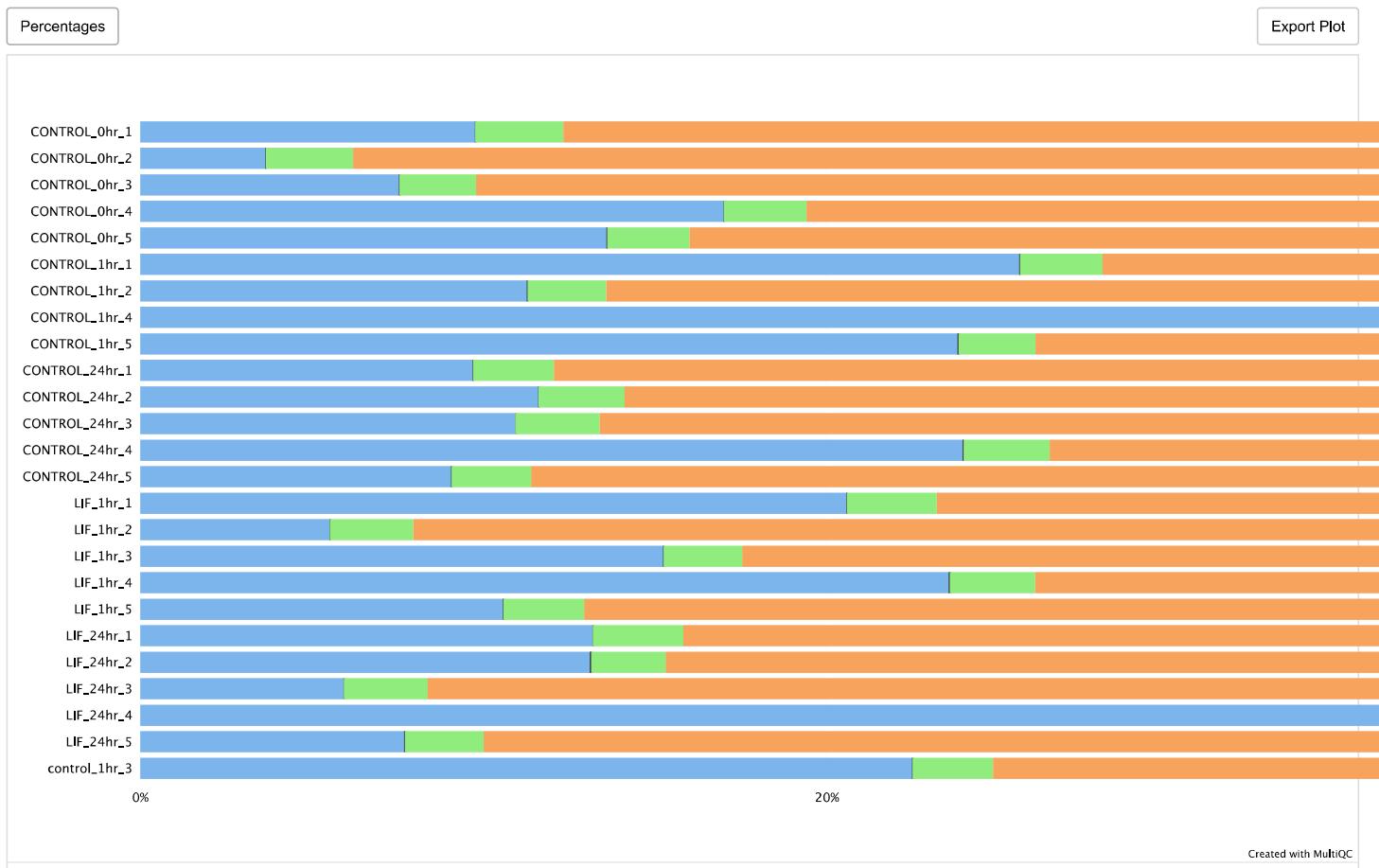
[Export Plot](#)

Picard

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

Mark Duplicates

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

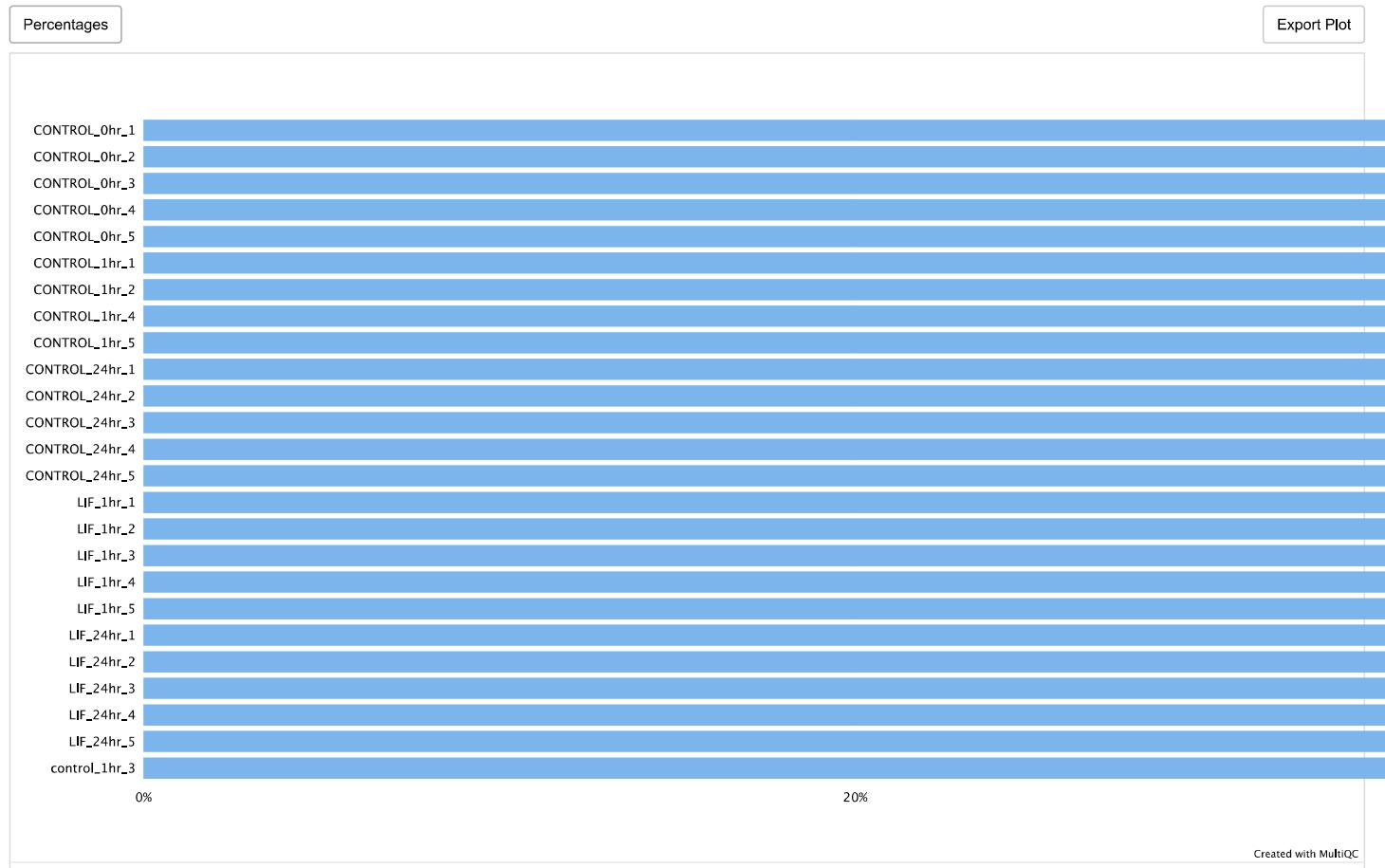


QualiMap

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

Genomic origin of reads

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.

