

Report for dataset 3prime_v4_GEX_hPBMCs_10k_Rep1-Cell1 (all samples)

Dataset e2deabef-c6d3-4939-a863-d9a254725a90

Summary

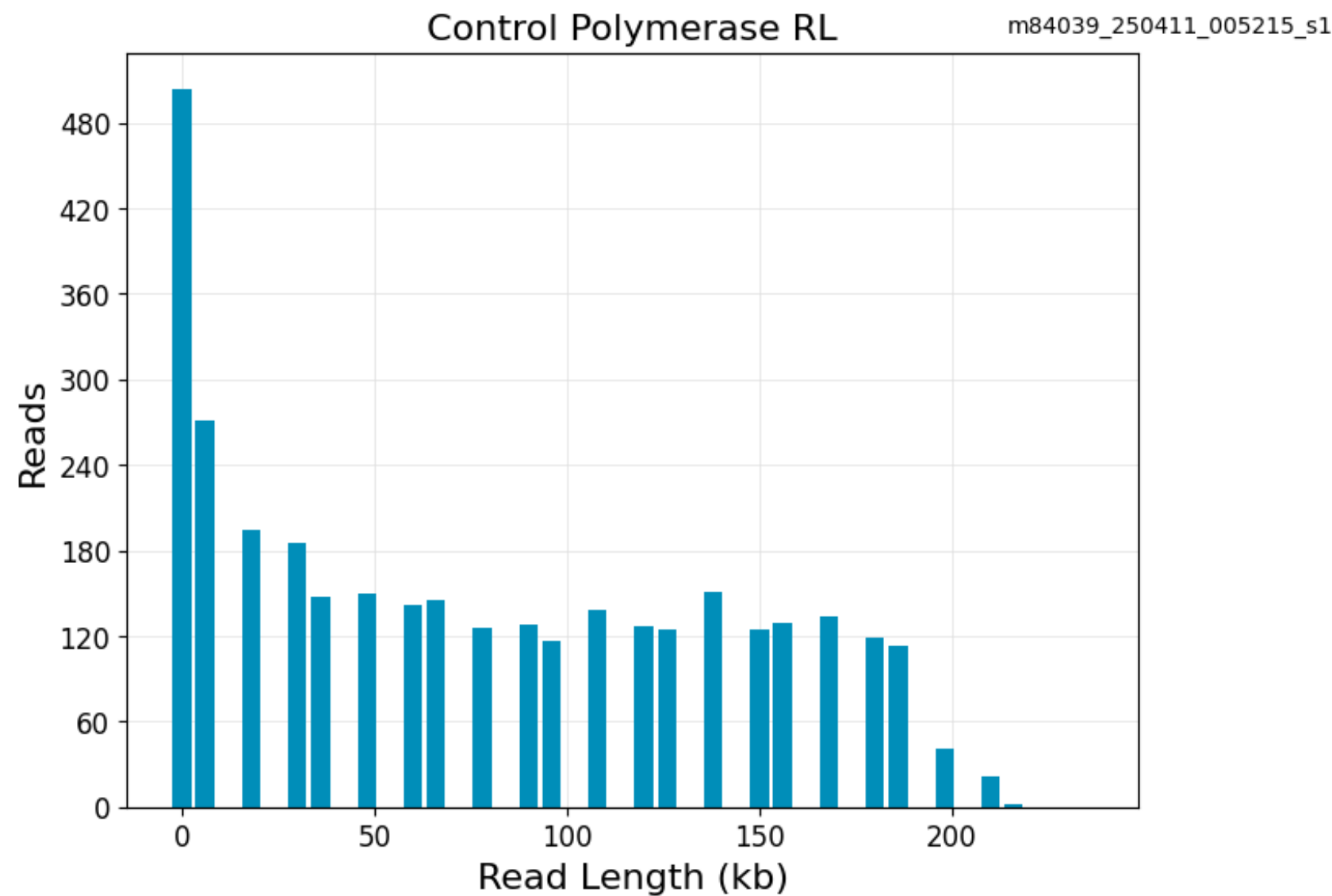
Name	3prime_v4_GEX_hPBMCs_10k_Rep1-Cell1 (all samples)
Created At	2025-04-12 15:15:34.529
Number of Records	9,709,186
Total Length	145,021,072,005
Movie Name	m84039_250411_005215_s1
ICS Version	13.3.0.253824
Well Sample	3prime_v4_GEX_hPBMCs_10k_Rep1
Biological Sample	3prime_v4_GEX_hPBMCs_10k_Rep1
Barcode Name	bcM0001--bcM0001

Control Report

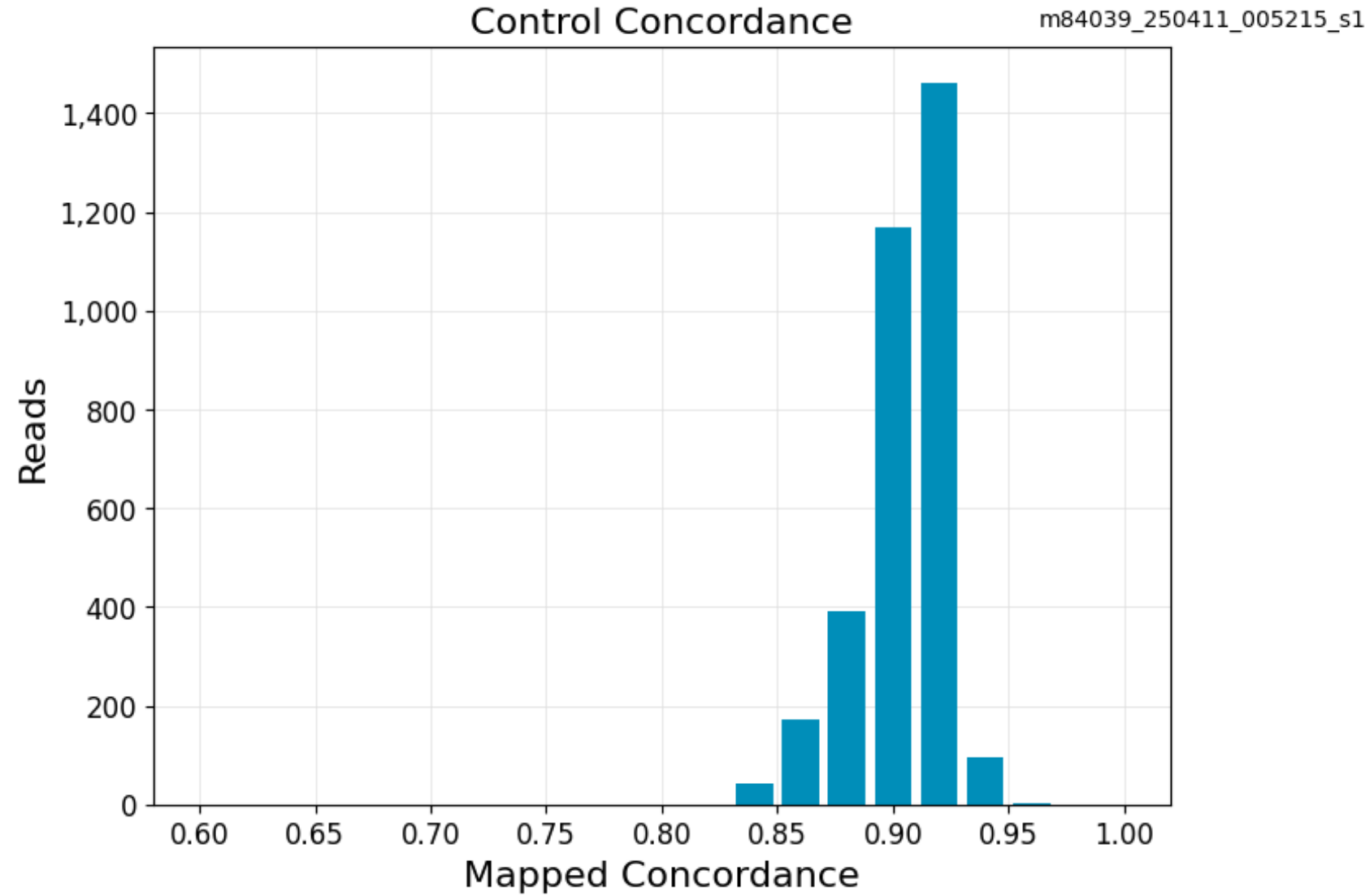
Summary

Number of Control Reads	3,335
Control Read Length Mean	79,391
Control Read Concordance Mean	0.91
Control Read Concordance Mode	0.93

Control Polymerase RL



Control Concordance



Barcodes

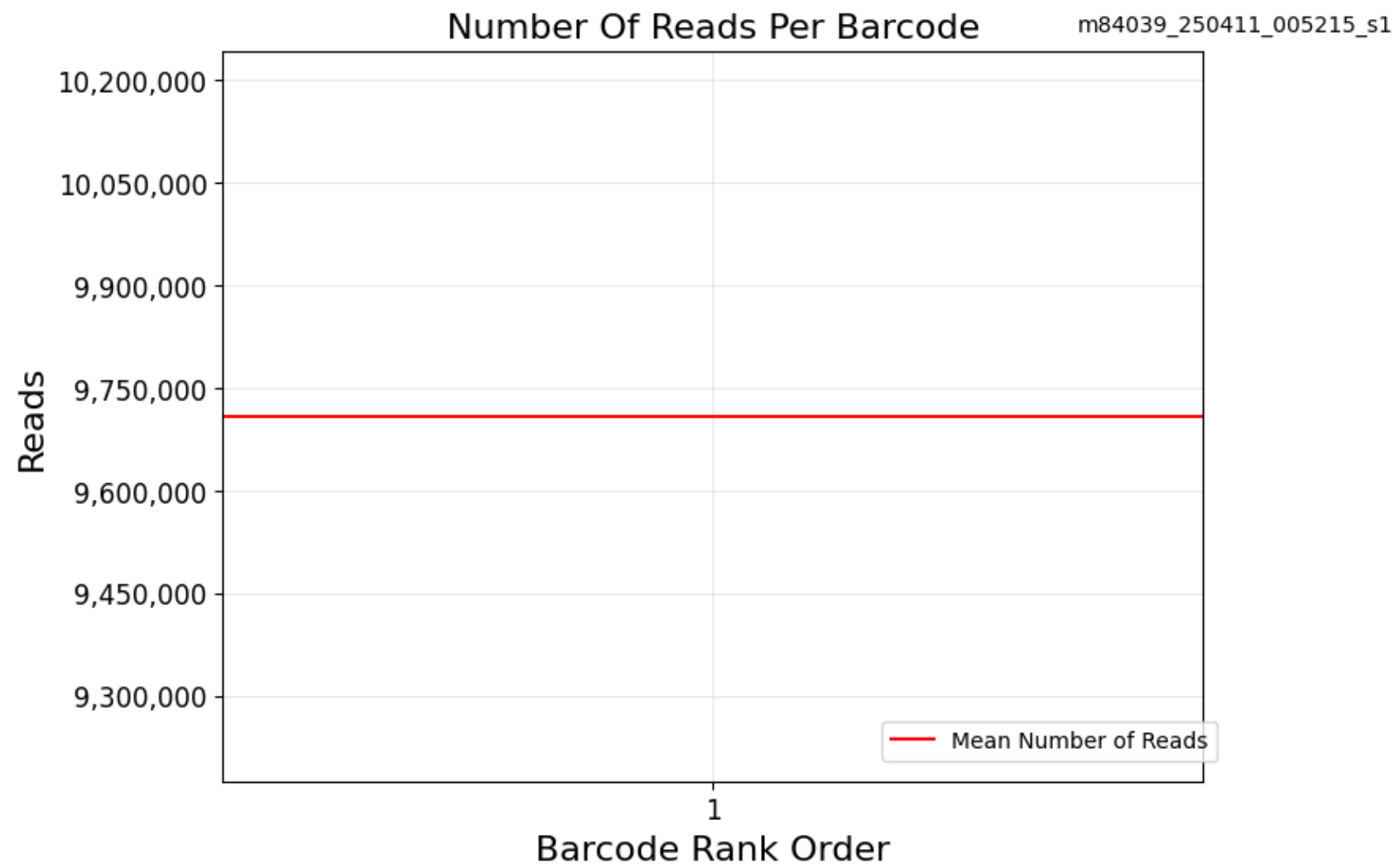
Summary

Unique Barcodes	1
Barcoded HiFi Reads	9,709,186
Unbarcoded HiFi Reads	12,179
Barcoded HiFi Reads (%)	99.87 %
Barcoded HiFi yield (Gb)	145.02 Gb
Unbarcoded HiFi yield (Gb)	0.14 Gb
Barcoded HiFi Yield (%)	99.90 %
Mean HiFi Reads per Barcode	9,709,186
Max. HiFi Reads per Barcode	9,709,186
Min. HiFi Reads per Barcode	9,709,186
Barcoded HiFi read length (mean, kb)	14.94 kb
Unbarcoded HiFi read length (mean, kb)	11.67 kb

