Report for dataset 3prime_v4_GEX_hPBMCs_20k_Rep1-Cell2 (all samples)

Dataset 0adf5da8-9938-4cc5-8c39-af4a98257573

Summary

Name	3prime_v4_GEX_hPBMCs_20k_Rep1-Cell2 (all samples)
Created At	2025-04-12 17:26:48.950
Number of Records	9,489,284
Total Length	145,665,994,491
Movie Name	m84039_250411_025511_s2
ICS Version	13.3.0.253824
Well Sample	3prime_v4_GEX_hPBMCs_20k_Rep1
Biological Sample	3prime_v4_GEX_hPBMCs_20k_Rep1
Barcode Name	bcM0002--bcM0002

CCS Analysis Report

Summary

HiFi reads	9.5 M
HiFi reads yield	146.15 Gb
HiFi reads length (mean)	15.34 kb
HiFi reads length (median, bp)	15,453
HiFi Read Length N50 (bp)	15,724
HiFi Read Quality (median)	Q29
HiFi Read Quality (median)	29
Base Quality Q30 (%)	95.34%
HiFi Number of Passes (mean)	8
Missing adapters (%)	4.08%

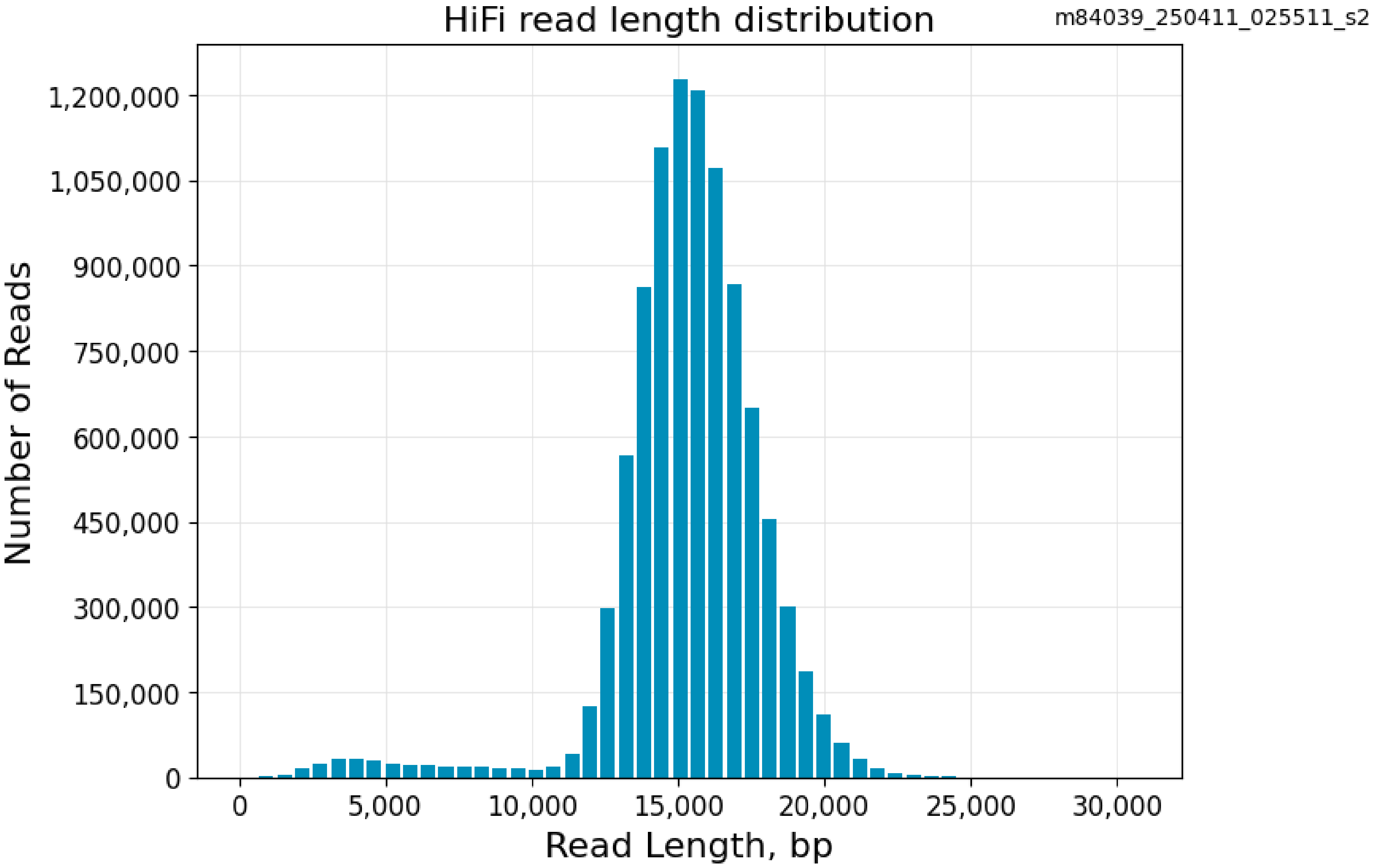
HiFi Read Length Summary

Read Length (kb)	Reads	Reads (%)	Yield (Gb)	Yield (%)
0	9529554	100	146.15 Gb	100
5,000	9382809	98	145.64 Gb	100
10,000	9225831	97	144.50 Gb	99
15,000	5680604	60	95.37 Gb	65
20,000	168350	2	3.52 Gb	2
25,000	391	0	0.01 Gb	0
30,000	15	0	0.00 Gb	0
35,000	2	0	0.00 Gb	0
40,000	0	0	0.00 Gb	0