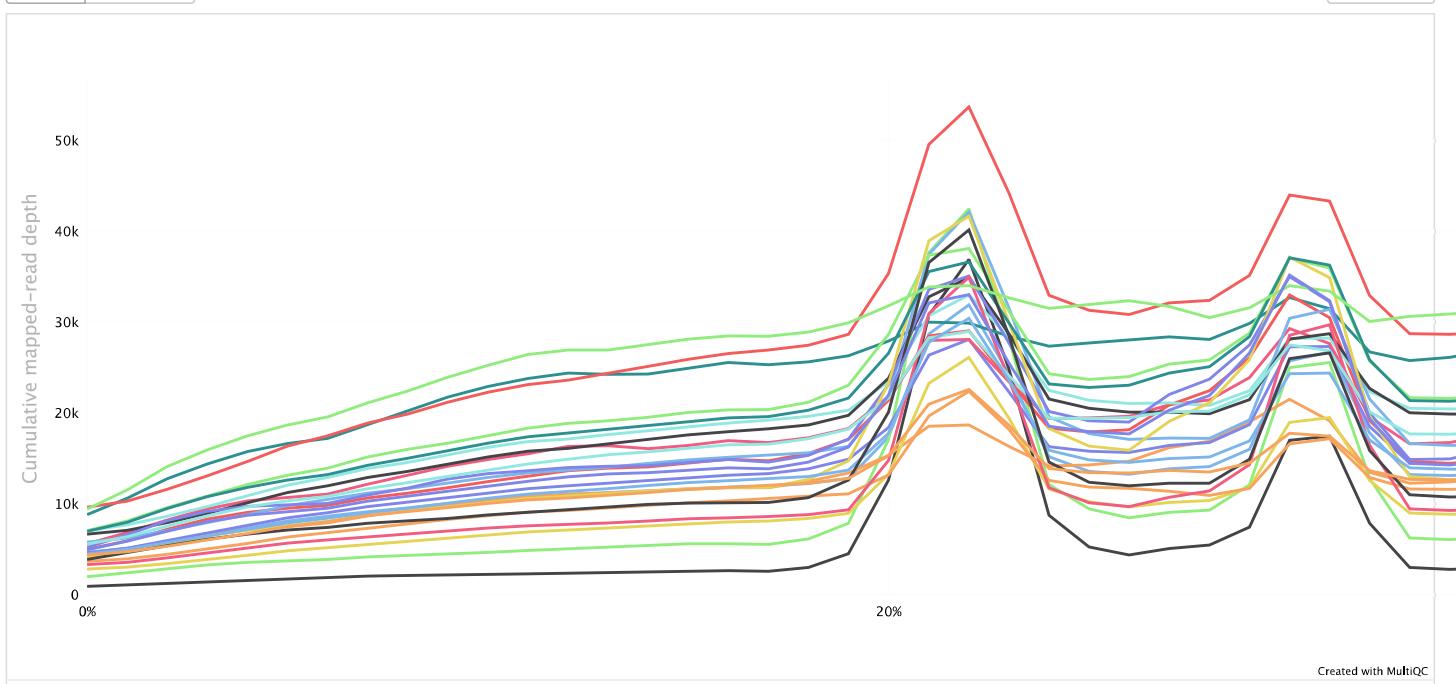


Gene Coverage Profile

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

Counts Normalised

Export Plot



Created with MultiQC

RSeQC

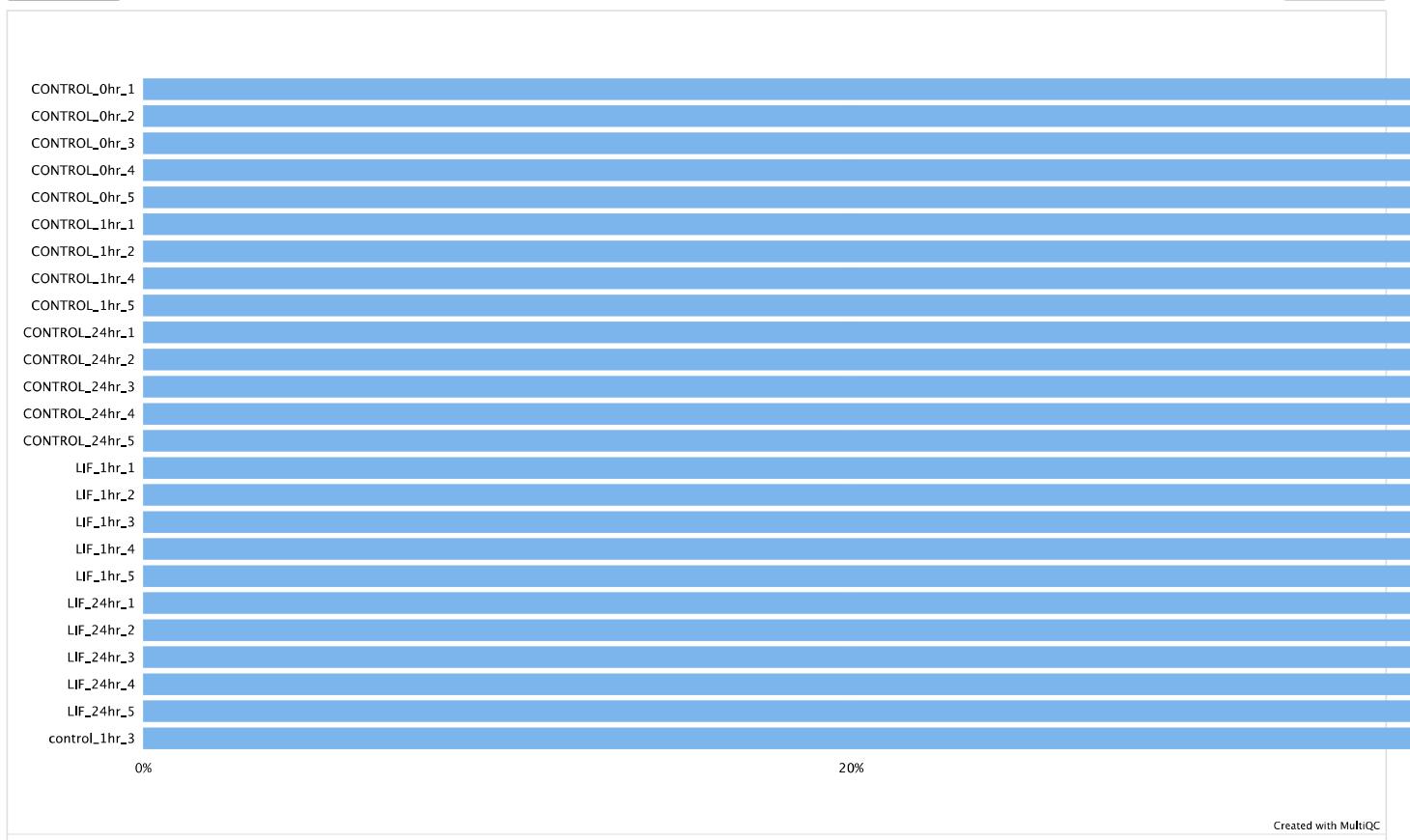
Evaluates high throughput RNA-seq data. URL: <http://rseqc.sourceforge.net> DOI: 10.1093/bioinformatics/bts356