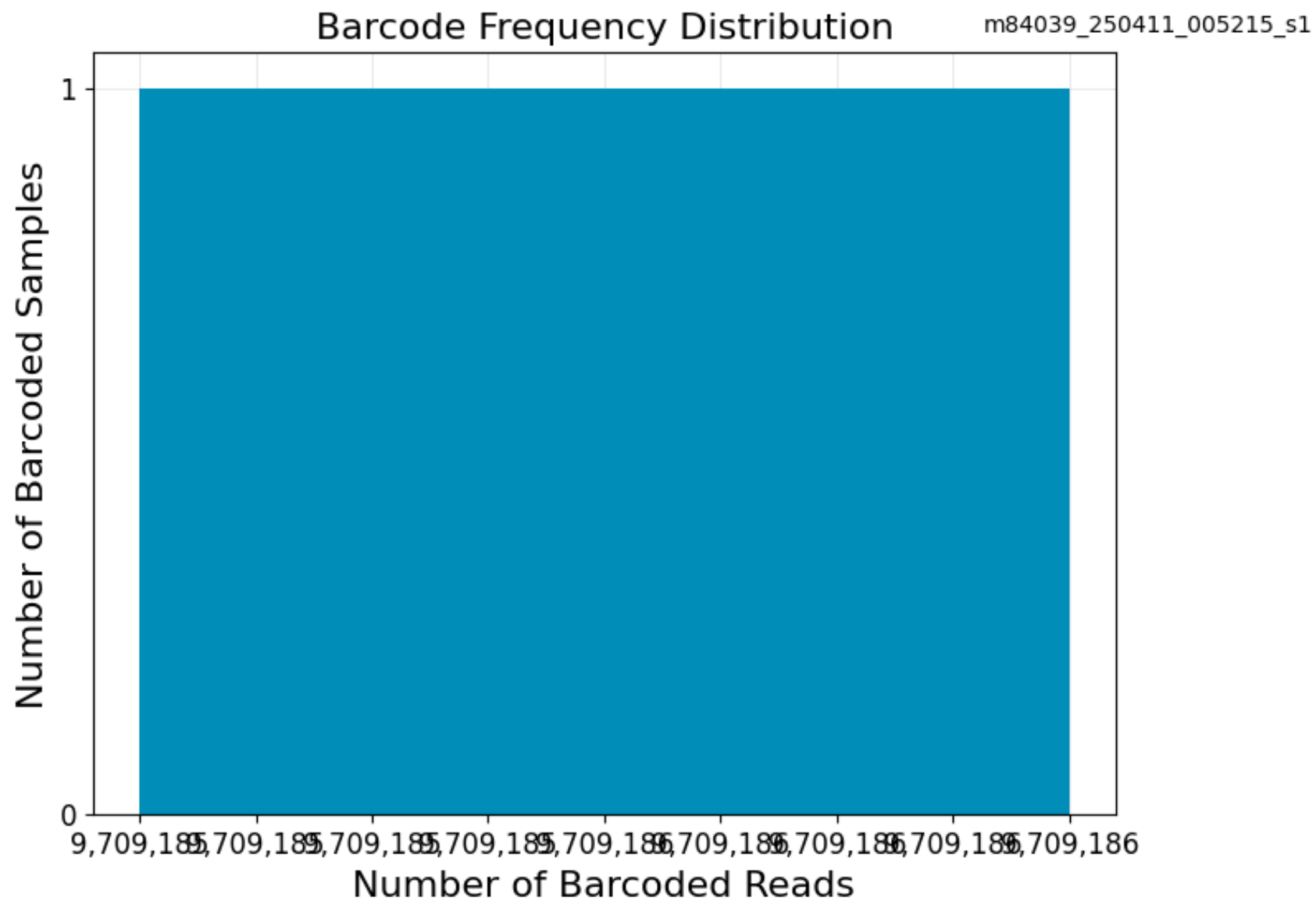
Barcode Data

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_10k_Rep1	bcM0001--bcM0001	97.7	9709186	14936	Q29	145021072005	142659
No Name	Not Barcoded	0.0	12179	11672	Q27	142155042	115714

