

Report for dataset 3prime_v4_GEX_hPBMCs_20k_Rep2-Cell4 (all samples)

Dataset afaf4300-43ef-44d1-80ab-556caa61ea0e

Summary

Name	3prime_v4_GEX_hPBMCs_20k_Rep2-Cell4 (all samples)
Created At	2025-04-13 04:04:20.489
Number of Records	9,923,093
Total Length	149,398,176,863
Movie Name	m84039_250411_070102_s4
ICS Version	13.3.0.253824
Well Sample	3prime_v4_GEX_hPBMCs_20k_Rep2
Biological Sample	3prime_v4_GEX_hPBMCs_20k_Rep2
Barcode Name	bcM0003--bcM0003

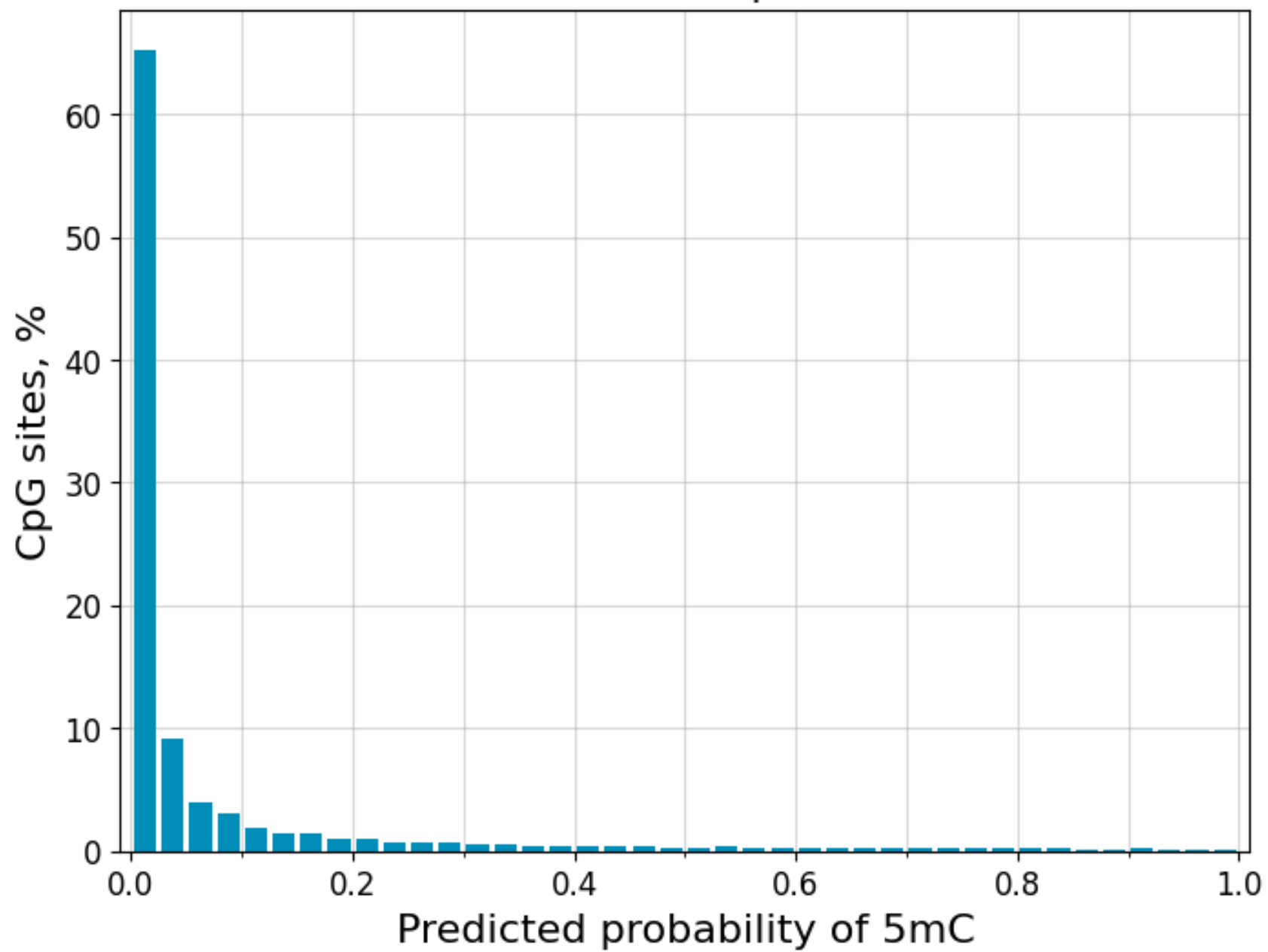
Methylation

Summary

Modification	Motif	Scored sites	Modified sites (Pr > 0.5)
5mC	CpG	98.4%	4.5%
6mA	A	0.5%	0.5%

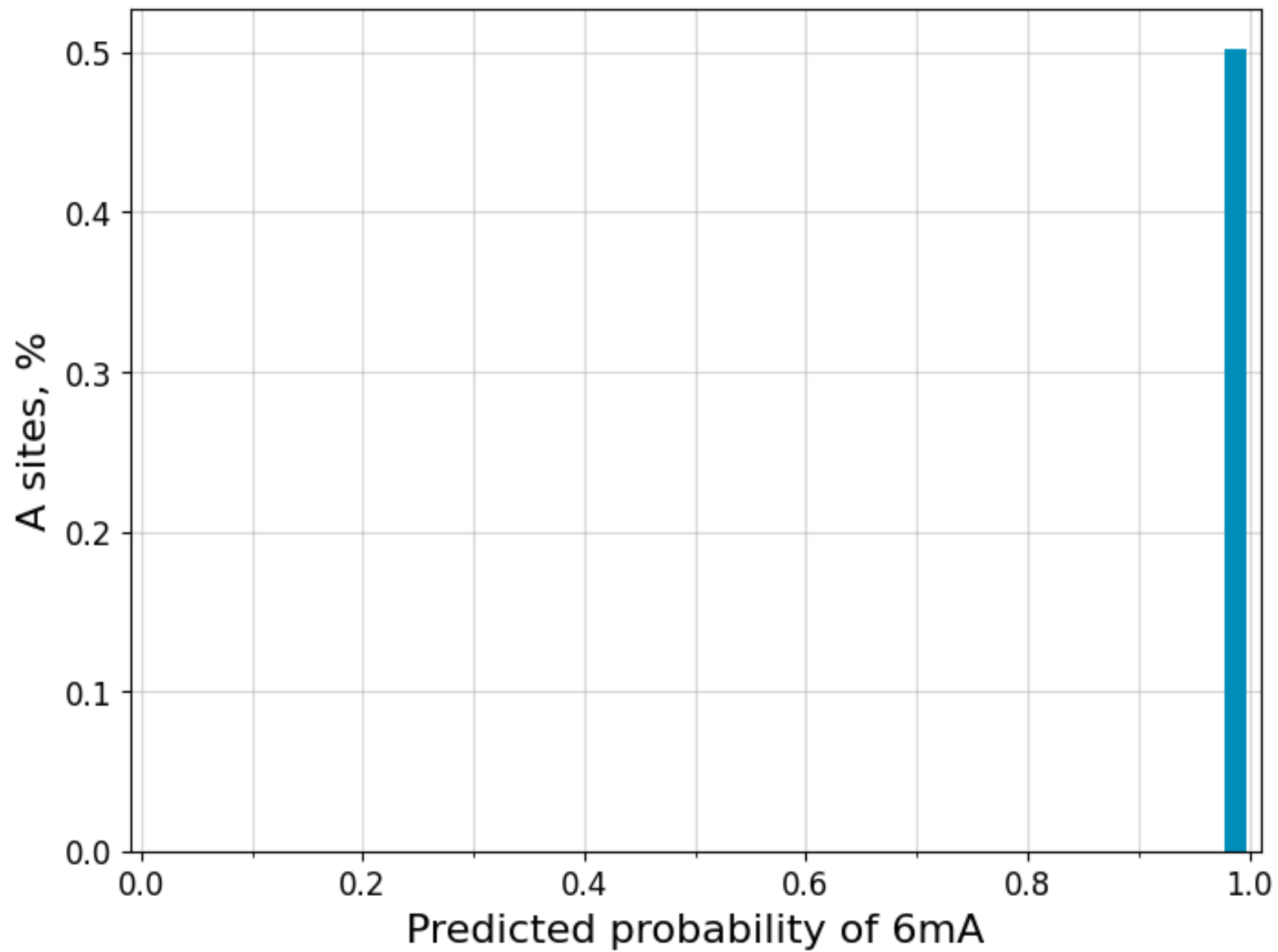
5mC methylation at CpG in reads

5mC at CpG



6mA methylation in reads

6mA

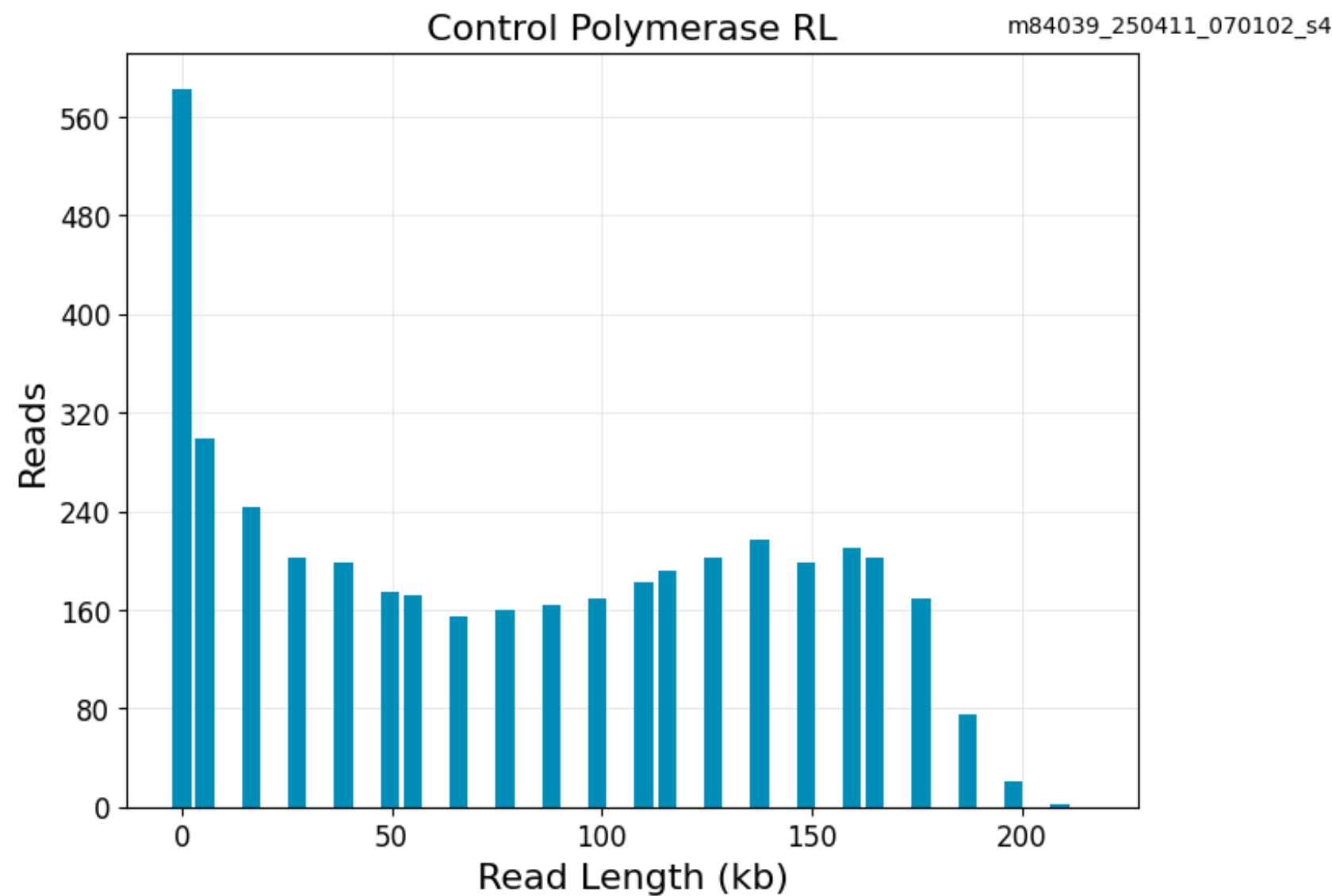


Control Report

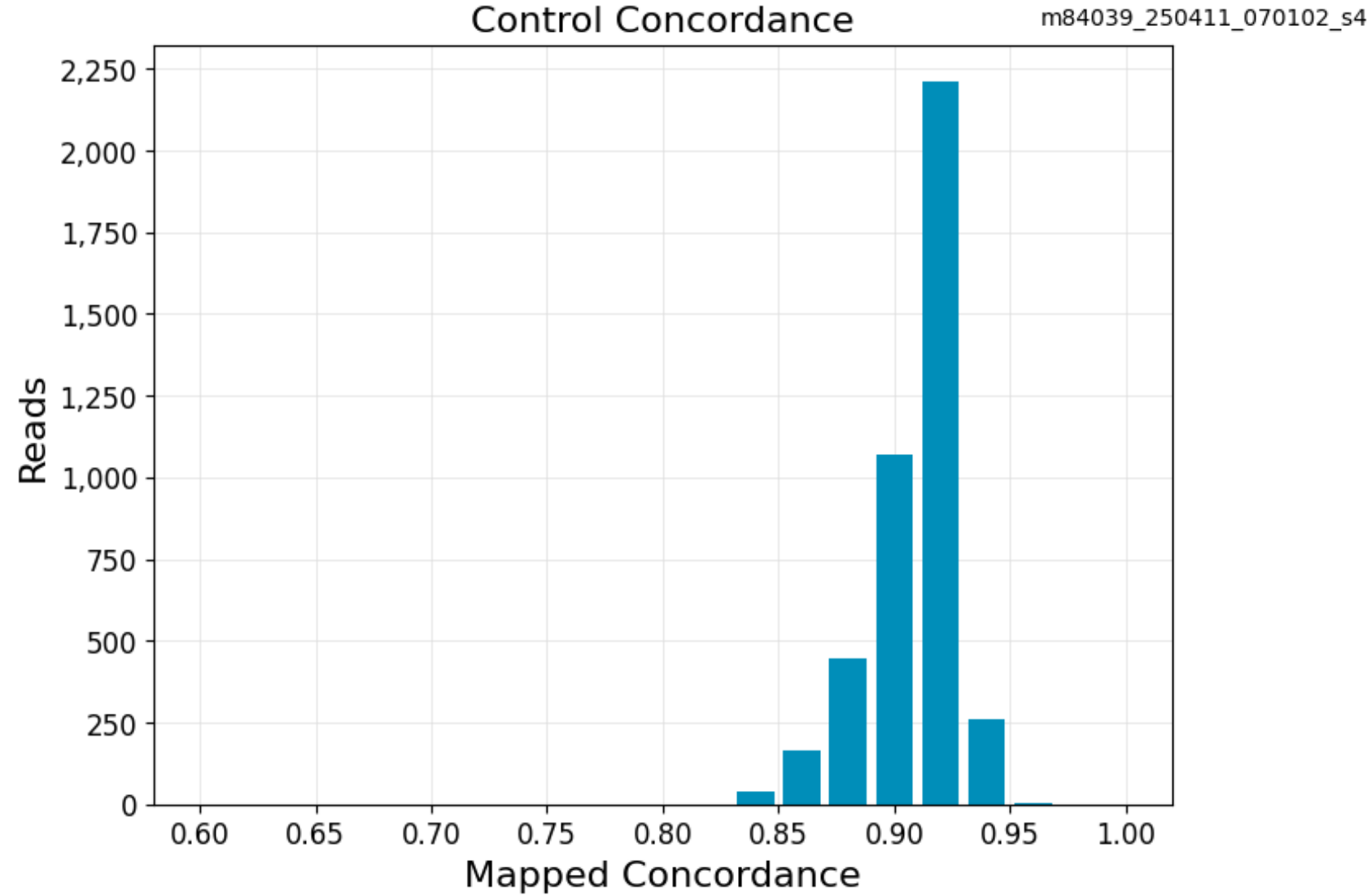
Summary

Number of Control Reads	4,195
Control Read Length Mean	82,248
Control Read Concordance Mean	0.92
Control Read Concordance Mode	0.93