Read Distribution

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

Percentages

Export Plot



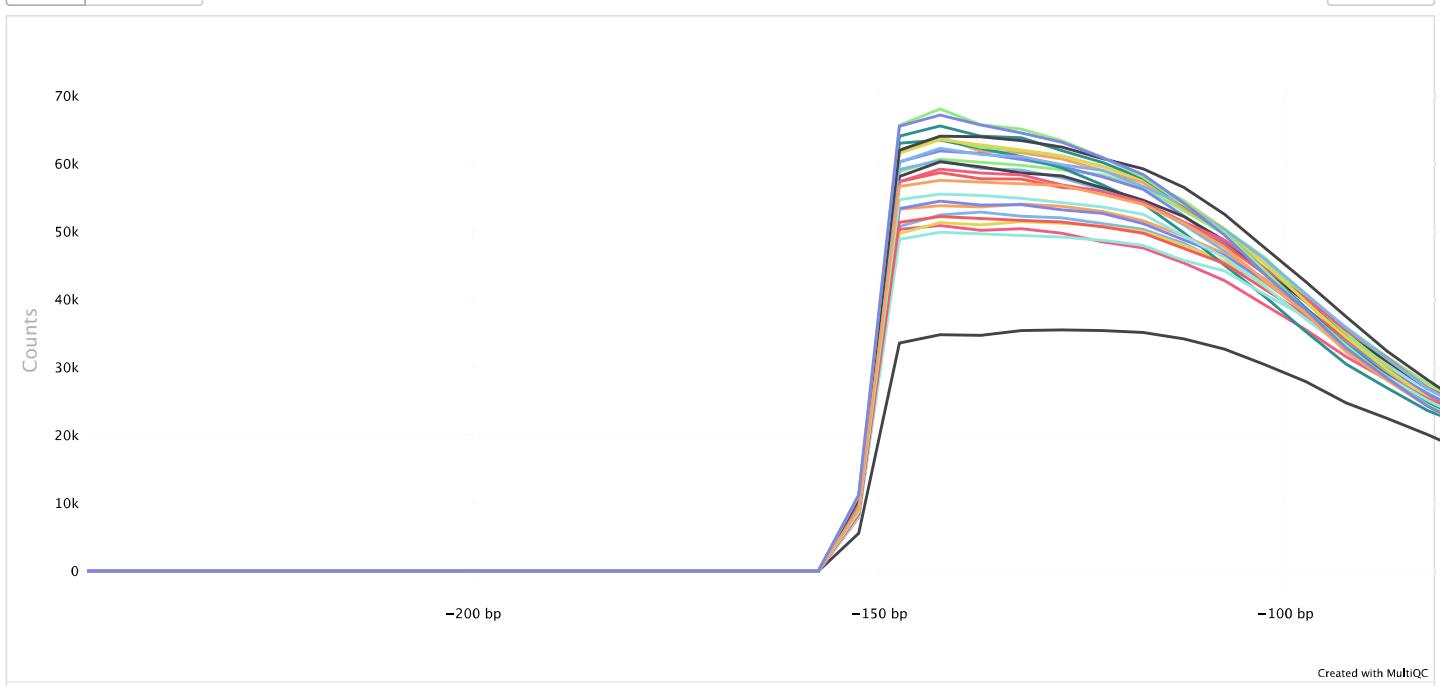
Created with MultiQC

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

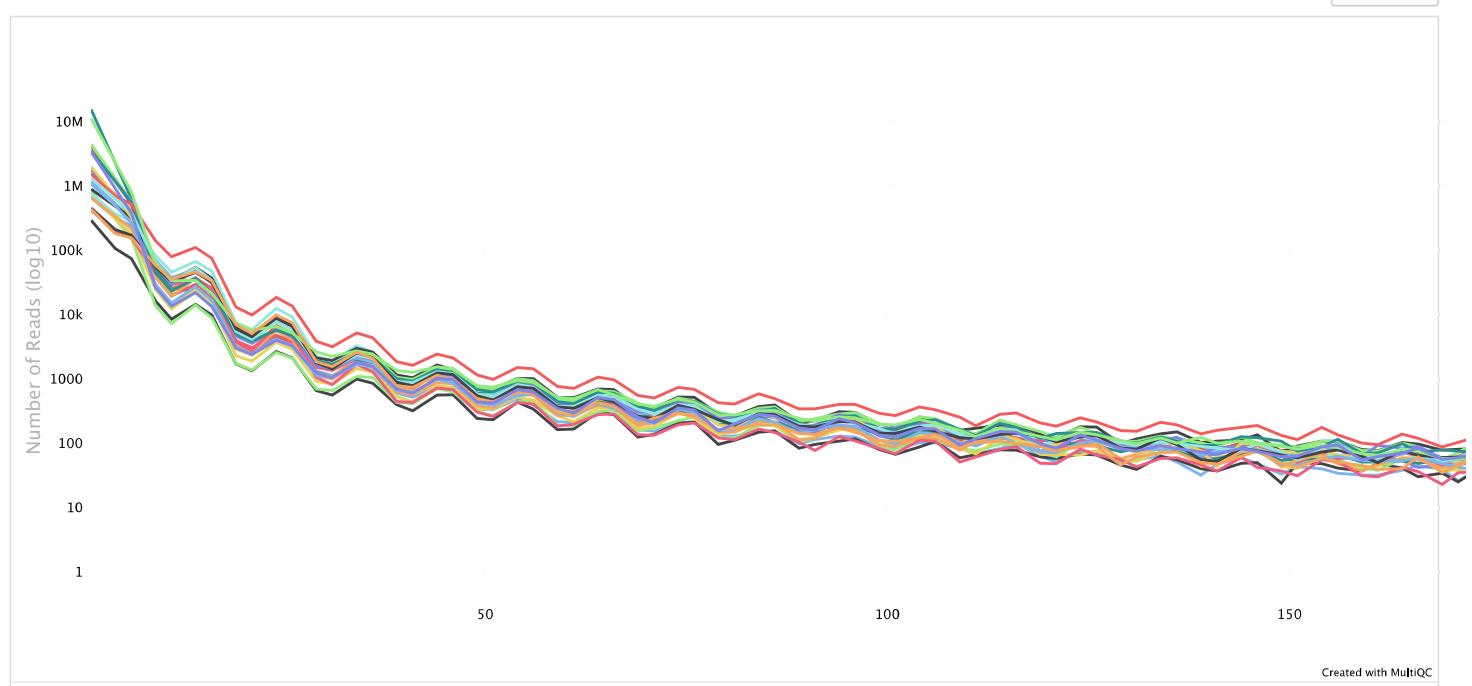
[Export Plot](#)



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.

[Export Plot](#)



