

# SERVICE REPORT

## Small RNA-Seq Data Report

Project SV00001

### SERVICE REPORT

Customer: Dr. John Doe  
Company/Institute: ABC Institution  
Date: January 1, 2020

Performed by:  
Norgen Biotek Corp.  
3430 Schmon Pkwy, Thorold, ON  
Canada L2V 4Y6

## TABLE OF CONTENTS

---

Report Summary .....	3
Small RNA-Seq Data Analysis Workflow Used .....	4
Sources of Small RNA Reference Sequences .....	4
Downloading Your Files .....	4
Structure of Files Provided.....	5
Sample Raw Read File Information .....	6
Quantification of RNA used for Small RNA-Seq .....	6
Mapping and Biotype Summary.....	7
Help and Technical Support.....	9
Small RNA Analysis (Additional Library Summary) .....	10

## Report Summary

<b>Client Name</b>	Dr. John Doe (ABC Institution)
<b>Norgen Order Number</b>	SV00001
<b>Number of Samples to Process</b>	10 Samples
<b>Sample Type</b>	RNA (Plasma-human)
<b>Sequencing Service Provided</b>	Small RNA-Seq
<b>Sequencing Platform</b>	Illumina NextSeq 500
<b>Sequencing Platform Reagent</b>	NextSeq 500/550 Mid Output Kit v2 (51 Cycles using a 75-Cycle Kit)
<b>Product Used for Library Preparation</b>	Norgen Biotek Small RNA Library Prep Kit (Cat. 63600)

**Date and Time of Release of Report:** January 1, 2020

**Report Prepared by:** Dr. Mohamed El-Mogy **Date:** January 1, 2020

**Report Approved by:** Dr. Yousef Haj-Ahmad **Date:** January 1, 2020

## Small RNA-Seq Data Analysis Workflow Used

exceRpt small RNA-seq Pipeline (v4.6.2)

Link: [http://genboree.org/theCommons/projects/exrna-tools-may2014/wiki/Small RNA-seq Pipeline](http://genboree.org/theCommons/projects/exrna-tools-may2014/wiki/Small_RNA-seq_Pipeline)

## Sources of Small RNA Reference Sequences

Small RNA Species	Reference or Database Sequences Used
miRNAs	miRBase version 22
tRNAs	gtRNadb
piRNAs	piRNadb, piRNAbank & RNAcentral
Genome	Gencode version 22 (hg38)

## Downloading Your Files

All files pertaining to this service have been uploaded to our servers. These files will be available for up to 1 month after the completion of the project. Please make sure to download all files before then.

To access your files, visit the link below and enter the user name and password provided. Click the [Download](#) button to start the download. If you have any issues accessing the files, please contact us for help.

Please note that all fields are case-sensitive.

<b>Download Link:</b>	<code>https:\\services.norgenbiotek.com\\sv00001</code>
<b>Username:</b>	user-sv0001
<b>Password:</b>	\$Abh64*2
<b>Expiration:</b>	April 1, 2020

## Structure of Files Provided


Inside the UBS hard drive provided, you will find the following file structure that contains the analysis results.

### Raw Reads


 Contains compressed original FASTQ file of each sample


### Small RNA Analysis

#### Intermediate Mapping Files


 Contains intermediate mapping files of each sample

#### Post Processed Results

 Contains excerpt-generated summary of read counts of each small RNA species (\*.txt format)

 Mapping summary provided in the file: excerpt\_readMappingSummary

#### Processed Files for Individual Samples

 Contains compressed folder of excerpt-generated data files of each sample

## Sample Raw Read File Information

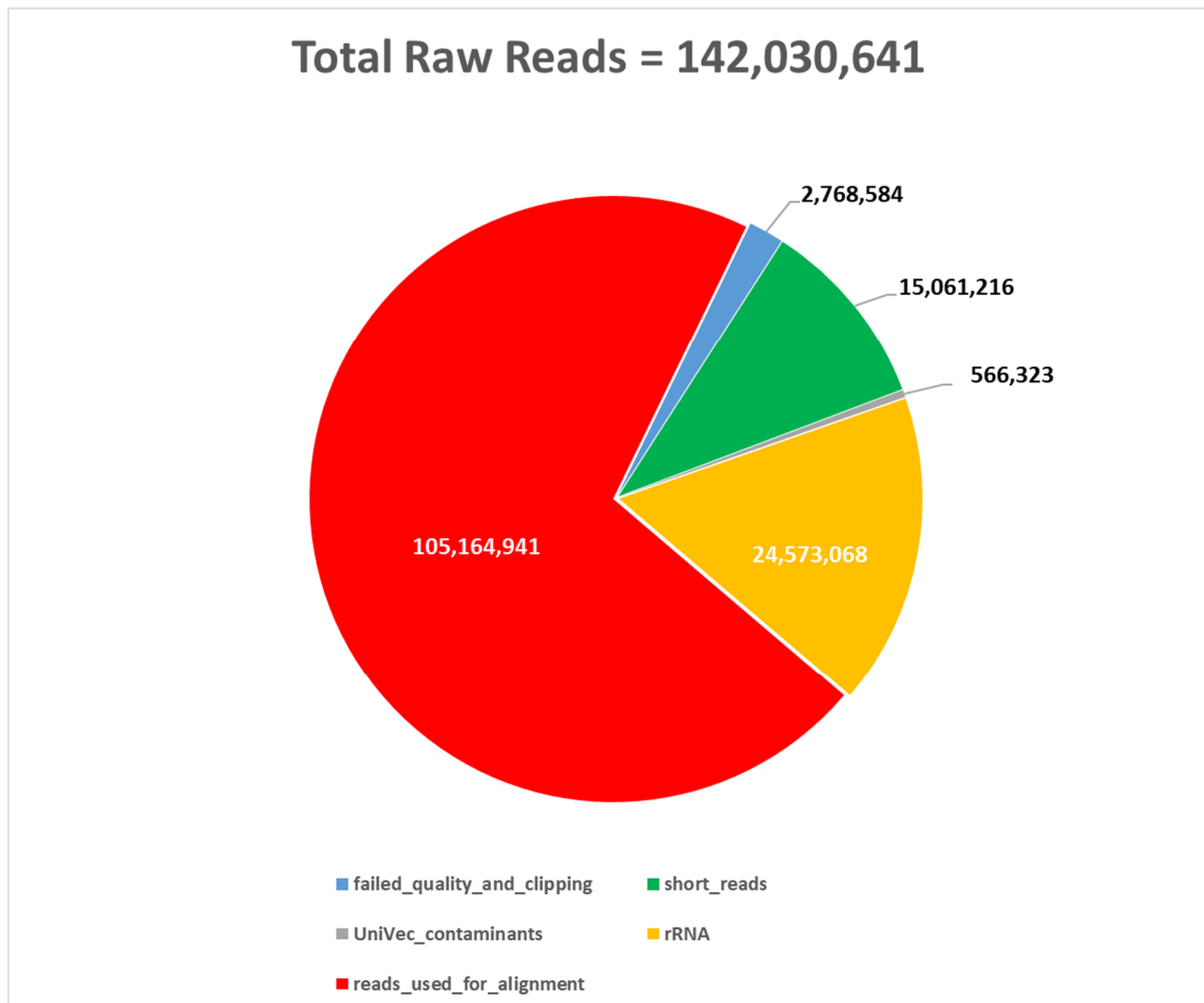
Norgen ID	Sample ID	Group ID	Index	Associated FASTQ File Name
SV00001_0001	1	Control	RPI1	SV00001_0001_S1_R1_001
SV00001_0002	2	Control	RPI2	SV00001_0002_S2_R1_001
SV00001_0003	3	Control	RPI3	SV00001_0003_S3_R1_001
SV00001_0004	4	Control	RPI4	SV00001_0004_S4_R1_001
SV00001_0005	5	Control	RPI5	SV00001_0005_S5_R1_001
SV00001_0006	6	Treatment	RPI6	SV00001_0006_S6_R1_001
SV00001_0007	7	Treatment	RPI7	SV00001_0007_S7_R1_001
SV00001_0008	8	Treatment	RPI8	SV00001_0008_S8_R1_001
SV00001_0009	9	Treatment	RPI9	SV00001_0009_S9_R1_001
SV00001_0010	10	Treatment	RPI10	SV00001_0010_S10_R1_001

## Quantification of RNA used for Small RNA-Seq

Sample ID	RNA ID	RNA Conc., pg/uL (Bioanalyzer)
SV00001_0001	SV00001_0001_0001-RNA	85
SV00001_0002	SV00001_0002_0002-RNA	121
SV00001_0003	SV00001_0003_0003-RNA	40
SV00001_0004	SV00001_0004_0004-RNA	78
SV00001_0005	SV00001_0005_0005-RNA	45
SV00001_0006	SV00001_0006_0006-RNA	72
SV00001_0007	SV00001_0007_0007-RNA	153
SV00001_0008	SV00001_0008_0008-RNA	60
SV00001_0009	SV00001_0009_0009-RNA	81
SV00001_0010	SV00001_0010_0010-RNA	67

Amount of RNA is in line with observed yield of similar RNA samples. It meets the minimum requirement for library prep and sequencing.