Control Polymerase RL



Control Concordance



Barcodes

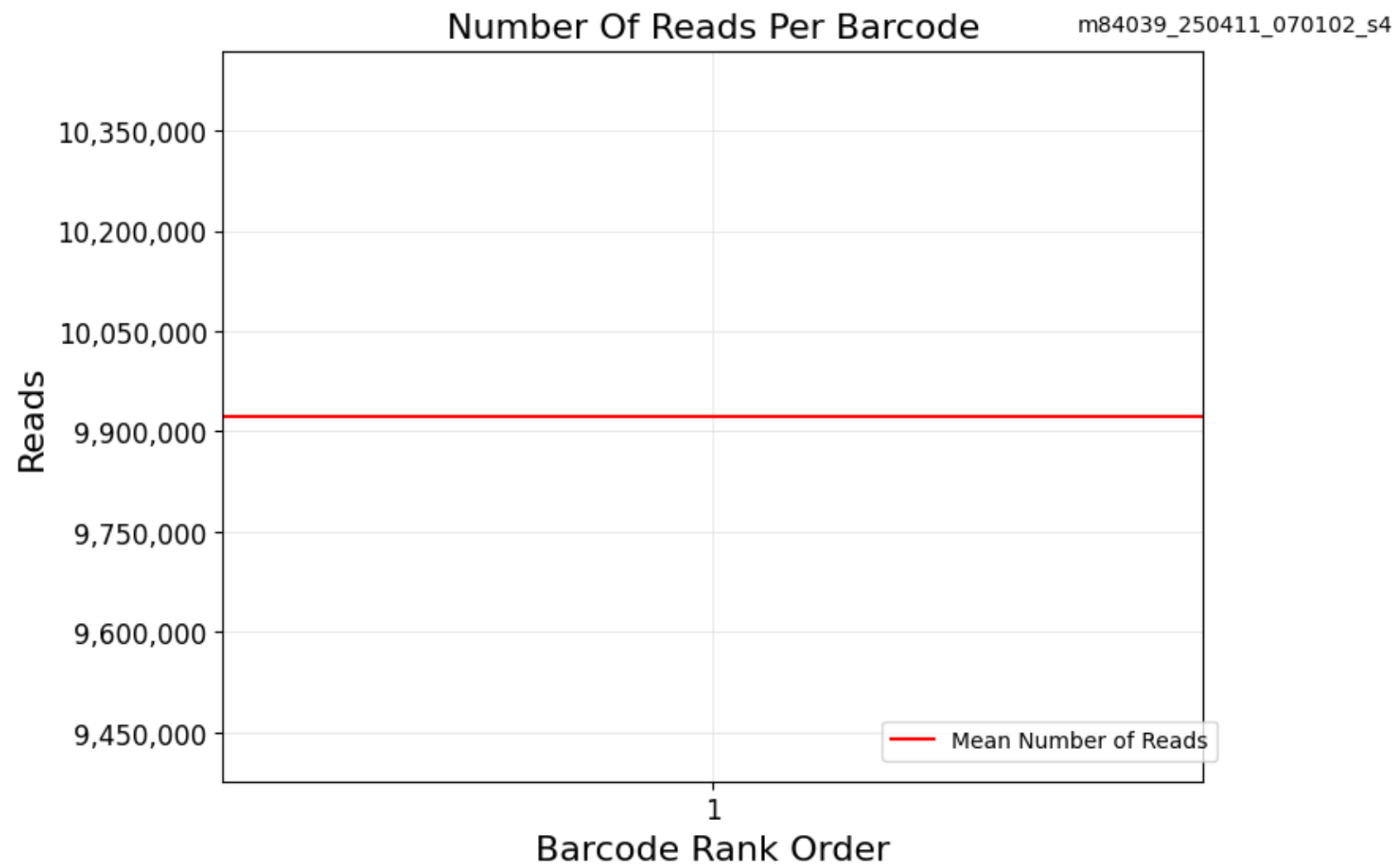
Summary

Unique Barcodes	1
Barcoded HiFi Reads	9,923,093
Unbarcoded HiFi Reads	19,911
Barcoded HiFi Reads (%)	99.80 %
Barcoded HiFi yield (Gb)	149.40 Gb
Unbarcoded HiFi yield (Gb)	0.23 Gb
Barcoded HiFi Yield (%)	99.85 %
Mean HiFi Reads per Barcode	9,923,093
Max. HiFi Reads per Barcode	9,923,093
Min. HiFi Reads per Barcode	9,923,093
Barcoded HiFi read length (mean, kb)	15.06 kb
Unbarcoded HiFi read length (mean, kb)	11.45 kb

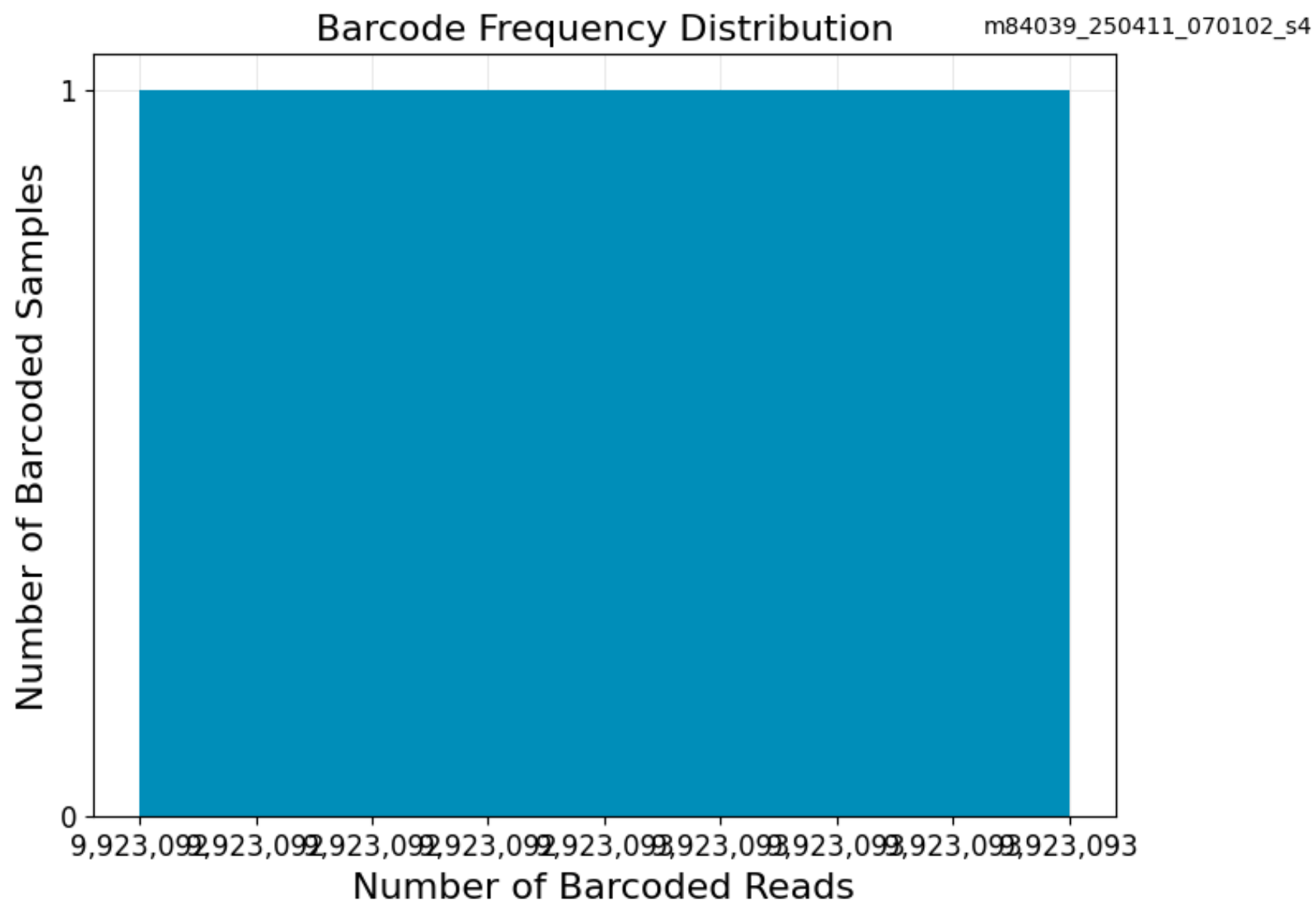
Barcode Data

Sample Name	Barcode	Barcode Quality	HiFi Reads	HiFi Read Length (mean, bp)	HiFi Read Quality (median, QV)	HiFi Yield (bp)	Polymerase Read Length (mean, bp)
3prime_v4_GEX_hPBMCs_20k_Rep2	bcM0003--bcM0003	97.3	9923093	15055	Q30	149398176863	141551
No Name	Not Barcoded	0.0	19911	11448	Q27	227956413	116616