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

Export Plot

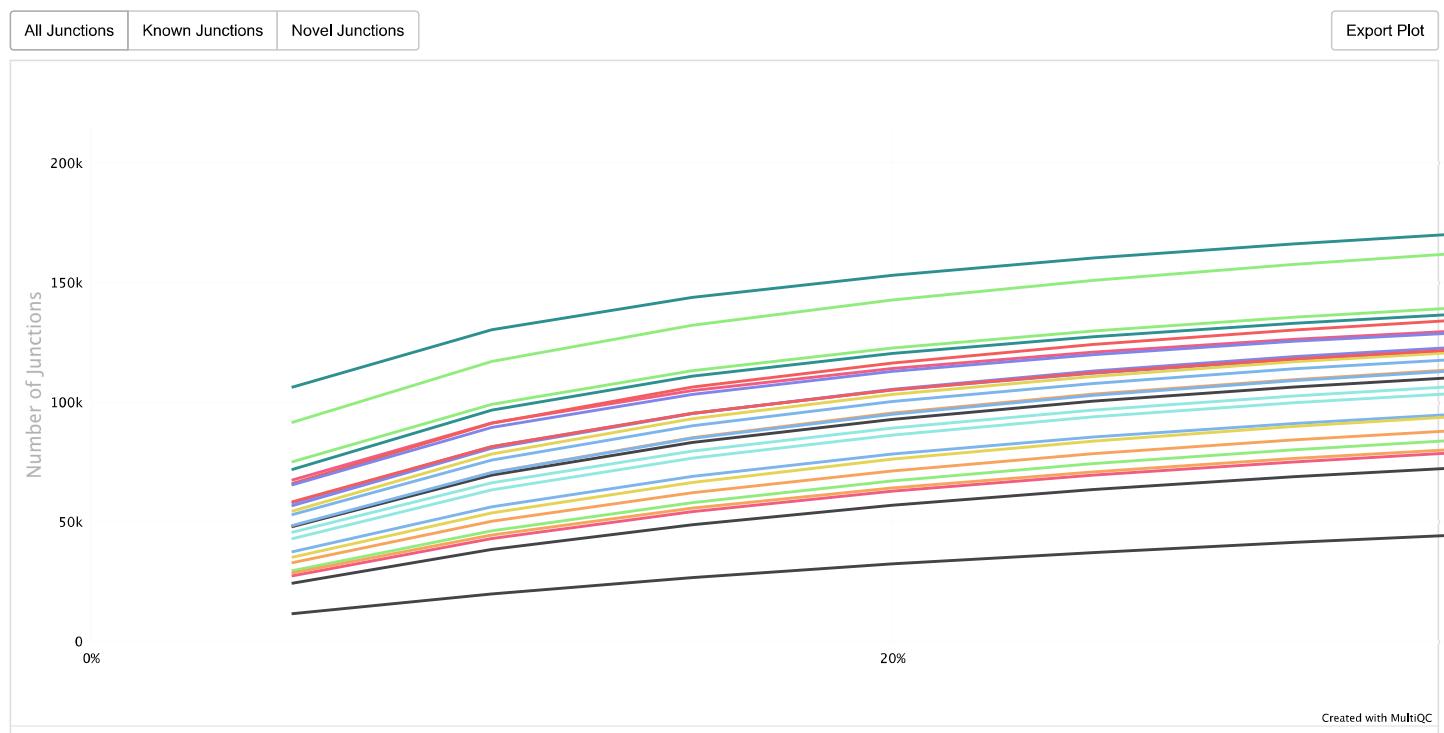


Created with MultiQC

Junction Saturation

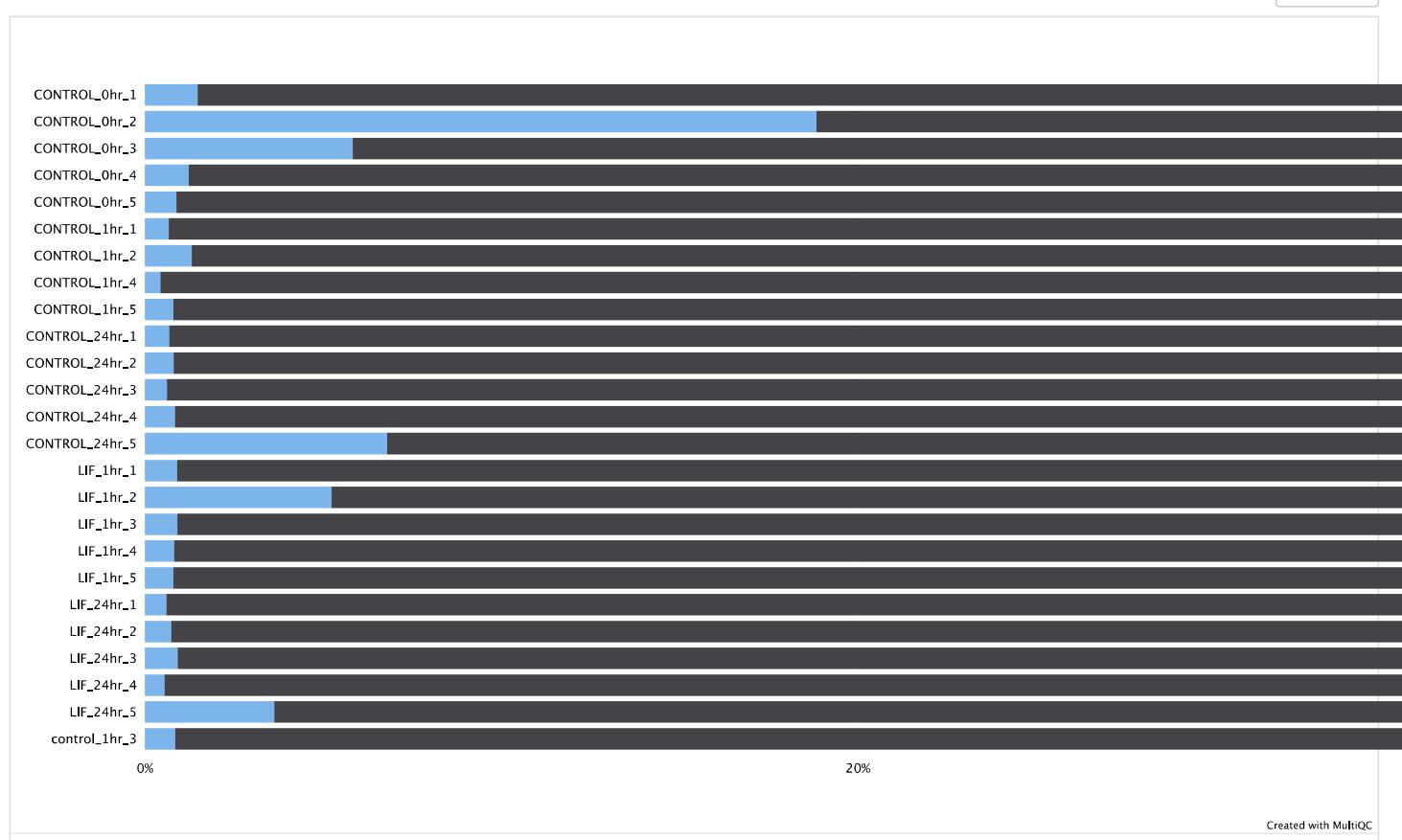
Junction Saturation counts the number of known splicing junctions that are observed in each dataset. If sequencing depth is sufficient, all (annotated) splice junctions should be rediscovered, resulting in a curve that reaches a plateau. Missing low abundance splice junctions can affect downstream analysis.

👉 Click a line to see the data side by side (as in the original RSeQC plot).



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).

[Export Plot](#)

Created with MultiQC

Bam Stat

All numbers reported in millions.