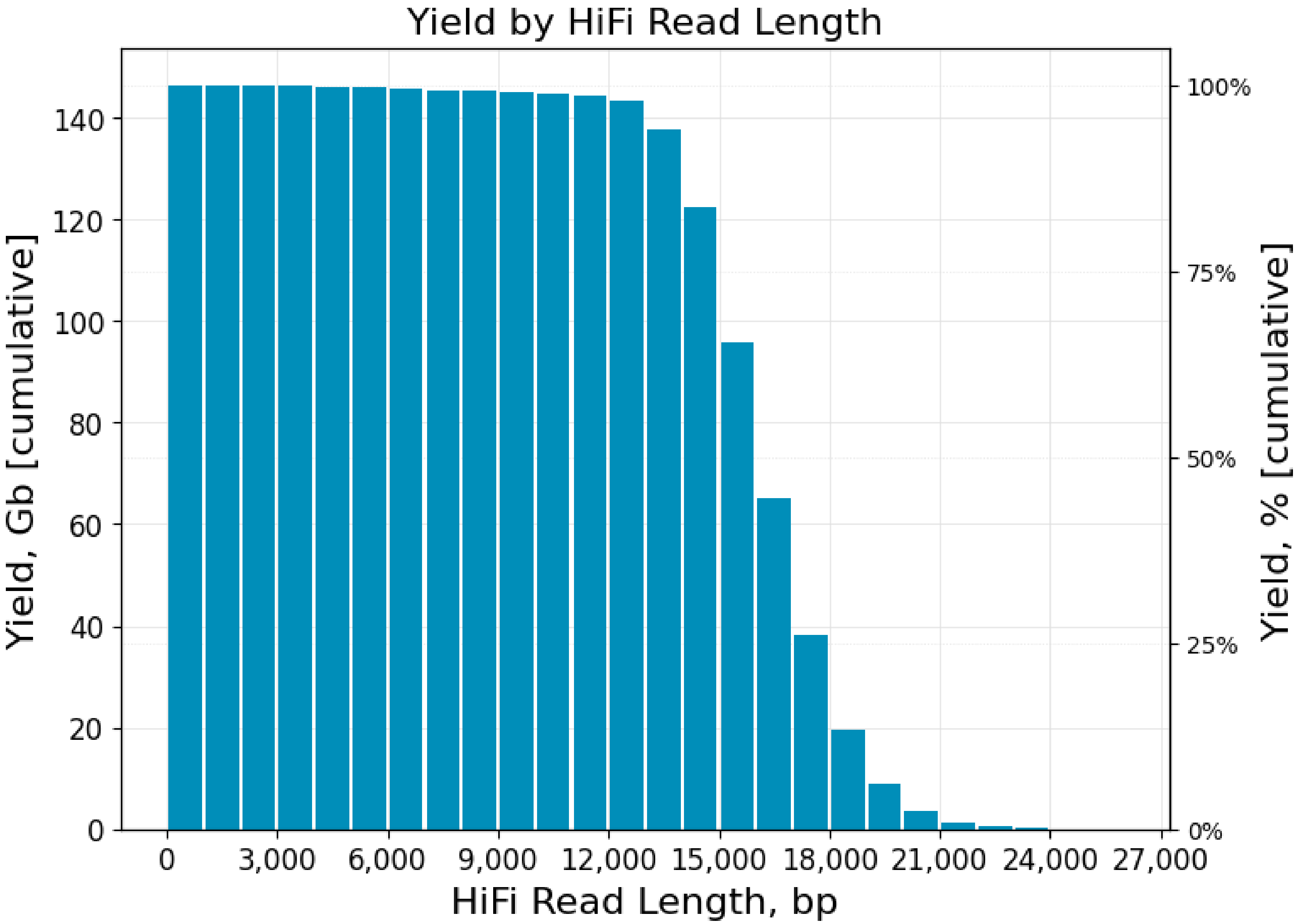
HiFi Read Quality Summary

Read Quality (Phred)	Reads	Reads (%)	Yield (Gb)	Yield (%)
Q20	9529554	100	146.15 Gb	100
Q30	4203807	44	63.29 Gb	43
Q40	86540	1	0.46 Gb	0
Q50	23172	0	0.08 Gb	0