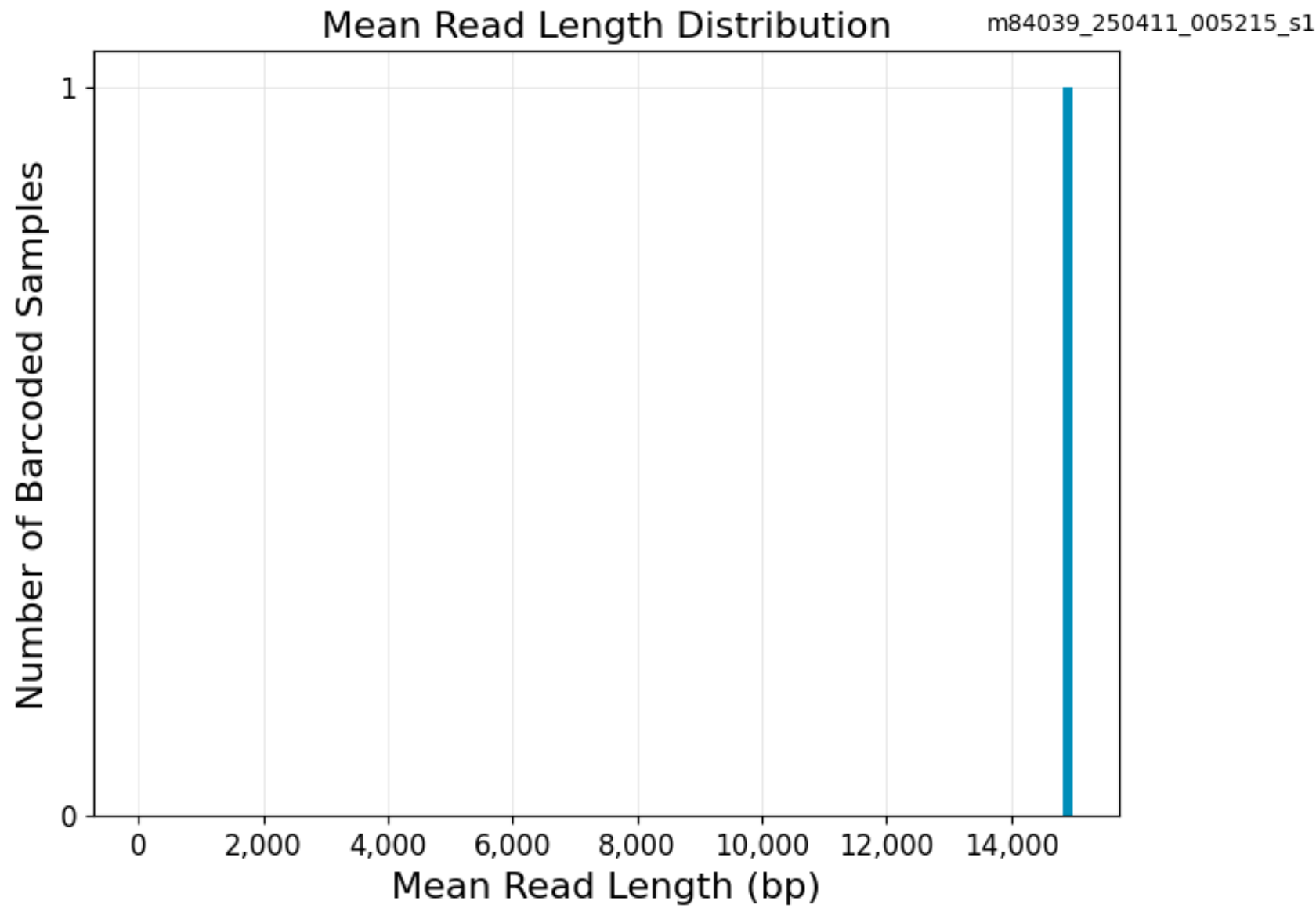
Number Of Reads Per Barcode



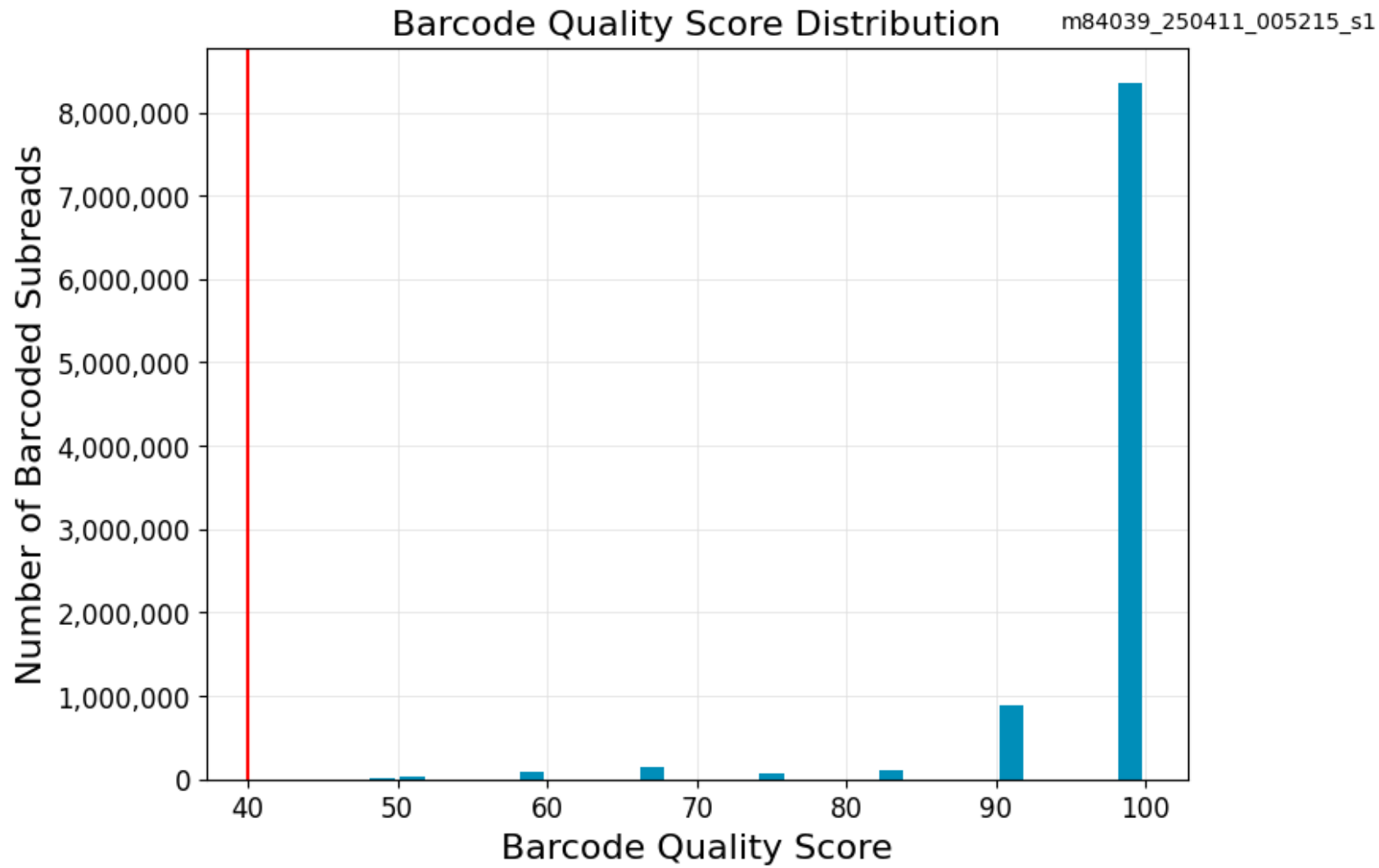
Barcode Frequency Distribution



Mean Read Length Distribution



Barcode Quality Score Distribution

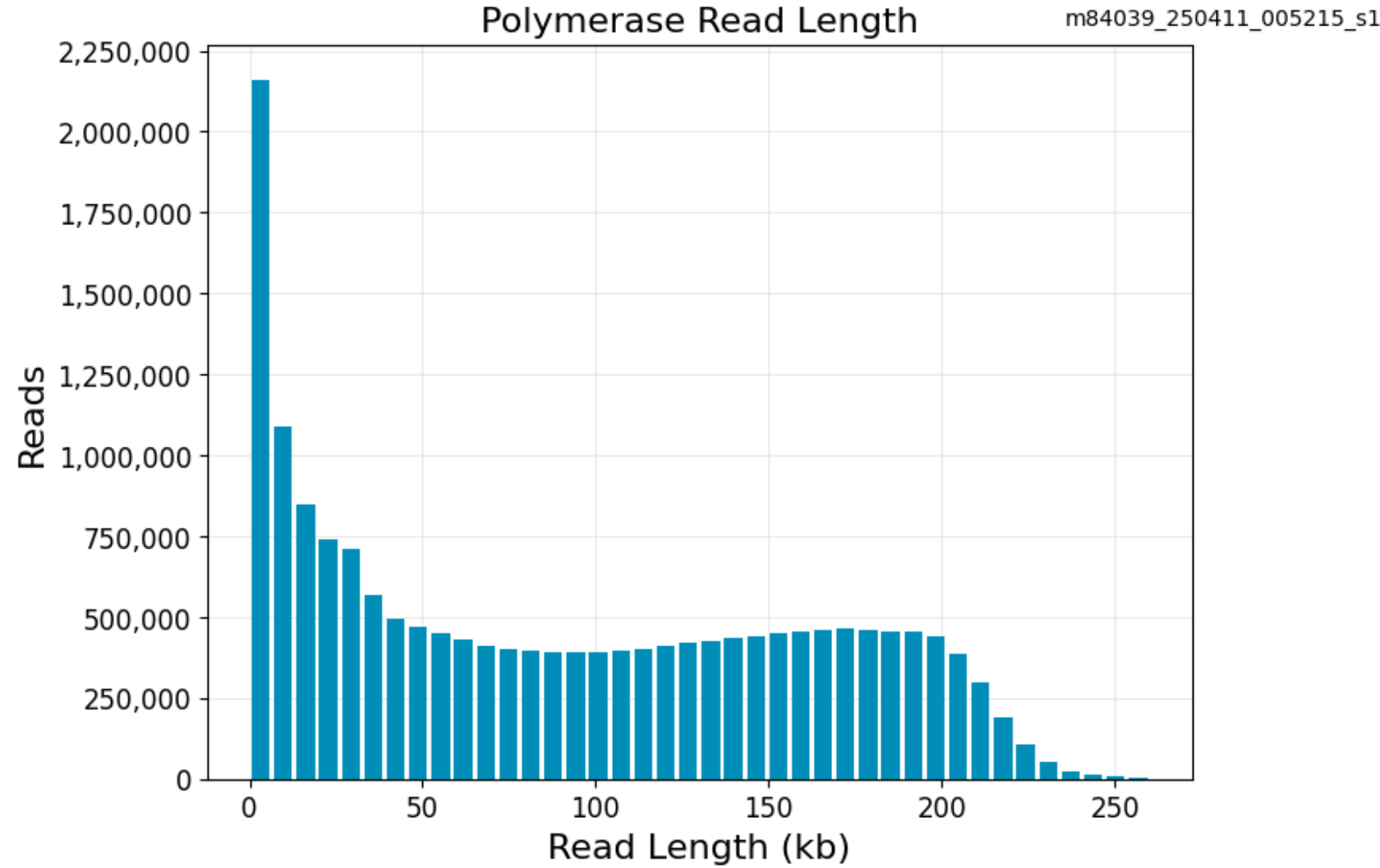


Raw Data Report

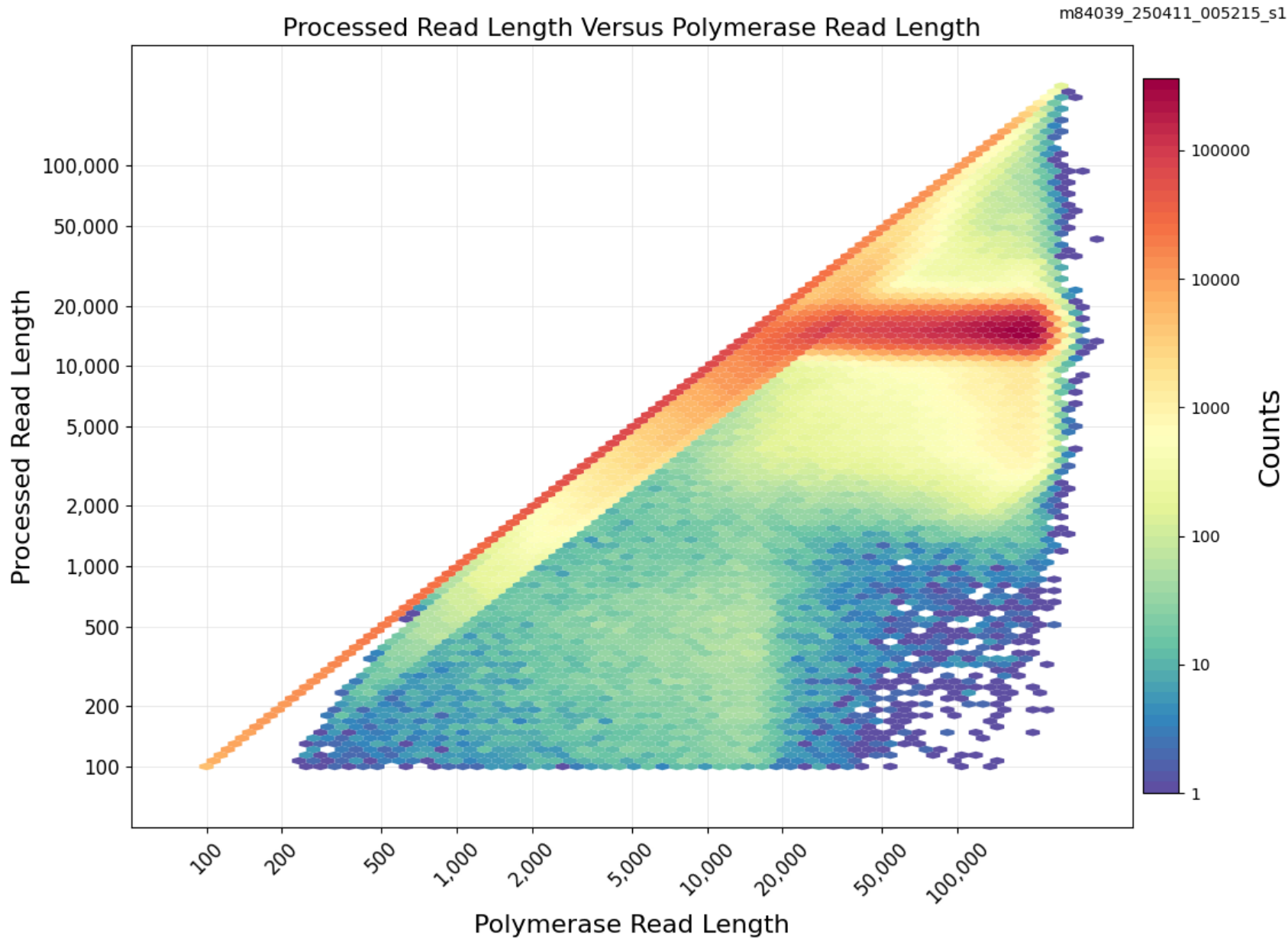
Summary

Polymerase Read Bases	1,568,076,744,192
Polymerase Reads	17,980,119
Polymerase Read Length (mean)	87.21 kb
Polymerase read length (N50)	154.75 kb
Polymerase read length longest subread length (mean)	15.72 kb
Polymerase read length longest subread length (N50)	16.25 kb
Unique Molecular Yield	273,036,918,784
Local Base Rate	2.53

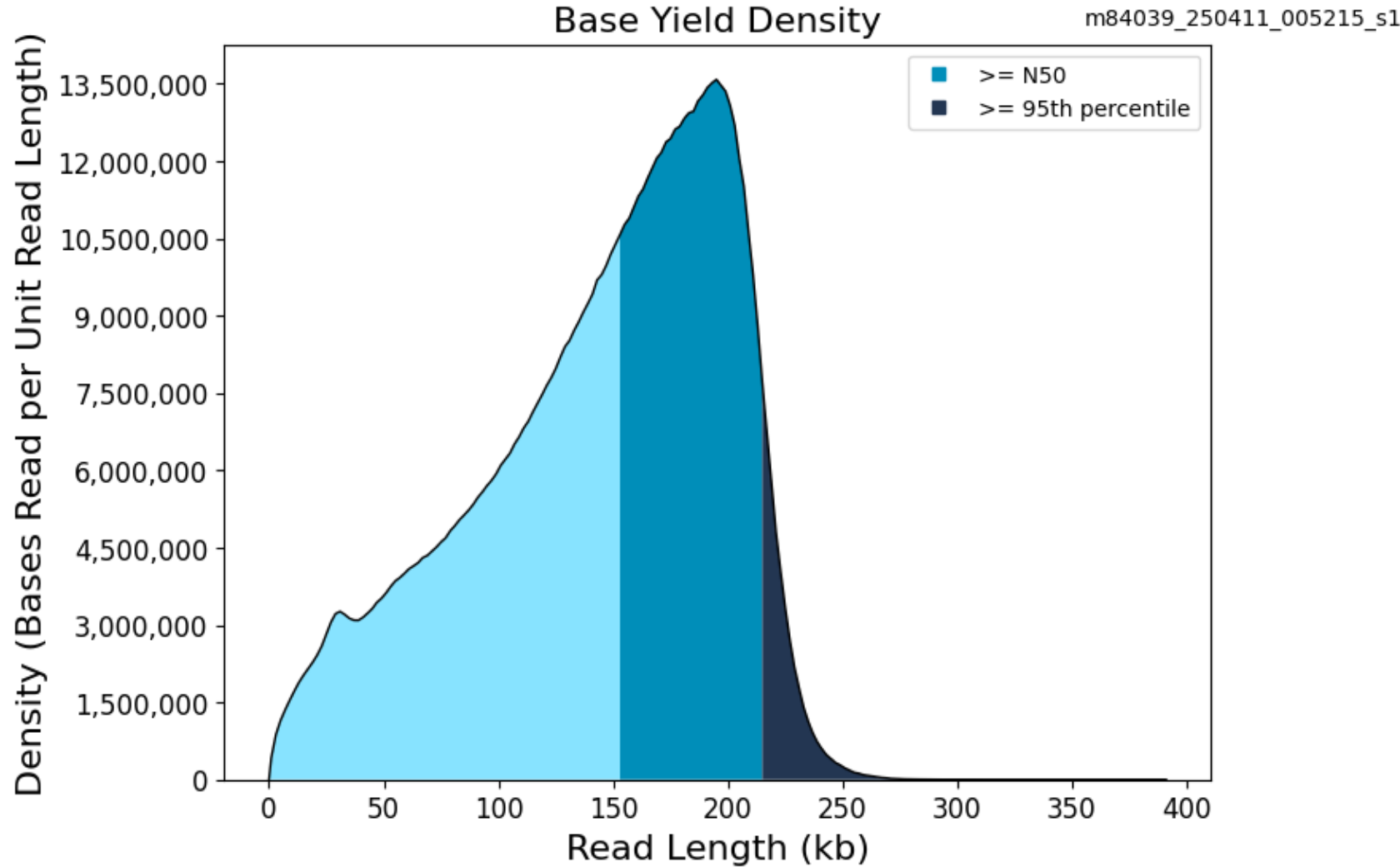
Polymerase Read Length



Longest Subread Length Versus Polymerase Read Length



Base Yield Density



Loading Report

Summary

Productive ZMWs	25,165,824
Productivity 0	7,088,910
Productivity 1	17,983,454
Productivity 2	93,460

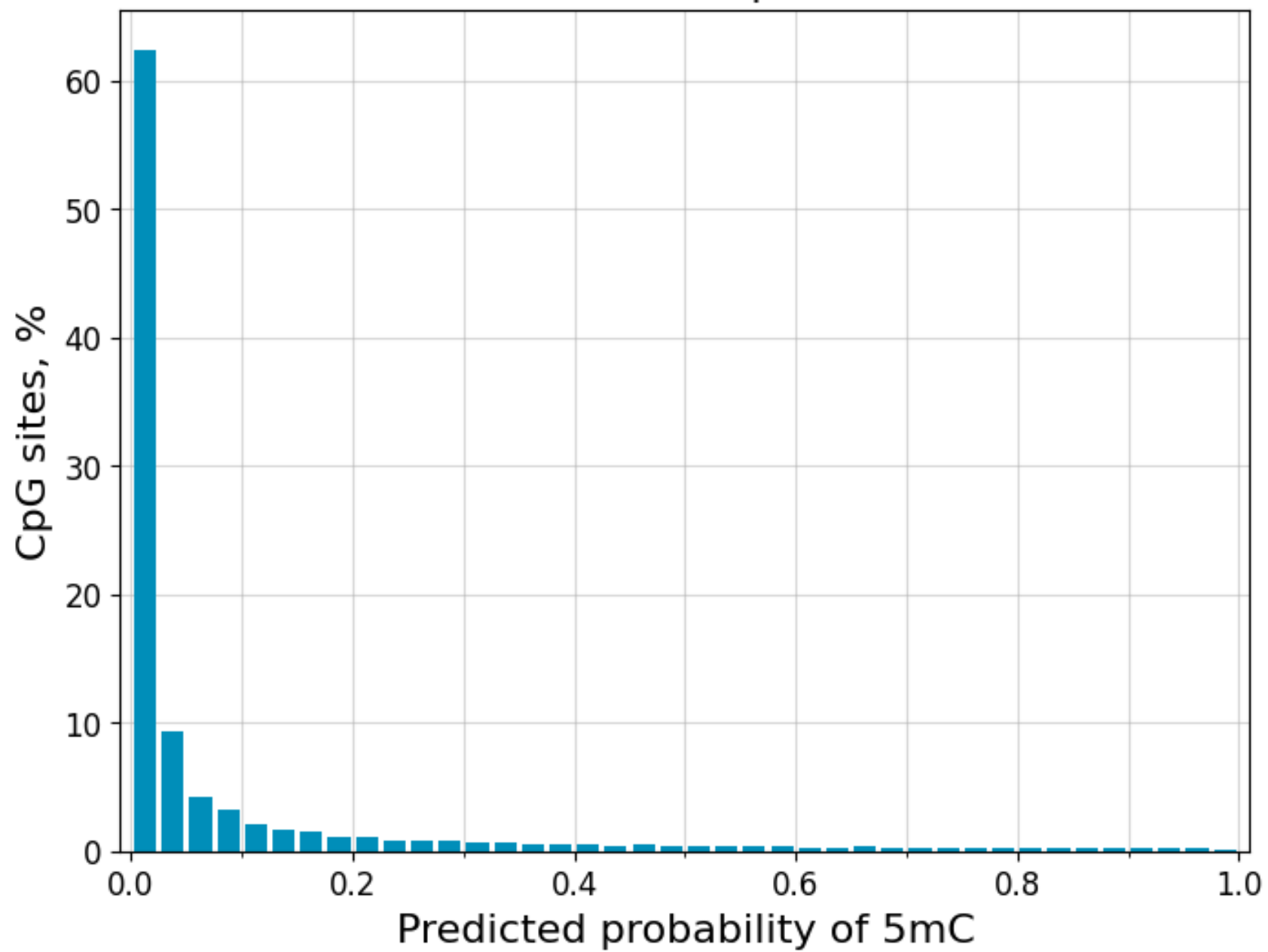
Methylation

Summary

Modification	Motif	Scored sites	Modified sites (Pr > 0.5)
5mC	CpG	98.8%	5.6%
6mA	A	0.5%	0.5%

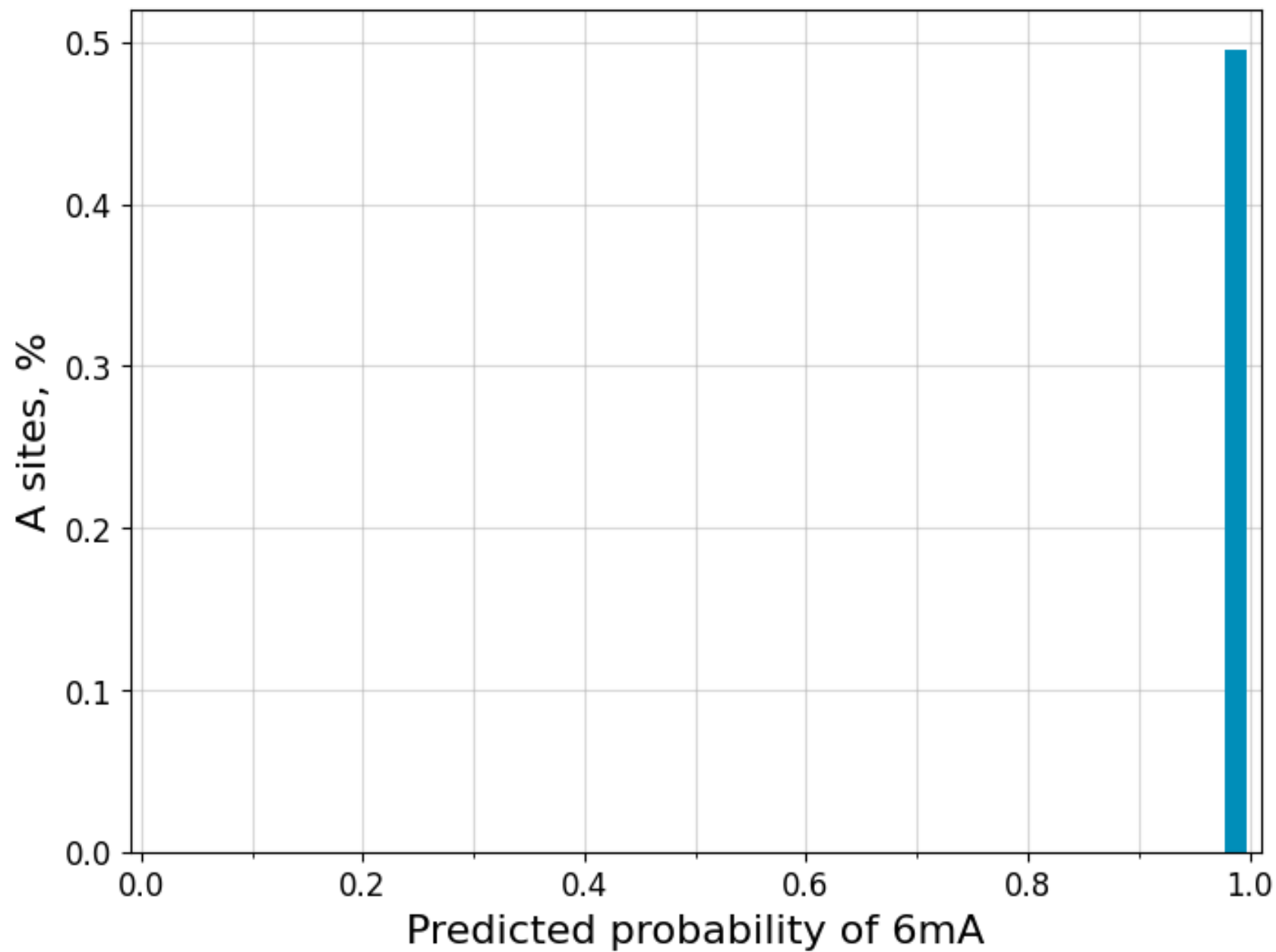
5mC methylation at CpG in reads

5mC at CpG



6mA methylation in reads

6mA



CCS Analysis Report

Summary

HiFi reads	9.7 M
HiFi reads yield	145.16 Gb
HiFi reads length (mean)	14.93 kb
HiFi reads length (median, bp)	14,935
HiFi Read Length N50 (bp)	15,183
HiFi Read Quality (median)	Q29
HiFi Read Quality (median)	29
Base Quality Q30 (%)	95.43%
HiFi Number of Passes (mean)	9
Missing adapters (%)	2.60%

