

1. Post-Alignment

- Select the sample and download the alignment report which shows mapping statistics of each fastq file within the sample.
- Configure settings and click the “View Bam Statistics” button to view aggregated mapping statistics of all the fastq files within the sample selected.
- The settings shown below act as a filter, reads which are filtered out are described as “excluded” reads in the charts.
(note: In the screenshots below for alignment scores, alignment lengths & edit distance, there is no difference between the 2 graphs as all reads pass the default filter settings.)

Example alignment report:

https://github.com/paigerollex/gene_cloud_omics/blob/main/output_data/alignment_report.pdf

Select FASTQ Sample

male_a

Download Alignment Report

Min Alignment Score (MAPQ)

0 255

0 26 52 78 104 130 156 182 208 234 255

Min Fragment Size

50

Max Fragment Size

250

Max Repeats

1

View Bam Statistics ← summary statistics

Fig 5.1: Plot settings



Fig 5.2: Reads aligned

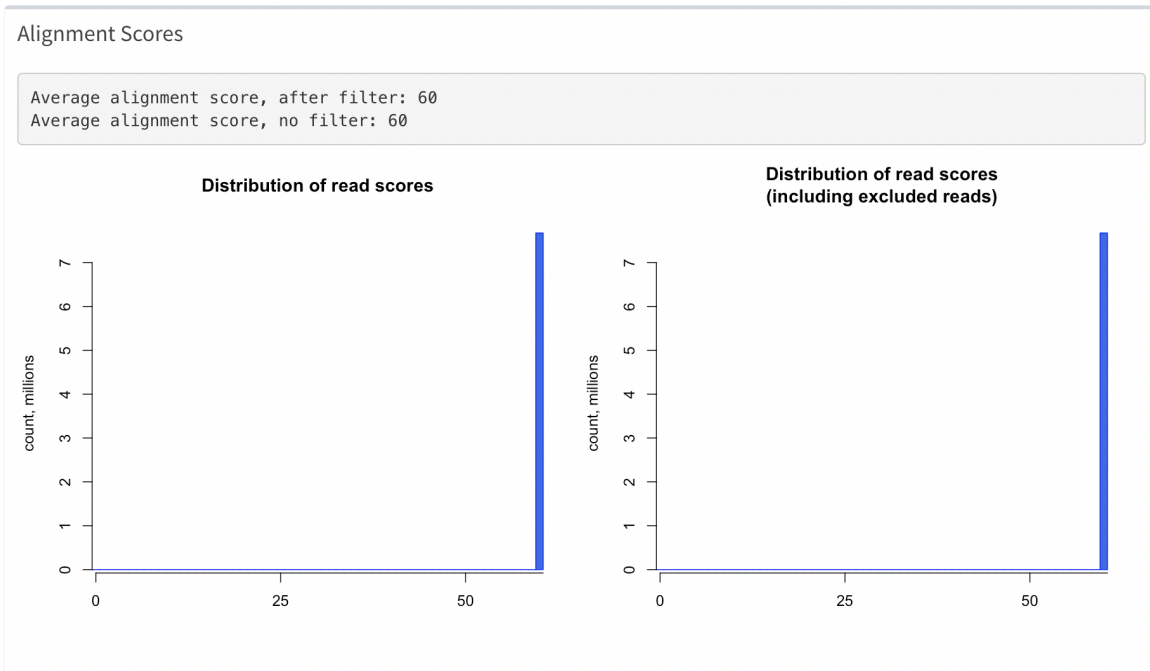


Fig 5.3: Alignment scores

Alignment Length

Average aligned length: 72
Average aligned length, no filter: 72

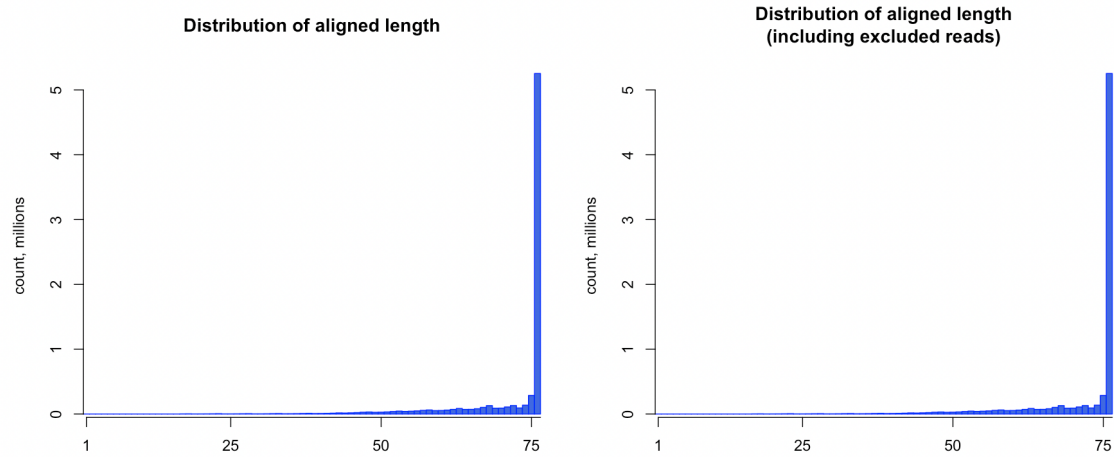


Fig 5.4: Alignment length

Edit Distance

Average edit distance, after filter: 0.249203492683443
Average edit distance, no filter: 0.249203492683443

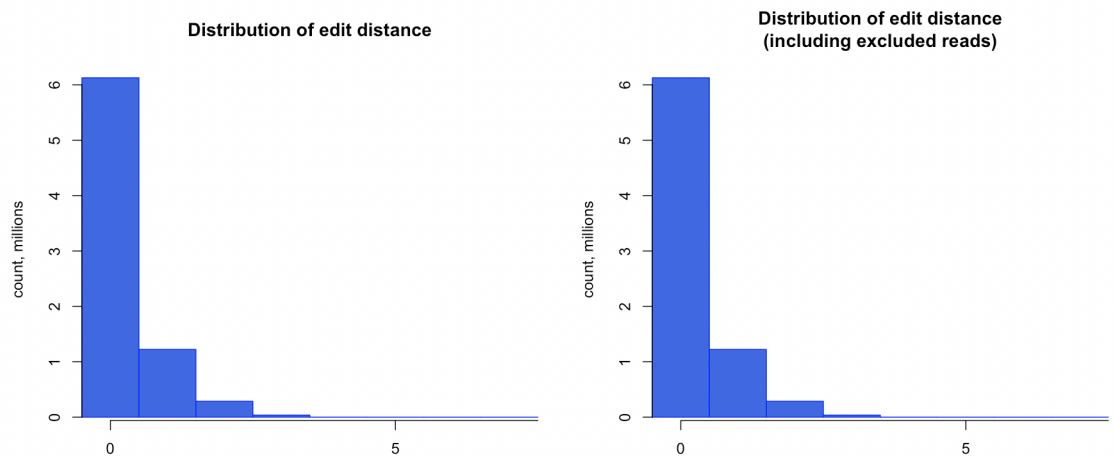


Fig 5.5: Edit distance

2. Genome Browser

- Select the sample to view coverage
- Select the chromosome from the gene annotation file and the range to plot. At times, some gene ids or transcripts ids of the coverage graph cannot be seen clearly, so you can extend the plot from the right or left.