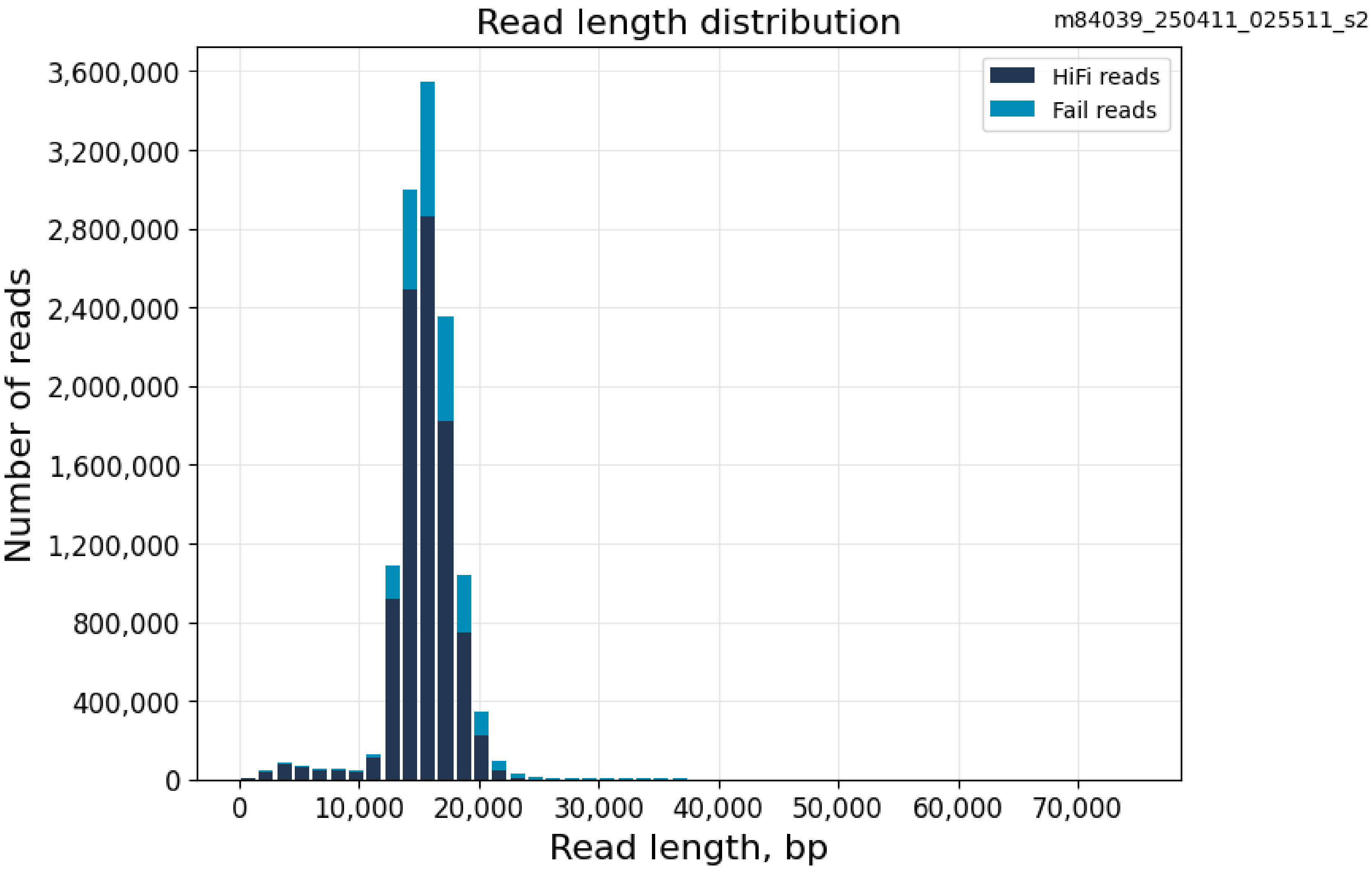
HiFi read length distribution



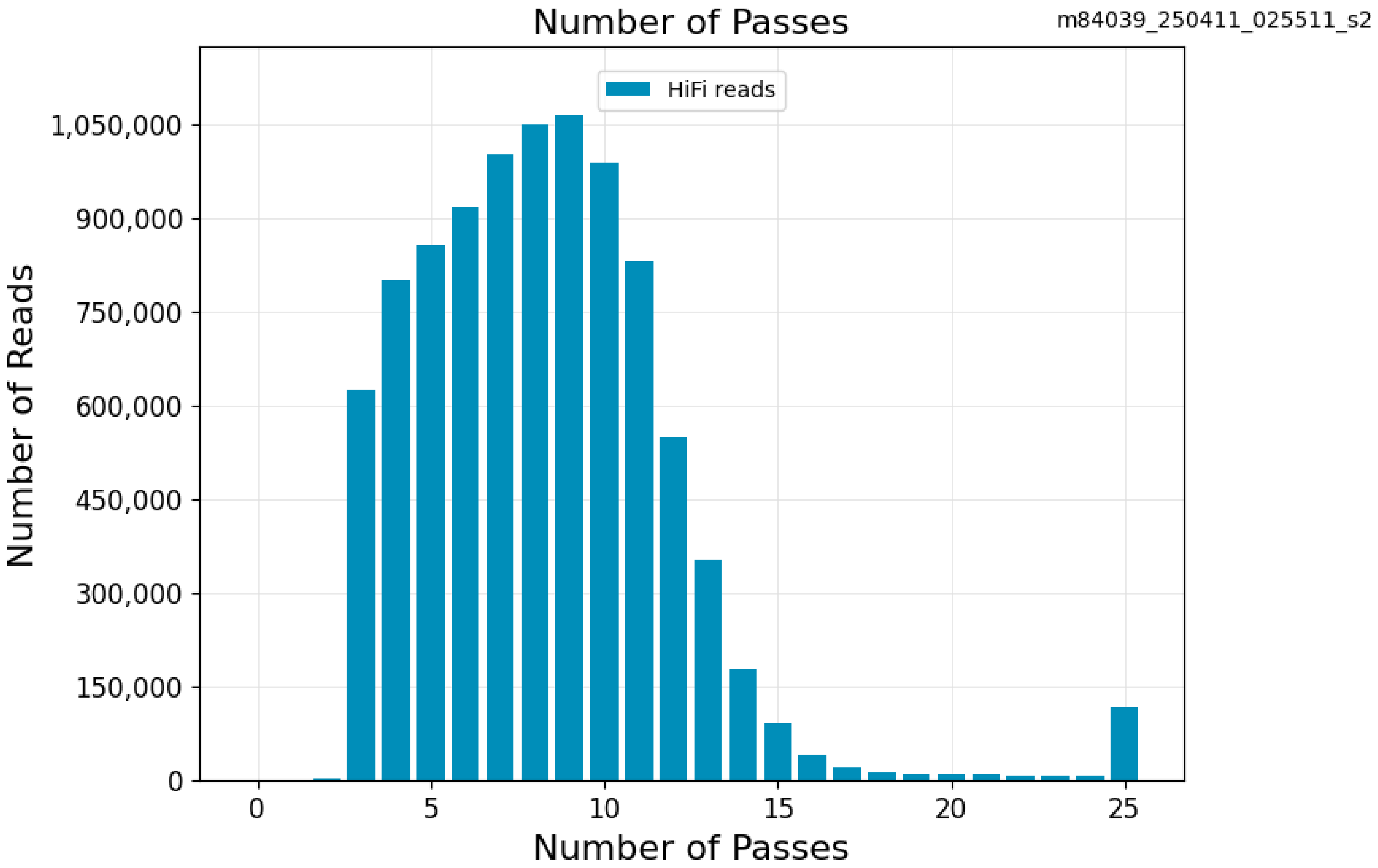
Yield by HiFi Read Length



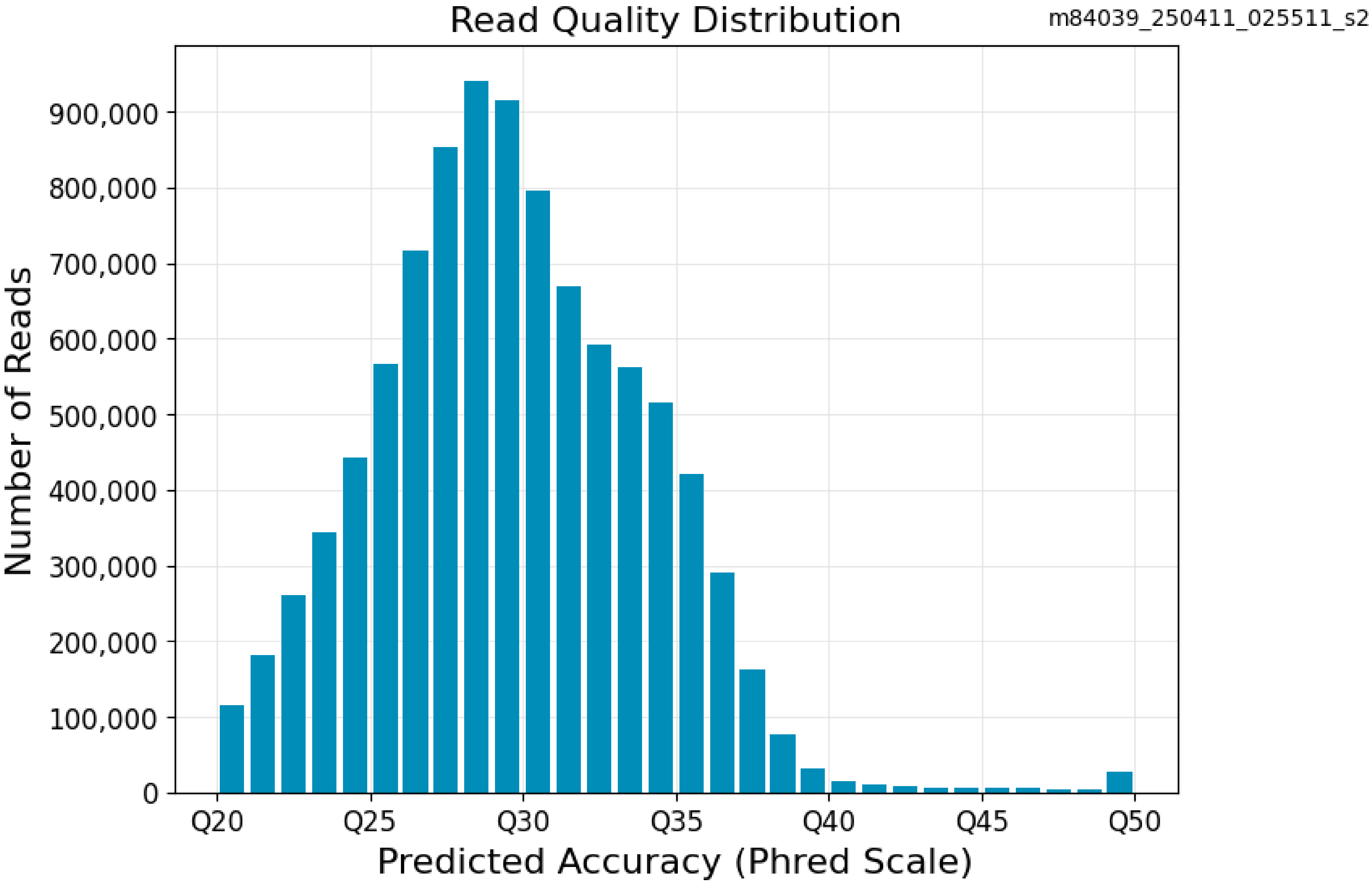
Read length distribution



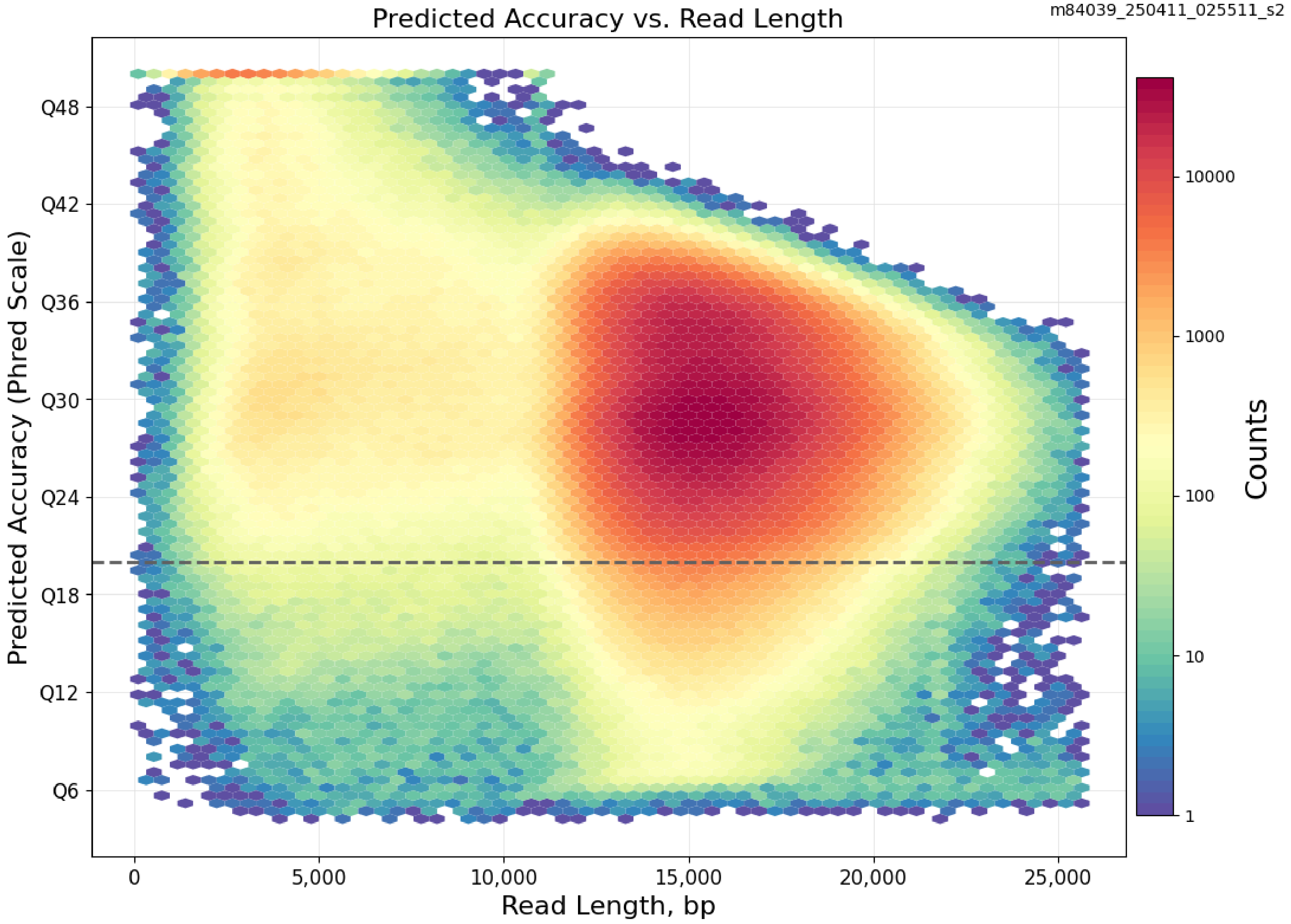