## Mapping and Biotype Summary



	Total Sample	SV00001_0001	SV00001_0002	SV00001_0003	SV00001_0004	SV00001_0005	SV00001_0006	SV00001_0007	SV00001_0008	SV00001_0009	SV00001_0010
<b>Total Raw Read</b>	<b>142,030,641</b>	13,865,213	17,635,243	12,457,851	14,657,892	16,578,455	13,457,841	12,145,412	13,545,221	14,235,411	13,452,102
failed_quality_and_clipping	2,768,584	203,451	302,621	348,621	260,132	386,542	234,612	164,587	204,562	362,345	301,111
short_reads	15,061,216	1,634,524	1,678,542	1,634,875	1,346,521	1,244,511	1,346,785	1,346,511	2,134,521	1,346,531	1,347,895
passed_initial_qc	129,738,009	12,434,140	16,259,322	11,171,597	13,571,503	15,720,486	12,345,668	10,963,488	11,615,262	13,251,225	12,405,318
rRNA	24,573,068	2,645,123	2,451,212	2,643,154	3,012,102	2,341,245	2,553,122	2,224,342	2,134,512	2,155,122	2,413,134
<b>Total Mappable Reads</b>	<b>105,164,941</b>	9,789,017	13,808,110	8,528,443	10,559,401	13,379,241	9,792,546	8,739,146	9,480,750	11,096,103	9,992,184
reads mapped to genome	73,744,864	7,546,213	8,451,213	7,546,213	6,723,142	7,665,431	6,998,763	5,032,103	7,874,541	9,875,123	6,032,122
not mapped to genome or libs	31,420,077	2,242,804	5,356,897	982,230	3,836,259	5,713,810	2,793,783	3,707,043	1,606,209	1,220,980	3,960,062
<b>miRNA_sense</b>	<b>59,874,241</b>	4,623,150	7,645,212	6,458,721	7,632,023	7,324,510	5,421,611	3,465,112	4,312,565	7,645,124	5,346,213
miRNA antisense	0	0	0	0	0	0	0	0	0	0	0
miRNAprecursor_sense	148,671	12,345	13,452	17,204	21,346	16,423	14,345	13,462	12,422	13,402	14,270
miRNAprecursor_antisense	578	103	34	24	24	35	93	102	34	103	26
tRNA_sense	1,805,661	249,762	246,134	213,462	163,403	134,652	103,421	134,061	134,021	210,314	216,431
tRNA_antisense	14,395	4,621	1,348	1,346	333	1,304	314	1,446	1,342	999	1,342
piRNA_sense	355,495	48,316	34,685	44,316	42,031	30,134	31,420	26,784	23,014	43,120	31,675
piRNA_antisense	199	4	55	24	34	26	4	6	7	20	19
gencode_sense	3,186,844	214,645	346,212	645,313	234,676	275,803	310,118	316,402	362,212	134,678	346,785
gencode_antisense	235,790	44,858	53,901	22,207	28,159	23,677	27,491	16,227	11,325	6,424	1,522
circularRNA_sense	36,839	8,645	7,643	4,316	1,634	2,134	2,136	457	1,875	3,291	4,708
circularRNA_antisense	26,960	4,316	4,872	2,340	1,642	2,317	4,312	2,103	1,945	1,686	1,426

Sample ID	SV00001_0001	SV00001_0002	SV00001_0003	SV00001_0004	SV00001_0005	SV00001_0006	SV00001_0007	SV00001_0008	SV00001_0009	SV00001_0010
Number of detected miRNAs at > 5 raw reads	232	342	365	332	365	234	264	302	340	260
Number of detected miRNAs at > 1 raw read	248	352	402	346	375	256	283	356	391	286

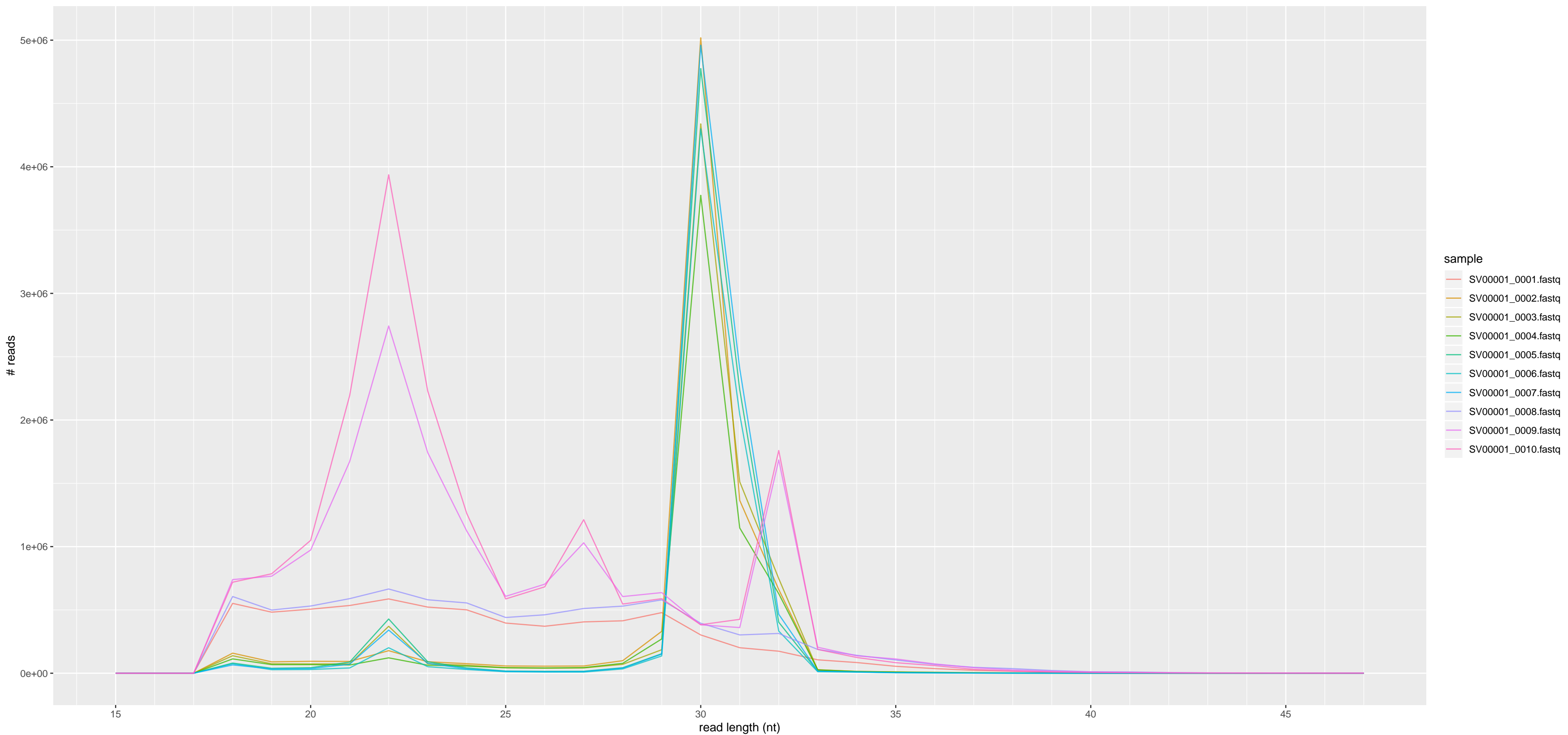


## Help and Technical Support

For any questions or additional help regarding this report, please contact our Technical Support Team between the hours of 9:00 AM and 5:30 PM (Eastern Standard Time) at (905) 227-8848 or Toll Free at 1-866-667-4362. Technical support can also be obtained through email at [techsupport@norgenbiotek.com](mailto:techsupport@norgenbiotek.com).

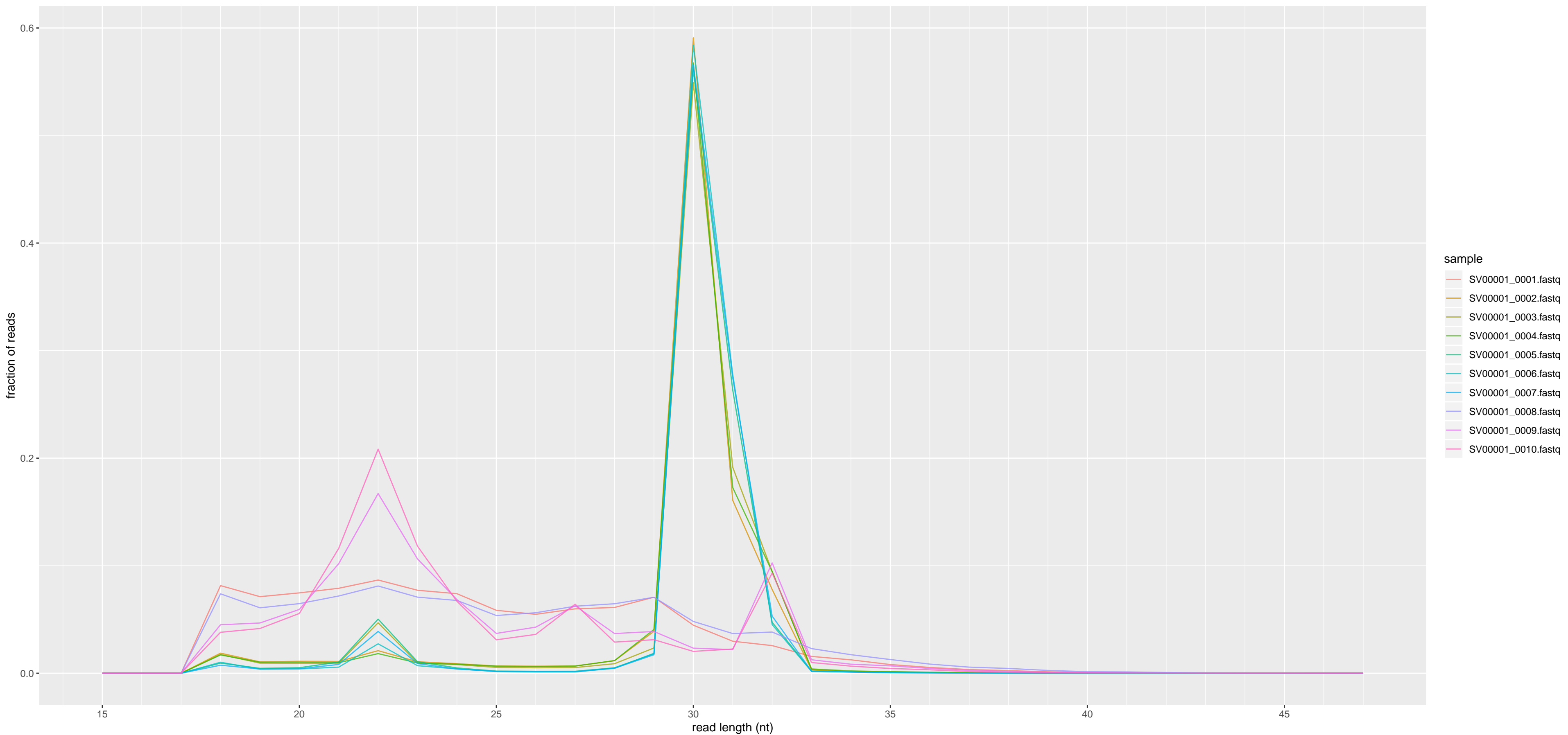
## Small RNA Analysis (Additional Library Summary)