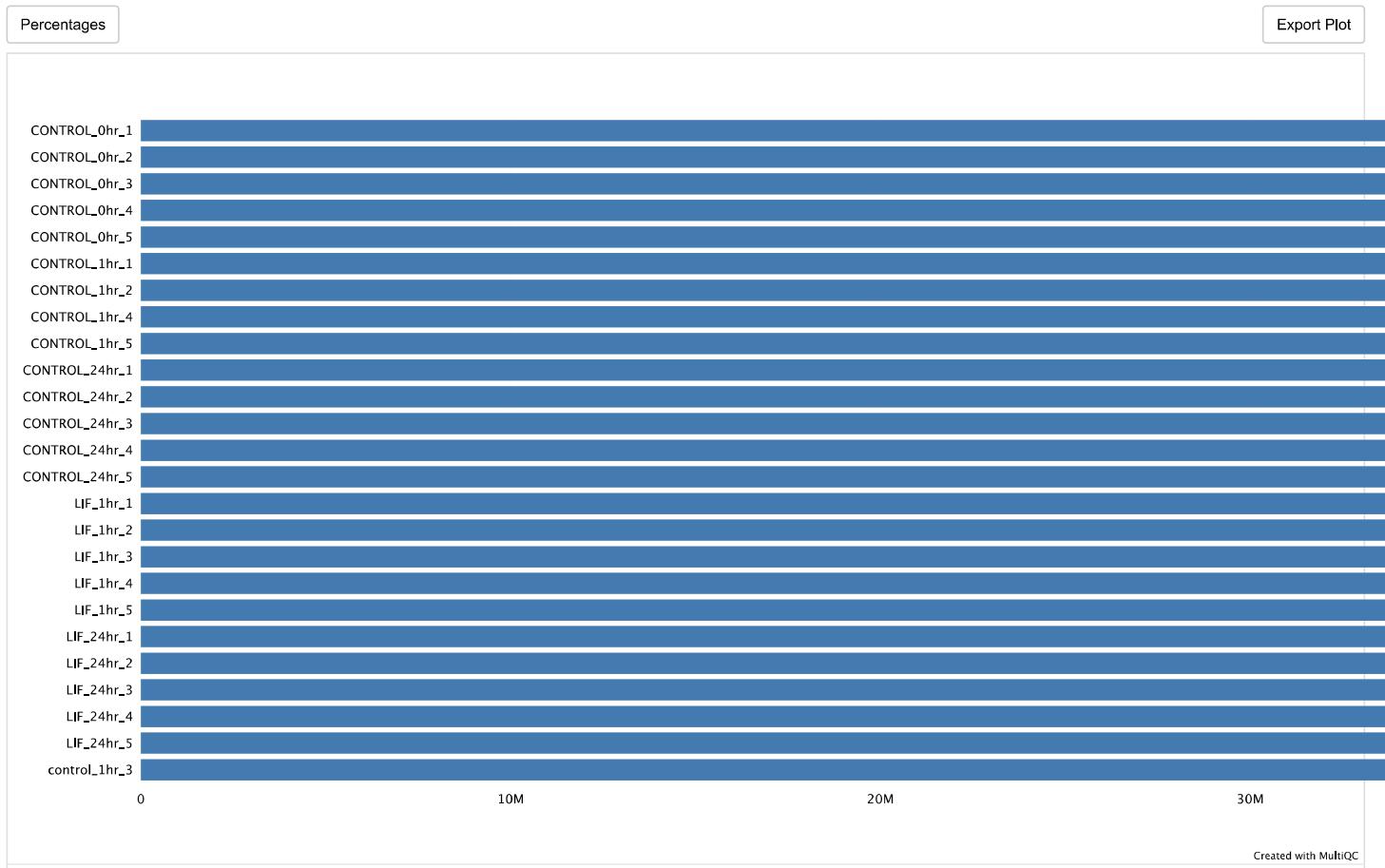
Created with MultiQC

Samtools

Toolkit for interacting with BAM/CRAM files. URL: <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.