Number of Passes



Read Quality Distribution



Predicted Accuracy vs. Read Length



Loading Report

Summary

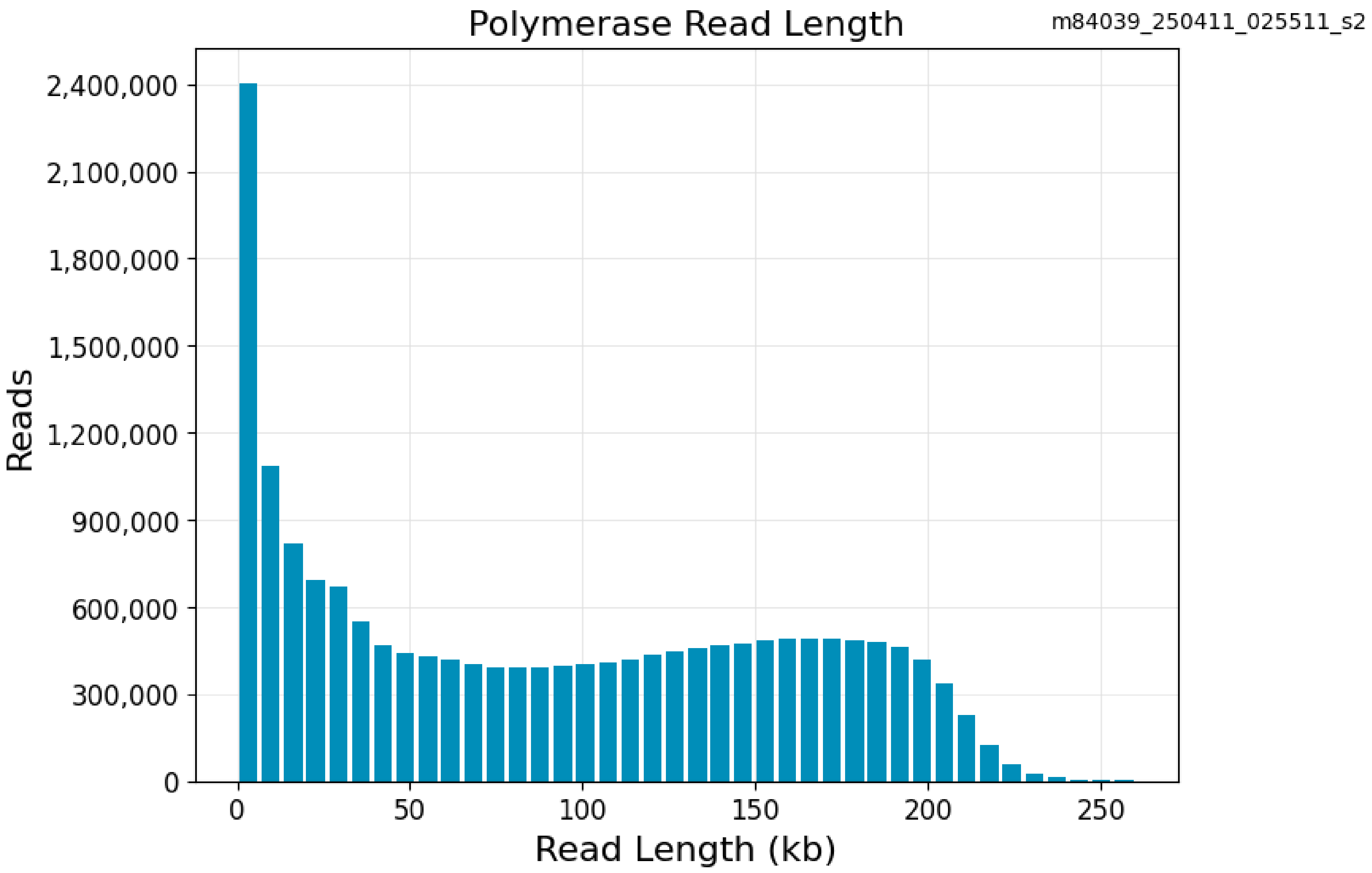
Productive ZMWs	25,165,824
Productivity 0	6,983,492
Productivity 1	18,078,462
Productivity 2	103,870