Read length distribution of sequenced samples based on number of raw reads



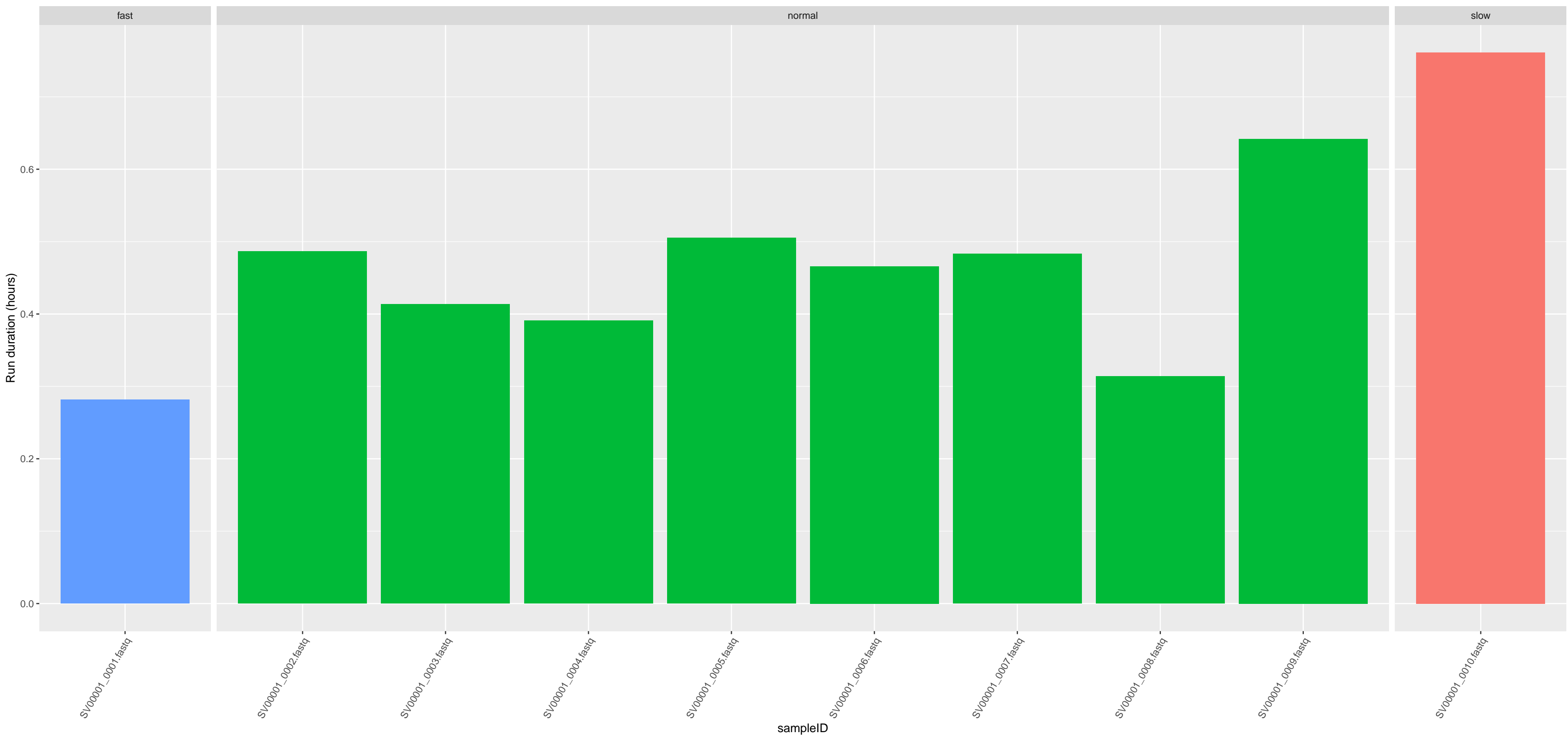
Read length is plotted against number of reads at each specific size (nucleotides) to indicate predominant molecules of nucleotide length corresponding to a specific biotype.

Read length distribution of sequenced samples based on normalized read fractions



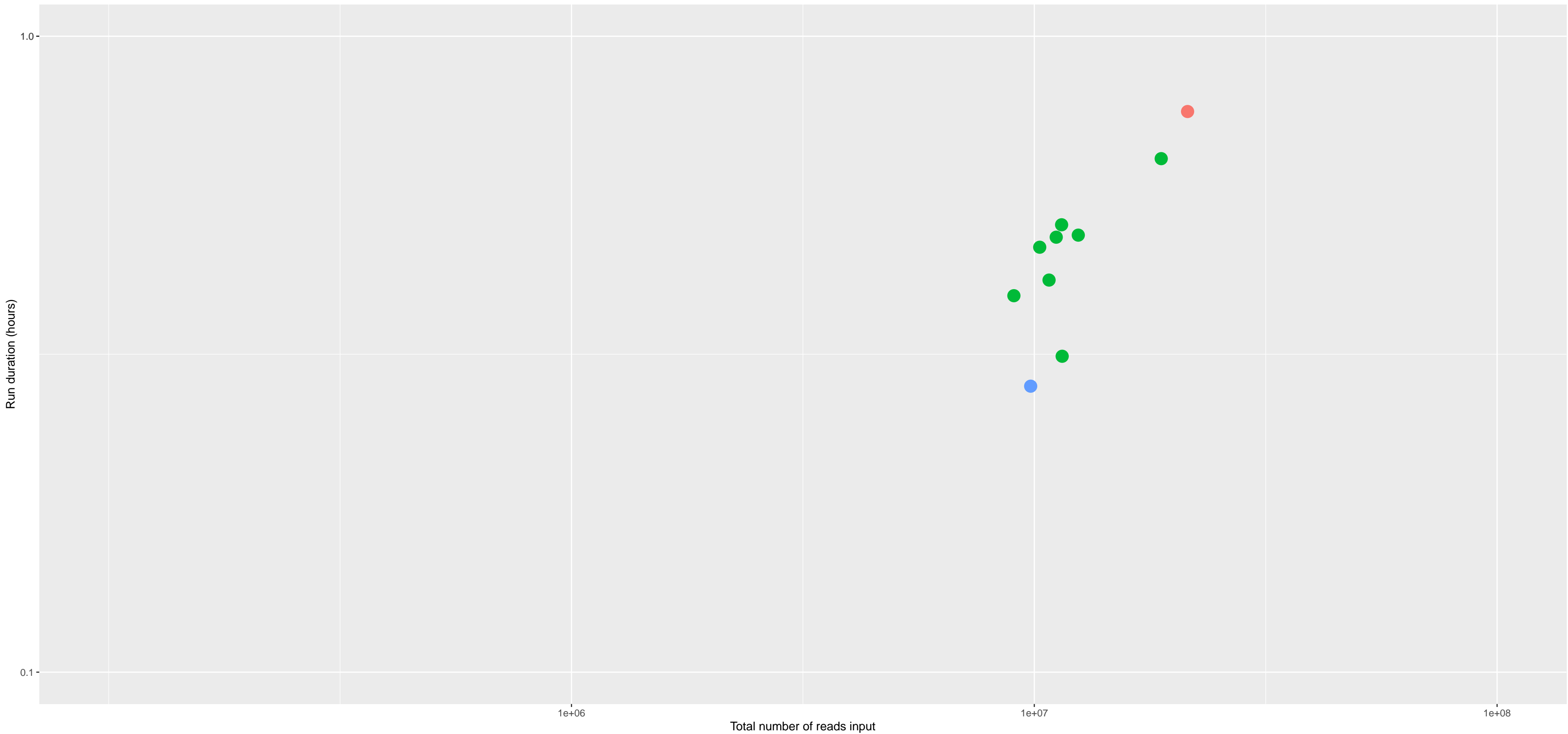
Read length is plotted against normalized reads at each specific size (nucleotides) to indicate predominant molecules of nucleotide length corresponding to a specific biotype.