(note: Range to plot changes with the chromosome selected to plot)

The figure shows two side-by-side screenshots of the Genome Browser interface. Both screenshots have a 'Select FASTQ Sample' dropdown set to 'male_a'. The left screenshot has 'Select Chromosome to View Coverage' set to 'NC_004353.4'. Its 'Range to plot' slider shows values 879, 318,599, 636,319, and 1,271,759. The right screenshot has 'Select Chromosome to View Coverage' set to 'NT_033777.3'. Its 'Range to plot' slider shows values 567,076, 8,442,417, 16,317,758, and 32,068,441. Both have 'Extend plot from the right' and 'Extend plot from the left' input fields set to 0. A 'View Coverage' button is at the bottom of each, with a blue arrow pointing to it from the text 'Click once you've selected range'.

Fig 5.6: Ranges of different chromosomes

- Based on the selected chromosome and range to plot, you can view additional information about the genes, transcripts and exons plotted in the graph.

The figure shows the 'View Ranges' tab in the Genome Browser. At the top are 'View Ranges' and 'View Coverage' tabs. Below is a 'Genomic Feature Type' section with radio buttons for 'gene' (selected), 'transcript', and 'exon'. A blue arrow points to the 'gene' button with the text 'Select type'. Below this is a 'Show 15 entries' dropdown and a 'Search:' input field. A table displays genomic features with columns: seqnames, start, end, width, strand, and gene_id. The table has 5 rows of data.

	seqnames	start	end	width	strand	gene_id
1	NC_004353.4	469409	488994	19586	-	Asator
2	NC_004353.4	583108	589348	6241	+	CG1909
3	NC_004353.4	314923	331387	16465	+	CG32850
4	NC_004353.4	467497	471476	3980	+	CG33941
5	NC_004353.4	445504	453947	8444	+	CaMKI