Raw Data Report

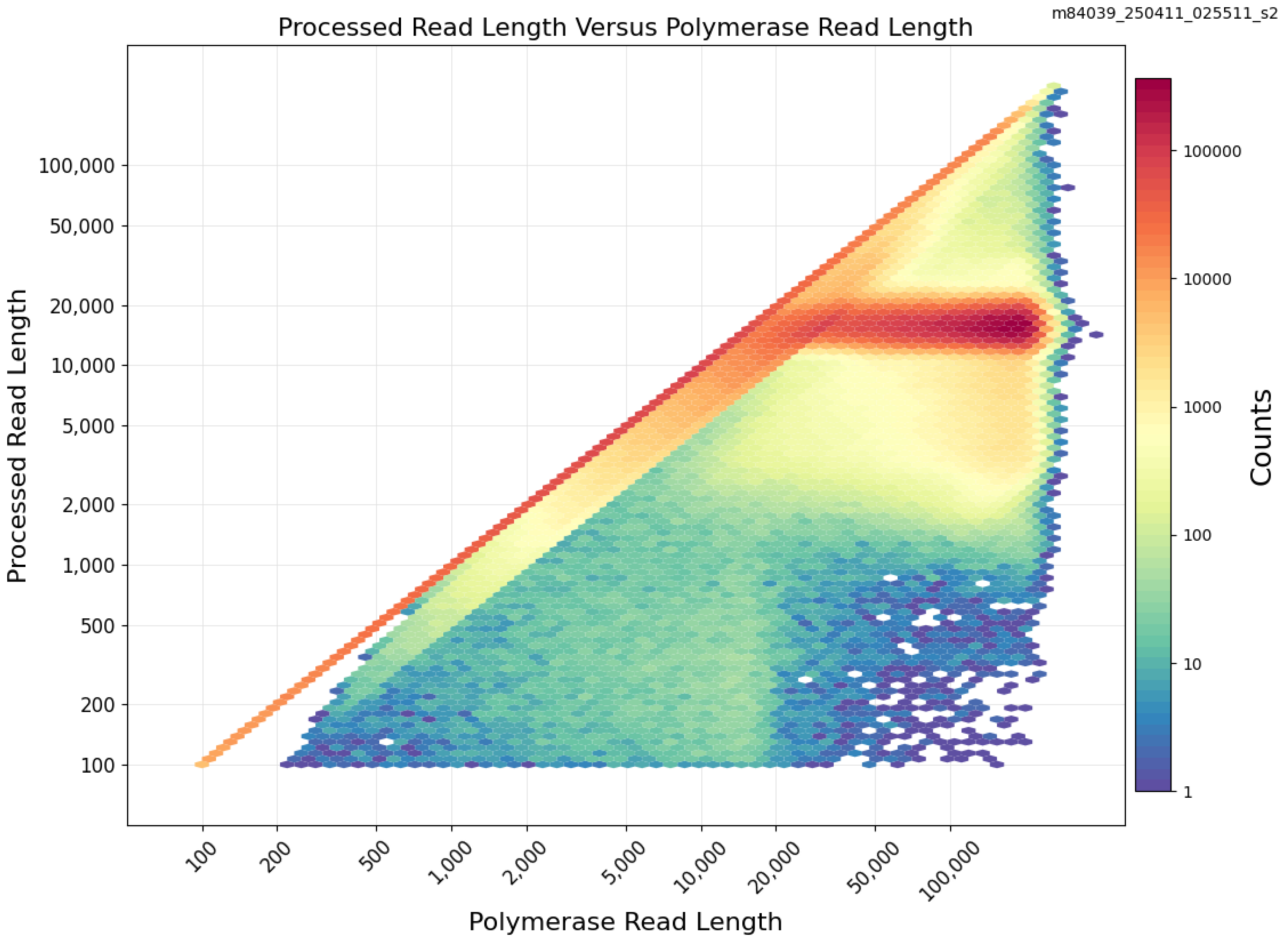
Summary

Polymerase Read Bases	1,549,397,275,731
Polymerase Reads	18,074,717
Polymerase Read Length (mean)	85.72 kb
Polymerase read length (N50)	152.25 kb
Polymerase read length longest subread length (mean)	16.17 kb
Polymerase read length longest subread length (N50)	17.25 kb
Unique Molecular Yield	279,911,301,120
Local Base Rate	2.42