Duration of analysis for each sample



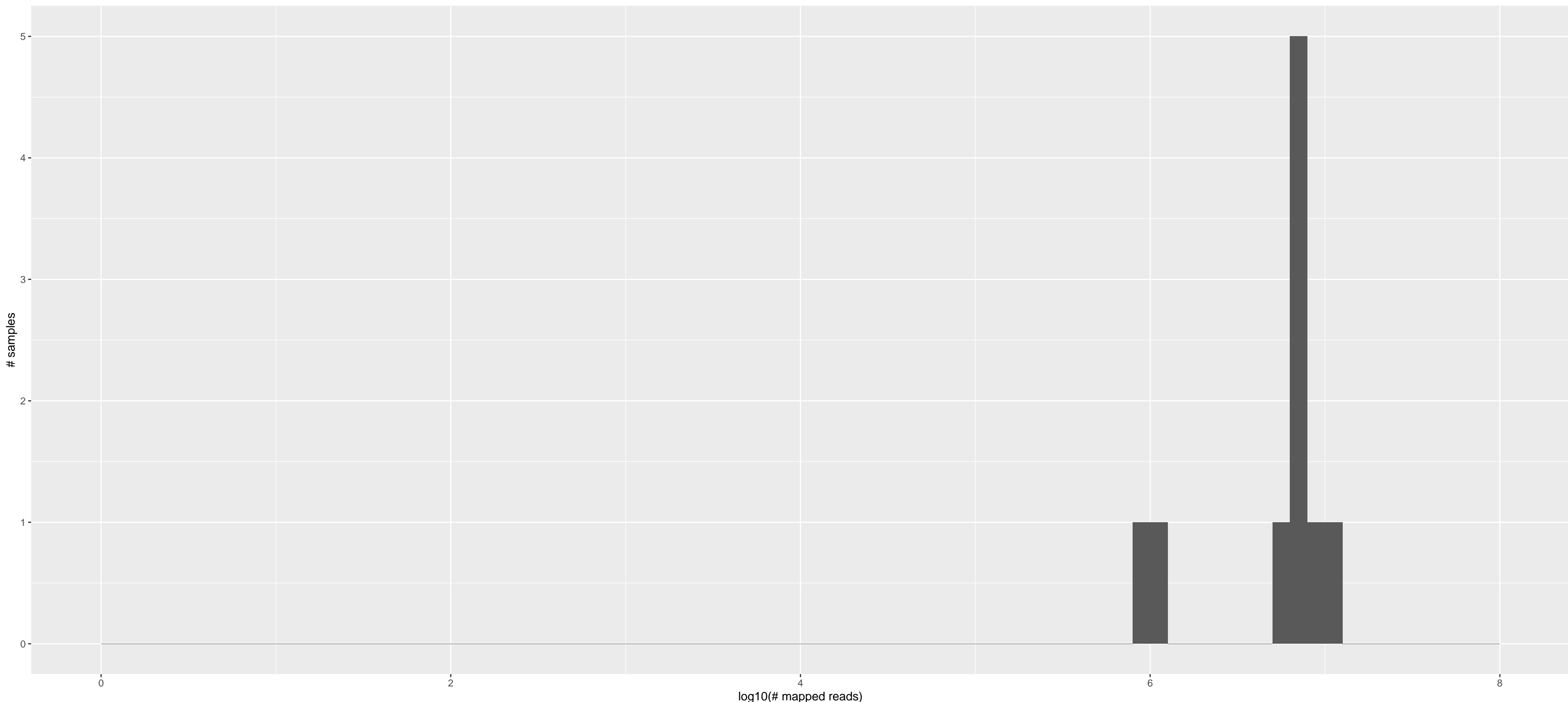
The run duration in hours is illustrated to indicate the time (hours) spent by the analysis pipeline to complete the analysis.

Duration of analysis per sequencing yield



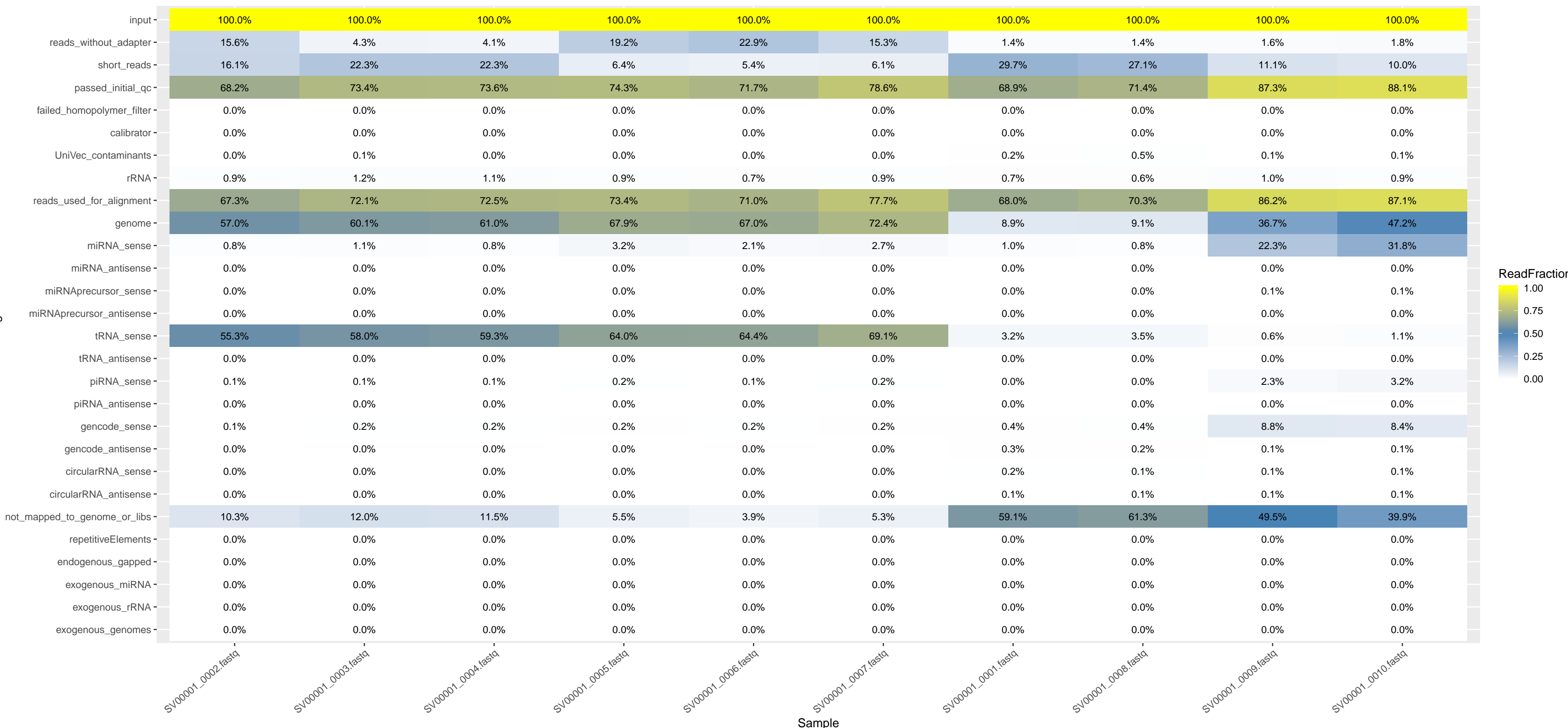
The run duration in hours is illustrated to indicate the time (hours) spent by the analysis pipeline to complete the analysis per total number of reads input.

Library size (all mapped reads)



A histogram of library sizes. This plot provides an easy way to visualize the distribution of samples according to the number of their mapped reads. Ideally a consistent and robust library preparation and sequencing run should result in a tight distribution provided that samples have the same biological origin and have been handled with the same procedures prior to their library preparation (collection, storage and isolation).

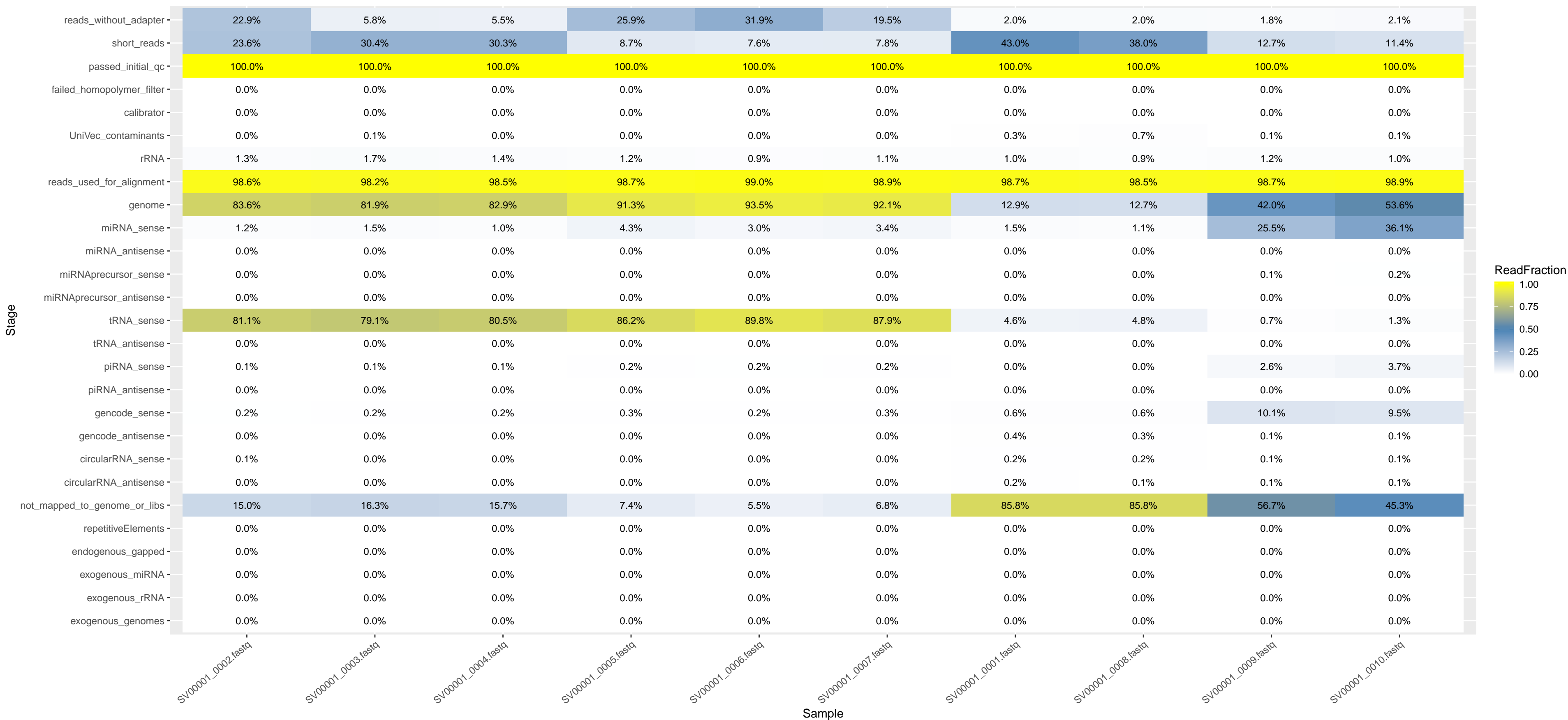
Fraction aligned reads normalized by number of input reads



The percentage of reads is illustrated in a graphical table with a colored scheme that indicates fraction size and to allow for comparing fractions within a sample or across samples to oversee sample variations and their relative biotype distribution.