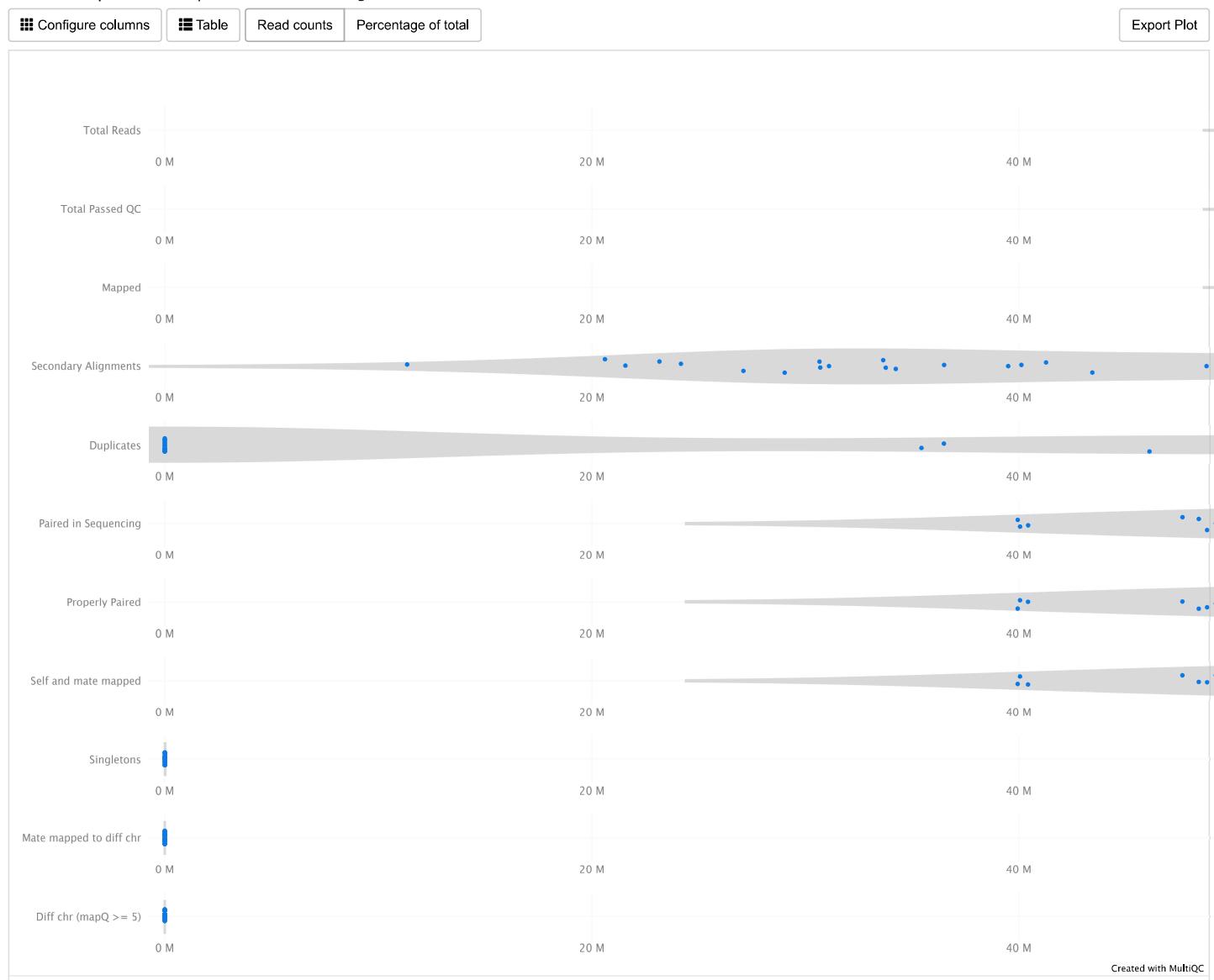
Alignment stats

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

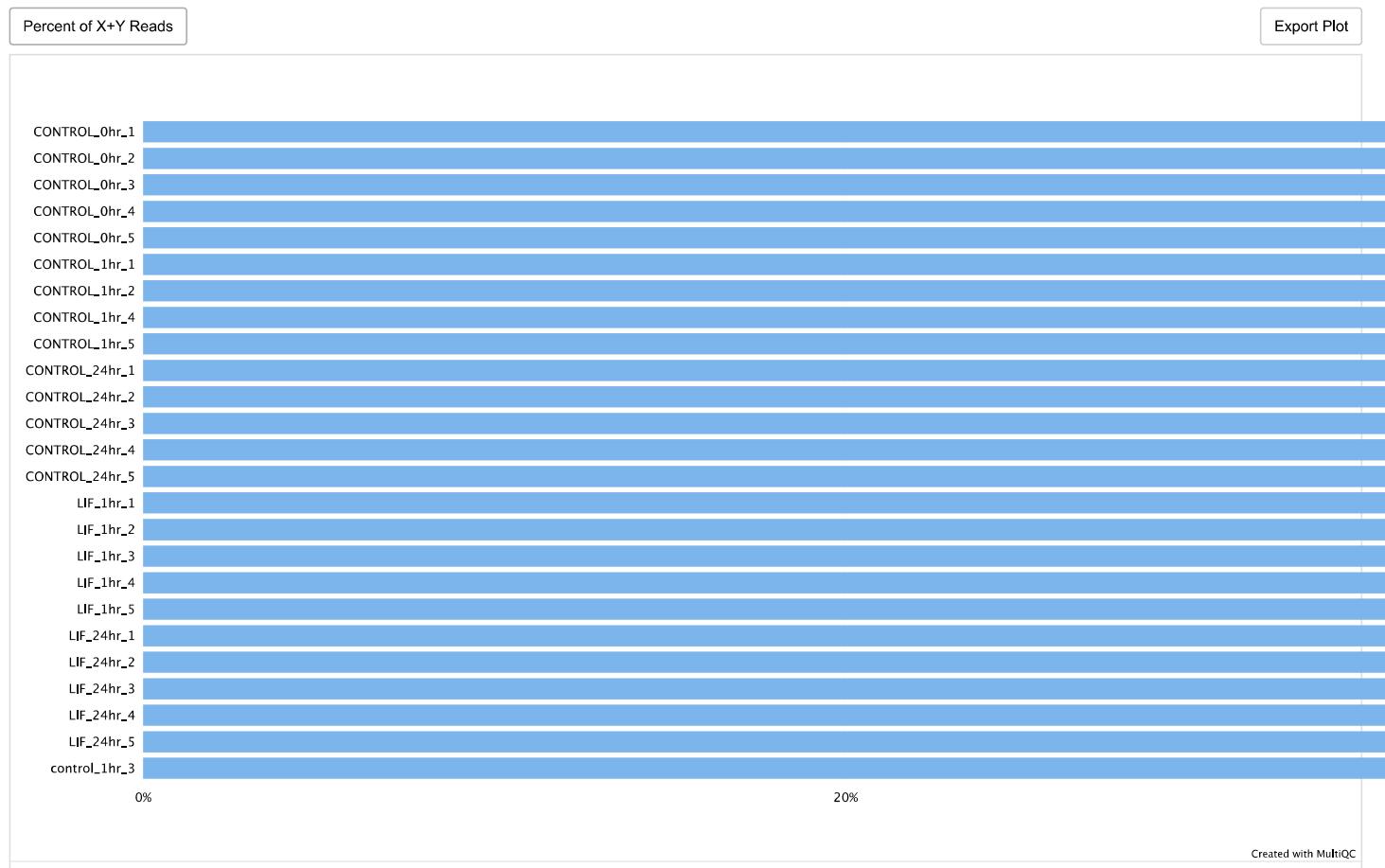


Flagstat

This module parses the output from `samtools flagstat`

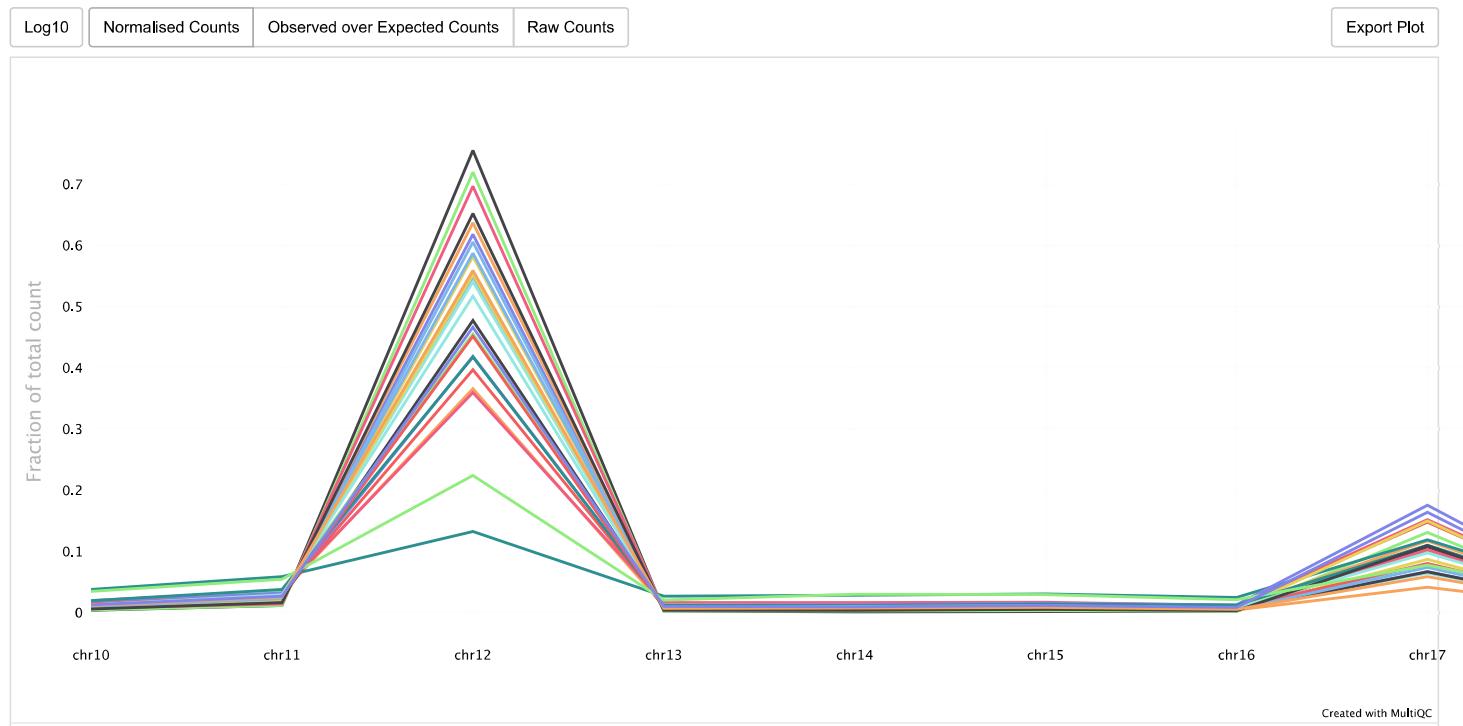


XY counts



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. URL: <https://github.com/alexdobin/STAR> DOI: 10.1093/bioinformatics/bts635