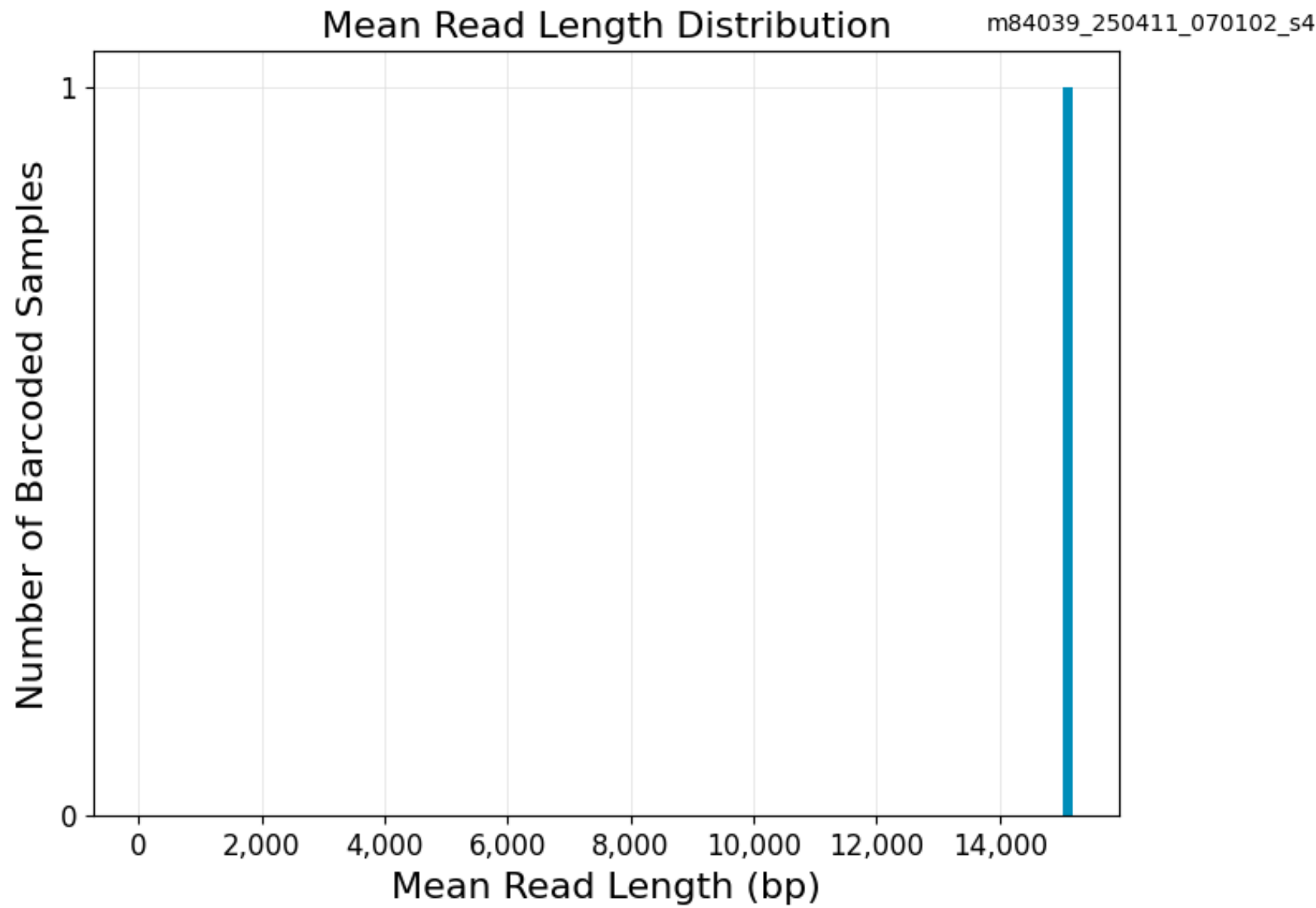
Number Of Reads Per Barcode



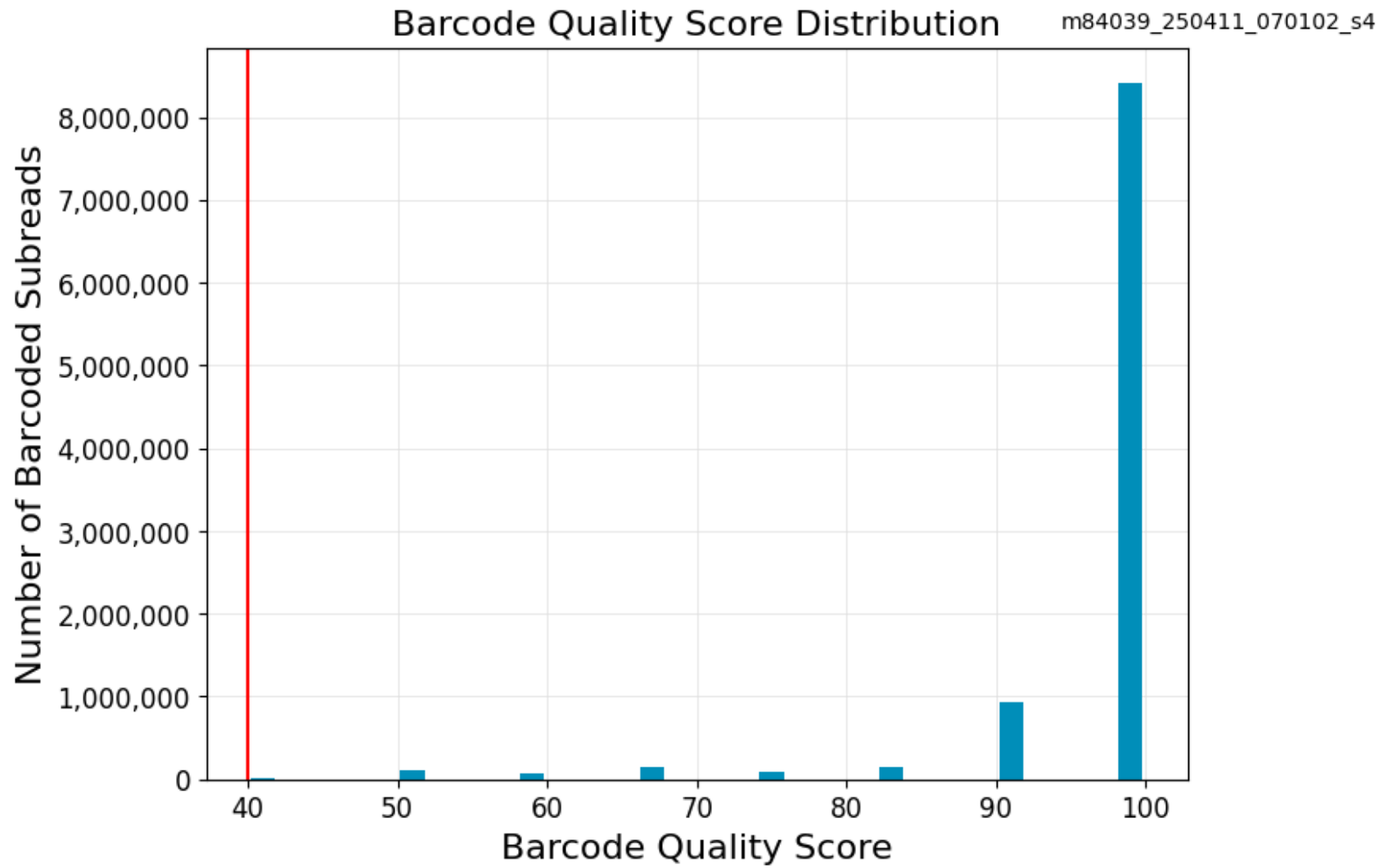
Barcode Frequency Distribution



Mean Read Length Distribution



Barcode Quality Score Distribution

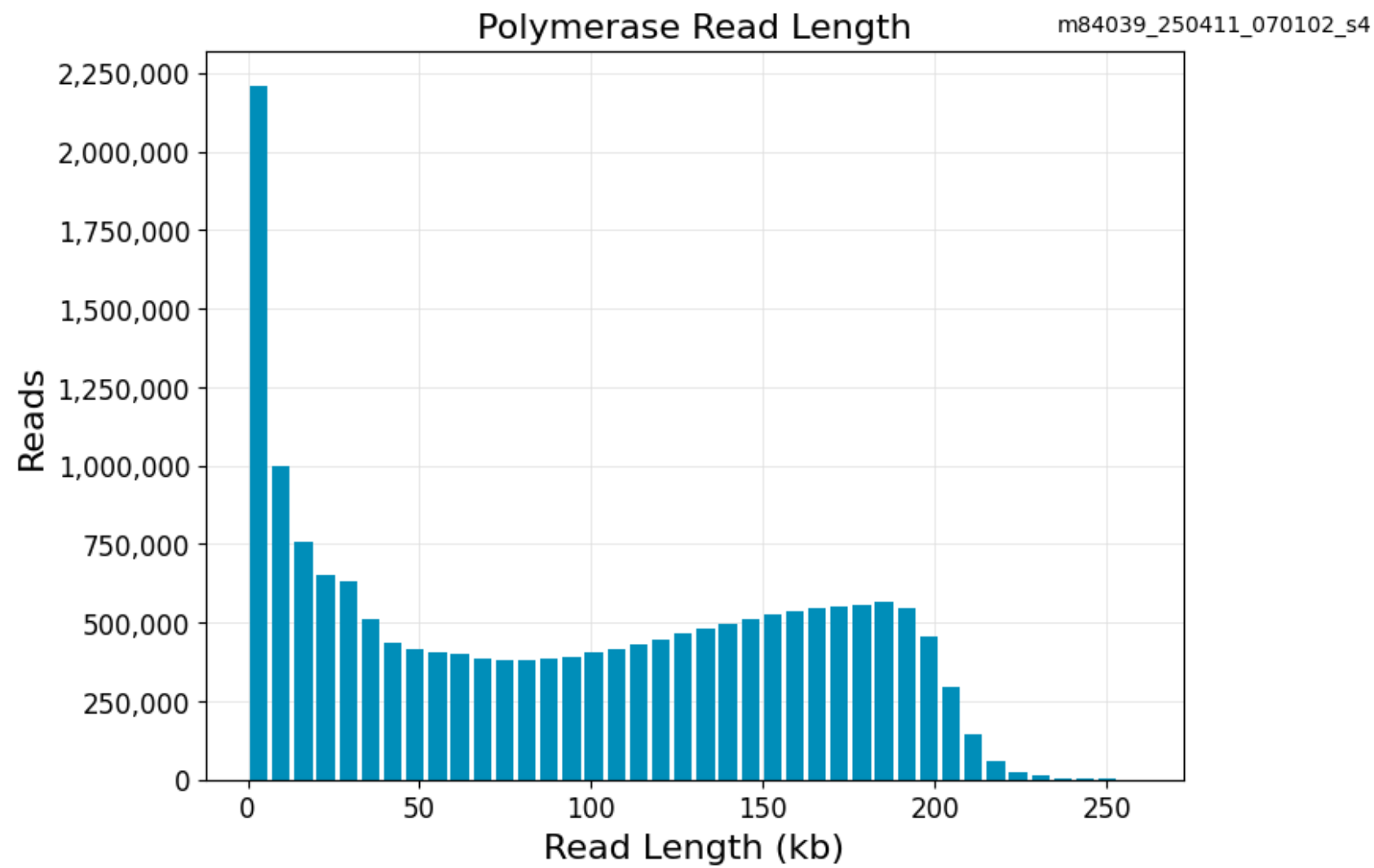


Raw Data Report

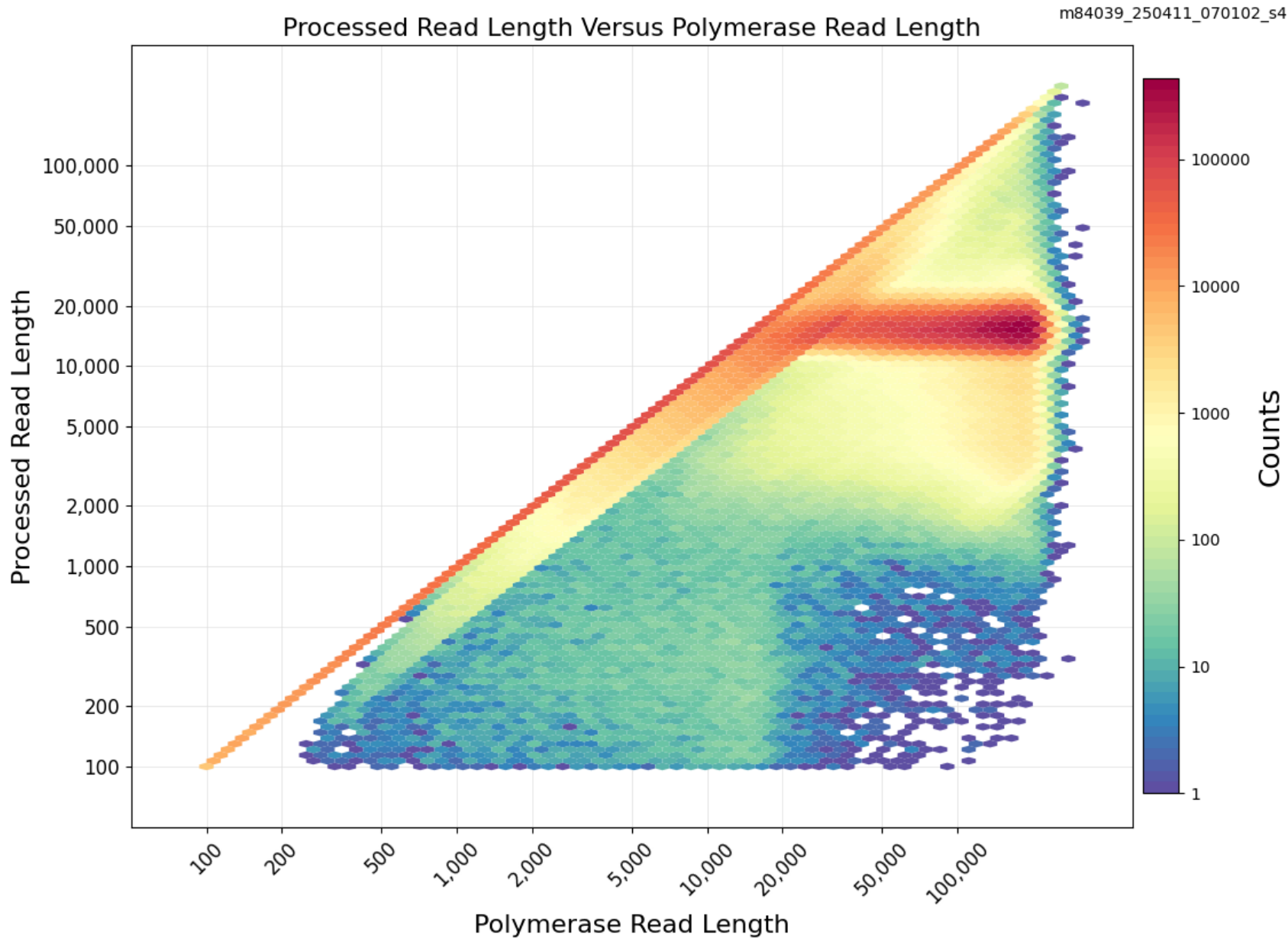
Summary

Polymerase Read Bases	1,579,126,659,511
Polymerase Reads	17,829,761
Polymerase Read Length (mean)	88.57 kb
Polymerase read length (N50)	153.25 kb
Polymerase read length longest subread length (mean)	15.98 kb
Polymerase read length longest subread length (N50)	16.75 kb
Unique Molecular Yield	274,690,605,056
Local Base Rate	2.36

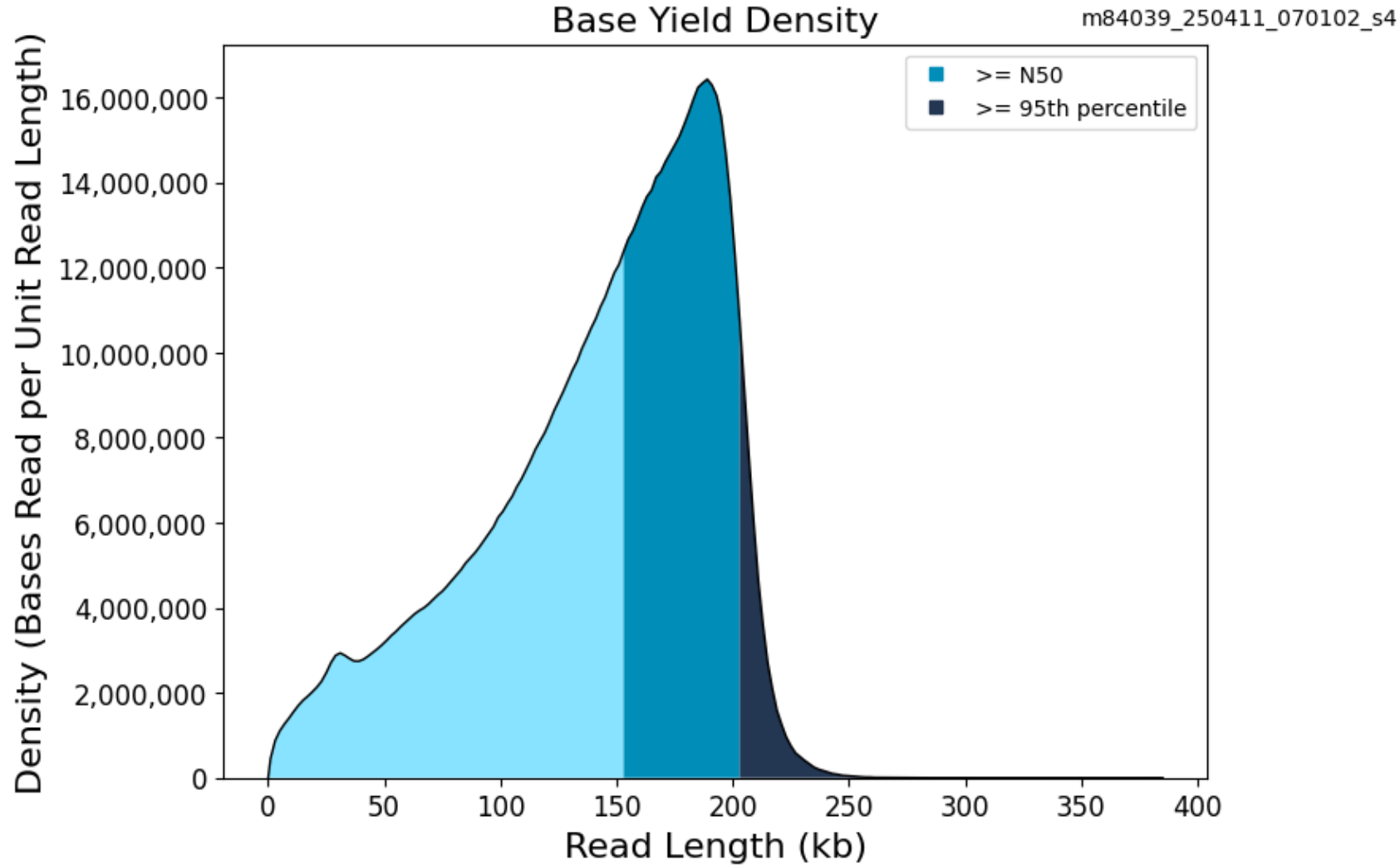
Polymerase Read Length



Longest Subread Length Versus Polymerase Read Length



Base Yield Density



CCS Analysis Report

Summary

HiFi reads	9.9 M
HiFi reads yield	149.63 Gb
HiFi reads length (mean)	15.05 kb
HiFi reads length (median, bp)	15,129
HiFi Read Length N50 (bp)	15,391
HiFi Read Quality (median)	Q29
HiFi Read Quality (median)	29
Base Quality Q30 (%)	95.75%
HiFi Number of Passes (mean)	9
Missing adapters (%)	3.66%