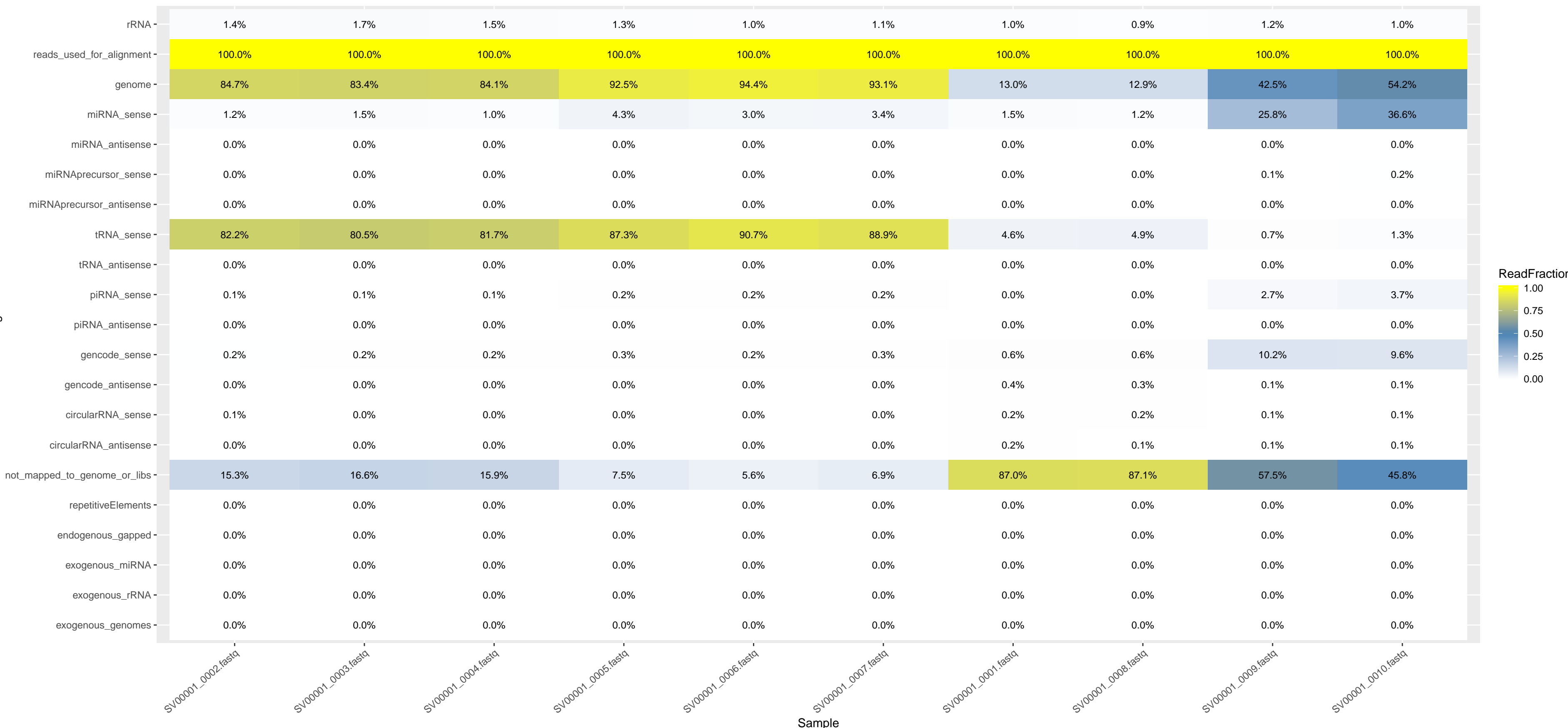


Fraction aligned reads normalized by number of adapter-clipped reads



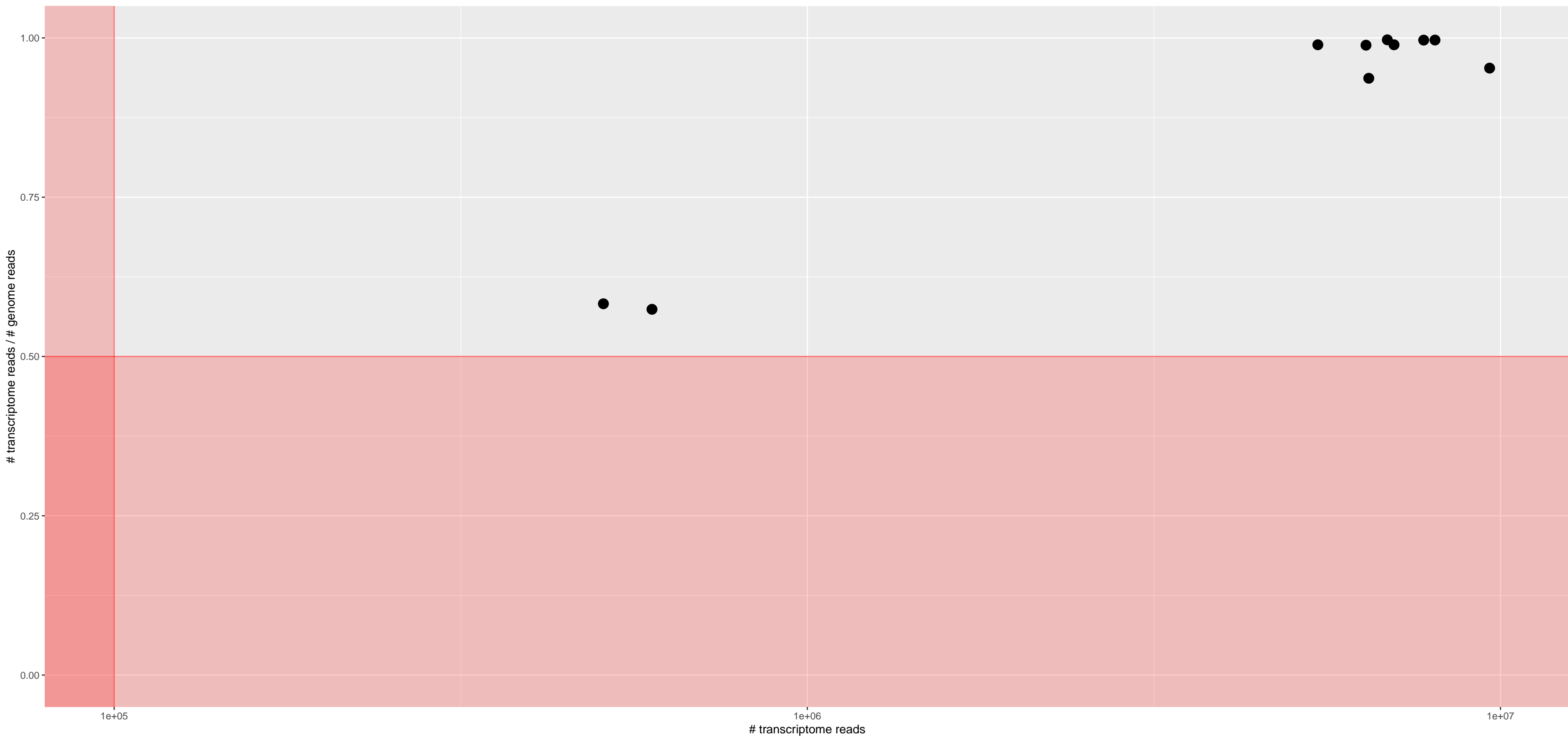
The percentage of reads is illustrated in a graphical table with a colored scheme that indicates fraction size and to allow for comparing fractions within a sample or across samples to oversee sample variations and their relative biotype distribution.

Fraction aligned reads normalized by number of non-contaminant reads



The percentage of reads is illustrated in a graphical table with a colored scheme that indicates fraction size and to allow for comparing fractions within a sample or across samples to oversee sample variations and their relative biotype distribution.