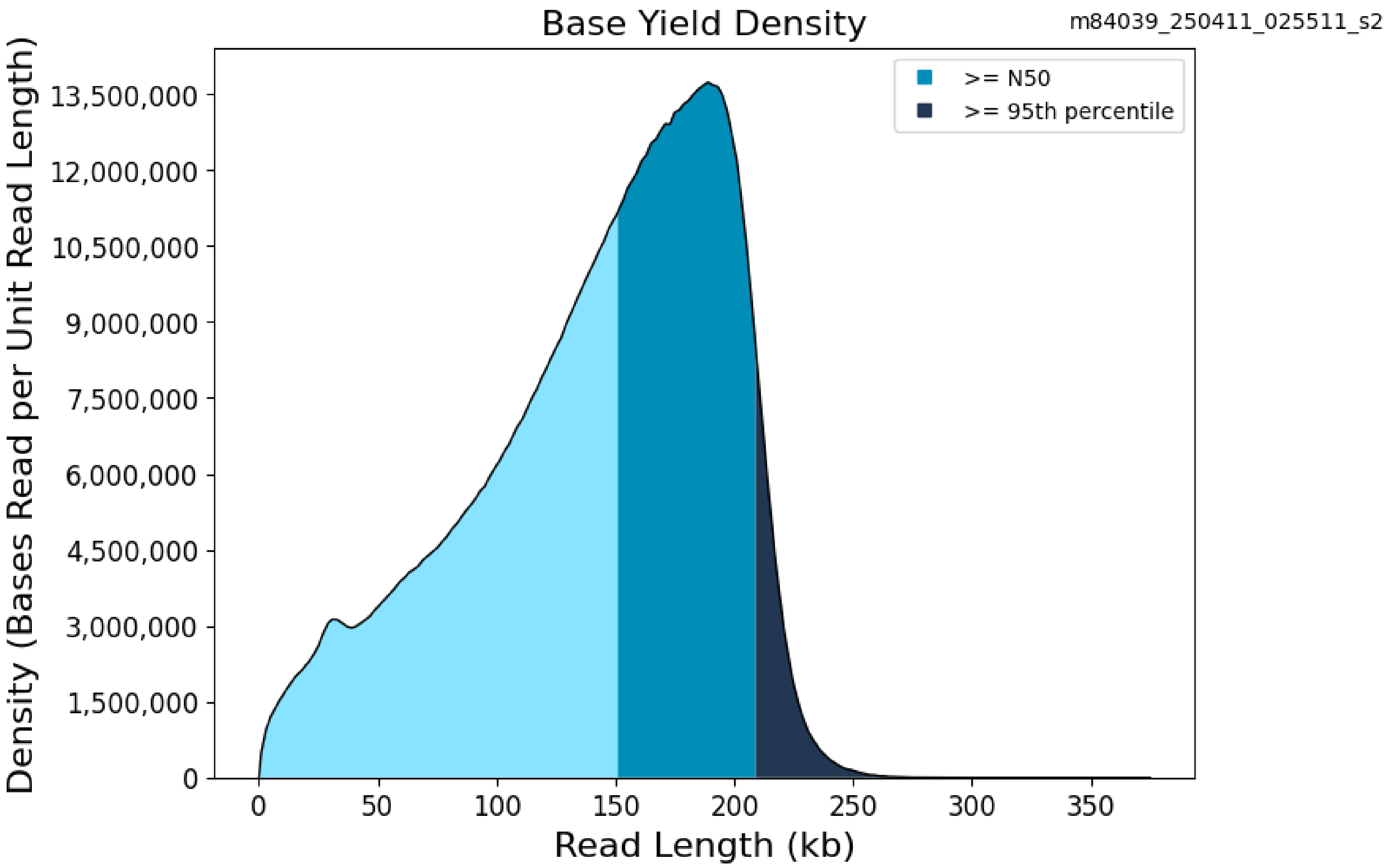
Polymerase Read Length



Longest Subread Length Versus Polymerase Read Length



Base Yield Density



Barcodes

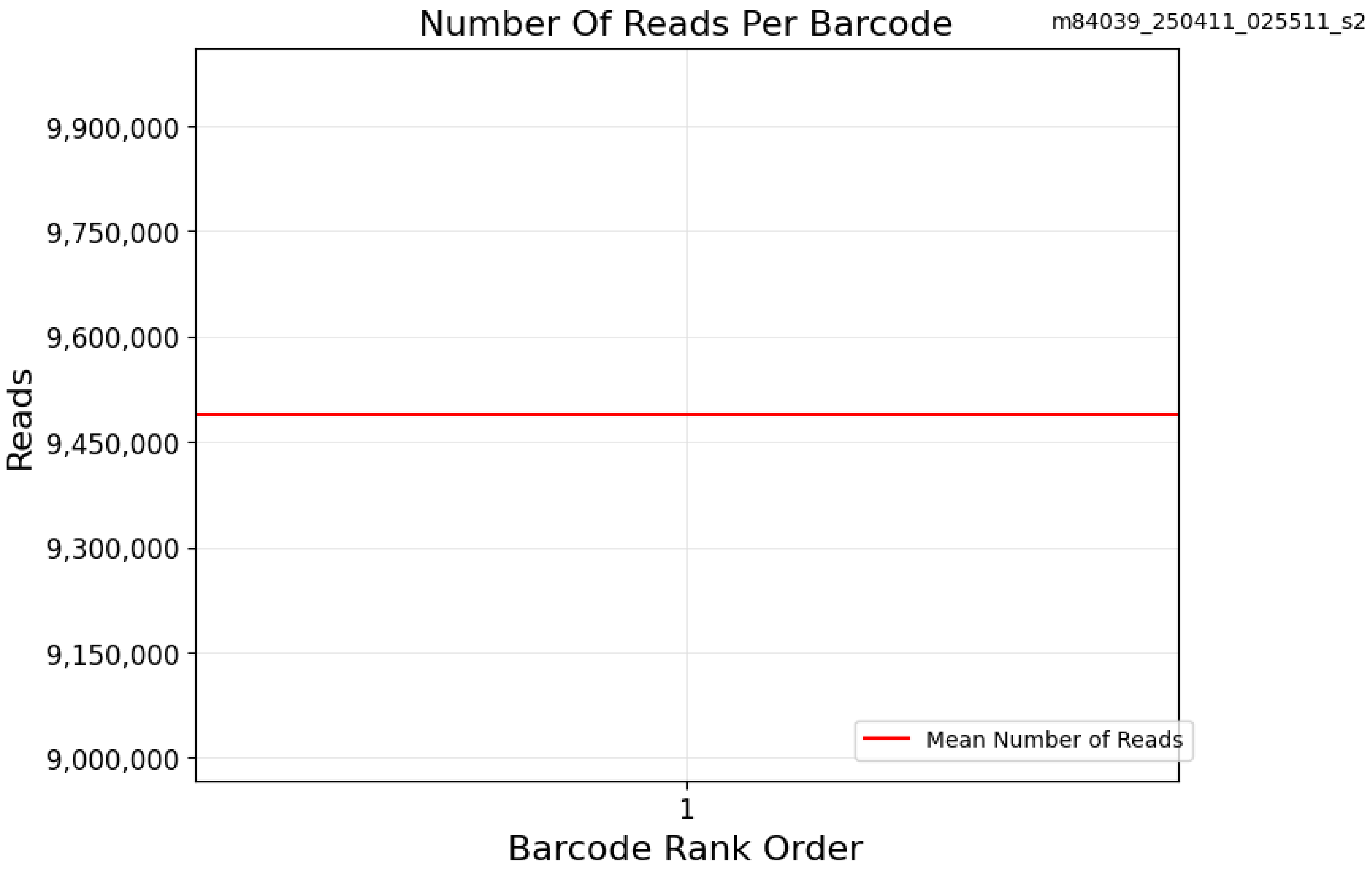
Summary

Unique Barcodes	1
Barcoded HiFi Reads	9,489,284
Unbarcoded HiFi Reads	40,270
Barcoded HiFi Reads (%)	99.58 %
Barcoded HiFi yield (Gb)	145.67 Gb
Unbarcoded HiFi yield (Gb)	0.48 Gb
Barcoded HiFi Yield (%)	99.67 %
Mean HiFi Reads per Barcode	9,489,284
Max. HiFi Reads per Barcode	9,489,284
Min. HiFi Reads per Barcode	9,489,284
Barcoded HiFi read length (mean, kb)	15.35 kb
Unbarcoded HiFi read length (mean, kb)	12.00 kb

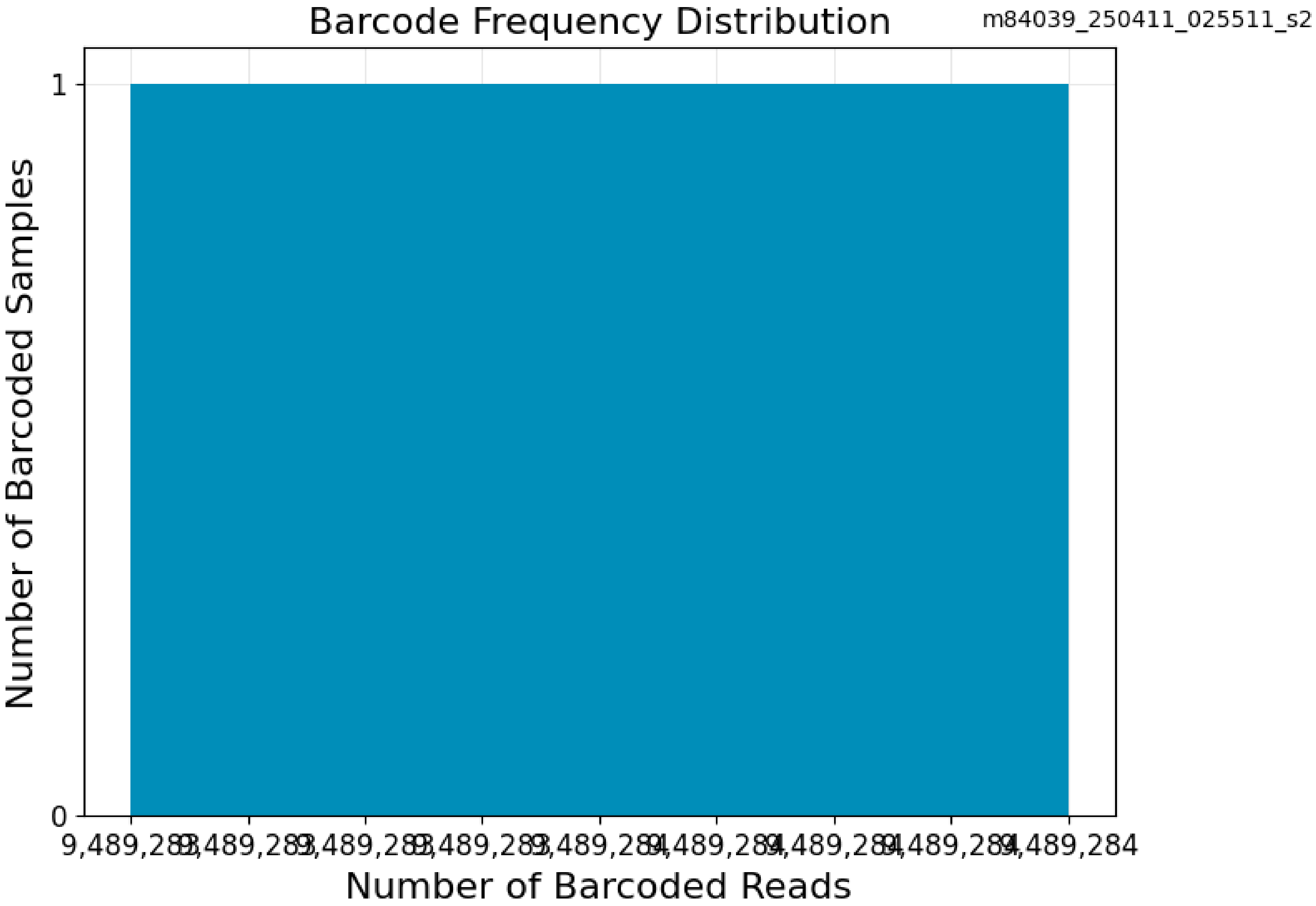
Barcode Data

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_Rep1	bcM0002--bcM0002	96.8	9489284	15350	Q29	145665994491	142076
No Name	Not Barcoded	0.0	40270	11999	Q27	483224143	120618