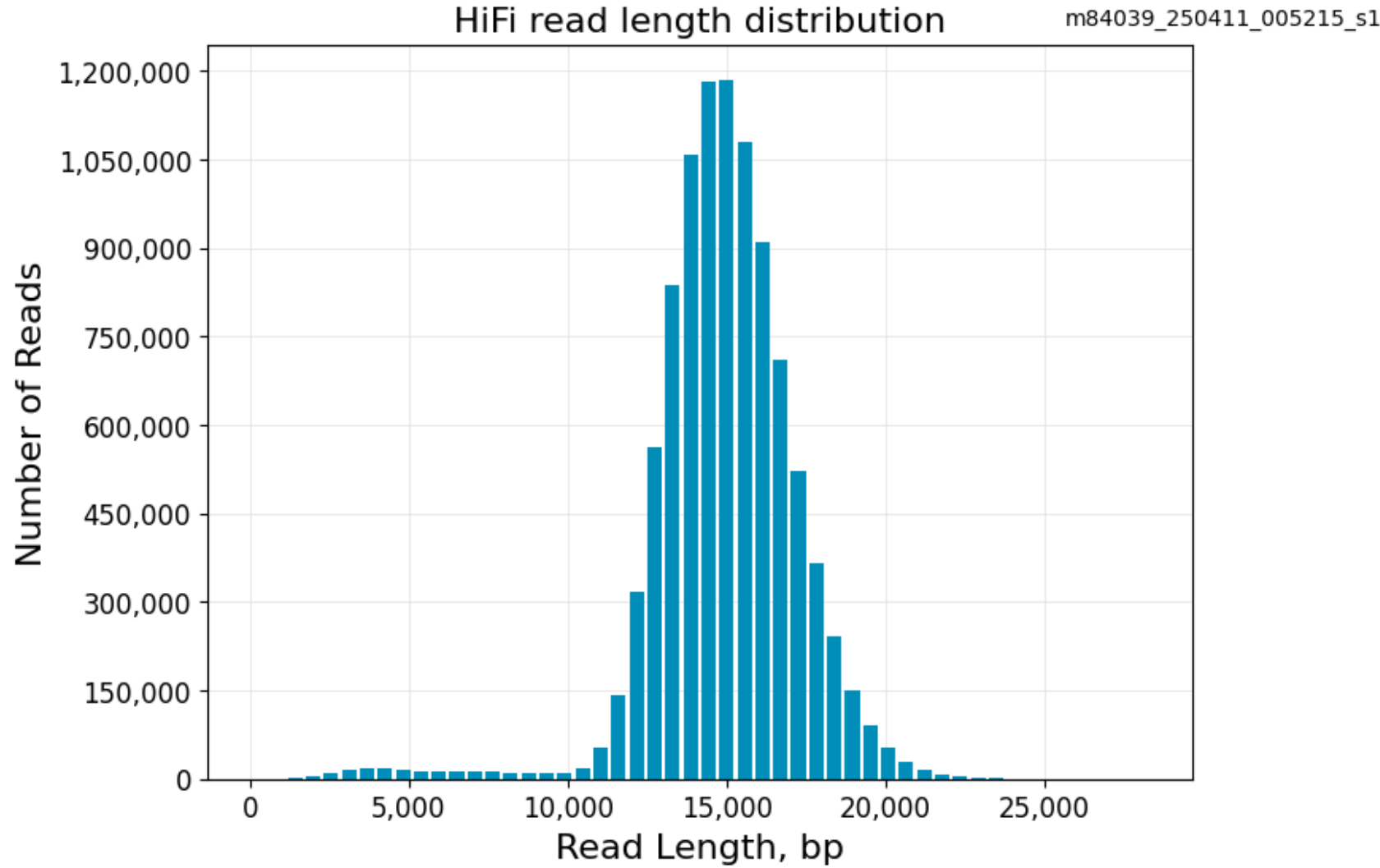
HiFi Read Length Summary

Read Length (kb)	Reads	Reads (%)	Yield (Gb)	Yield (%)
0	9721365	100	145.16 Gb	100
5,000	9638325	99	144.87 Gb	100
10,000	9533935	98	144.10 Gb	99
15,000	4724349	49	78.32 Gb	54
20,000	89192	1	1.86 Gb	1
25,000	180	0	0.00 Gb	0
30,000	7	0	0.00 Gb	0
35,000	1	0	0.00 Gb	0
40,000	1	0	0.00 Gb	0

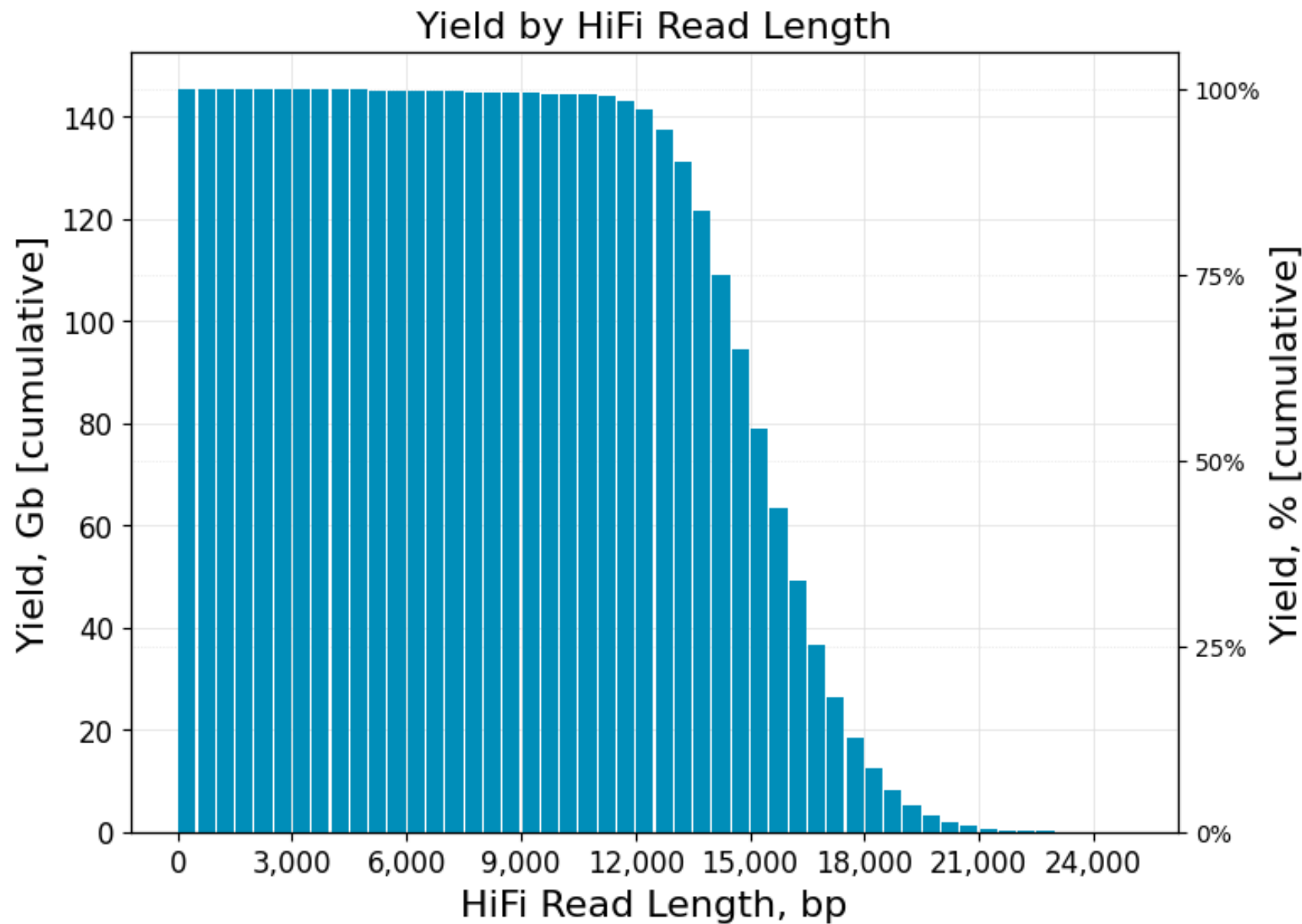
HiFi Read Quality Summary

Read Quality (Phred)	Reads	Reads (%)	Yield (Gb)	Yield (%)
Q20	9721365	100	145.16 Gb	100
Q30	4177466	43	61.47 Gb	42
Q40	58848	1	0.40 Gb	0
Q50	12302	0	0.04 Gb	0