# Overall QC Results of Small RNA Analysis



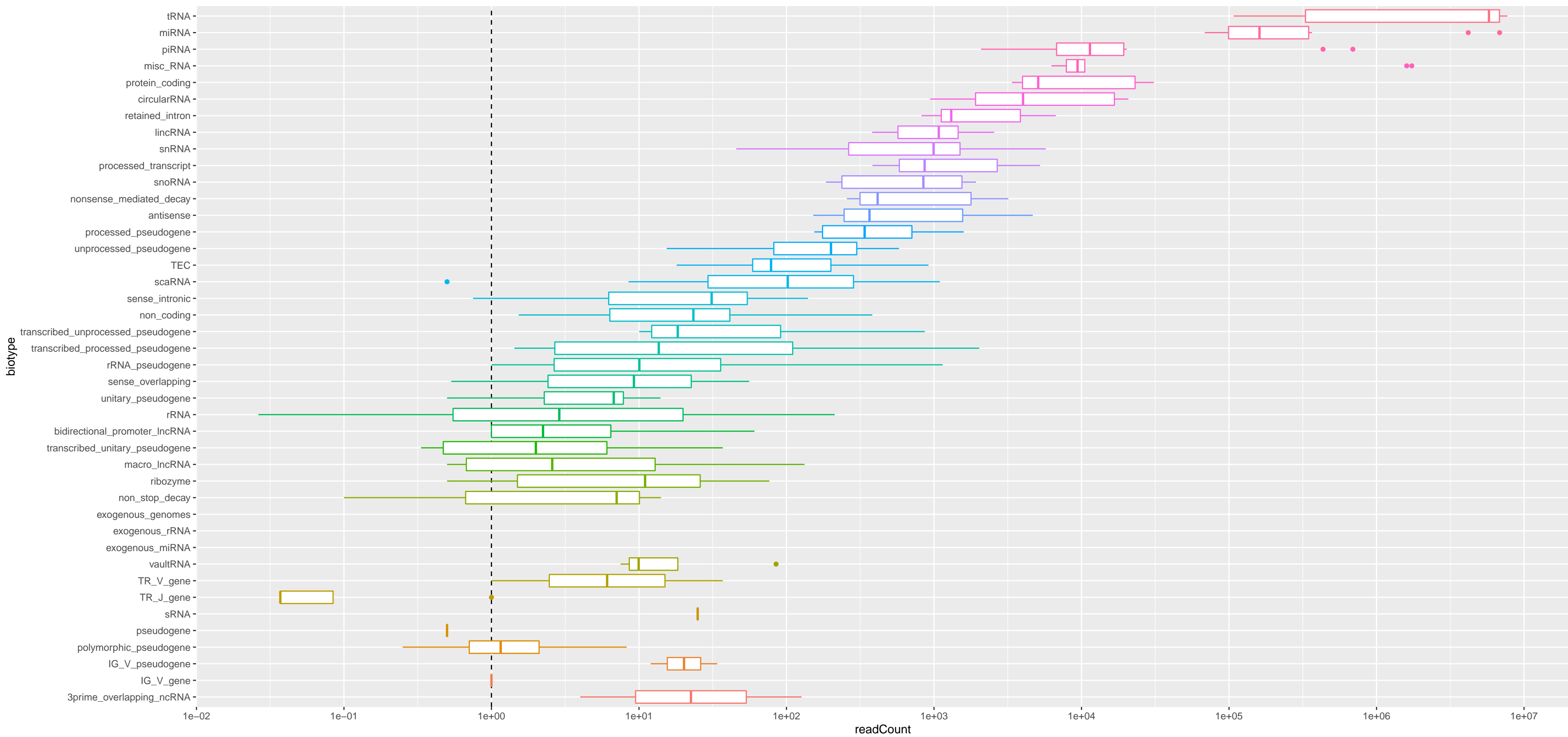
The graph indicates sample quality based on the small RNA analysis pipeline. Samples that pass QC have a number of transcriptome reads that exceeds 100,000 and transcriptome to genome ratio of > 0.5.

### Small RNA analysis QC results per sample

Sample	InputReads	GenomeReads	TranscriptomeReads	TranscriptomeGenomeRatio	TranscriptomeComplexity
SV00001_0010.fastq	21438811	10124865	9645137	0.95	0.003
SV00001_0009.fastq	18803262	6892775	6456391	0.94	0.004
SV00001_0008.fastq	11487537	1039777	596897	0.57	0.015
SV00001_0001.fastq	9820858	871565	507934	0.58	0.018
SV00001_0007.fastq	11152723	8071705	8044879	1	0.001
SV00001_0006.fastq	10273181	6887537	6867012	1	0.002
SV00001_0005.fastq	11455227	7775213	7748530	1	0.002
SV00001_0004.fastq	9034105	5510881	5451768	0.99	0.002
SV00001_0003.fastq	10761376	6470866	6397040	0.99	0.002
SV00001_0002.fastq	12445649	7097874	7021502	0.99	0.002

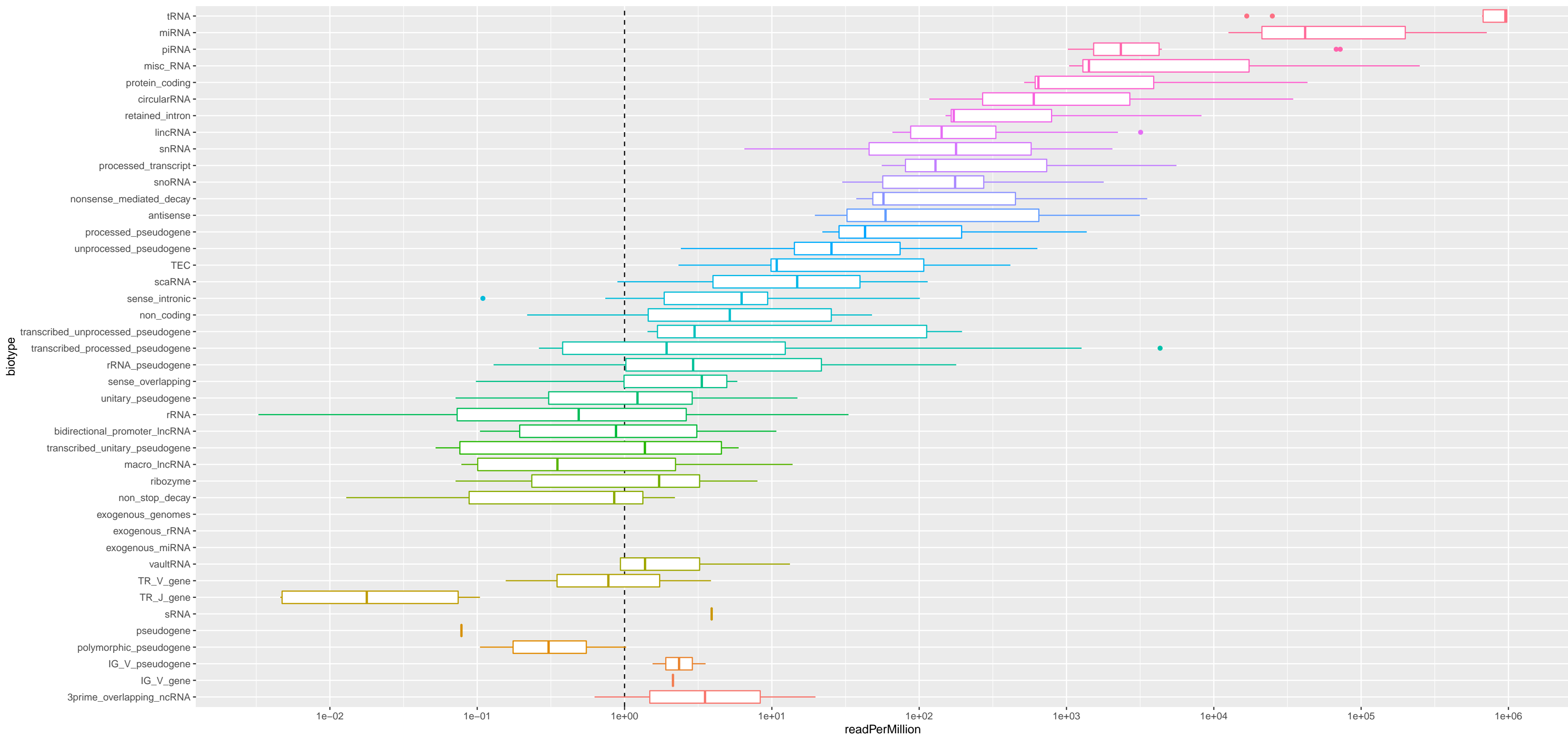
The graph provides per sample analysis in an illustrated table where transcriptome reads and transcriptome genome ratio in addition to transcriptome complexity. These values are highlighted to indicate if they do pass or fail QC.

Overall biotype distribution within the NGS library based on raw read counts



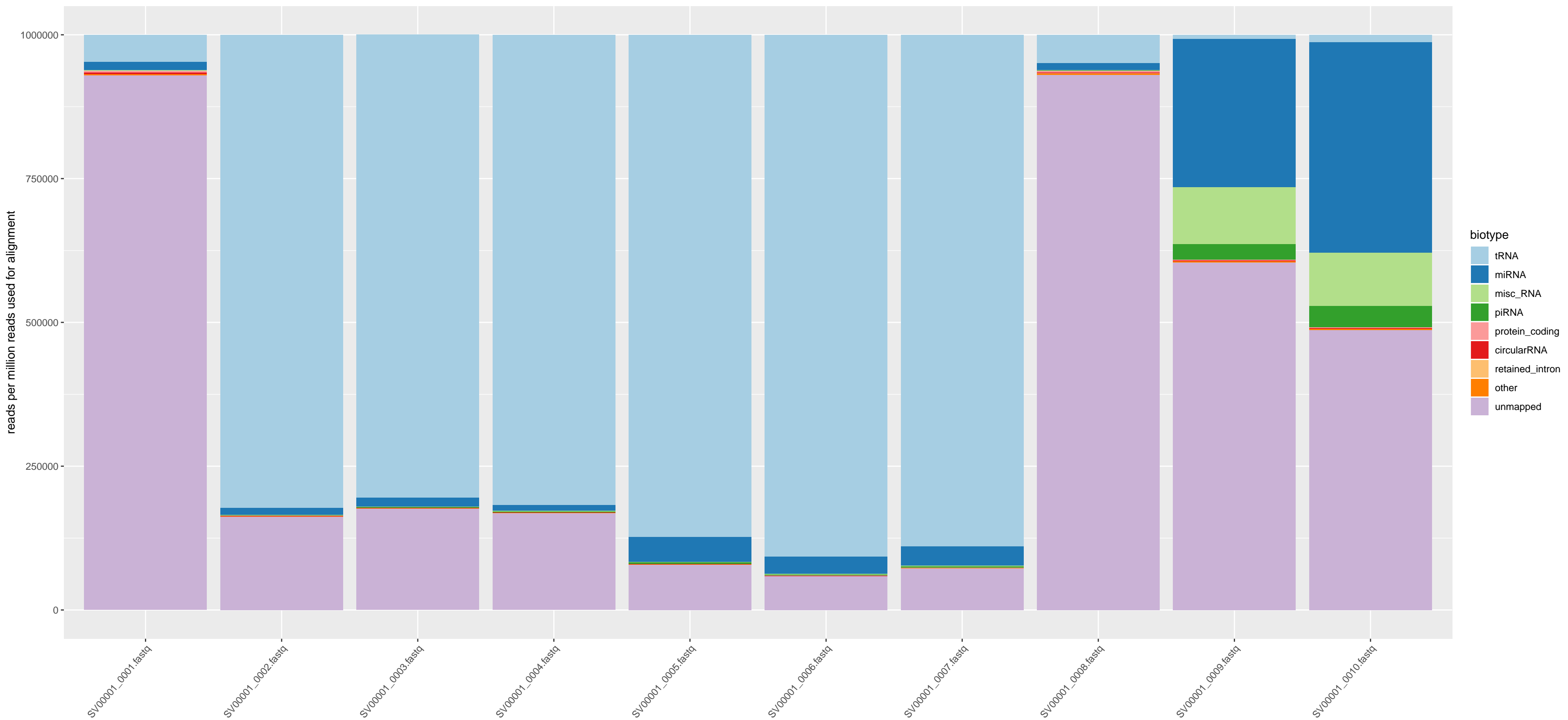
This analysis provides biotype fractions and their proportions across all samples in the NGS library.