Fig 5.7: Genes plotted

View Ranges

View Coverage

Genomic Feature Type

☐ gene
☒ transcript
☐ exon

Show

15

entries

Search:

	seqnames	start	end	width	strand	tx_id	tx_name
	All	All	All	All	All	All	All
1	NC_004353.4	314923	321385	6463	+	54	NM_166753.3
2	NC_004353.4	314923	331387	16465	+	55	NM_001272124.1
3	NC_004353.4	359938	372784	12847	+	56	NM_001258491.3
4	NC_004353.4	359938	372784	12847	+	57	NM_079882.4
5	NC_004353.4	359938	372784	12847	+	58	NM_166757.3

Fig 5.8: Transcripts plotted

- d. Click “View Coverage” button
- e. Coverage plot:
 - i. On the left gray rectangle are fastq identifiers within the sample.
 - ii. Empty transcript names in gene annotation file: coverage plot will only show genes within the selected range.
(note: Often the case when downloading directly from NCBI)
 - iii. No empty transcript names in gene annotation file: coverage plot will show both genes and transcripts within the selected range.
(note: Often the case when downloading directly from Ensembl)

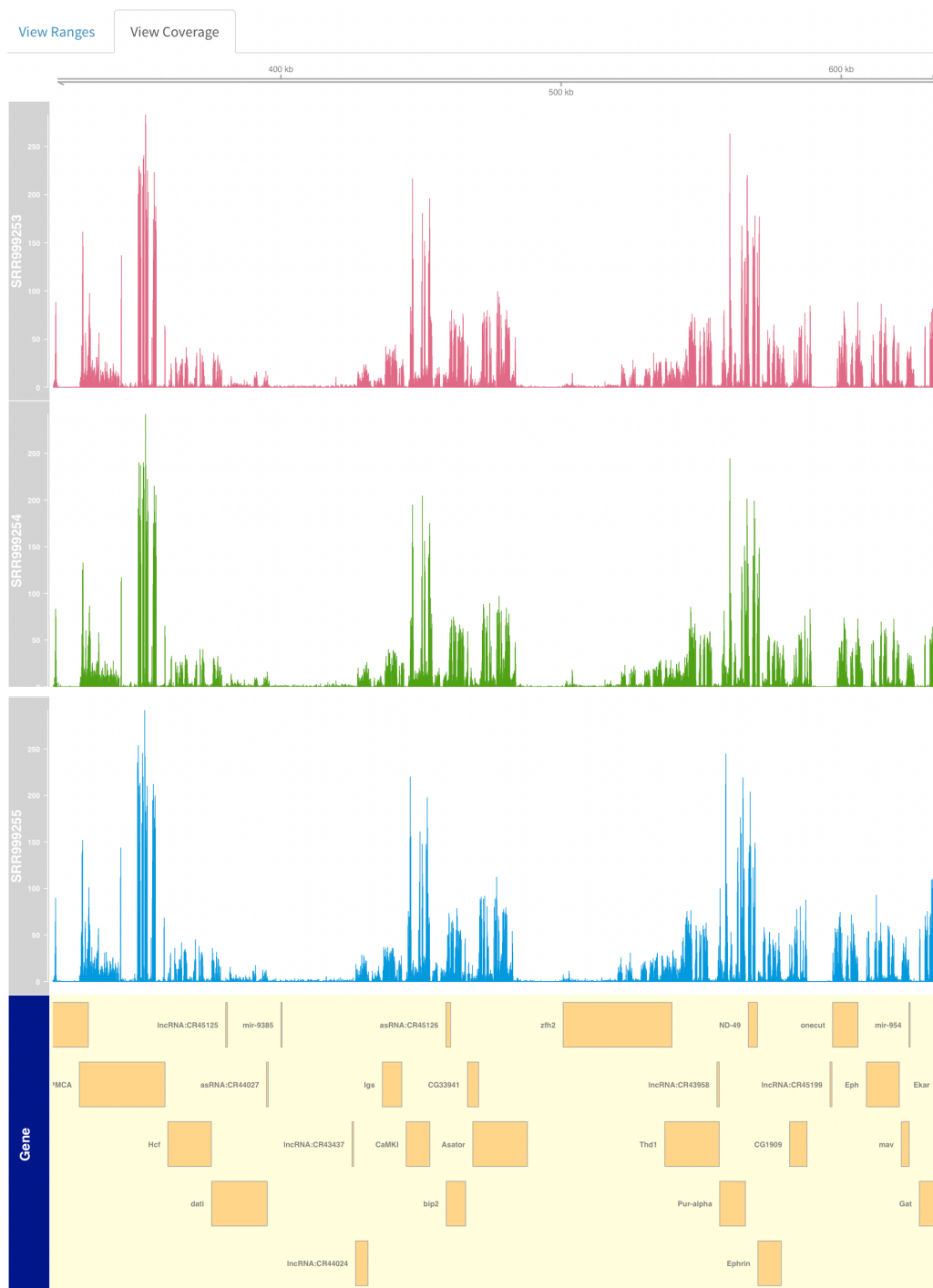


Fig 5.9: Coverage plot with genes (NCBI files)