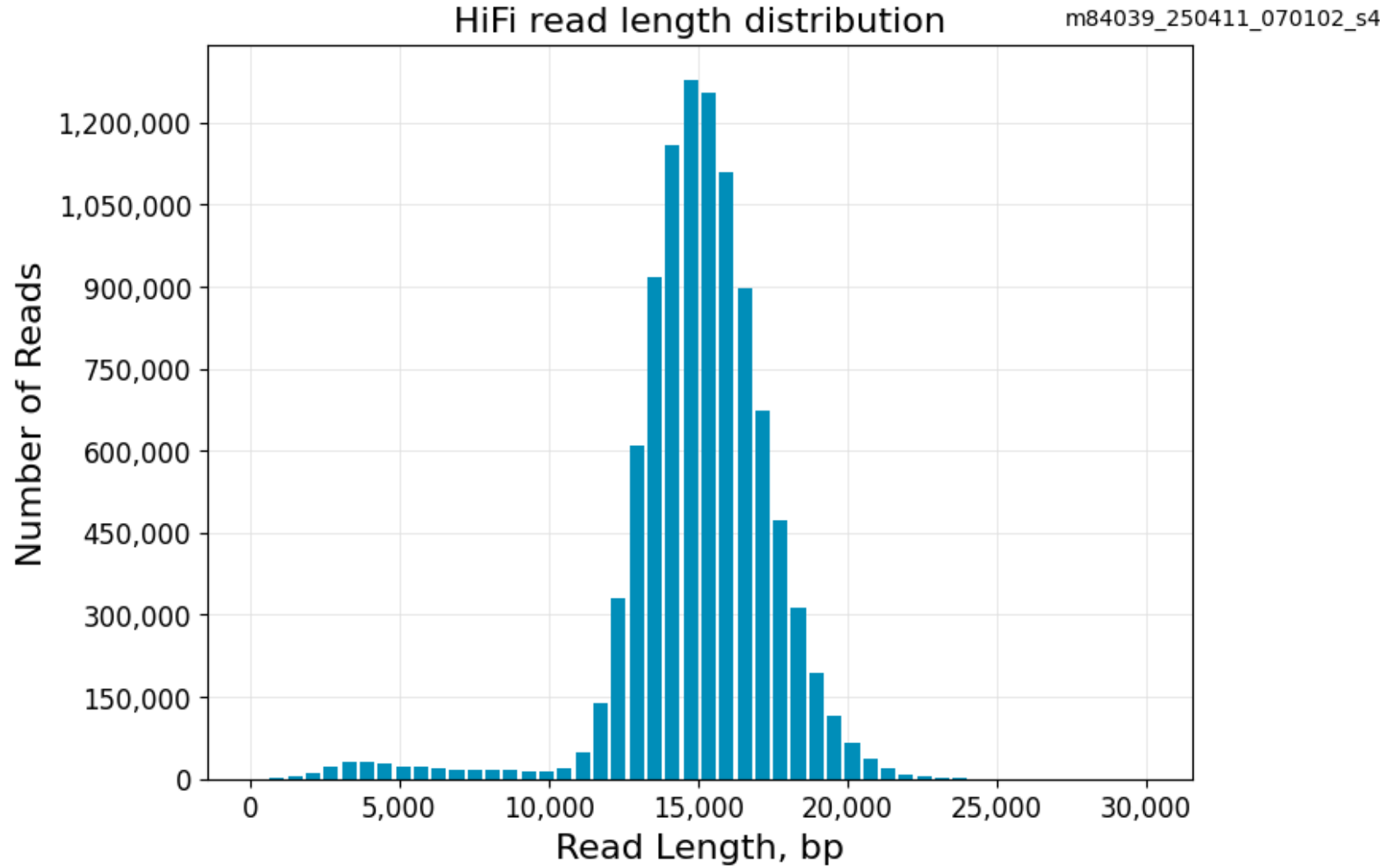
HiFi Read Length Summary

Read Length (kb)	Reads	Reads (%)	Yield (Gb)	Yield (%)
0	9943004	100	149.63 Gb	100
5,000	9807350	99	149.15 Gb	100
10,000	9659383	97	148.07 Gb	99
15,000	5248291	53	87.41 Gb	58
20,000	118939	1	2.48 Gb	2
25,000	273	0	0.01 Gb	0
30,000	16	0	0.00 Gb	0
35,000	1	0	0.00 Gb	0
40,000	1	0	0.00 Gb	0

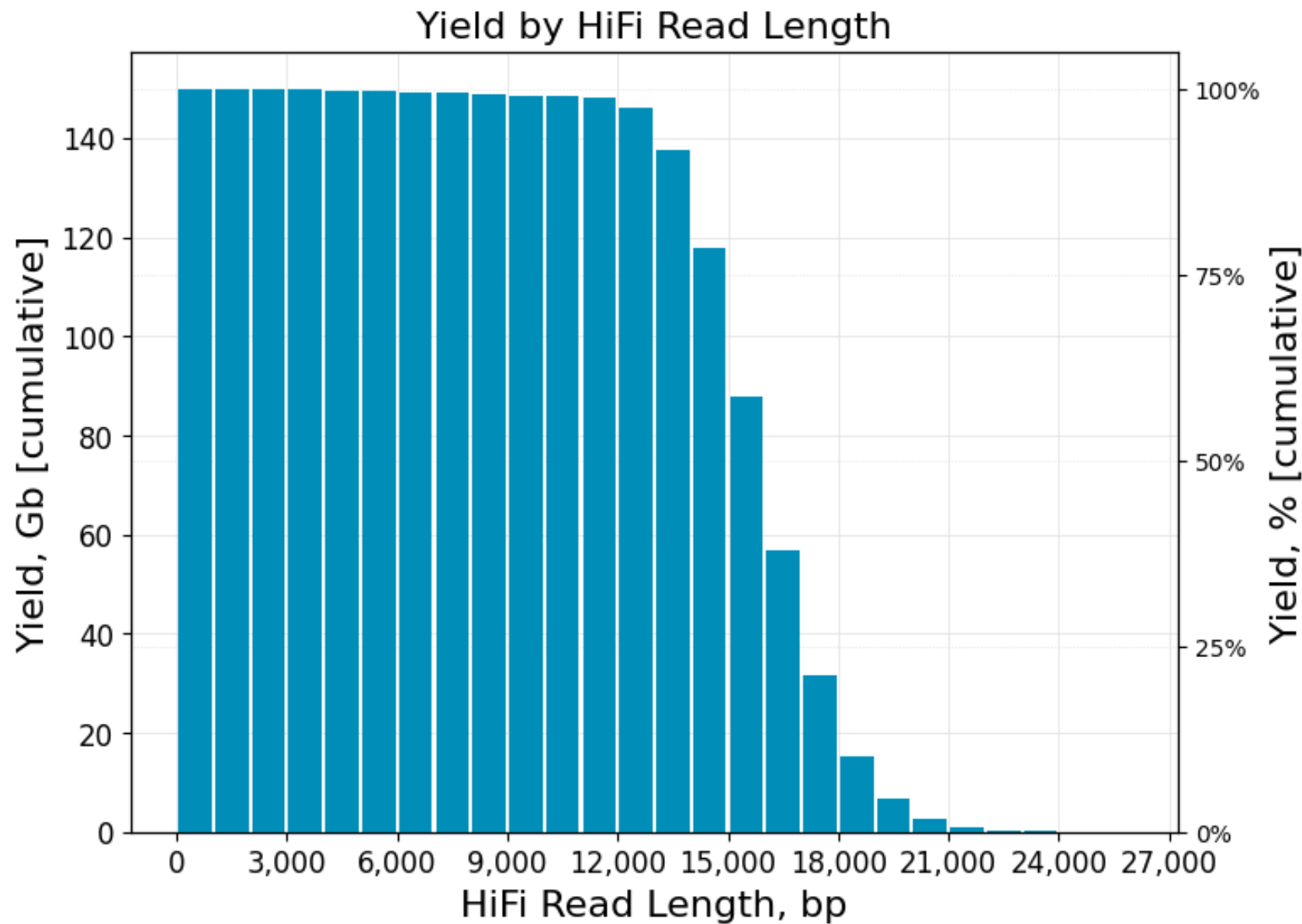
HiFi Read Quality Summary

Read Quality (Phred)	Reads	Reads (%)	Yield (Gb)	Yield (%)
Q20	9943004	100	149.63 Gb	100
Q30	4497738	45	66.61 Gb	45
Q40	82521	1	0.45 Gb	0
Q50	21417	0	0.07 Gb	0