Overall biotype distribution within the NGS library based on normalized counts



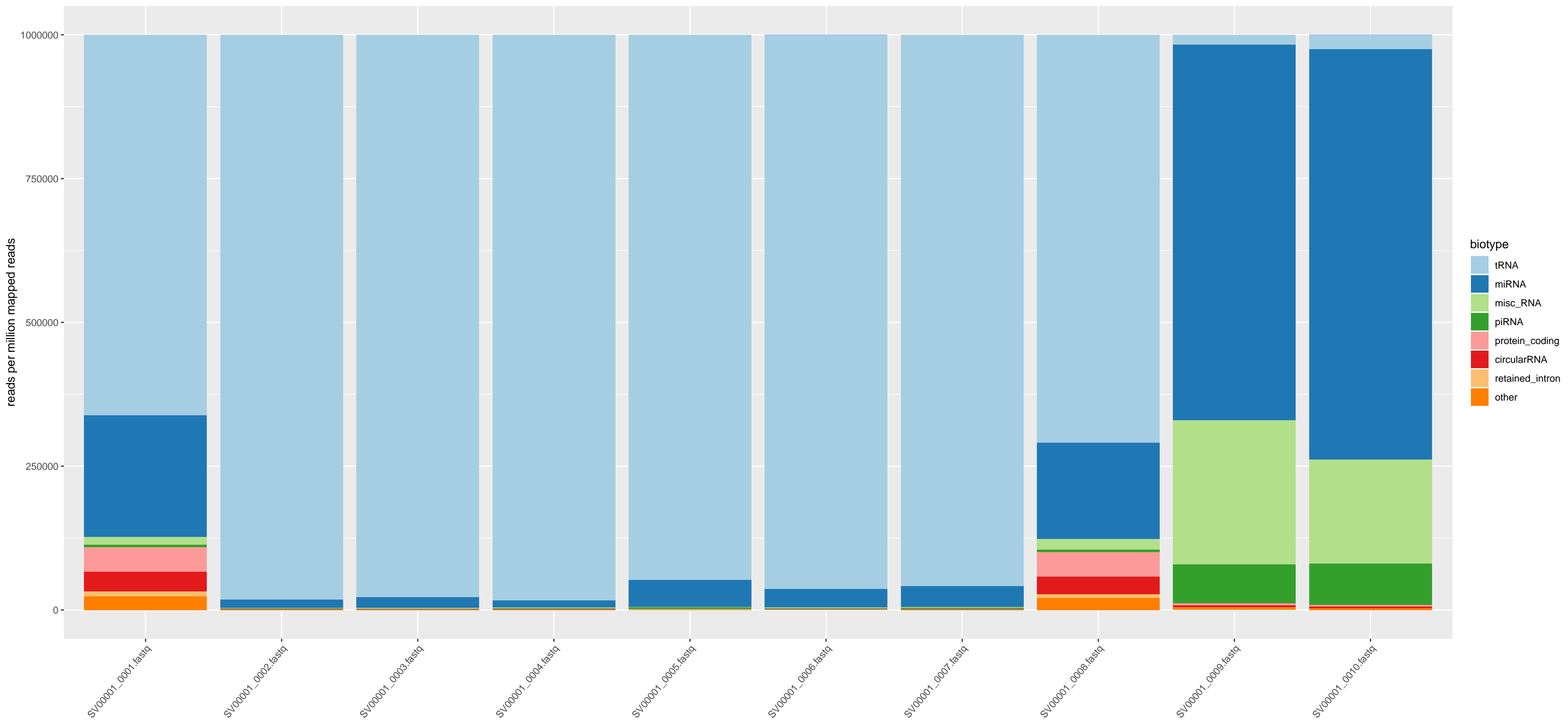
This analysis provides biotype fractions and their proportions across all samples in the NGS library.

Biotype distribution within the NGS library samples relative to reads used for alignment



This graph provides biotype fractions within each sample as well as between samples.

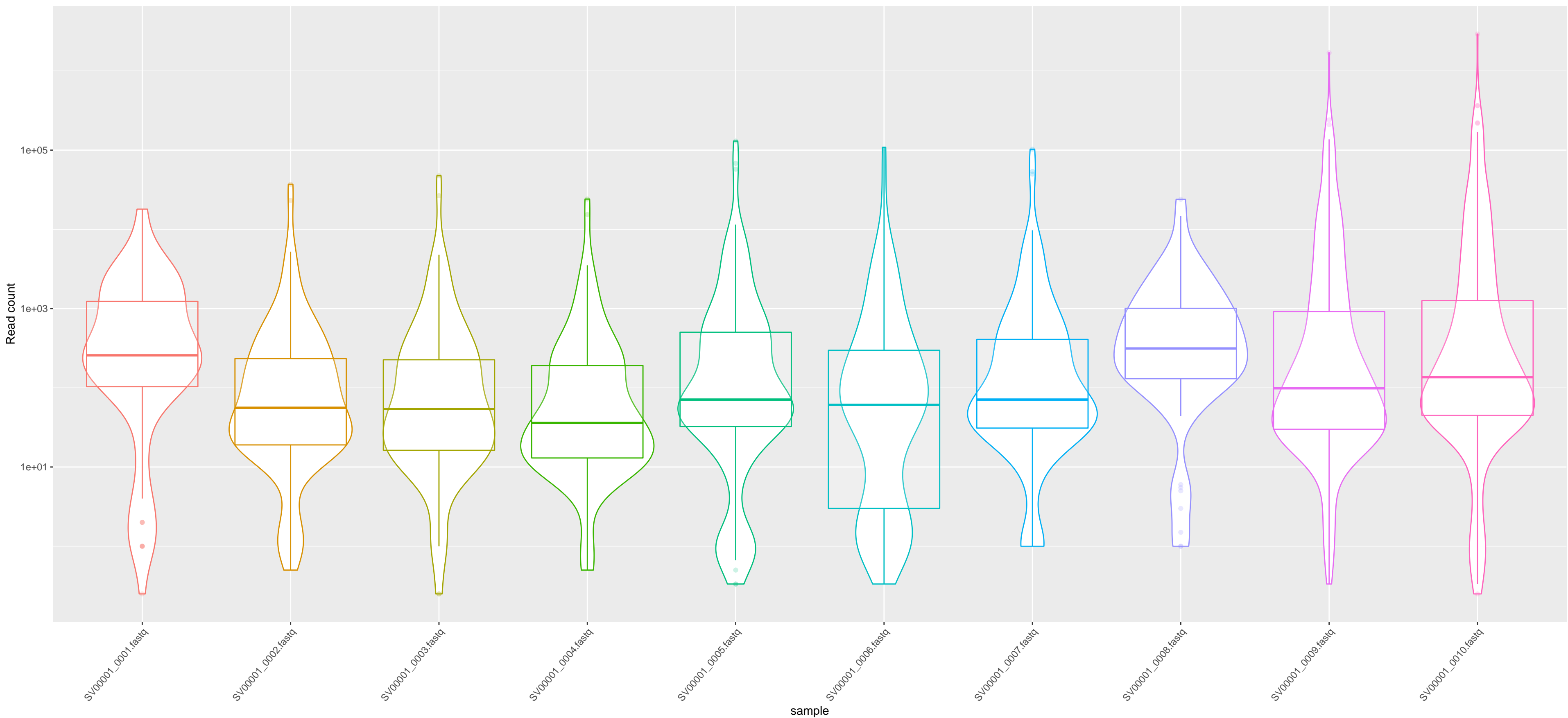
Biotype distribution within the NGS library samples relative to mapped reads



This graph provides biotype fractions within each sample as well as between samples.