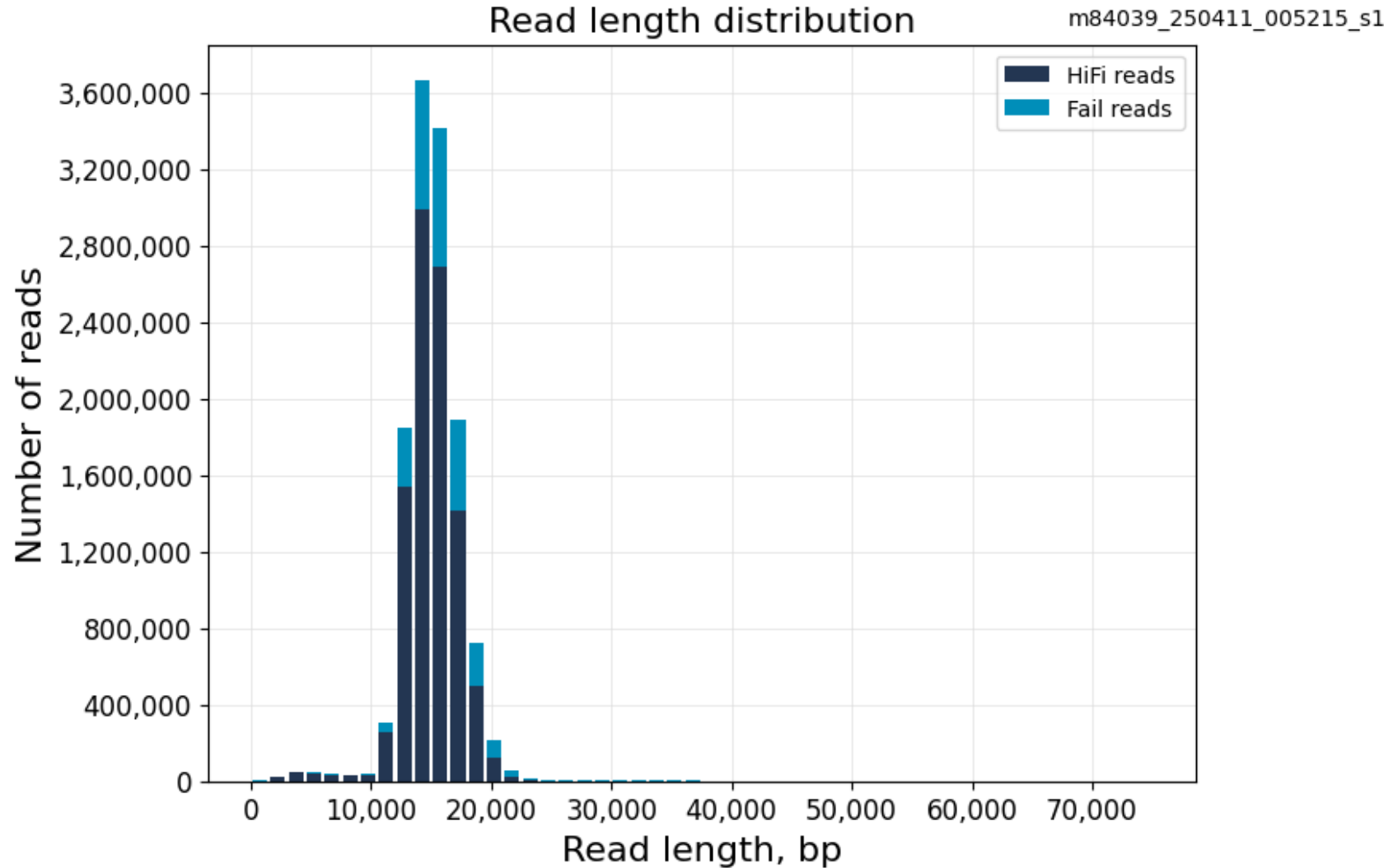
HiFi read length distribution



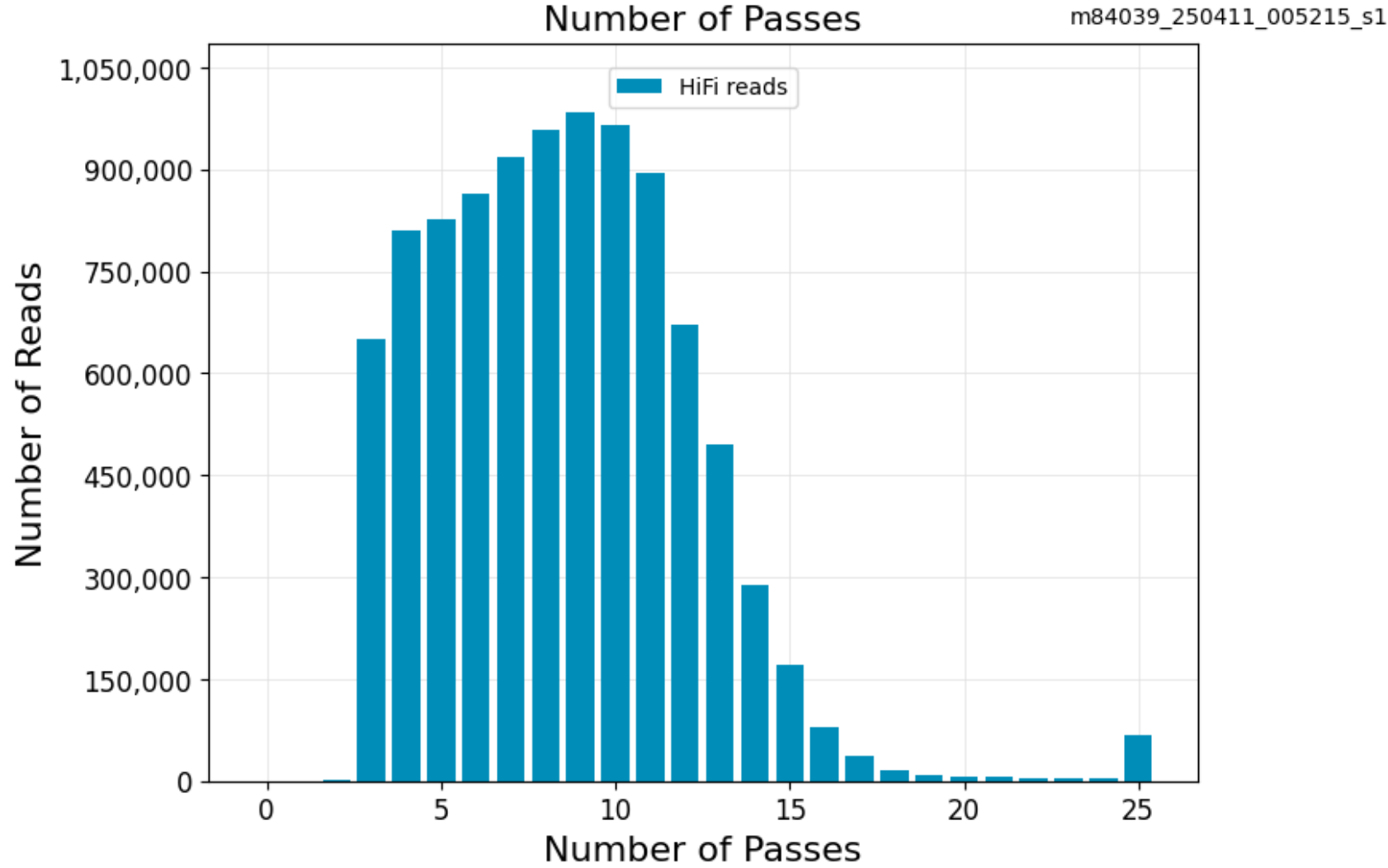
Yield by HiFi Read Length



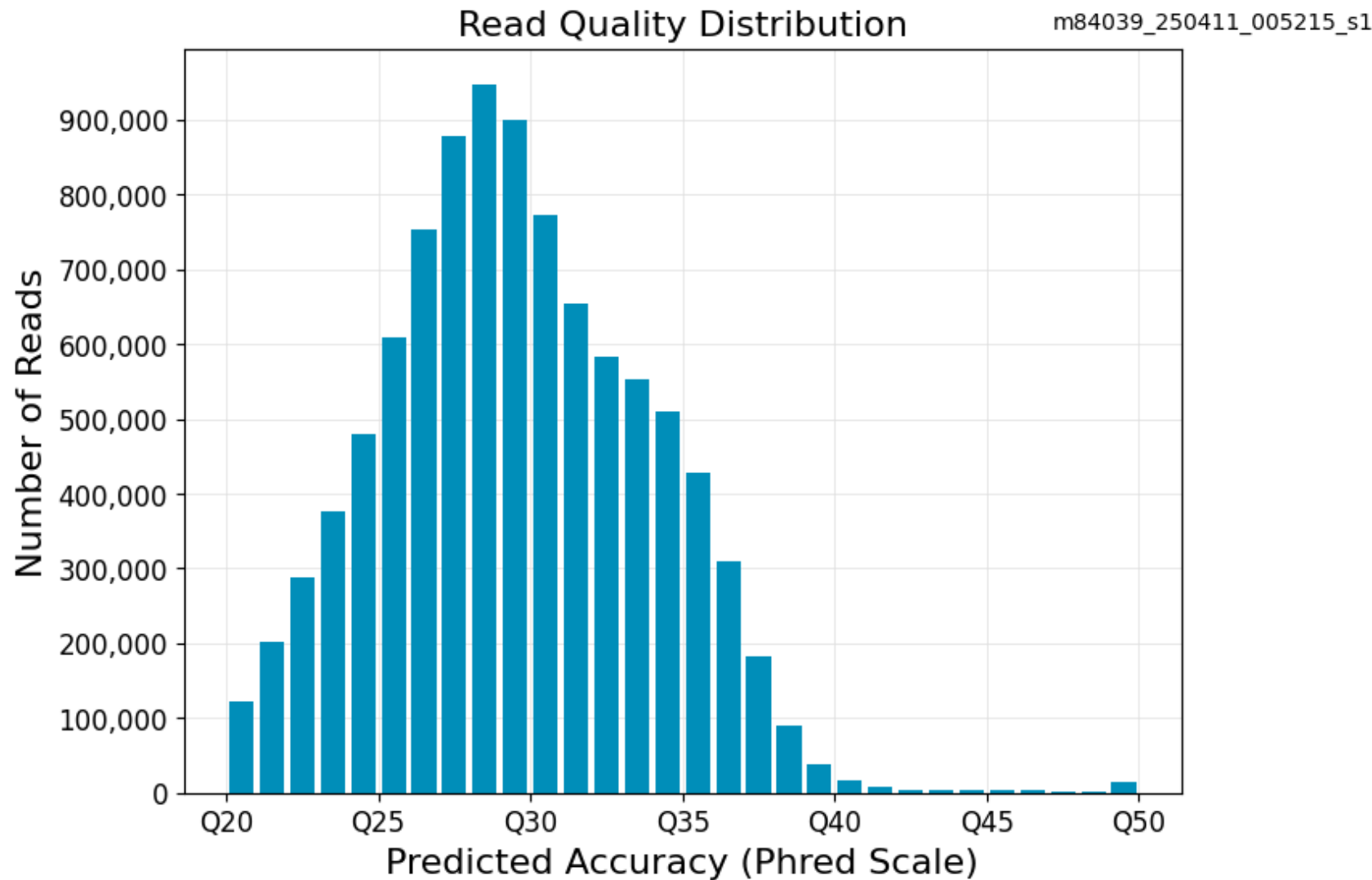
Read length distribution



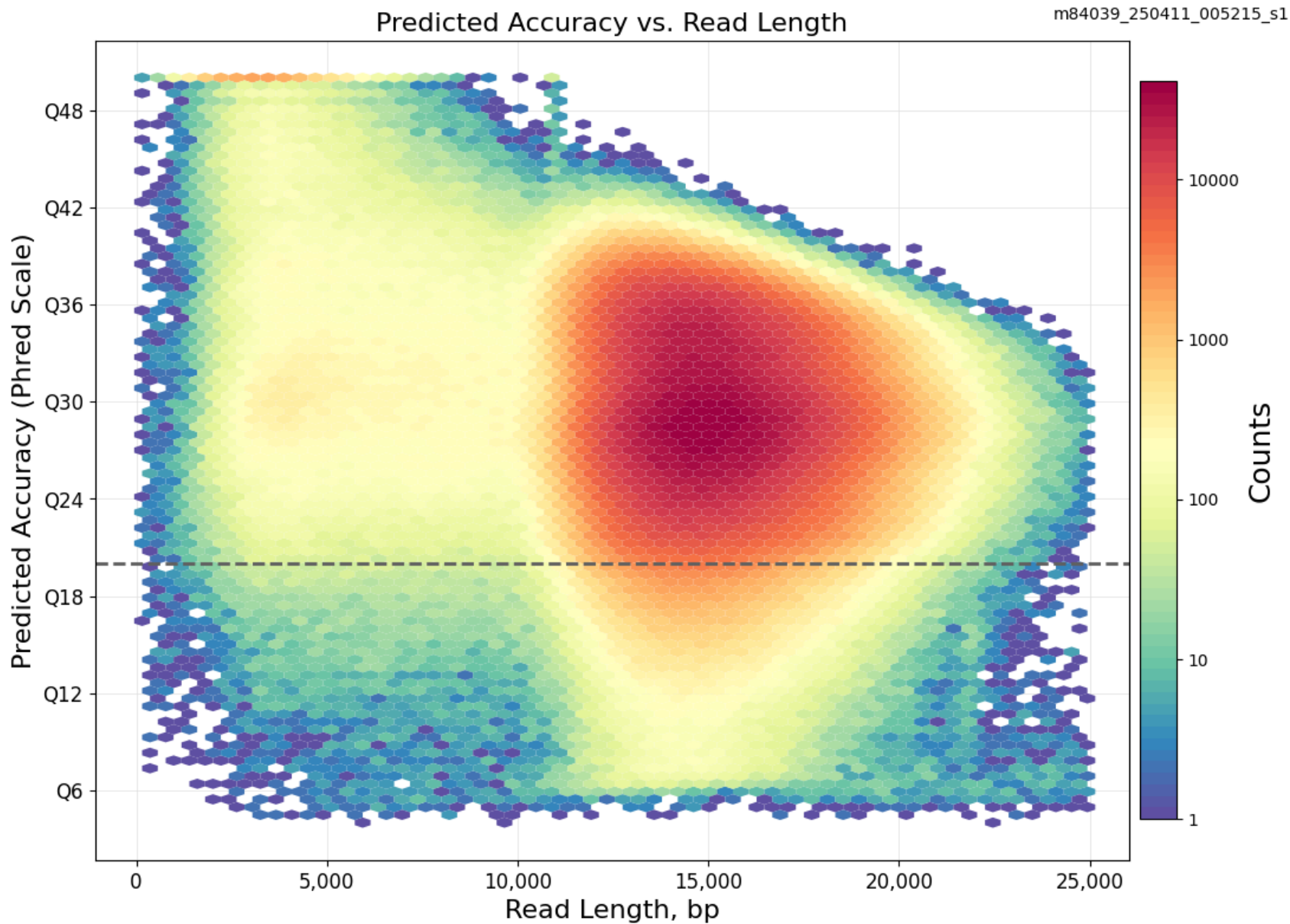
Number of Passes



Read Quality Distribution



Predicted Accuracy vs. Read Length



No Sample Setup found

No instrument run found

Parent jobs (1)

Job 124481

Summary

Job Type	import-dataset
Name	import-dataset
Comments	Description for job Import PacBio DataSet
Created At	2025-08-01 13:50:49.621
SMRT Link Version	25.3.0.273777

Child jobs (1)

Job 124624

Summary

Job Type	analysis
Pipeline	cromwell.workflows.pb_segment_reads_and_sc_isoseq
Name	KinnexRelease-10x-3pGEX-PBMC-10k-Revio-largeMem
Comments	Description for job Run Analysis Application
Created At	2025-08-02 13:22:21.949
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	10x Chromium single cell 3' cDNA primers	3332ee2d-56dc-07b0-fbcc-d7f6158a7a65	2
BarcodeSet	MAS-Seq Adapter v1 (MAS16)	10ce52cd-91c5-4bb9-4cac-2233832ecb12	17

Output files

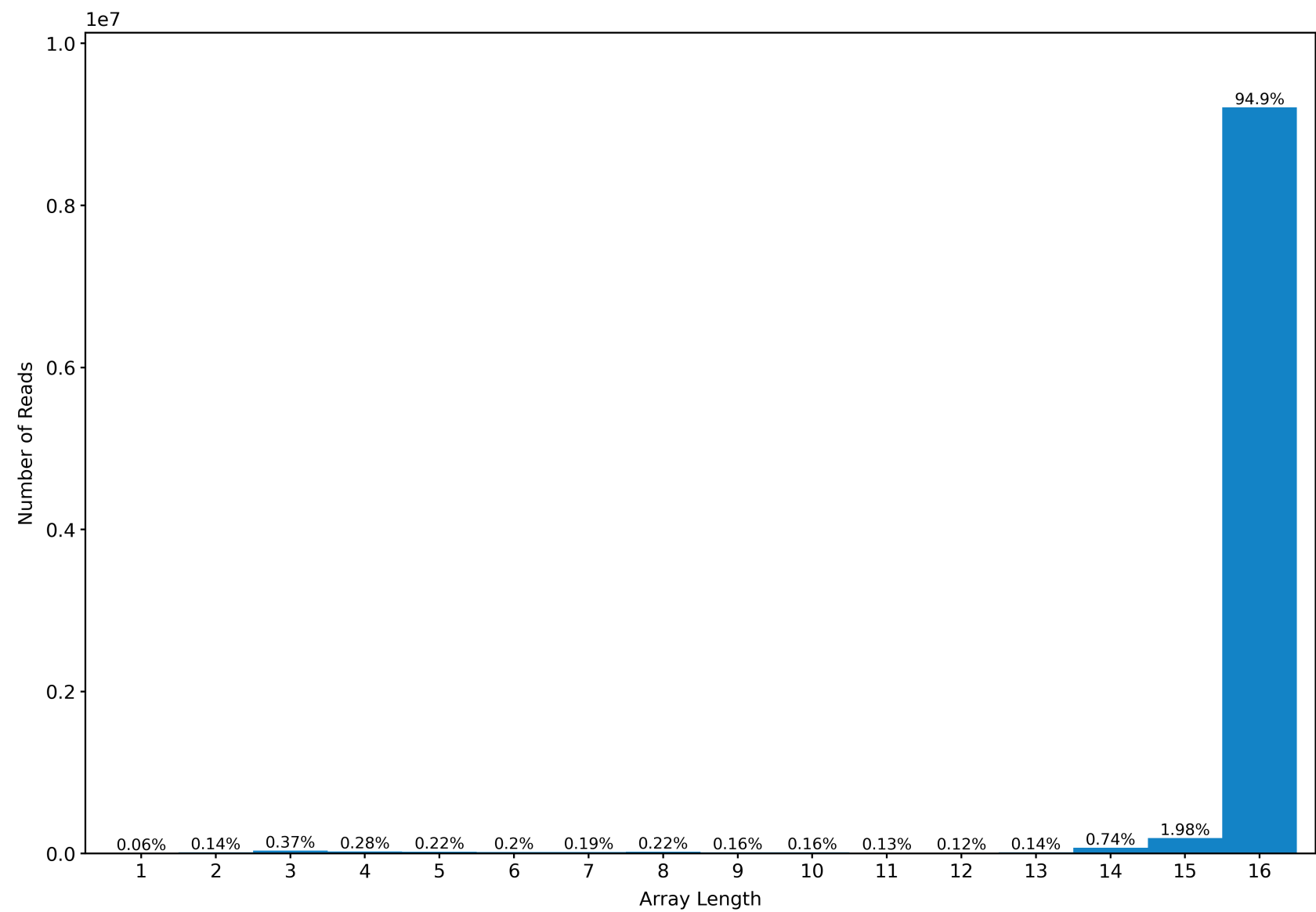
File name	Size	Type
3prime_v4_GEX_hPBMCs_10k_Rep1-Cell1 (all samples) Segmented Reads	19,636	ConsensusReadSet
Deduplicated reads after cell barcode correction, unmapped, BAM	30,994,077,155	bam
Single-cell isoform and gene matrix, tar-gzipped	3,164,855,166	tgz
Segmented Reads, passing, unaligned	95,007,878,808	bam
Non-passing reads, unaligned	3,056,016,137	bam
Unique mapped transcripts, GFF	2,042,640,042	gff
Unique mapped transcripts, filtered, GFF	889,086,439	gff
Deduplicated reads after cell barcode correction, mapped, BAM	16,491,703,086	bam
Deduplicated reads after cell barcode correction, mapped, BAM index	21,471,016	bam_bai
Unique mapped transcripts, junctions TXT	1,447,991,716	txt
Unique mapped transcripts, filtered, junctions TXT	818,668,217	txt
Deduplicated transcripts	22,646,149,931	Fasta
Collapsed transcript groups	1,895,493,653	txt
Unique mapped transcripts, classification TXT	1,459,541,085	txt
Unique mapped transcripts, filtered, classification TXT	492,053,901	txt
bcstats_report.tsv.gz	11,400,600	tsv
Analysis Log	76,908	log
737K_august_2016.txt.gz	2,238,609	txt
3M_3pgex_may_2023.REVCOMP.txt.gz	19,830,613	txt
SMRT Link Log	14,656	log