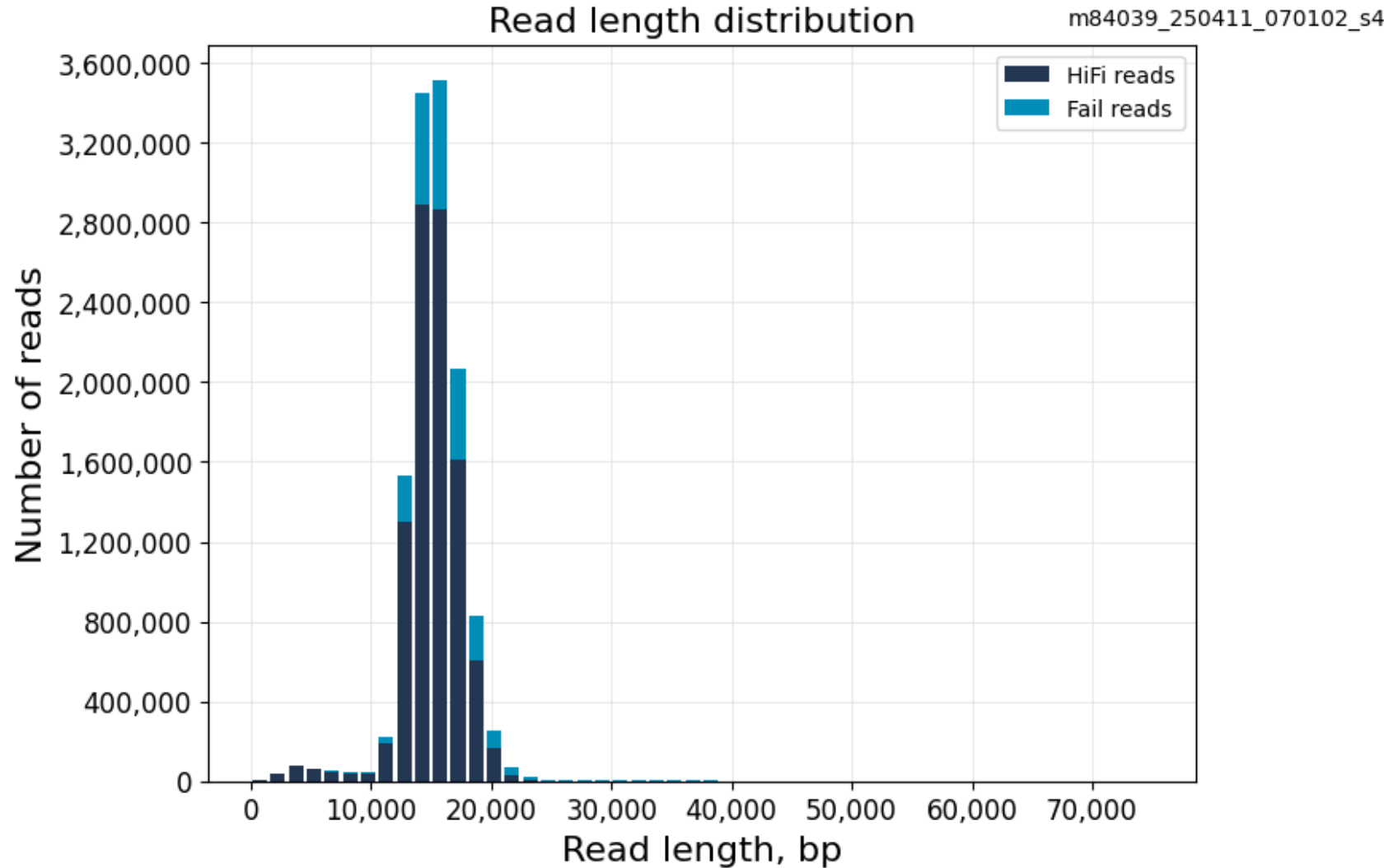
HiFi read length distribution



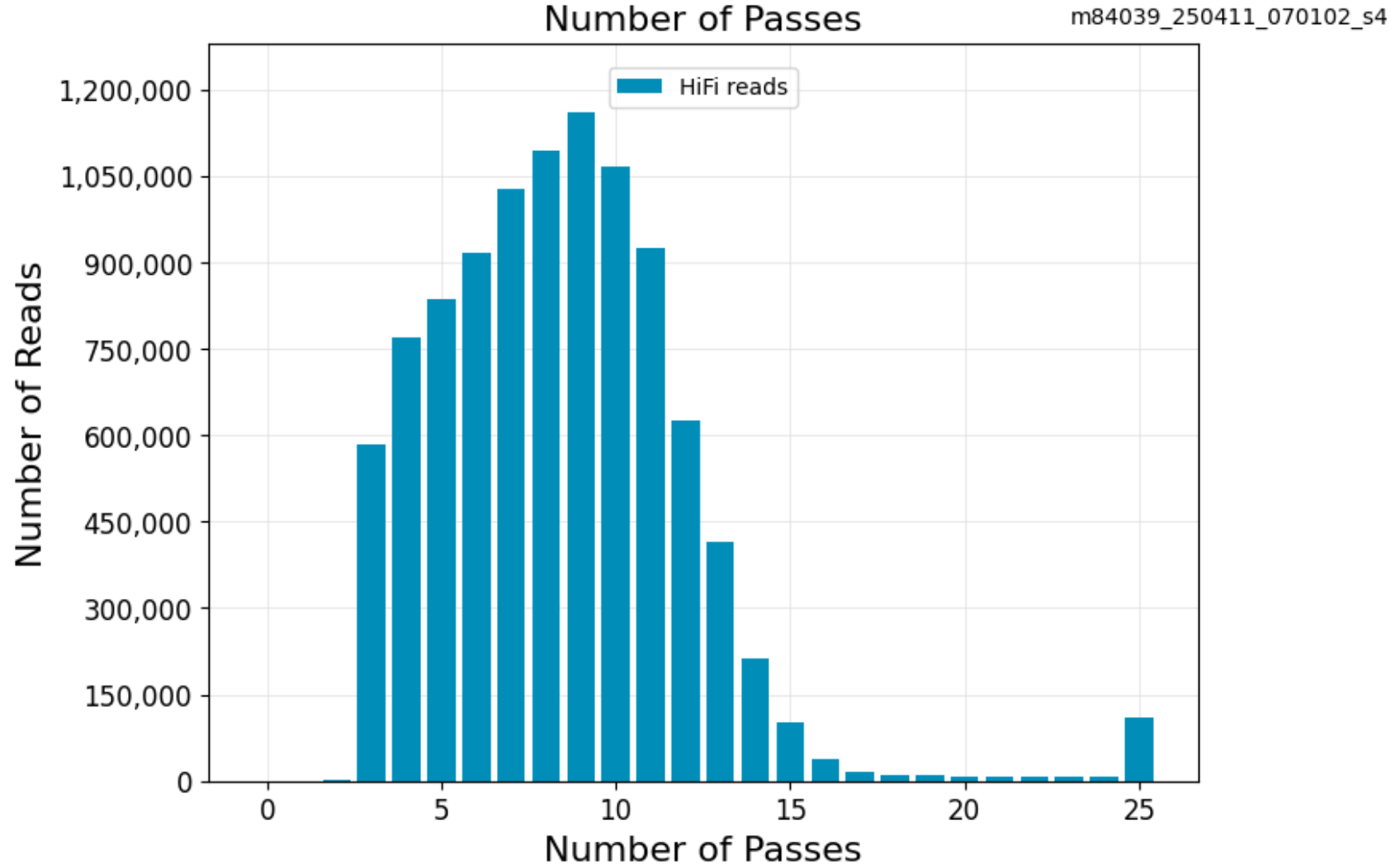
Yield by HiFi Read Length



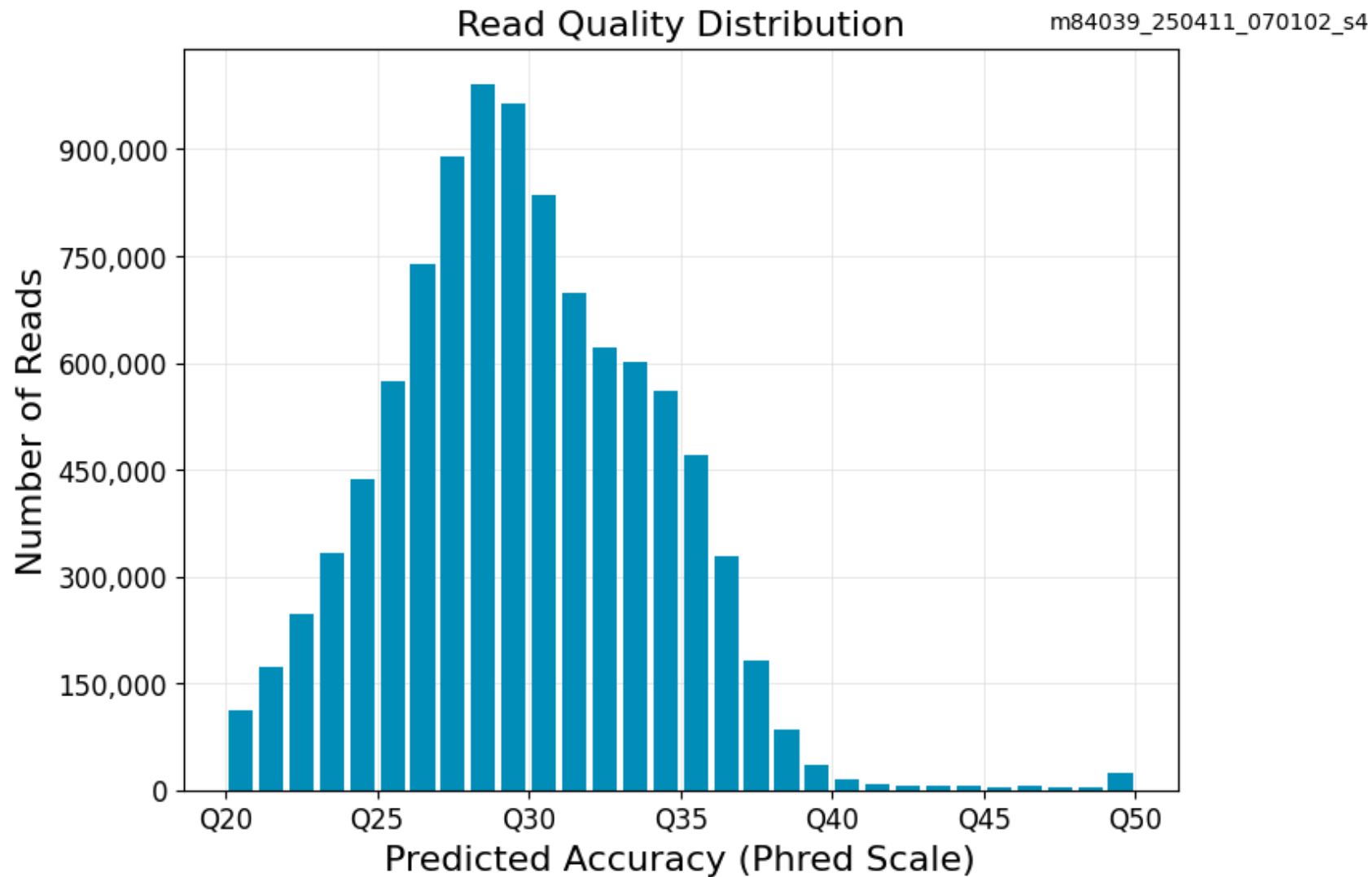
Read length distribution



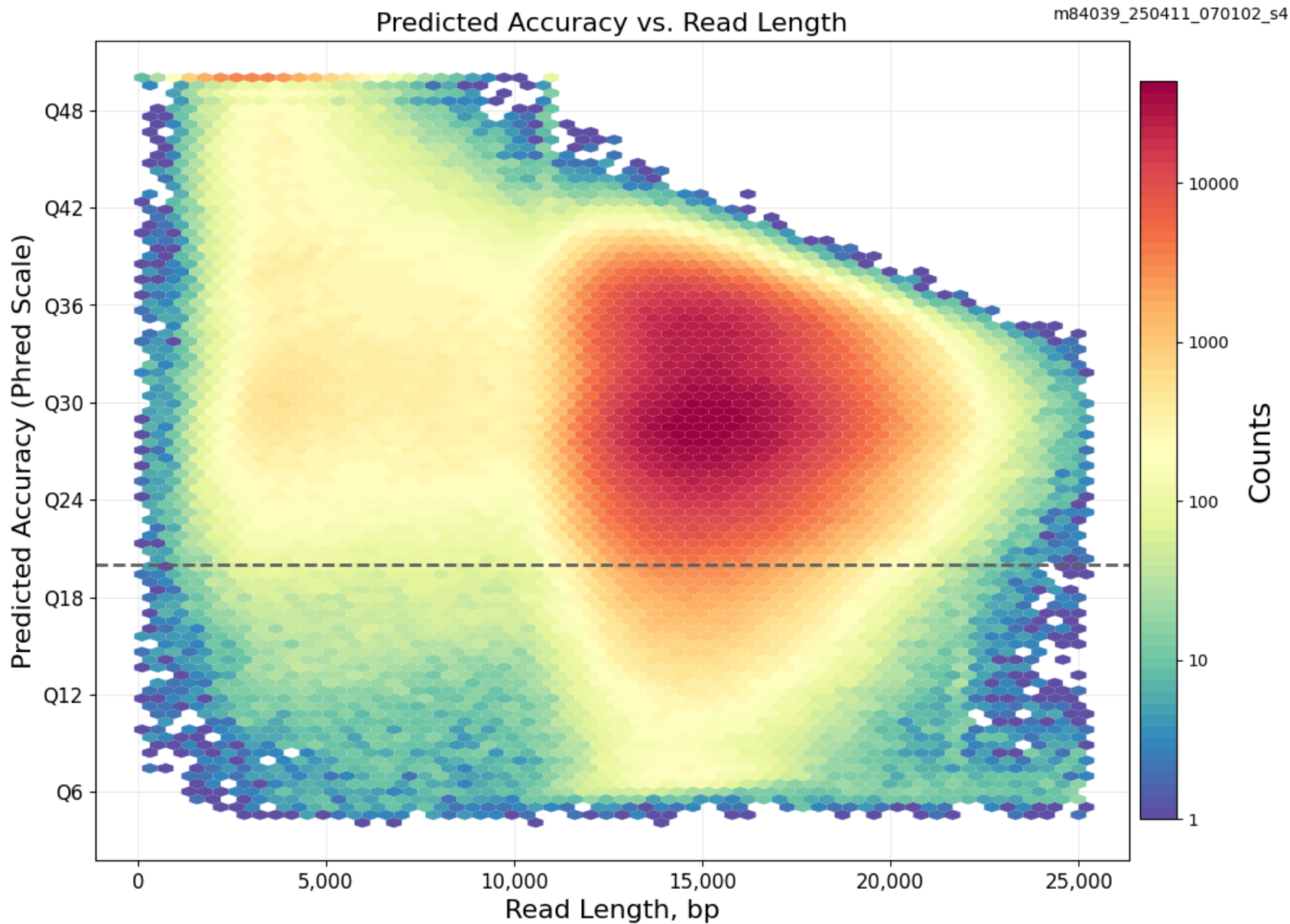
Number of Passes



Read Quality Distribution



Predicted Accuracy vs. Read Length



Loading Report

Summary

Productive ZMWs	25,165,824
Productivity 0	7,239,555
Productivity 1	17,833,956
Productivity 2	92,313

No Sample Setup found

No instrument run found

Parent jobs (1)

Job 124484

Summary

Job Type	import-dataset
Name	import-dataset
Comments	Description for job Import PacBio DataSet
Created At	2025-08-01 13:55:43.935
SMRT Link Version	25.3.0.273777

Child jobs (1)

Job 124626

Summary

Job Type	analysis
Pipeline	cromwell.workflows.pb_segment_reads_and_sc_isoseq
Name	KinnexRelease-10x-3pGEX-PBMC-20k-rep2-2RevioCell-largeMem
Comments	Description for job Run Analysis Application
Created At	2025-08-02 13:27:51.373
SMRT Link Version	25.3.0.273777

Other inputs

File Type	Name	Dataset ID	Number of Records
ReferenceSet	Human Genome hg38, with Gencode v39 annotations	ba3866bf-2aba-7c99-0570-0d6709174e4a	195
BarcodeSet	MAS-Seq Adapter v1 (MAS16)	10ce52cd-91c5-4bb9-4cac-2233832ecb12	17
BarcodeSet	10x Chromium single cell 3' cDNA primers	3332ee2d-56dc-07b0-fbcc-d7f6158a7a65	2
ConsensusReadSet	3prime_v4_GEX_hPBMCs_20k_Rep2-Cell5 (all samples)	4f092d77-35f9-46b0-9f69-83cdaae50f04	9,933,667

Output files

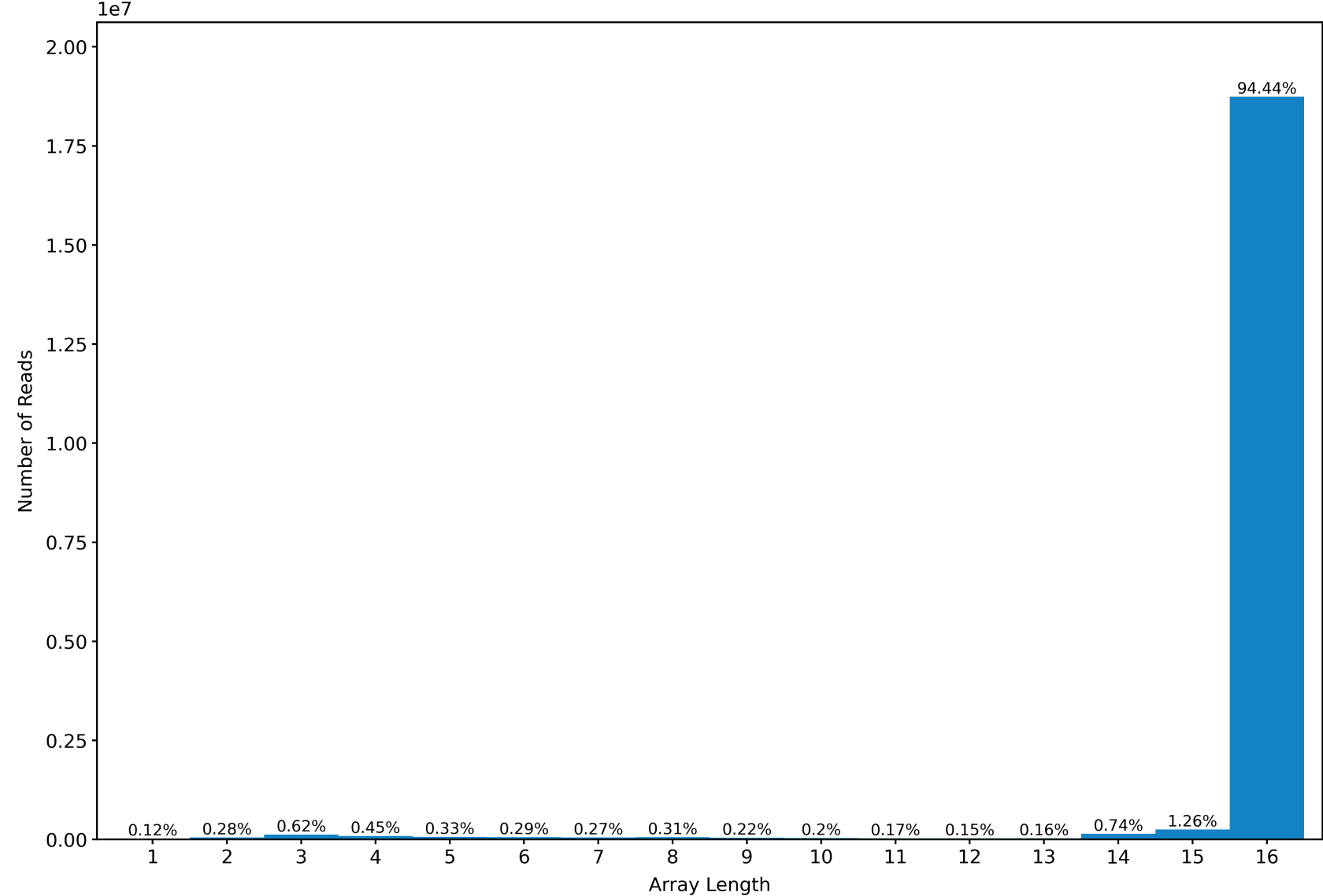
File name	Size	Type
KinnexRelease-10x-3pGEX-PBMC-20k-rep2-2RevioCell-largeMem Segmented Reads	47,593	ConsensusReadSet
Deduplicated reads after cell barcode correction, unmapped, BAM	66,363,407,179	bam
Single-cell isoform and gene matrix, tar-gzipped	6,241,434,204	tgz
Segmented Reads, passing, unaligned	96,816,532,776	bam
Segmented Reads, passing, unaligned	97,416,025,798	bam
Non-passing reads, unaligned	2,806,813,423	bam
Non-passing reads, unaligned	2,861,753,534	bam
Unique mapped transcripts, GFF	3,775,021,387	gff
Unique mapped transcripts, filtered, GFF	1,510,597,544	gff
Deduplicated reads after cell barcode correction, mapped, BAM	37,014,094,479	bam
Deduplicated reads after cell barcode correction, mapped, BAM index	47,528,224	bam_bai
Unique mapped transcripts, junctions TXT	2,685,857,336	txt
Unique mapped transcripts, filtered, junctions TXT	1,372,277,856	txt
Deduplicated transcripts	49,105,598,235	Fasta
Collapsed transcript groups	3,882,175,269	txt
Unique mapped transcripts, classification TXT	2,851,792,814	txt
Unique mapped transcripts, filtered, classification TXT	930,326,848	txt
bcstats_report.tsv.gz	17,866,210	tsv
Analysis Log	93,709	log
737K_august_2016.txt.gz	2,238,609	txt
3M_3pgex_may_2023.REVCOMP.txt.gz	19,830,613	txt
Analysis Log	4,091	log
SMRT Link Log	20,052	log