Summary Statistics

Summary statistics from the STAR alignment

Sample Name	Total reads	Aligned	Uniq aligned	Avg-mapped len	Annotated splices	Mismatch rate	Del rate	Del len	Ins rate	Ins len
CONTROL_0	22.2 M	91.3 %	48.8 %	278.1 bp	7.3 M	0.4 %	0.0 %	1.4 bp	0.0 %	3.3 bp
CONTROL_0	25.9 M	92.0 %	26.0 %	275.2 bp	2.3 M	0.8 %	0.1 %	1.4 bp	0.0 %	2.6 bp
CONTROL_0	31.3 M	91.0 %	32.3 %	275.4 bp	4.3 M	0.7 %	0.1 %	1.5 bp	0.0 %	2.3 bp
CONTROL_0	29.2 M	89.0 %	59.2 %	277.6 bp	6.9 M	0.3 %	0.0 %	2.0 bp	0.0 %	1.8 bp
CONTROL_0	35.0 M	93.0 %	41.9 %	277.8 bp	8.4 M	0.3 %	0.0 %	1.5 bp	0.0 %	1.9 bp
CONTROL_1	35.2 M	90.9 %	65.5 %	281.6 bp	9.6 M	0.3 %	0.1 %	2.3 bp	0.0 %	2.1 bp
CONTROL_1	21.5 M	92.9 %	42.7 %	275.7 bp	5.2 M	0.4 %	0.0 %	1.3 bp	0.0 %	2.2 bp
CONTROL_1	42.4 M	93.4 %	84.0 %	283.3 bp	15.1 M	0.2 %	0.0 %	2.5 bp	0.0 %	2.1 bp
CONTROL_1	38.8 M	90.0 %	59.5 %	279.9 bp	8.9 M	0.4 %	0.0 %	1.8 bp	0.0 %	1.6 bp
CONTROL_2	28.8 M	94.0 %	49.2 %	277.8 bp	10.5 M	0.3 %	0.0 %	1.4 bp	0.0 %	2.0 bp
CONTROL_2	34.1 M	92.4 %	41.9 %	277.1 bp	9.7 M	0.4 %	0.0 %	1.3 bp	0.0 %	2.6 bp
CONTROL_2	33.6 M	93.4 %	50.0 %	278.4 bp	10.8 M	0.3 %	0.0 %	1.5 bp	0.0 %	1.7 bp
CONTROL_2	42.9 M	93.1 %	56.1 %	275.4 bp	11.3 M	0.3 %	0.0 %	1.8 bp	0.0 %	2.2 bp
CONTROL_2	25.2 M	96.2 %	39.9 %	273.9 bp	6.5 M	0.3 %	0.0 %	1.4 bp	0.0 %	1.4 bp
LIF_1hr_1	39.5 M	87.8 %	60.5 %	278.5 bp	8.9 M	0.4 %	0.1 %	2.1 bp	0.0 %	2.0 bp
LIF_1hr_2	26.8 M	91.0 %	35.7 %	271.6 bp	6.5 M	0.5 %	0.0 %	1.3 bp	0.0 %	2.7 bp
LIF_1hr_3	34.2 M	90.4 %	46.8 %	277.6 bp	7.0 M	0.5 %	0.0 %	1.6 bp	0.0 %	2.1 bp
LIF_1hr_4	37.9 M	90.1 %	57.4 %	278.0 bp	12.0 M	0.3 %	0.0 %	1.8 bp	0.0 %	2.0 bp
LIF_1hr_5	51.0 M	92.3 %	55.1 %	278.9 bp	16.8 M	0.3 %	0.0 %	1.6 bp	0.0 %	1.6 bp
LIF_24hr_1	31.6 M	93.8 %	51.5 %	278.8 bp	8.7 M	0.3 %	0.0 %	1.9 bp	0.0 %	2.0 bp
LIF_24hr_2	26.3 M	93.5 %	43.5 %	275.8 bp	7.3 M	0.3 %	0.0 %	1.4 bp	0.0 %	2.0 bp
LIF_24hr_3	31.5 M	90.3 %	39.9 %	274.6 bp	6.5 M	0.5 %	0.0 %	1.5 bp	0.0 %	1.8 bp
LIF_24hr_4	43.7 M	96.8 %	81.0 %	279.2 bp	15.0 M	0.2 %	0.0 %	2.2 bp	0.0 %	1.8 bp
LIF_24hr_5	21.7 M	92.3 %	44.6 %	275.0 bp	6.9 M	0.3 %	0.0 %	1.5 bp	0.0 %	1.4 bp
control_1hr	34.2 M	87.0 %	52.8 %	277.9 bp	8.4 M	0.4 %	0.1 %	1.8 bp	0.0 %	1.7 bp