Read Segmentation

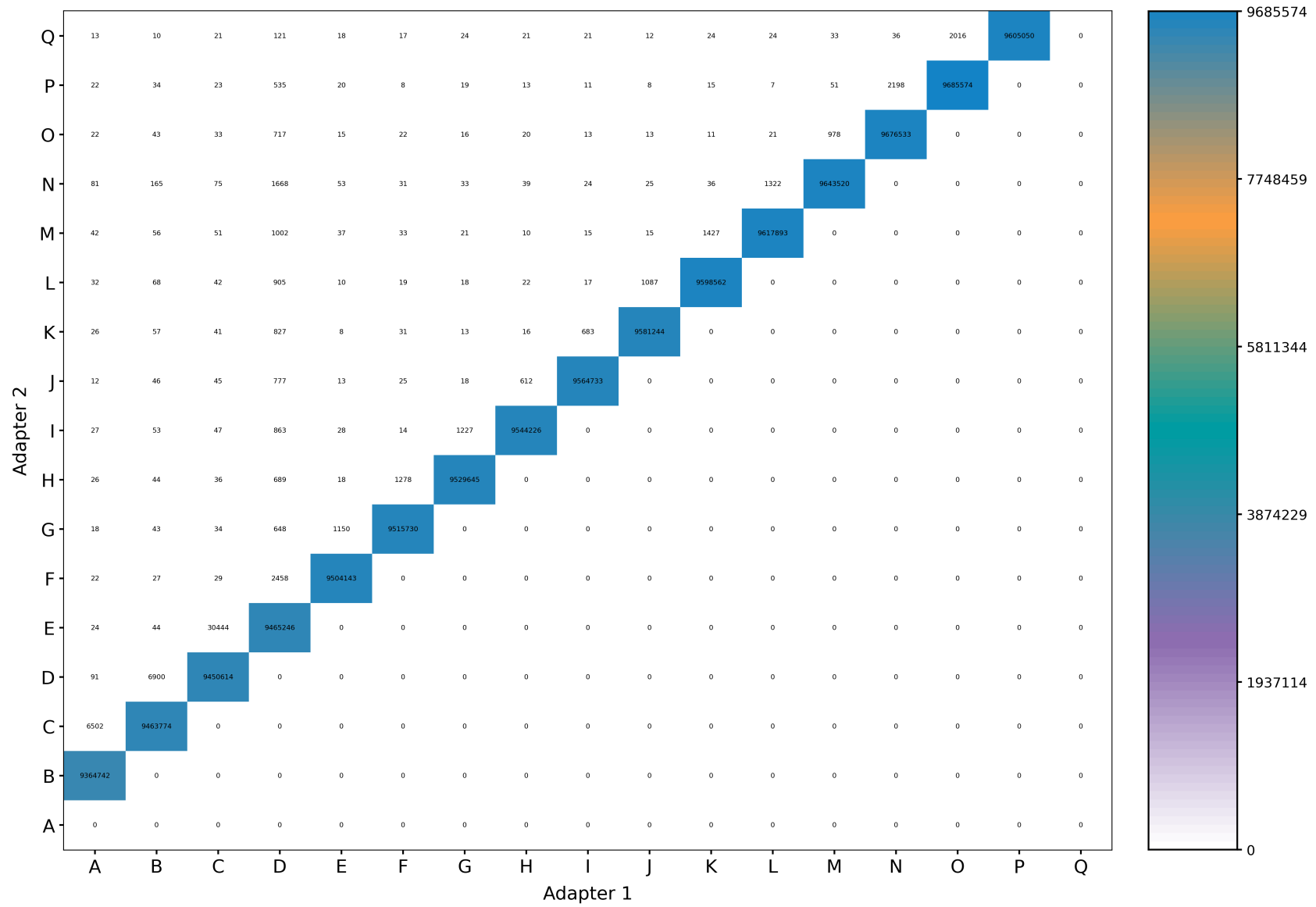
Summary

Reads	9,709,186
Segmented reads (S-reads)	152,811,229
Mean length of S-reads	928
Percent of reads with full arrays	94.89 %
Mean array size (concatenation factor)	15.74