Read Segmentation

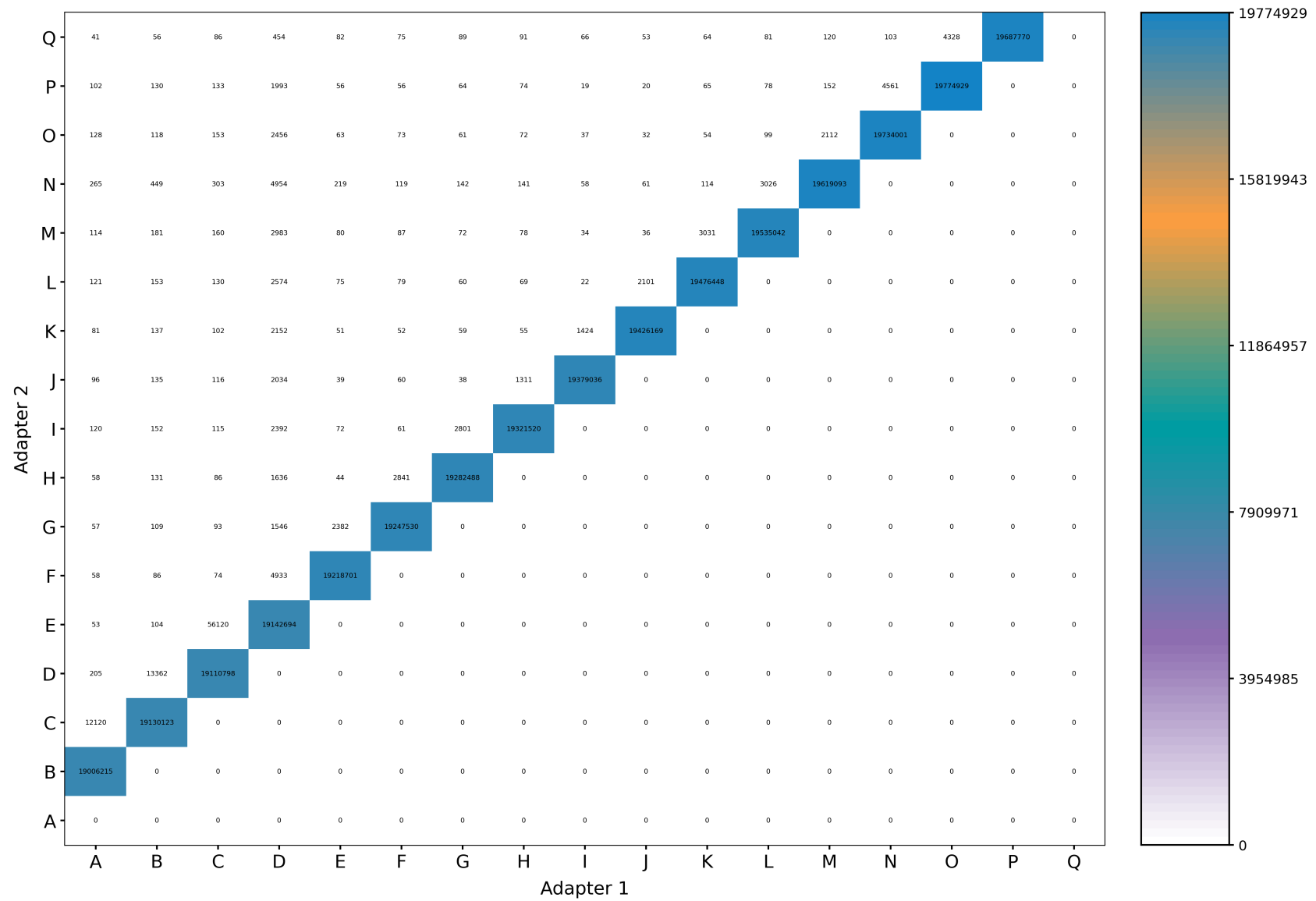
Summary

Reads	19,856,760
Segmented reads (S-reads)	310,092,557
Mean length of S-reads	941.5002639808604
Percent of reads with full arrays	94.42 %
Mean array size (concatenation factor)	15.62