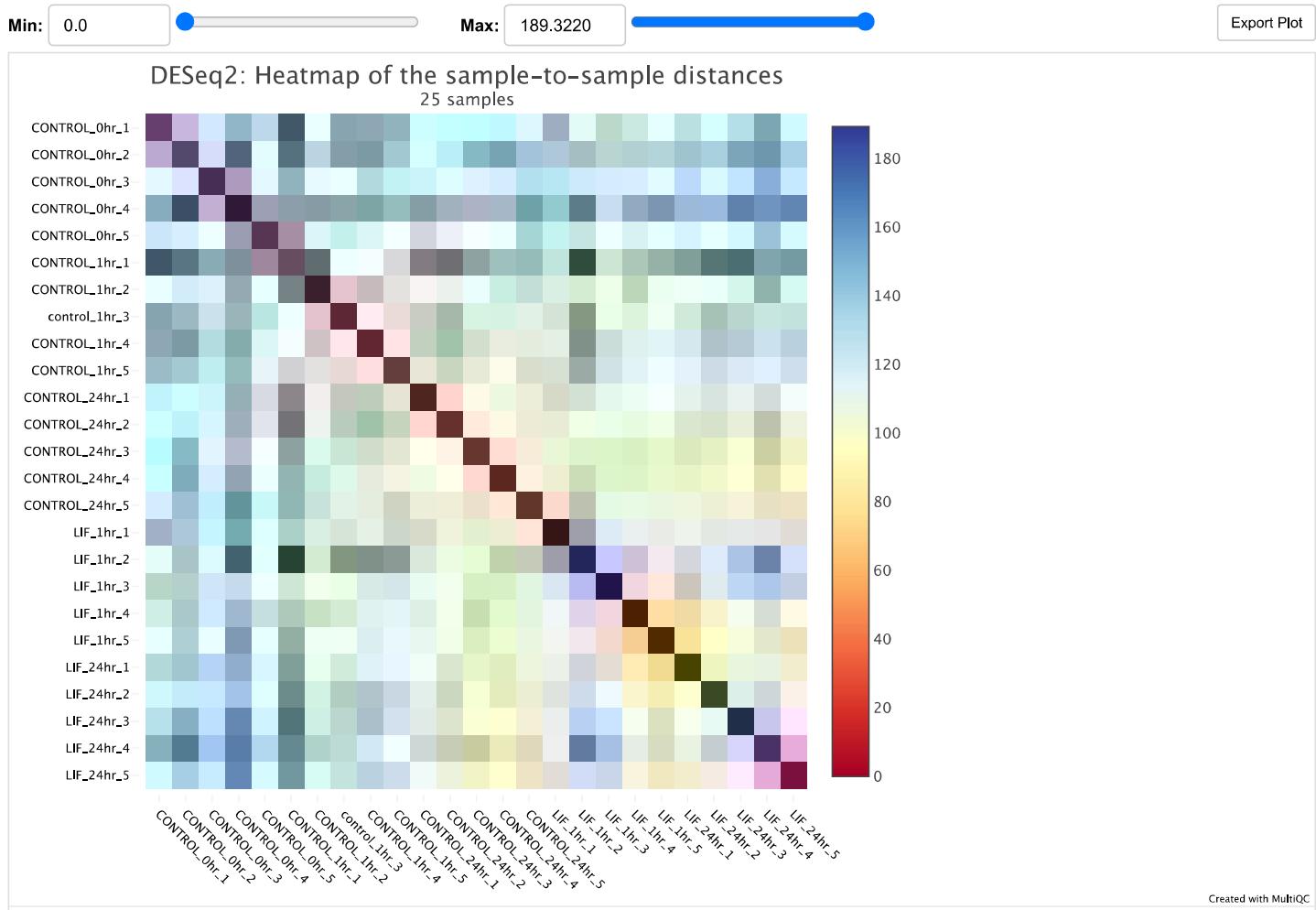
Alignment Scores



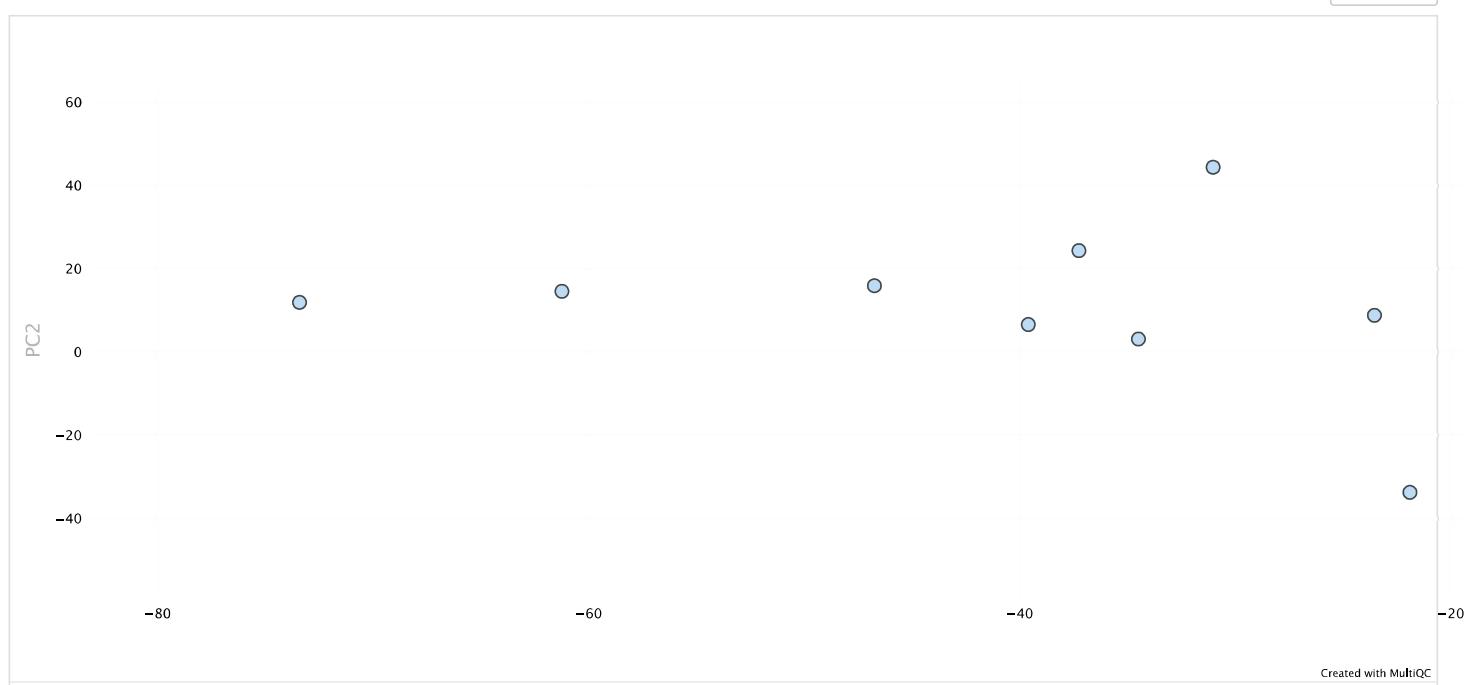
Sample relationships

Plots interrogating sample relationships, based on final count matrices.

STAR_SALMON DESeq2 sample similarity



STAR_SALMON DESeq2 PCA plot

[Export Plot](#)

Software Versions

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

[Copy table](#)

Group	Software	Version
BEDTOOLS_GENOMECOV_FW	bedtools	2.31.1
CUSTOM_GETCHROMSIZES	getchromsizes	1.2
CUSTOM_TX2GENE	python	3.9.5
DESEQ2_QC_STAR_SALMON	bioconductor-deseq2	1.28.0
	r-base	4.0.3
DupRadar	bioconductor-dupradar	1.32.0
FASTQC	fastqc	0.12.1
FQ_SUBSAMPLE	fq	0.9.1 (2022-02-22)
GTF_FILTER	python	3.9.5
MAKE_TRANSCRIPTS_FASTA	rsem	1.3.1
	star	2.7.10a
PICARD_MARKDUPPLICATES	picard	3.1.1
QUALIMAP_RNASEQ	qualimap	2.3
RSEQC_BAMSTAT	rseqc	5.0.2
RSEQC_INFERENCEEXPERIMENT	rseqc	5.0.2
RSEQC_INNERDISTANCE	rseqc	5.0.2
RSEQC_JUNCTIONANNOTATION	rseqc	5.0.2
RSEQC_JUNCTIONSATURATION	rseqc	5.0.2
RSEQC_READDISTRIBUTION	rseqc	5.0.2
RSEQC_READDUPLICATION	rseqc	5.0.2
SALMON_INDEX	salmon	1.10.1
SALMON_QUANT	salmon	1.10.1
SAMTOOLS_FLAGSTAT	samtools	1.2
SAMTOOLS_IDXSTATS	samtools	1.2
SAMTOOLS_INDEX	samtools	1.2
SAMTOOLS_SORT	samtools	1.2
SAMTOOLS_STATS	samtools	1.2
SE_GENE	bioconductor-summarizedexperiment	1.32.0
STAR_ALIGN_IGENOMES	gawk	5.1.0
	samtools	1.1
	star	2.6.1d
STRINGTIE_STRINGTIE	stringtie	2.2.1
TRIMGALORE	cutadapt	3.4
	trimgalore	0.6.7
TXIMETA_TXIMPORT	bioconductor-tximeta	1.20.1
UCSC_BEDCLIP	ucsc	377
UCSC_BEDGRAPHTOBIGWIG	ucsc	445
Workflow	Nextflow	24.04.4

Group	Software	Version
	nf-core/rnaseq	v3.16.1-g1f3f64d

nf-core/rnaseq Methods Description

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

Methods

Data was processed using nf-core/rnaseq v3.16.1 (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 v24.04.4 (Di Tommaso *et al.*, 2017) with the following command:

```
nextflow run nf-core/rnaseq -r 3.16.1 -profile docker -resume -params-file nf-params.json
```

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. URL: <https://github.com/nf-core/rnaseq>

Input/output options

input	/media/jochum00/Aagaard_Raid1/jochum00/k_pennington/LIF_project/BaseSpace/edwards-8184-434104856/cdna_samplesheet.csv
outdir	/media/jochum00/Aagaard_Raid1/jochum00/k_pennington/LIF_project/BaseSpace/edwards-8184-434104856/nfcore

Reference genome options

fasta	s3://ngi-igenomes/igenomes//Mus_musculus/UCSC/mm10/Sequence/WholeGenomeFasta/genome.fa
gene_bed	s3://ngi-igenomes/igenomes//Mus_musculus/UCSC/mm10/Annotation/Genes/genes.bed
genome	mm10
gtf	s3://ngi-igenomes/igenomes//Mus_musculus/UCSC/mm10/Annotation/Genes/genes.gtf
star_index	s3://ngi-igenomes/igenomes//Mus_musculus/UCSC/mm10/Sequence/STARIndex/

Alignment options

min_mapped_reads	5
-------------------------	---

Core Nextflow options

configFiles	N/A
containerEngine	docker
launchDir	/media/jochum00/Aagaard_Raid3/jochum00/k_pennington/LIF_project/BaseSpace/edwards-8184-434104856
profile	docker
projectDir	/home/jochum00/.nextflow/assets/nf-core/rnaseq
revision	3.16.1
runName	friendly_tuckerman
userName	jochum00
workDir	/media/jochum00/Aagaard_Raid3/jochum00/k_pennington/LIF_project/BaseSpace/edwards-8184-434104856/work