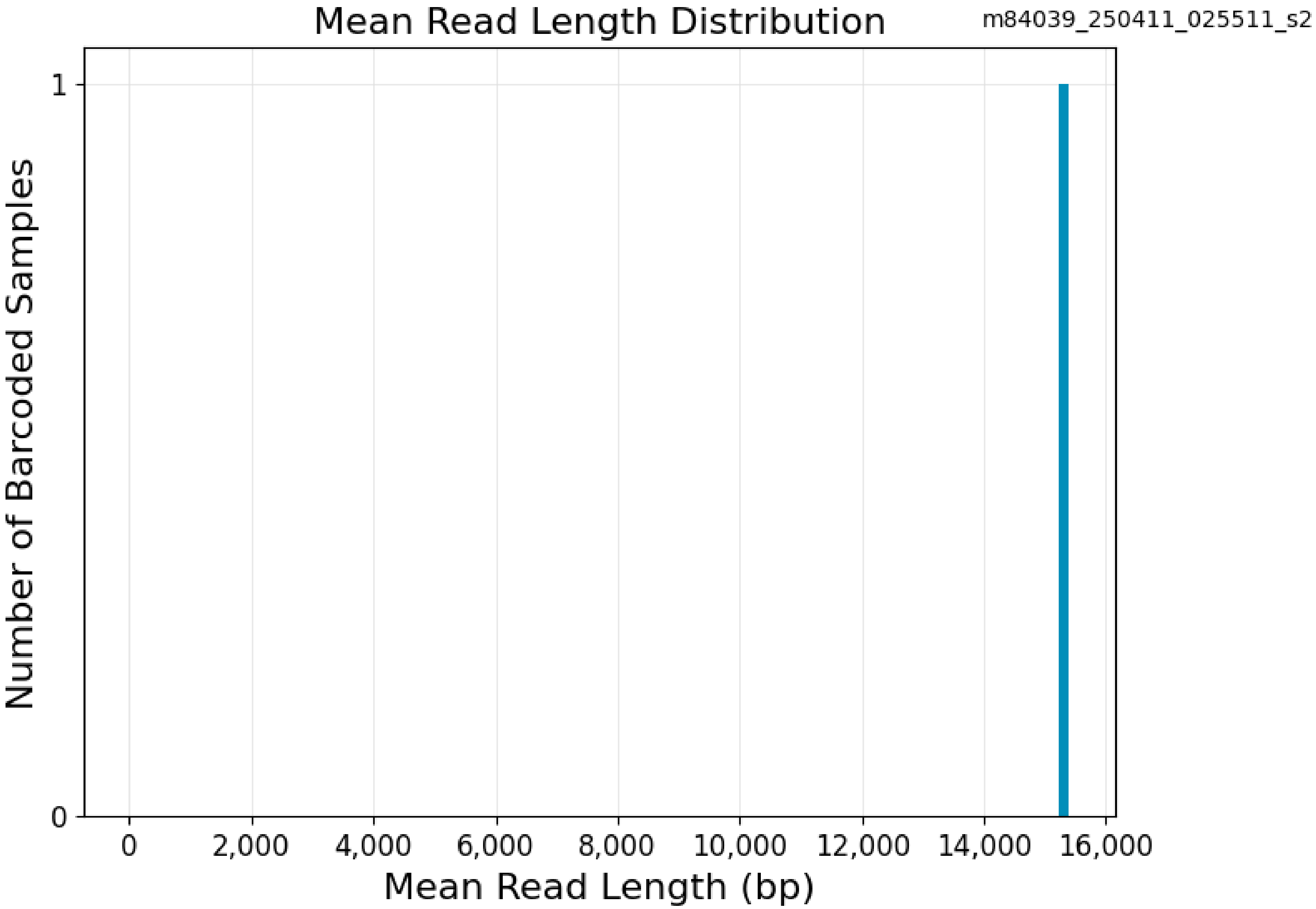
Number Of Reads Per Barcode



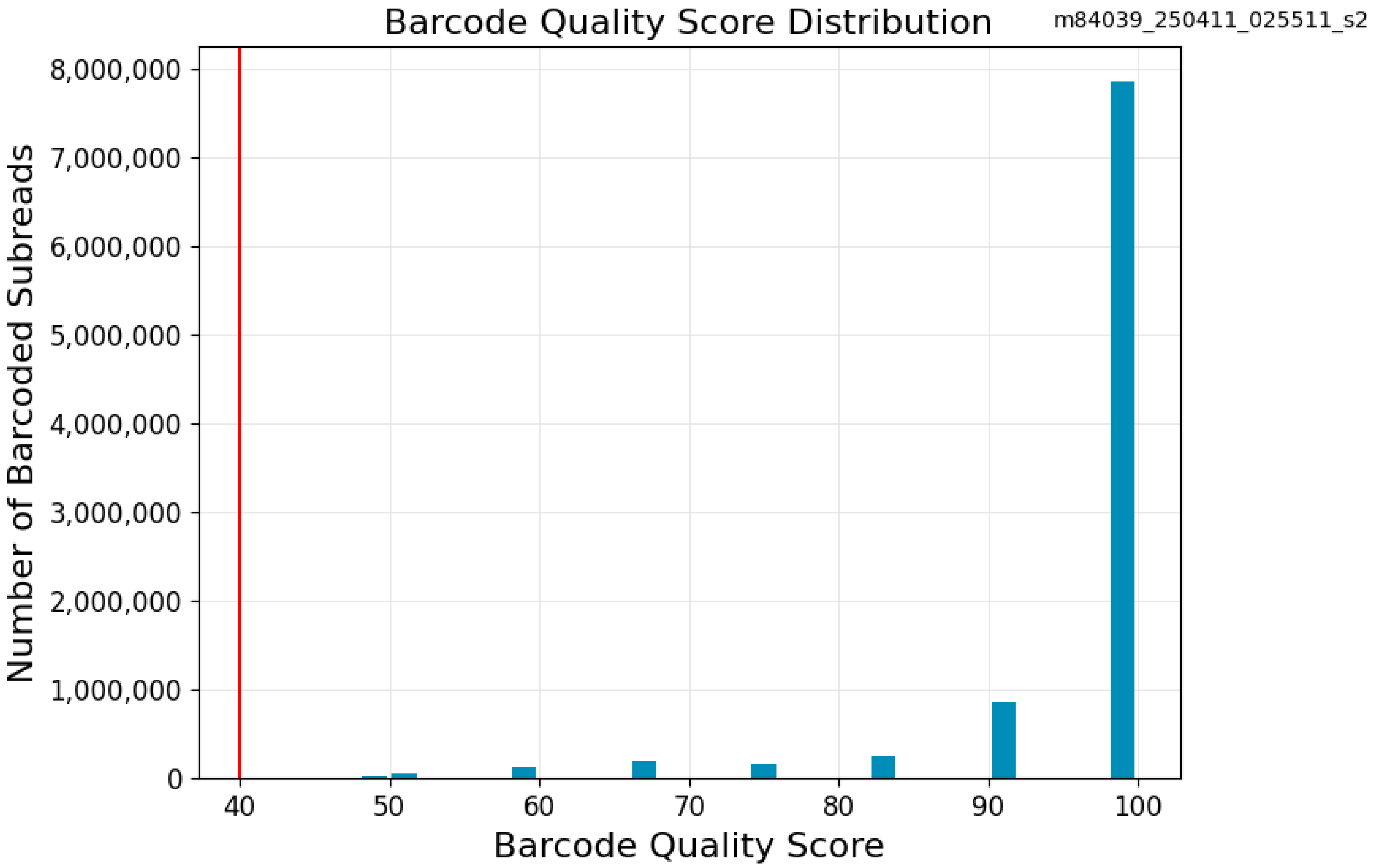
Barcode Frequency Distribution



Mean Read Length Distribution



Barcode Quality Score Distribution

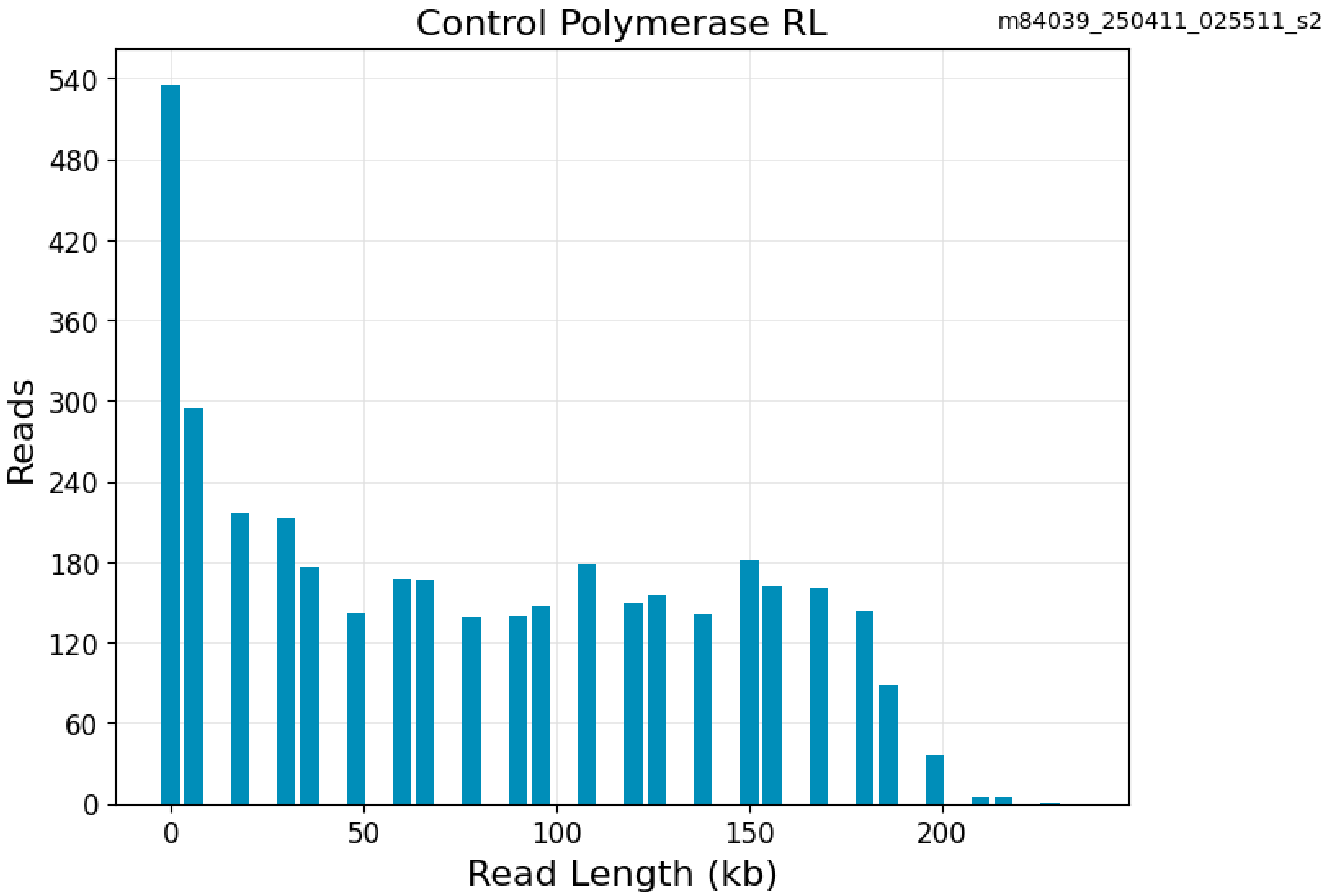