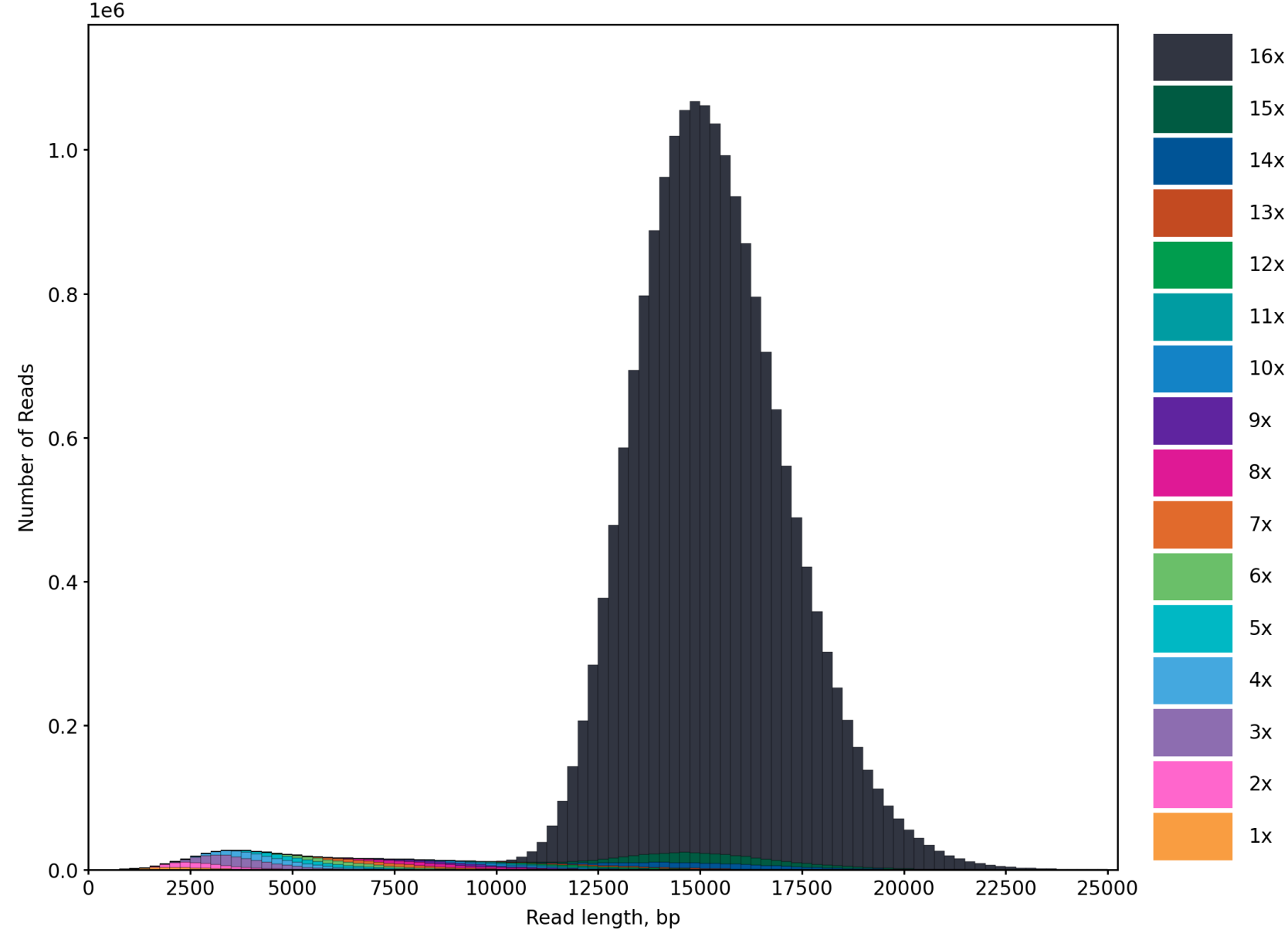
Histogram of concatemers



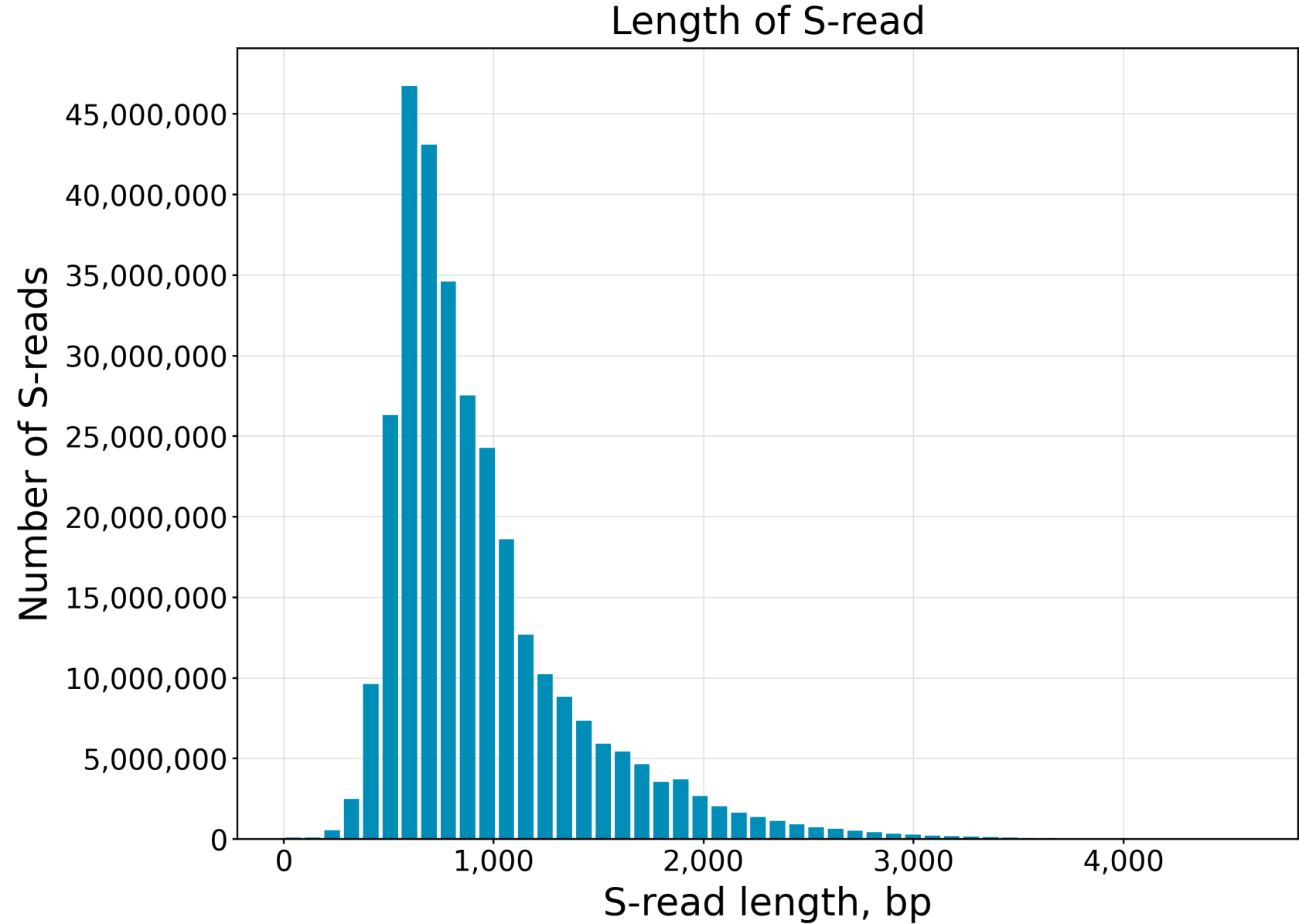
Heatmap of concatemer ligations



Length of Reads



Length of S-reads

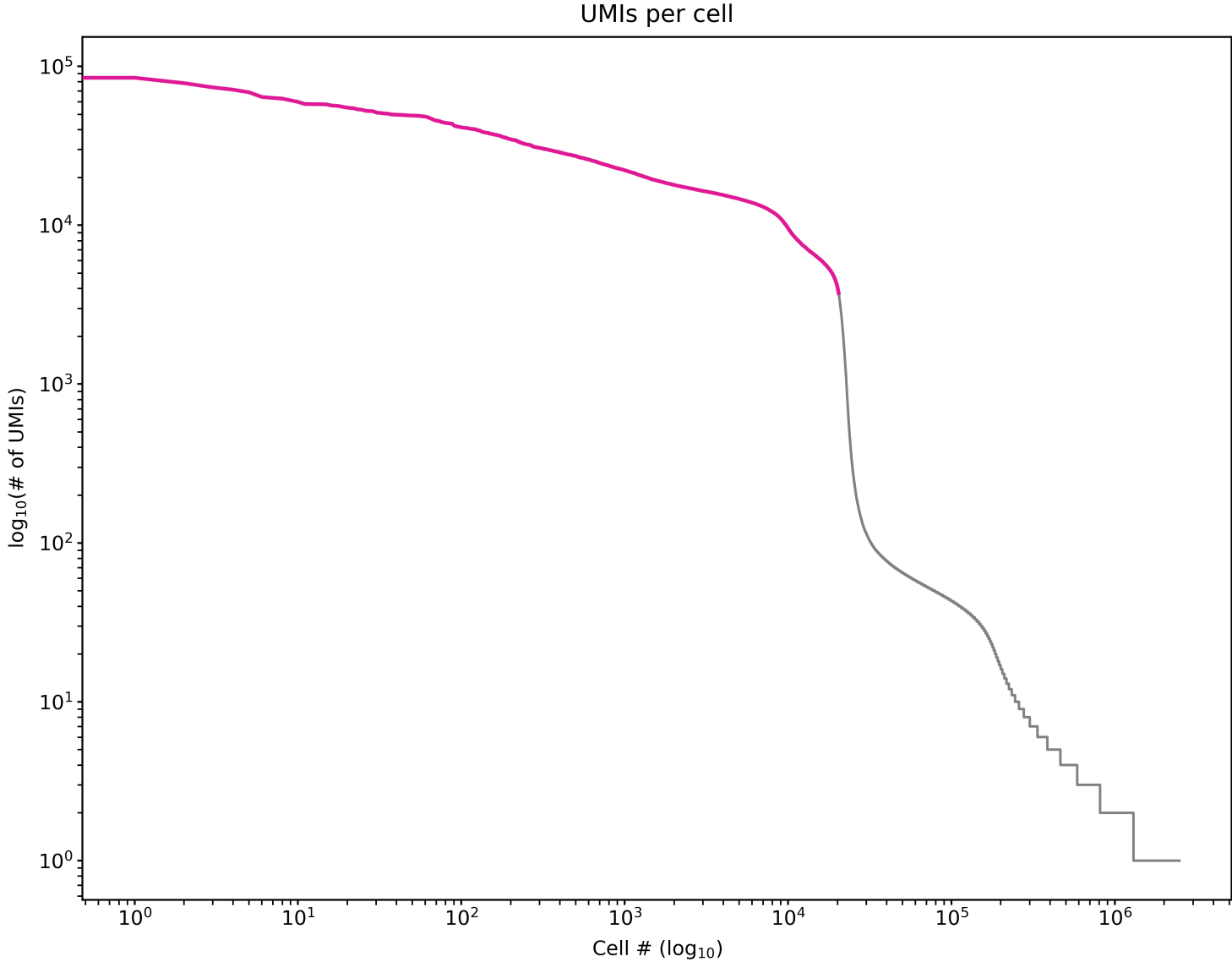


Cell Statistics

Summary

Estimated Number of Cells	20,218
Reads in Cells	91.80%
Mean Reads per Cell	13,632
Median UMIs per Cell	9,397

Barcode Rank Plot



Read Statistics

Summary

Reads	310,092,557
Read Type	SEGMENT
Reads with 5' and 3' Primers with extracted UMIs and Barcodes	307,623,491
Non-Concatamer Reads with 5' and 3' Primers and Poly-A Tail (FLNC reads)	300,266,603
FLNC Reads with Valid Barcodes	294,426,429
FLNC Reads with Valid Barcodes, corrected	299,602,036
Reads after Barcode Correction and UMI Deduplication	225,778,491

Transcript Statistics

Summary

FLNC Reads Mapped Confidently to Genome	242,519,132
FLNC Reads Mapped Confidently to Transcriptome	147,827,734
Total Unique Genes	3,494,509
Total Unique Genes, filtered	234,431
Total Unique Genes, known genes only	43,472
Total Unique Genes, filtered, known genes only	33,117
Total Unique Transcripts	8,344,200
Total Unique Transcripts, filtered	1,858,418
Total Unique Transcripts, known transcripts only	117,728
Total Unique Transcripts, filtered, known transcripts only	108,246

Transcript Summary

Value	Analysis Metric
3,729	Median Genes per Cell
2,586	Median Genes per Cell, known genes only
5,003	Median Transcripts per Cell
2,105	Median Transcripts per Cell, known transcripts only
3,494,509	Total Unique Genes
43,472	Total Unique Genes, known genes only
8,344,200	Total Unique Transcripts
117,728	Total Unique Transcripts, known transcripts only

Transcript Summary, filtered

Value	Analysis Metric
1,863	Median Genes per Cell
1,816	Median Genes per Cell, known genes only
2,444	Median Transcripts per Cell
1,802	Median Transcripts per Cell, known transcripts only
234,431	Total Unique Genes
33,117	Total Unique Genes, known genes only
1,858,418	Total Unique Transcripts
108,246	Total Unique Transcripts, known transcripts only

Transcript Classification

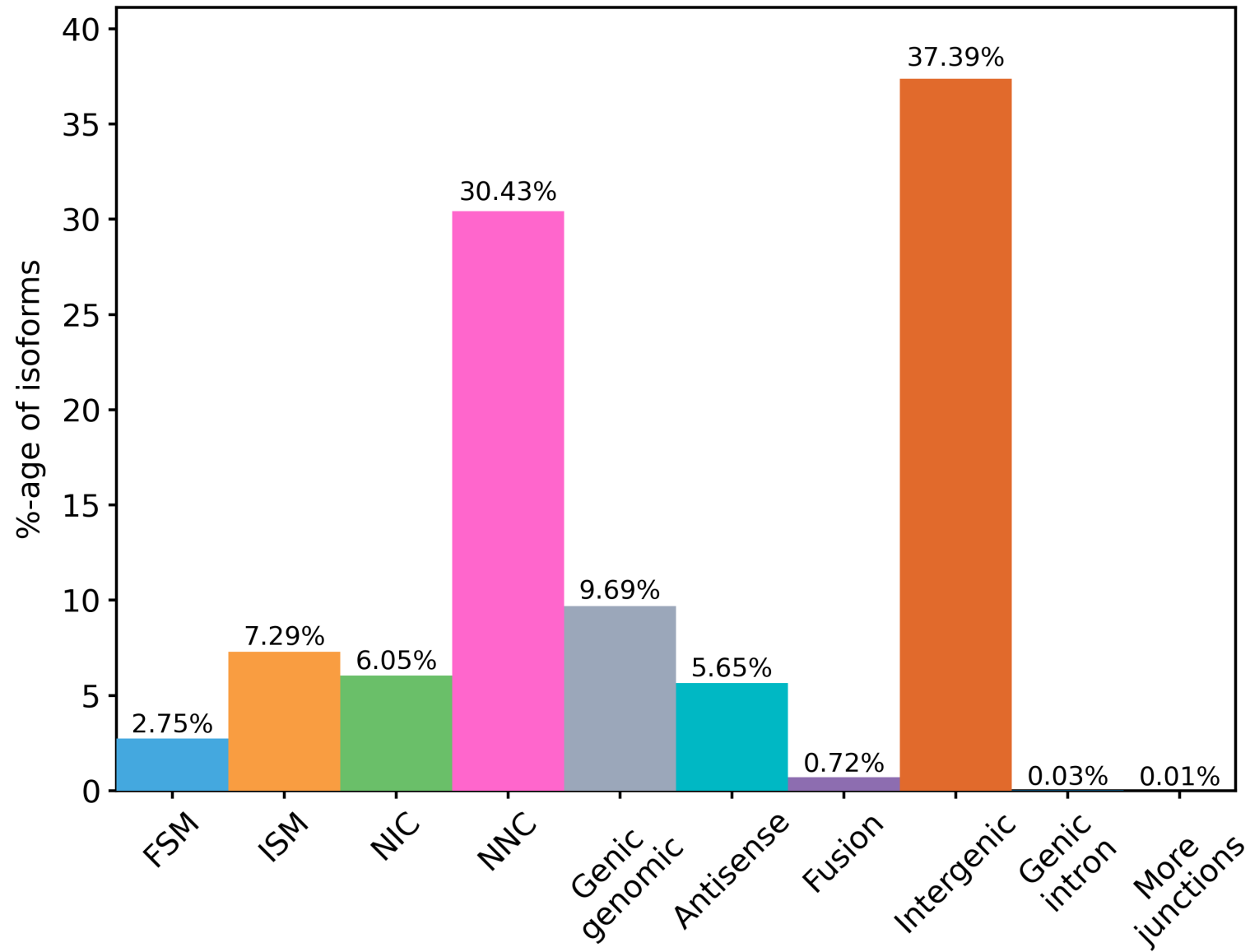
Category	Count	CAGE Detected	CAGE Detected (%)	polyA Motif Detected	polyA Motif Detected (%)
FSM	251486	141312	56.19%	71539	28.45%
ISM	666493	266062	39.92%	191159	28.68%
NIC	553284	354213	64.02%	172208	31.12%
NNC	2782904	1713835	61.58%	989545	35.56%
Antisense	517075	17598	3.40%	133611	25.84%
Fusion	65406	36305	55.51%	24163	36.94%
More junctions	912	436	47.81%	255	27.96%
Genic intron	2390	6	0.25%	668	27.95%
Genic genomic	885675	217877	24.60%	257045	29.02%
Intergenic	3418826	31704	0.93%	1114062	32.59%

Transcript Classification, filtered

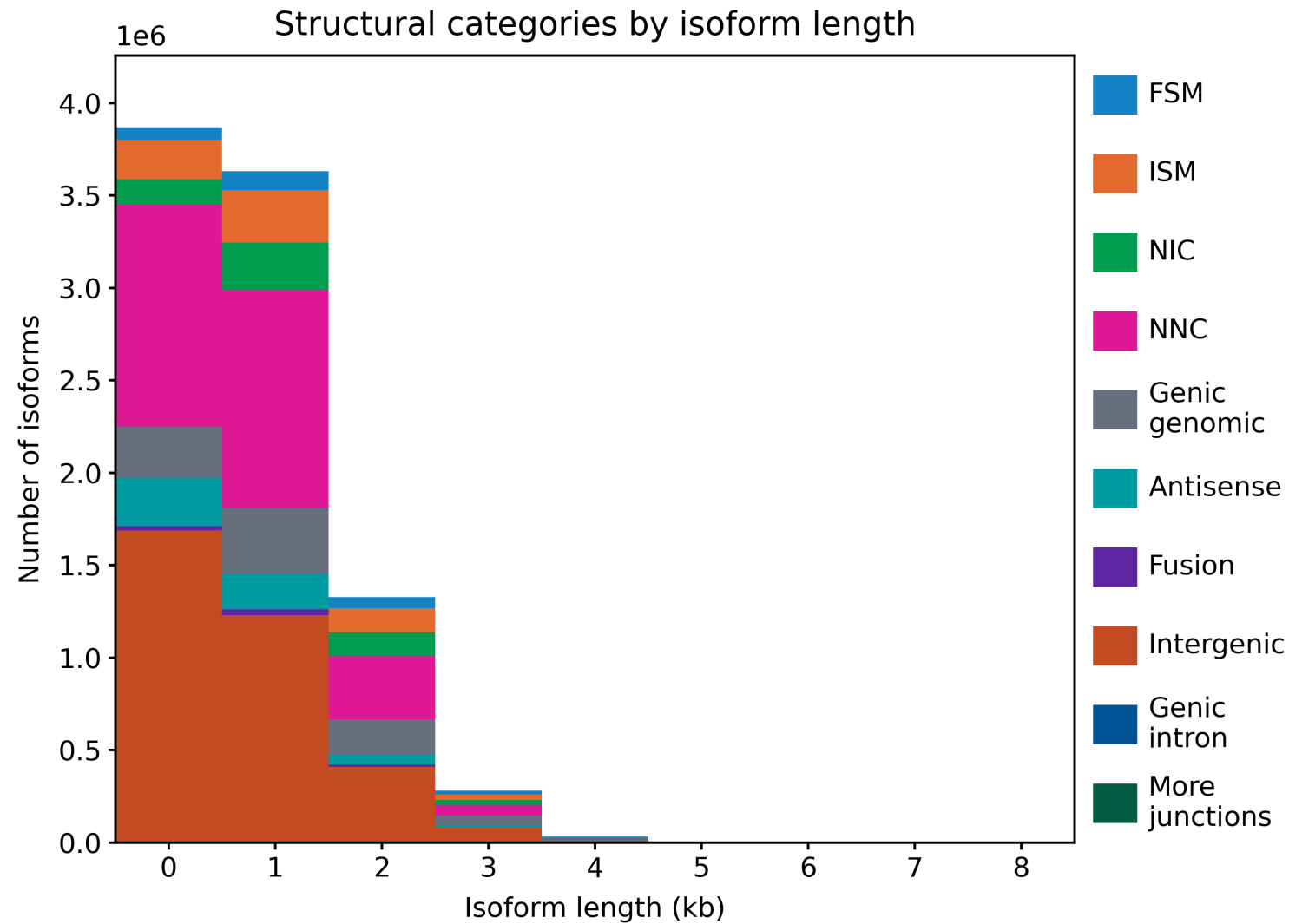
Category	Count	CAGE Detected	CAGE Detected, (%)	polyA Detected	polyA Detected, (%)
FSM	214619	123903	57.73%	71539	33.33%
ISM	479019	211101	44.07%	169356	35.35%
NIC	447260	288119	64.42%	170929	38.22%
NNC	992116	626669	63.16%	437476	44.10%
Antisense	42994	3803	8.85%	24608	57.24%
Fusion	30875	17615	57.05%	14027	45.43%
More junctions	547	262	47.90%	193	35.28%
Genic intron	0	0	0.00%	0	0.00%
Genic genomic	44856	20221	45.08%	20011	44.61%
Intergenic	191524	4418	2.31%	128022	66.84%