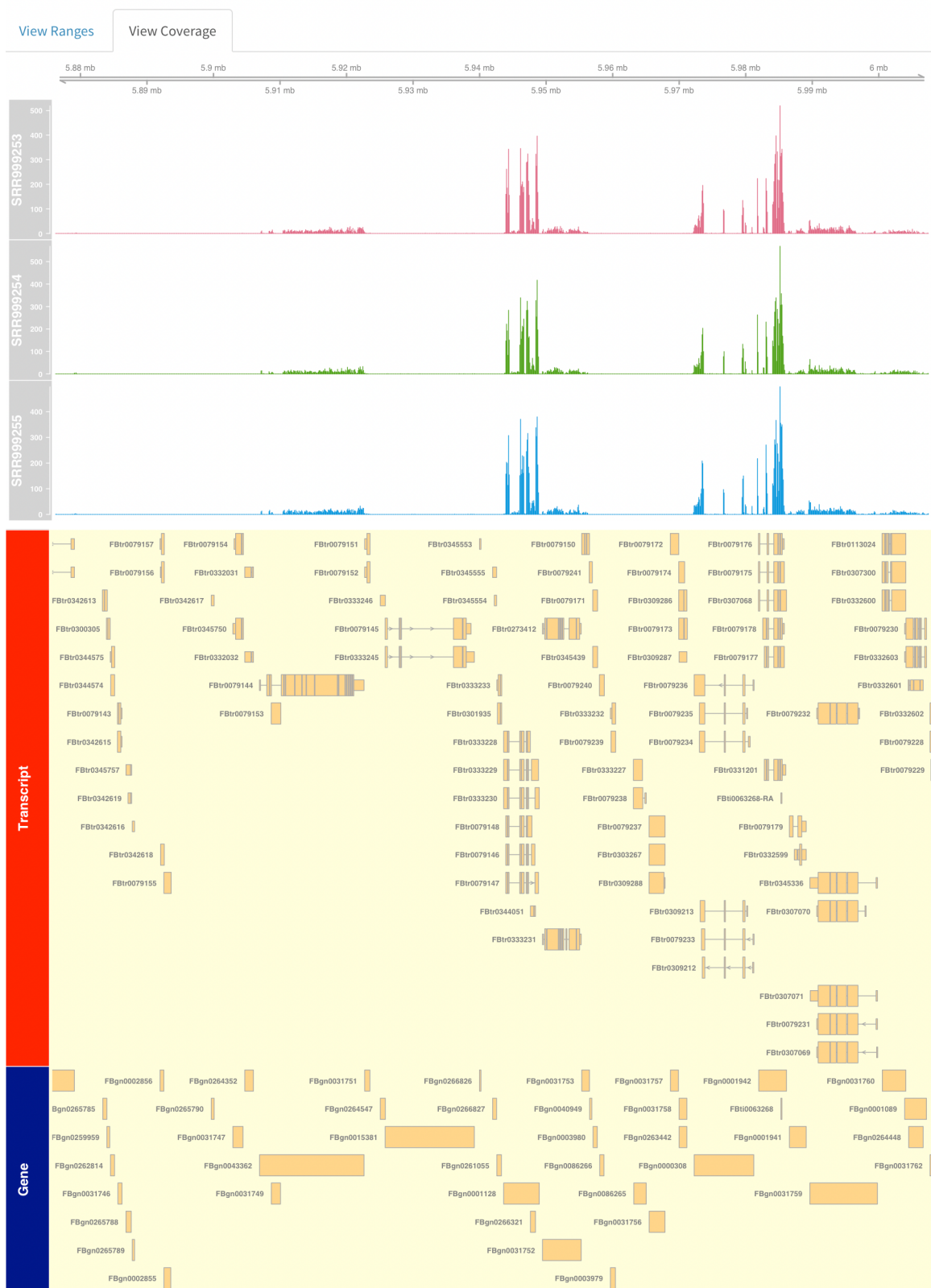


Fig 5.10: Coverage plot with genes & transcripts (Ensembl files)