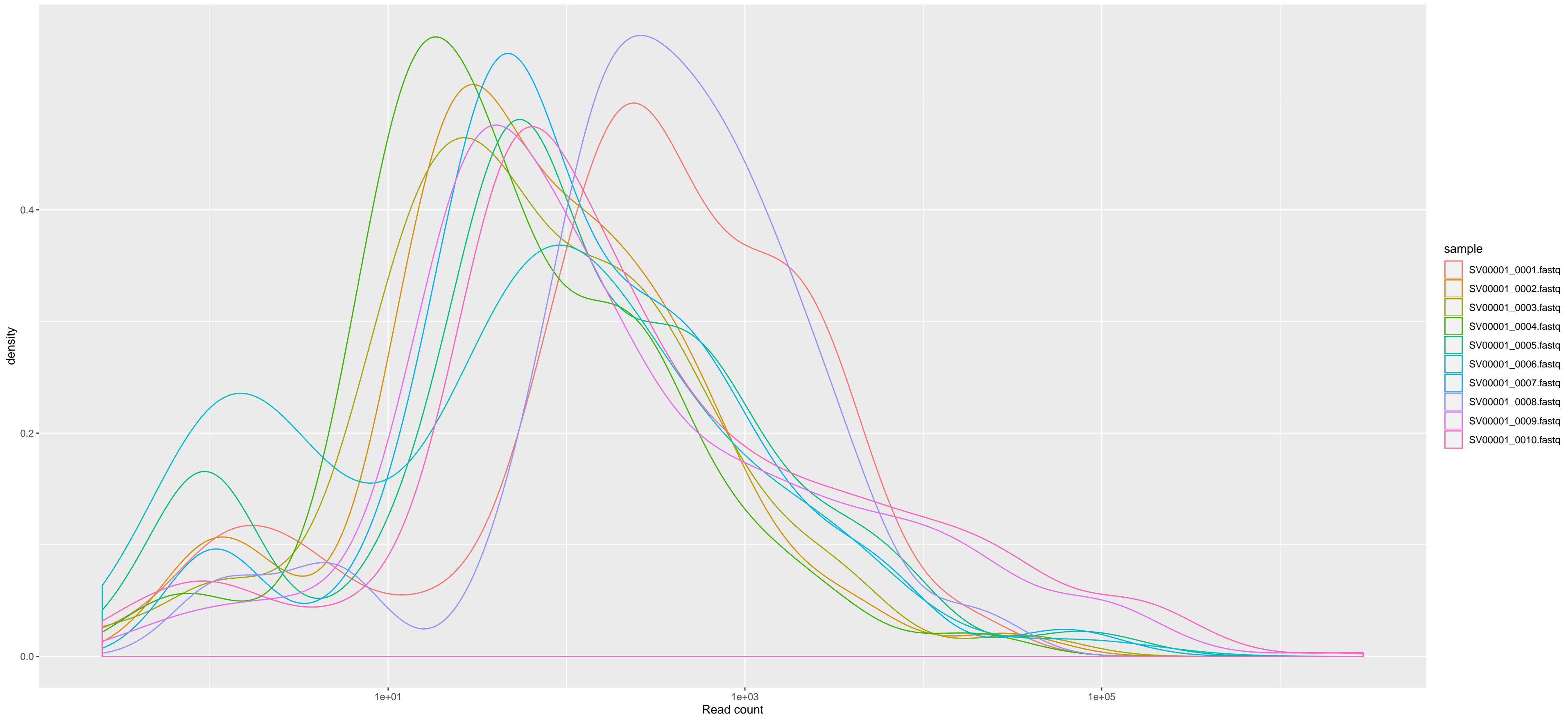


# Raw counts-based miRNA abundance distribution



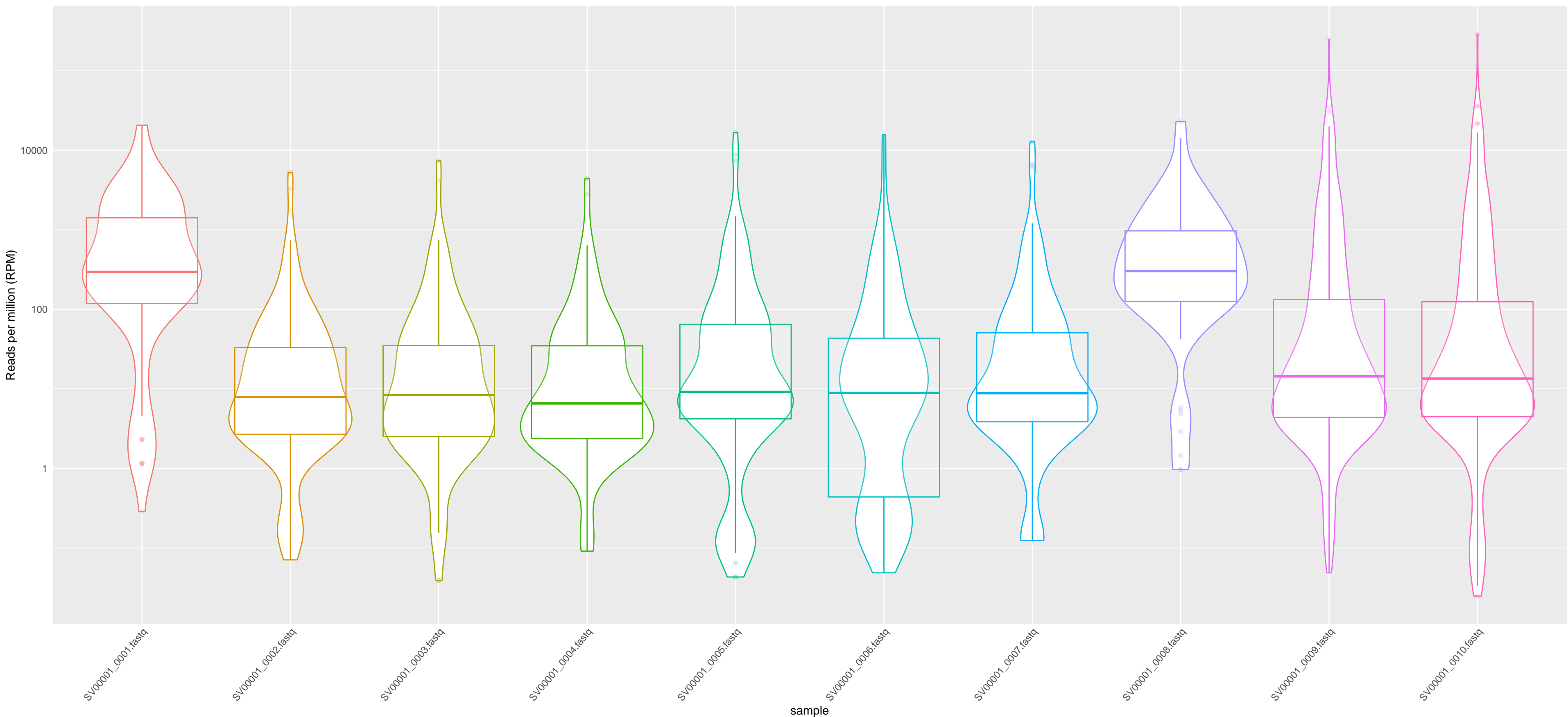
Abundance of miRNA categorizes the raw count of miRNAs within each sample, based on the number of miRNAs at each count. The normal distribution of the number of miRNAs in a sample is reversibly proportional to read counts.

# Normalized counts-based miRNA abundance distribution



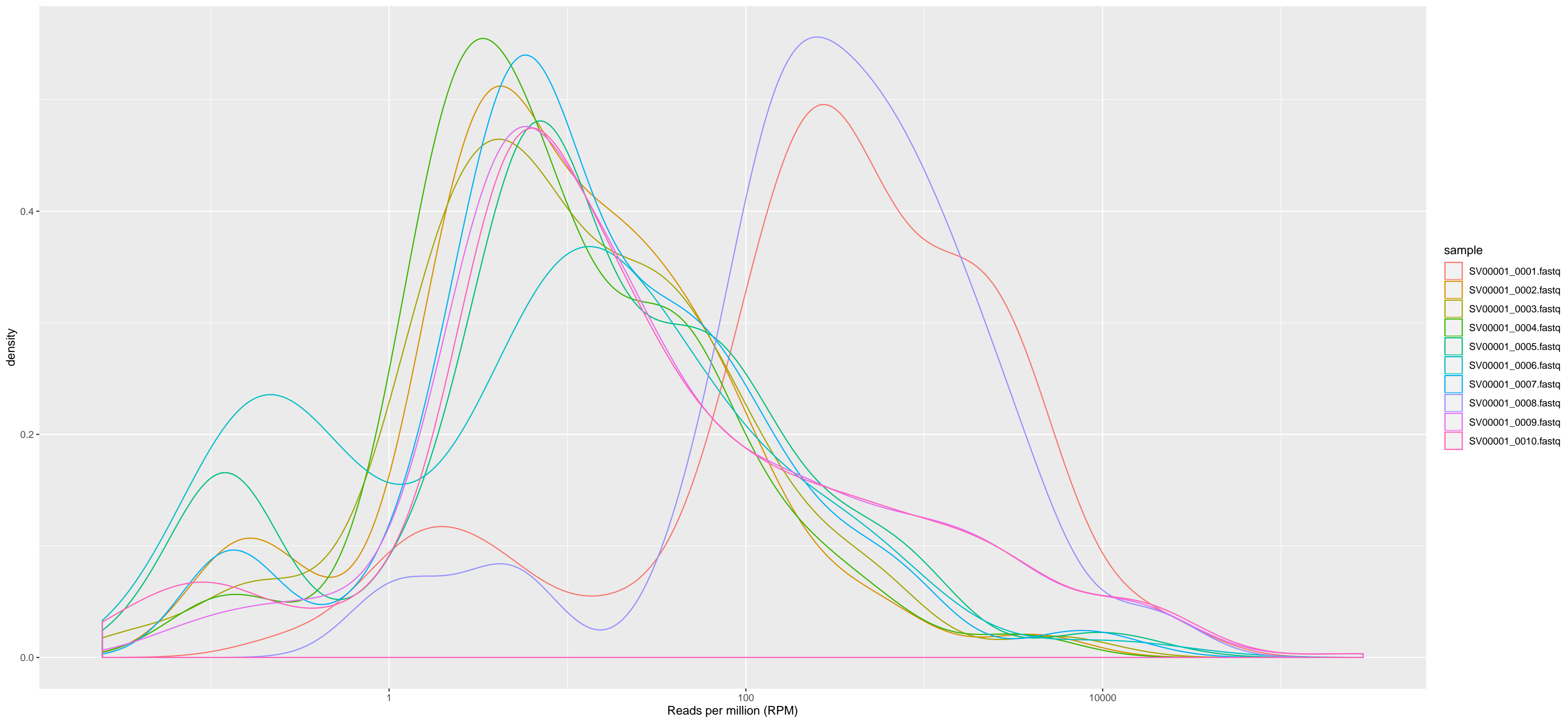
Abundance of miRNA categorizes the RPM normalized counts of miRNAs within each sample, based on the number of miRNAs at each count. The normal distribution of the number of miRNAs in a sample is reversibly proportional to read counts.

Density of miRNA abundance distribution based on raw counts



miRNA density in a sample is the number of miRNAs at a particular read count over the total number of miRNAs in a sample. Most miRNAs in a sample have low read counts and the density decreases at higher read counts.

Density of miRNA abundance distribution based on normalized counts



miRNA density in a sample is the number of miRNAs at a particular read count over the total number of miRNAs in a sample. Most miRNAs in a sample have low read counts and the density decreases at higher read counts.