Isoform Distribution Across Categories

Isoform distribution across structural categories

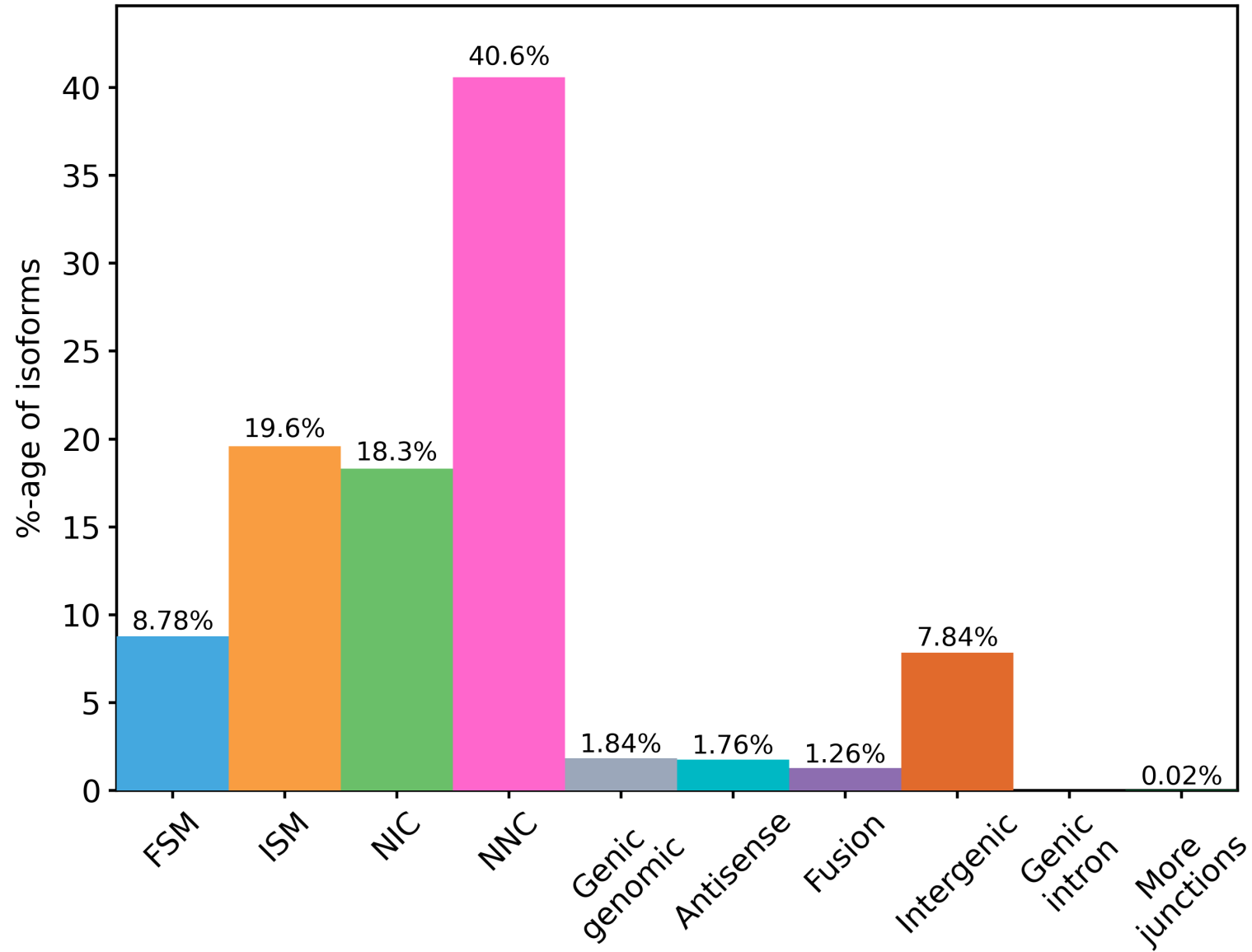


Structural Categories by Transcript Length

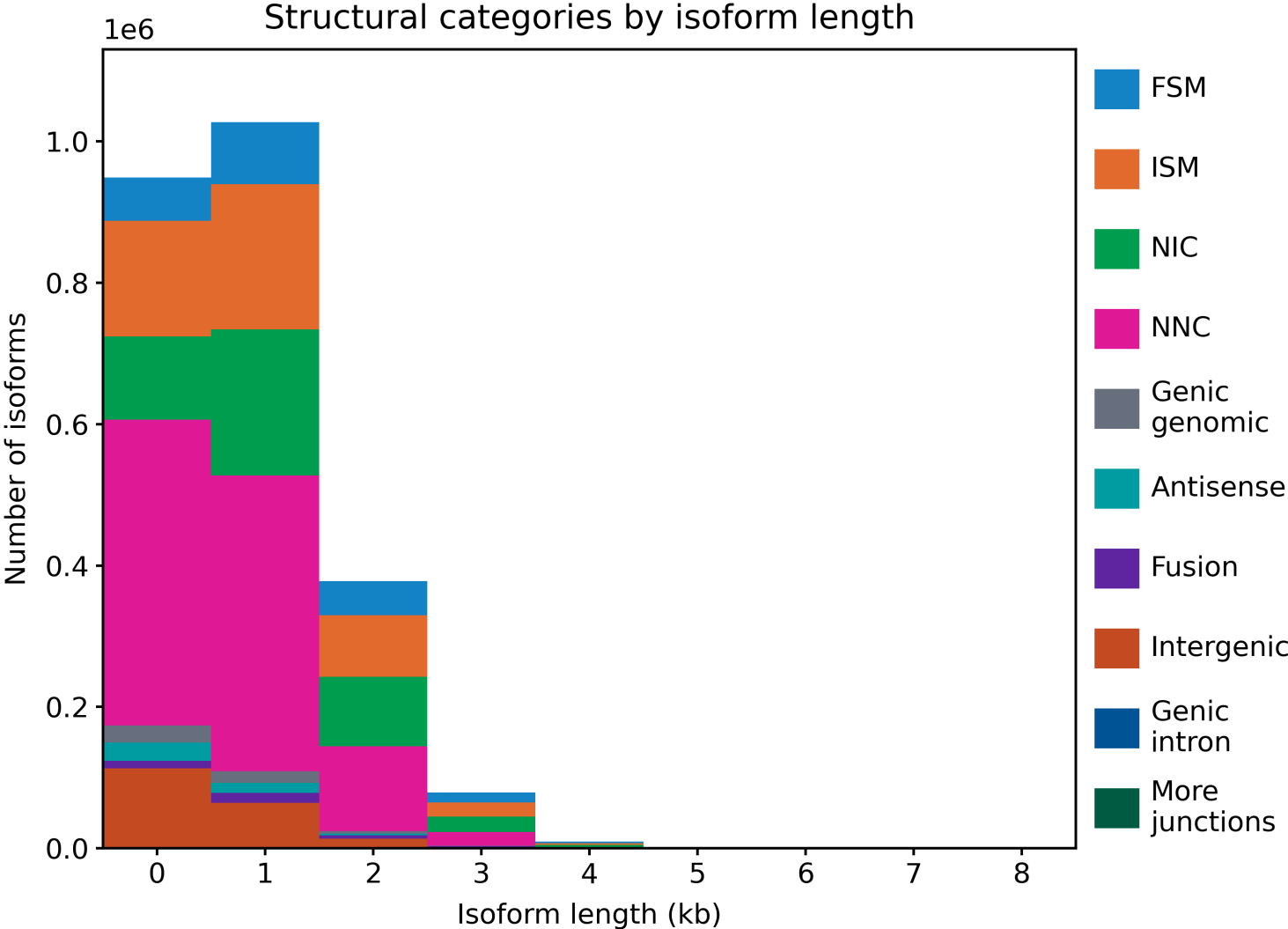


Isoform Distribution Across Categories

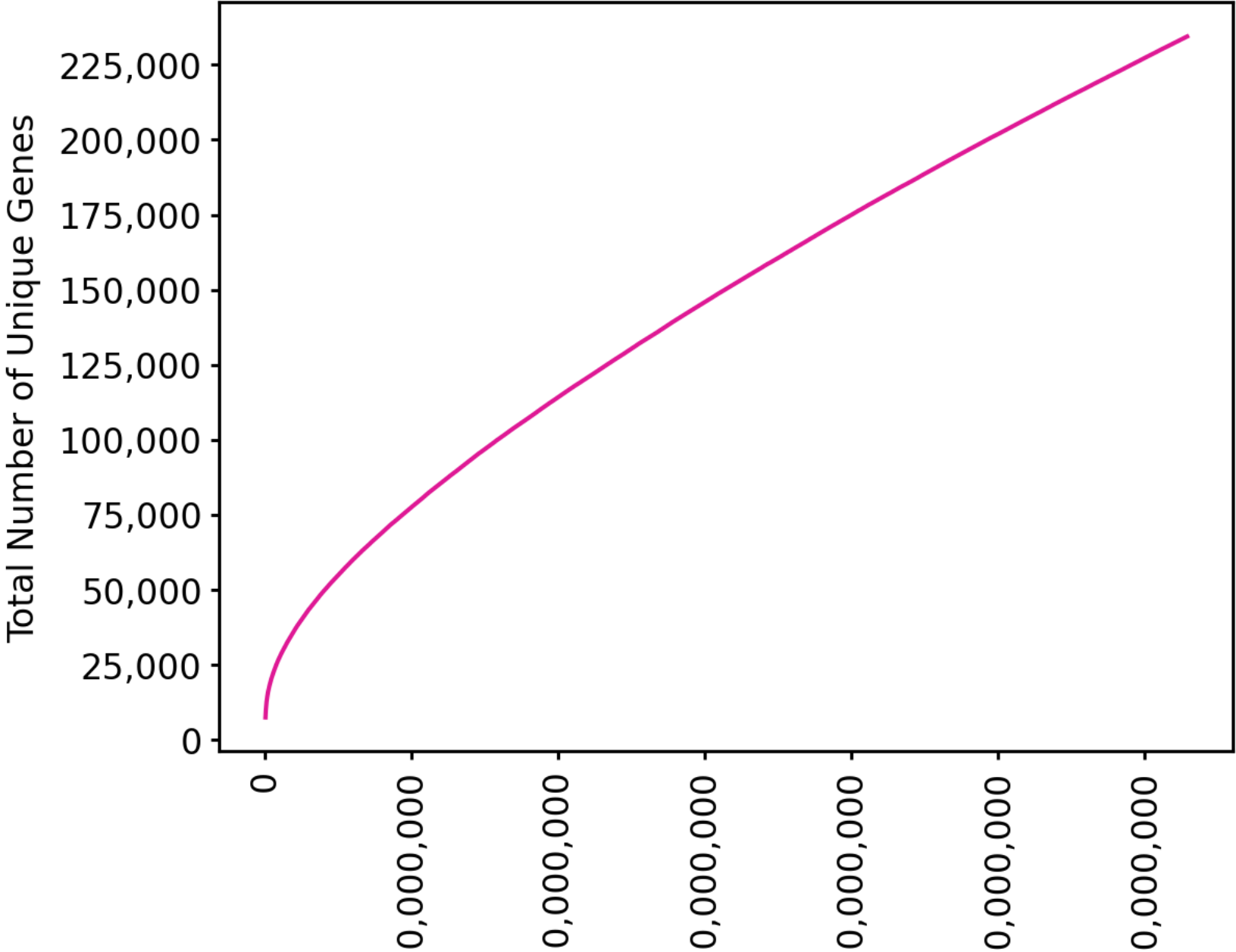
Isoform distribution across structural categories



Structural Categories by Transcript Length



Gene Saturation, all genes, filtered



2

4

6

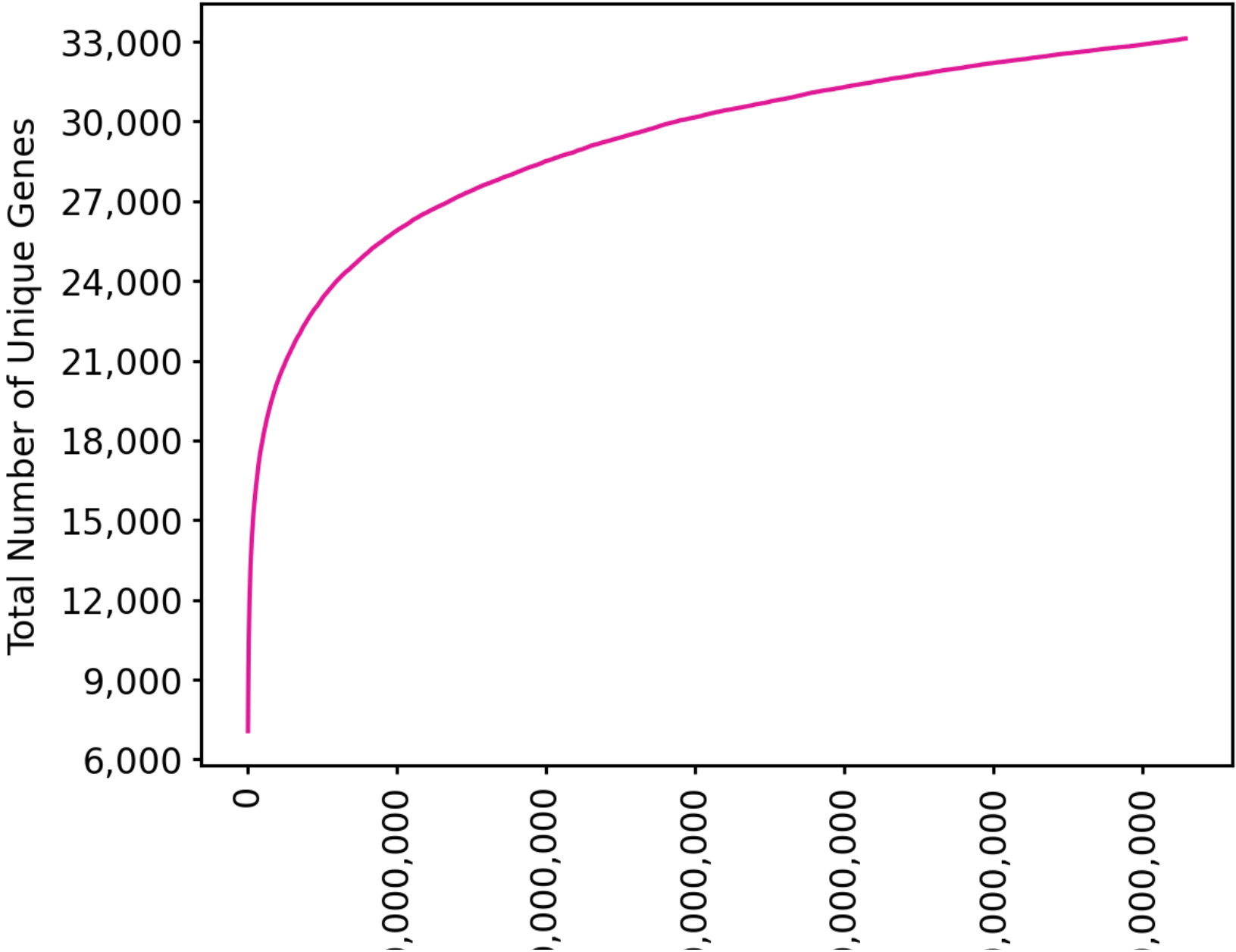
8

10

12

Total Reads

Gene Saturation, known genes only, filtered



20

40

60

80

100

120

Total Reads