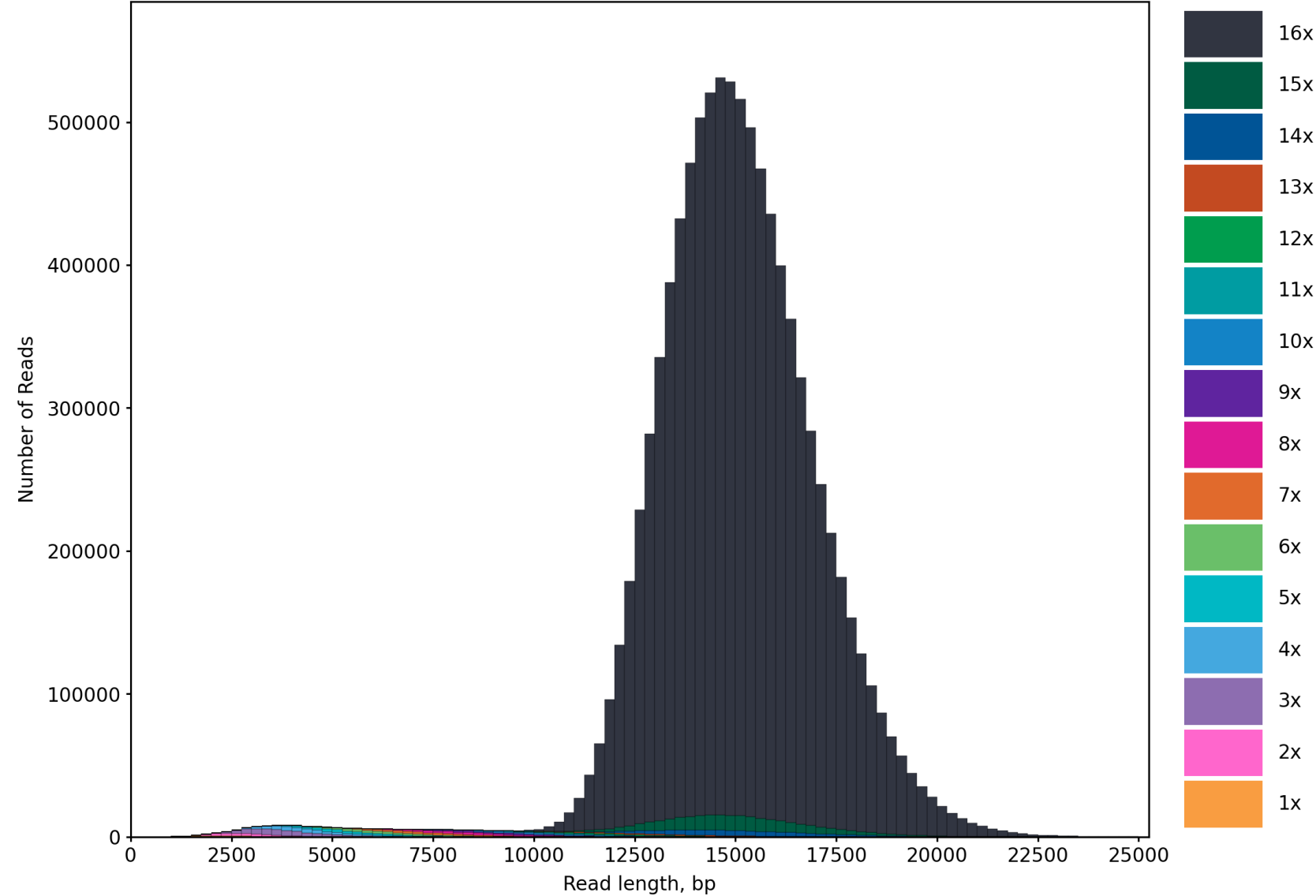
Histogram of concatemers



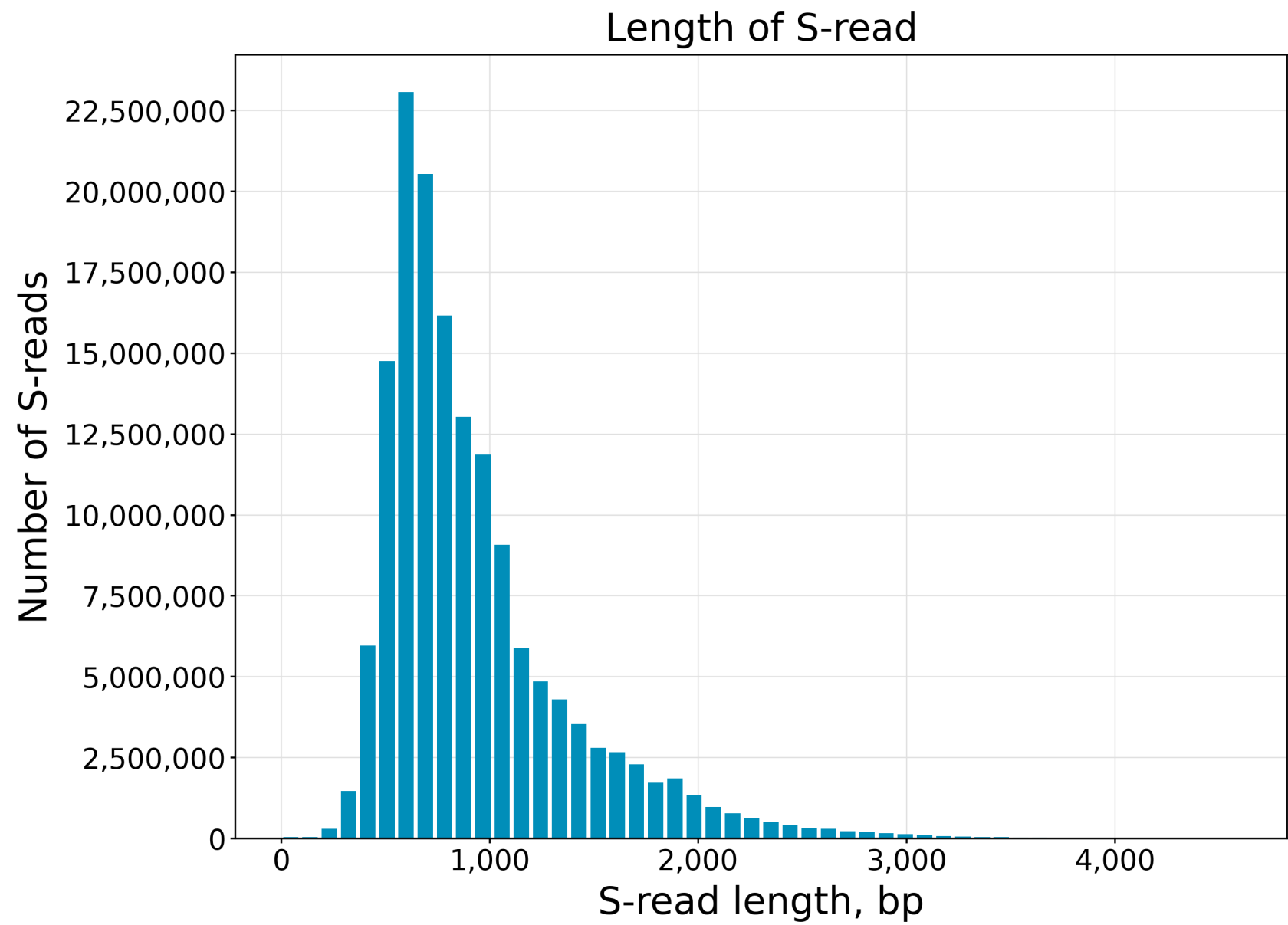
Heatmap of concatemer ligations



Length of Reads



Length of S-reads

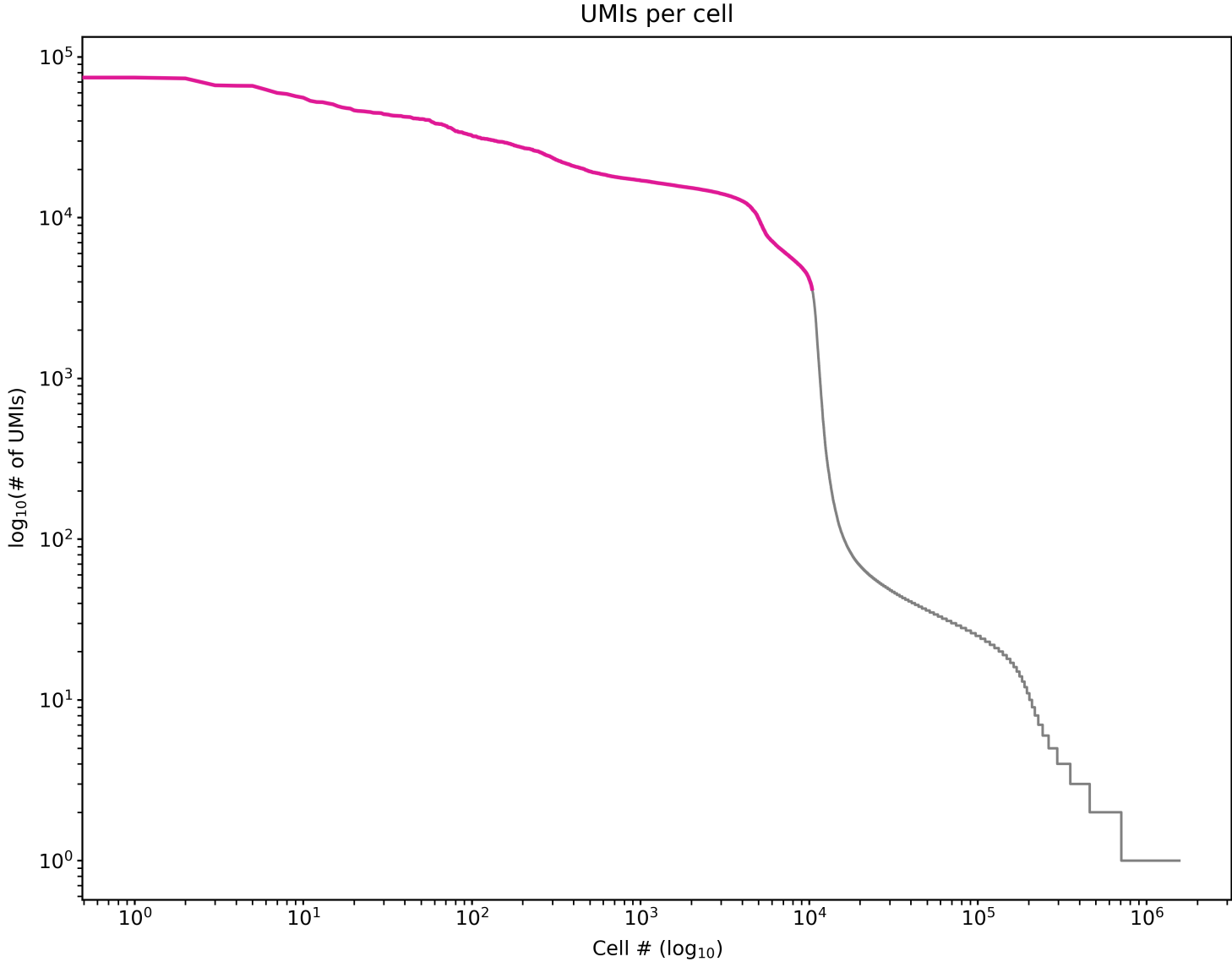


Cell Statistics

Summary

Estimated Number of Cells	10,324
Reads in Cells	90.65%
Mean Reads per Cell	12,978
Median UMIs per Cell	9,194

Barcode Rank Plot



Read Statistics

Summary

Reads	152,811,229
Read Type	SEGMENT
Reads with 5' and 3' Primers with extracted UMIs and Barcodes	151,565,748
Non-Concatamer Reads with 5' and 3' Primers and Poly-A Tail (FLNC reads)	147,805,910
FLNC Reads with Valid Barcodes	144,807,139
FLNC Reads with Valid Barcodes, corrected	147,466,710
Reads after Barcode Correction and UMI Deduplication	110,681,049

Transcript Statistics

Summary

FLNC Reads Mapped Confidently to Genome	120,736,189
FLNC Reads Mapped Confidently to Transcriptome	79,264,395
Total Unique Genes	1,902,966
Total Unique Genes, filtered	98,700
Total Unique Genes, known genes only	37,474
Total Unique Genes, filtered, known genes only	28,955
Total Unique Transcripts	4,399,569
Total Unique Transcripts, filtered	999,042
Total Unique Transcripts, known transcripts only	103,849
Total Unique Transcripts, filtered, known transcripts only	93,952

Transcript Summary

Value	Analysis Metric
3,369	Median Genes per Cell
2,476	Median Genes per Cell, known genes only
4,522	Median Transcripts per Cell
2,145	Median Transcripts per Cell, known transcripts only
1,902,966	Total Unique Genes
37,474	Total Unique Genes, known genes only
4,399,569	Total Unique Transcripts
103,849	Total Unique Transcripts, known transcripts only

Transcript Summary, filtered

Value	Analysis Metric
1,822	Median Genes per Cell
1,787	Median Genes per Cell, known genes only
2,389	Median Transcripts per Cell
1,839	Median Transcripts per Cell, known transcripts only
98,700	Total Unique Genes
28,955	Total Unique Genes, known genes only
999,042	Total Unique Transcripts
93,952	Total Unique Transcripts, known transcripts only

Transcript Classification

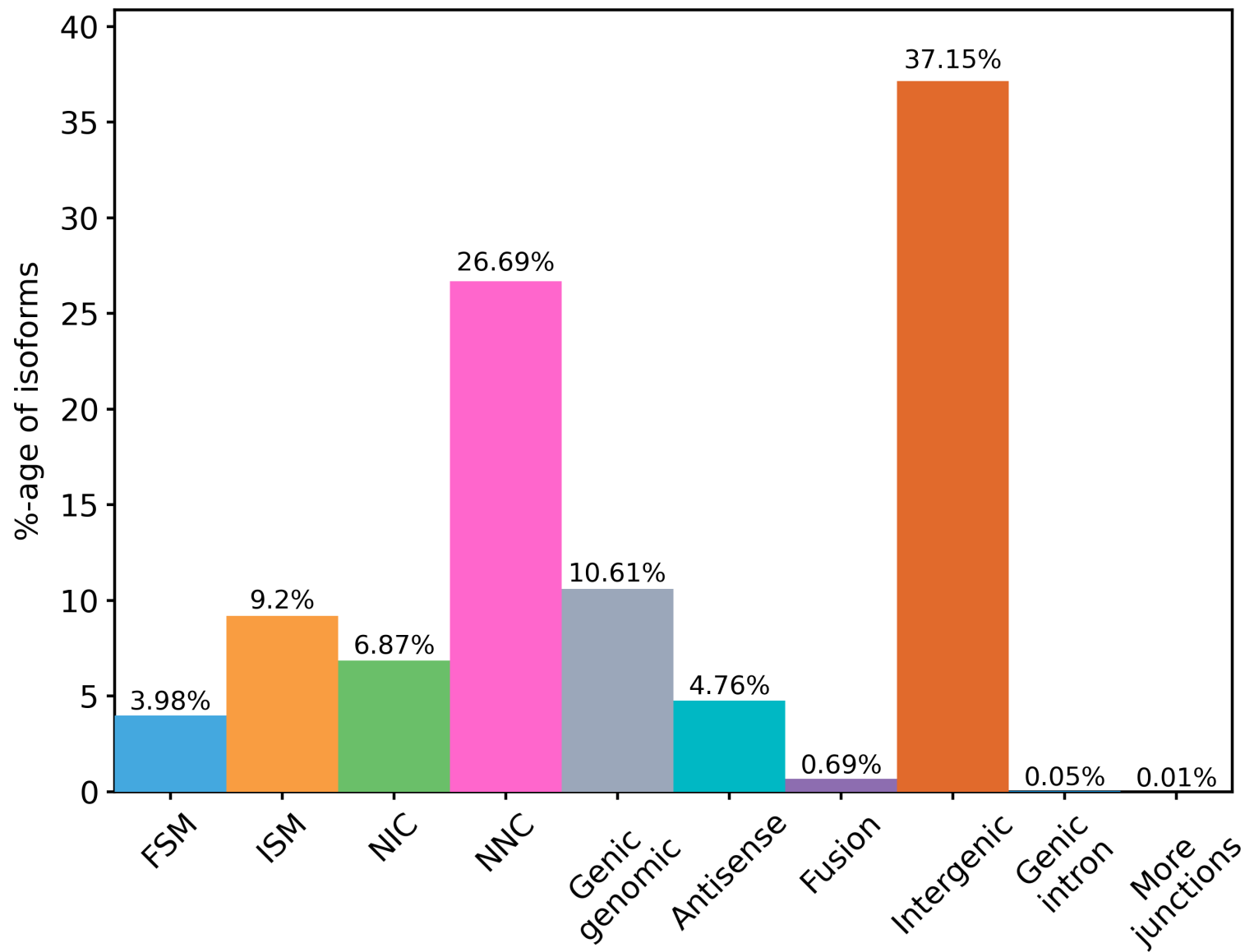
Category	Count	CAGE Detected	CAGE Detected (%)	polyA Motif Detected	polyA Motif Detected (%)
FSM	196828	119402	60.66%	58617	29.78%
ISM	455016	204144	44.87%	131503	28.90%
NIC	339815	235082	69.18%	110259	32.45%
NNC	1320505	874184	66.20%	491387	37.21%
Antisense	235426	10877	4.62%	63174	26.83%
Fusion	34017	20041	58.91%	13312	39.13%
More junctions	457	246	53.83%	151	33.04%
Genic intron	2254	12	0.53%	721	31.99%
Genic genomic	525027	138280	26.34%	156378	29.78%
Intergenic	1838219	21953	1.19%	615293	33.47%

Transcript Classification, filtered

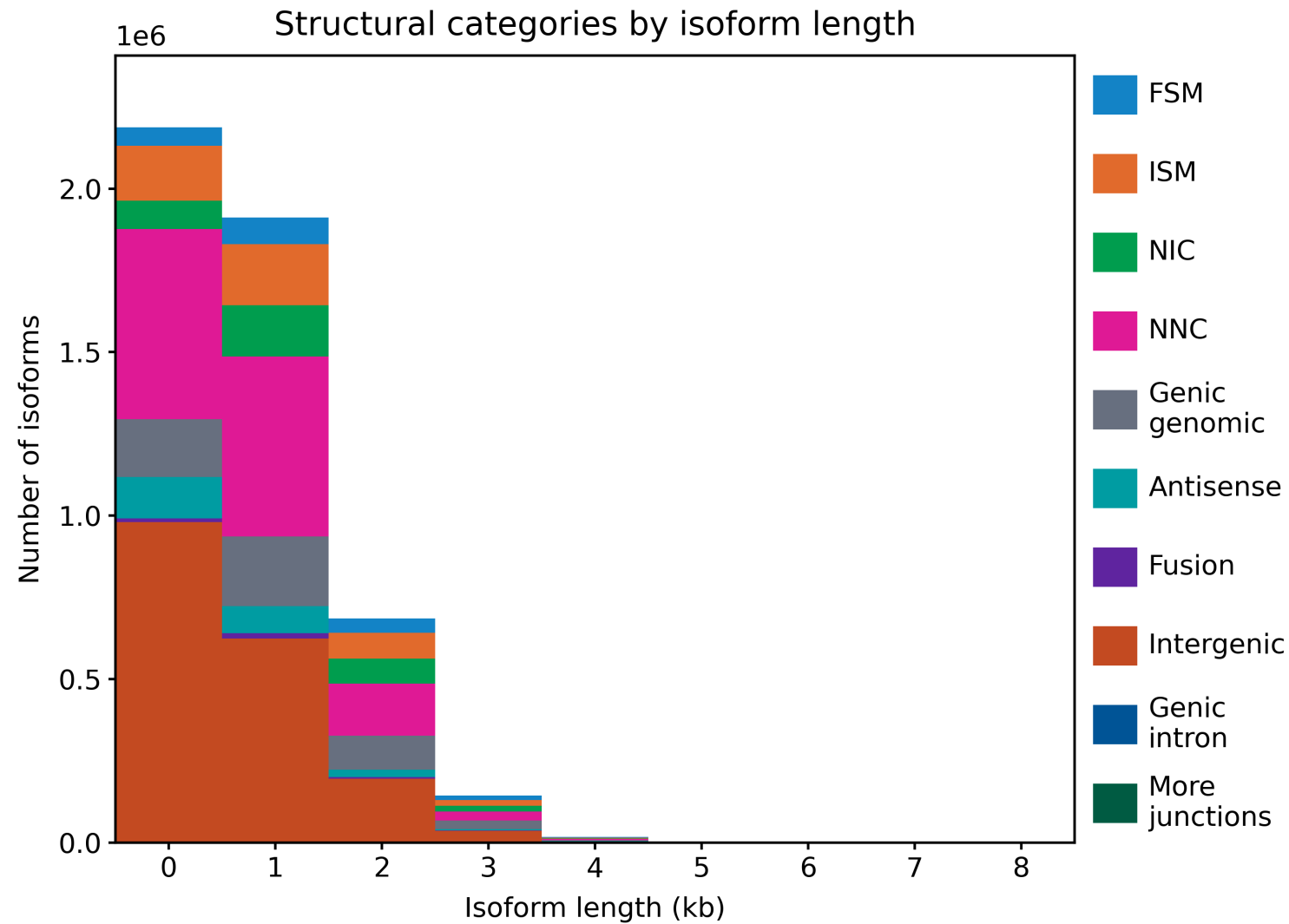
Category	Count	CAGE Detected	CAGE Detected, (%)	polyA Detected	polyA Detected, (%)
FSM	167573	103997	62.06%	58617	34.98%
ISM	314765	158530	50.36%	111958	35.57%
NIC	273808	190384	69.53%	109290	39.91%
NNC	514400	345302	67.13%	232608	45.22%
Antisense	17604	2071	11.76%	9938	56.45%
Fusion	16945	10215	60.28%	8019	47.32%
More junctions	282	158	56.03%	127	45.04%
Genic intron	0	0	0.00%	0	0.00%
Genic genomic	18361	9601	52.29%	7788	42.42%
Intergenic	63690	2479	3.89%	43701	68.62%

Isoform Distribution Across Categories

Isoform distribution across structural categories

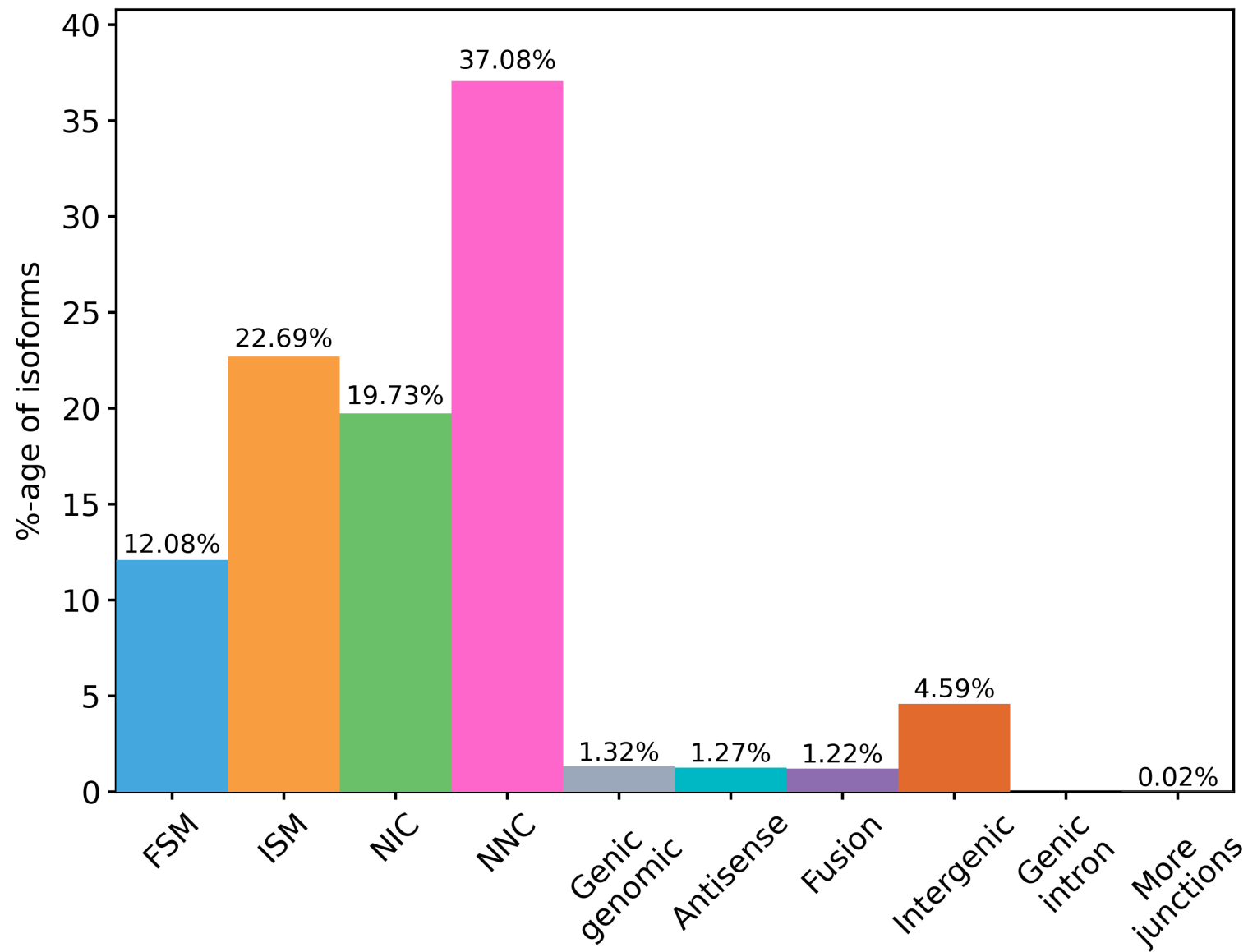


Structural Categories by Transcript Length

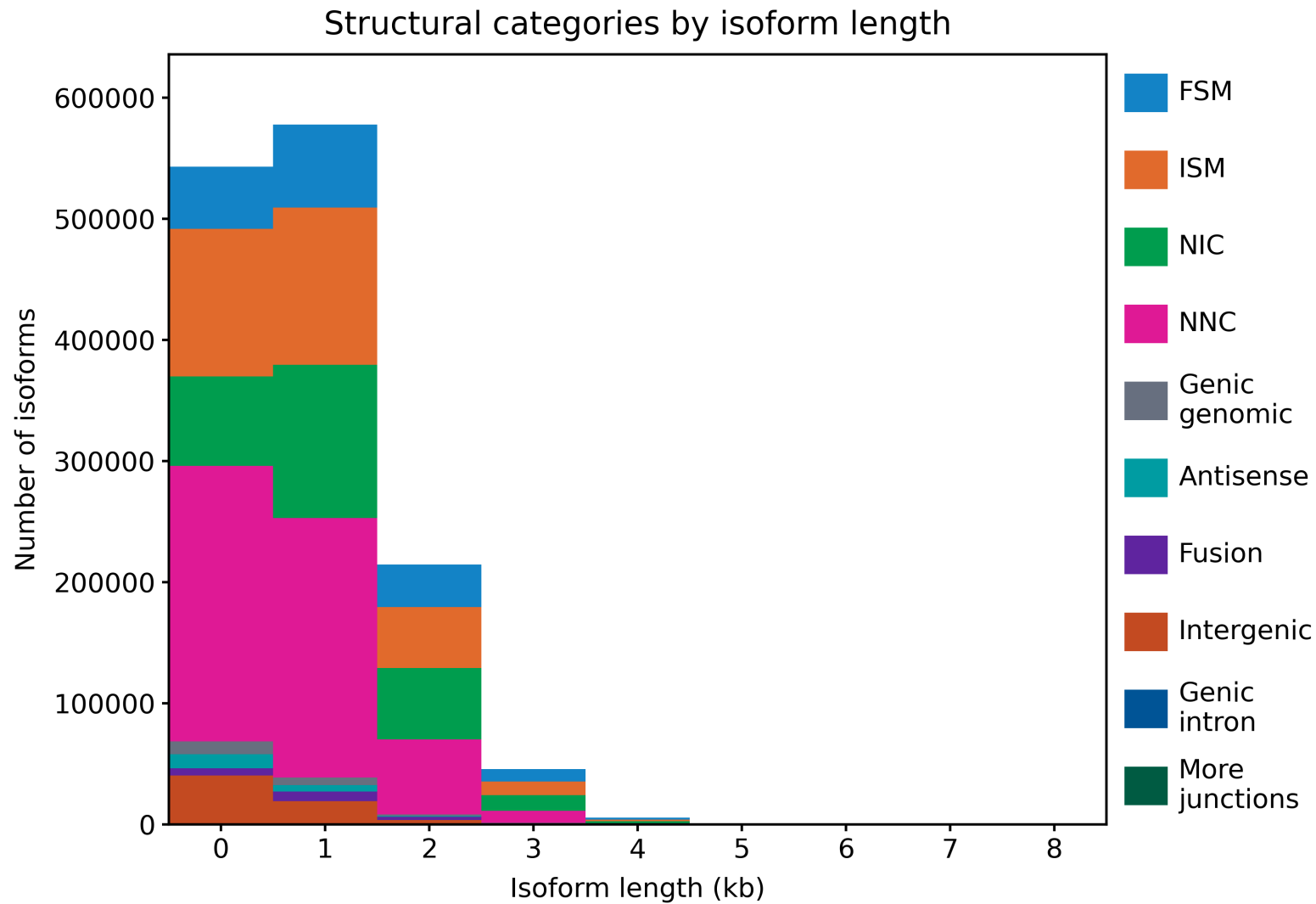


Isoform Distribution Across Categories

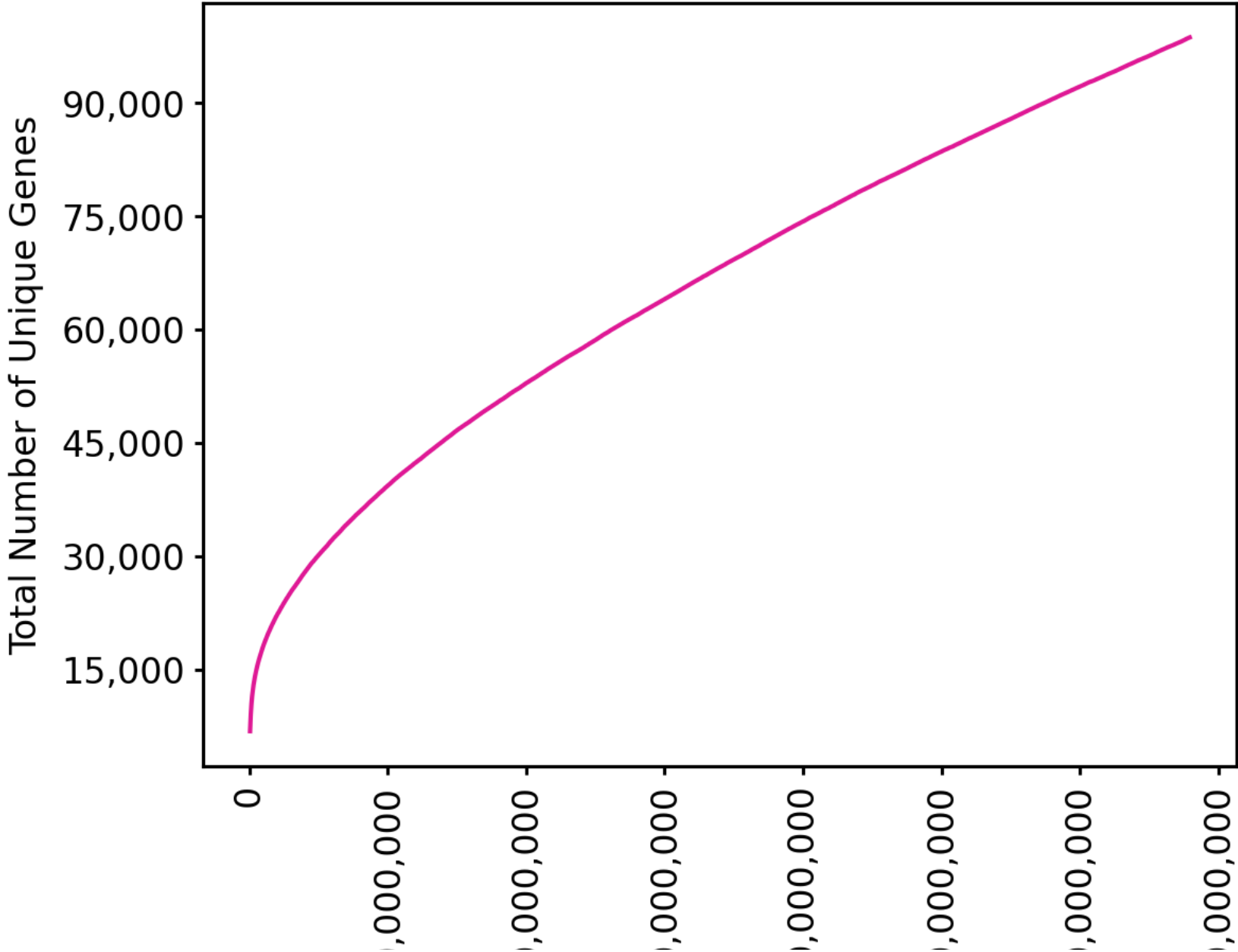
Isoform distribution across structural categories



Structural Categories by Transcript Length



Gene Saturation, all genes, filtered



10

20

30

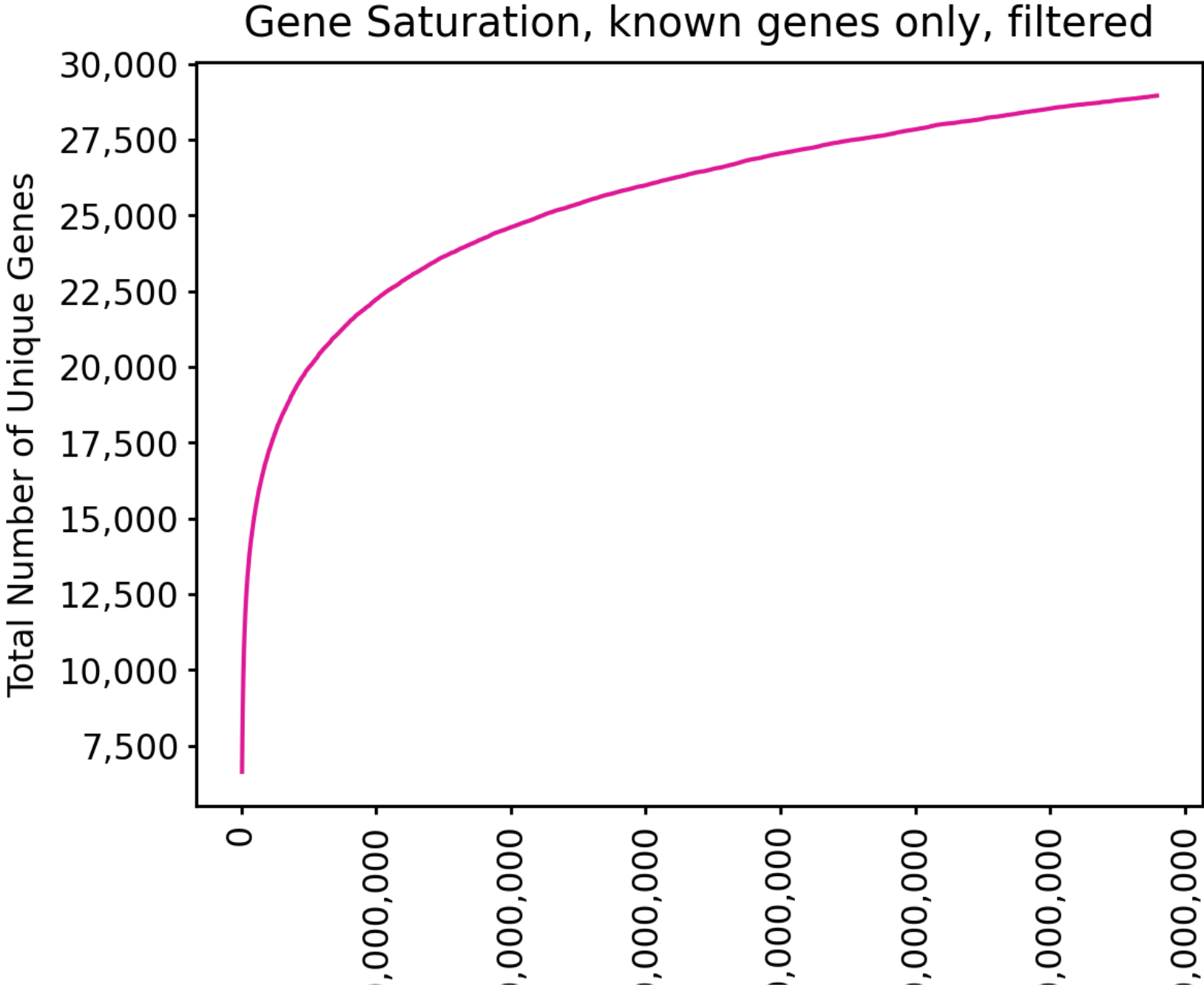
40

50

60

70

Total Reads



10

20

30

40

50

60

70

Total Reads