Control Report

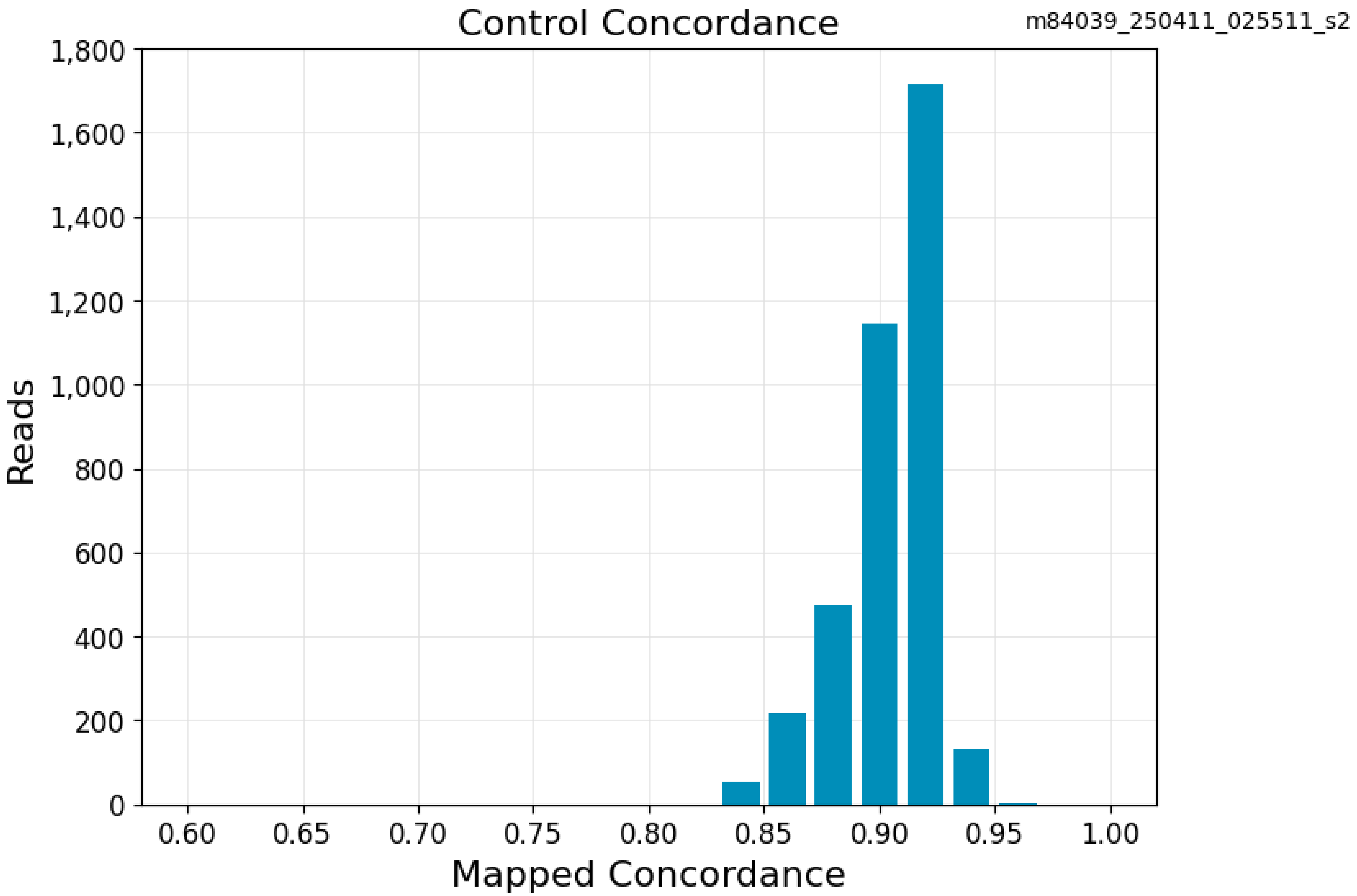
Summary

Number of Control Reads	3,745
Control Read Length Mean	79,974
Control Read Concordance Mean	0.91
Control Read Concordance Mode	0.93

Control Polymerase RL



Control Concordance

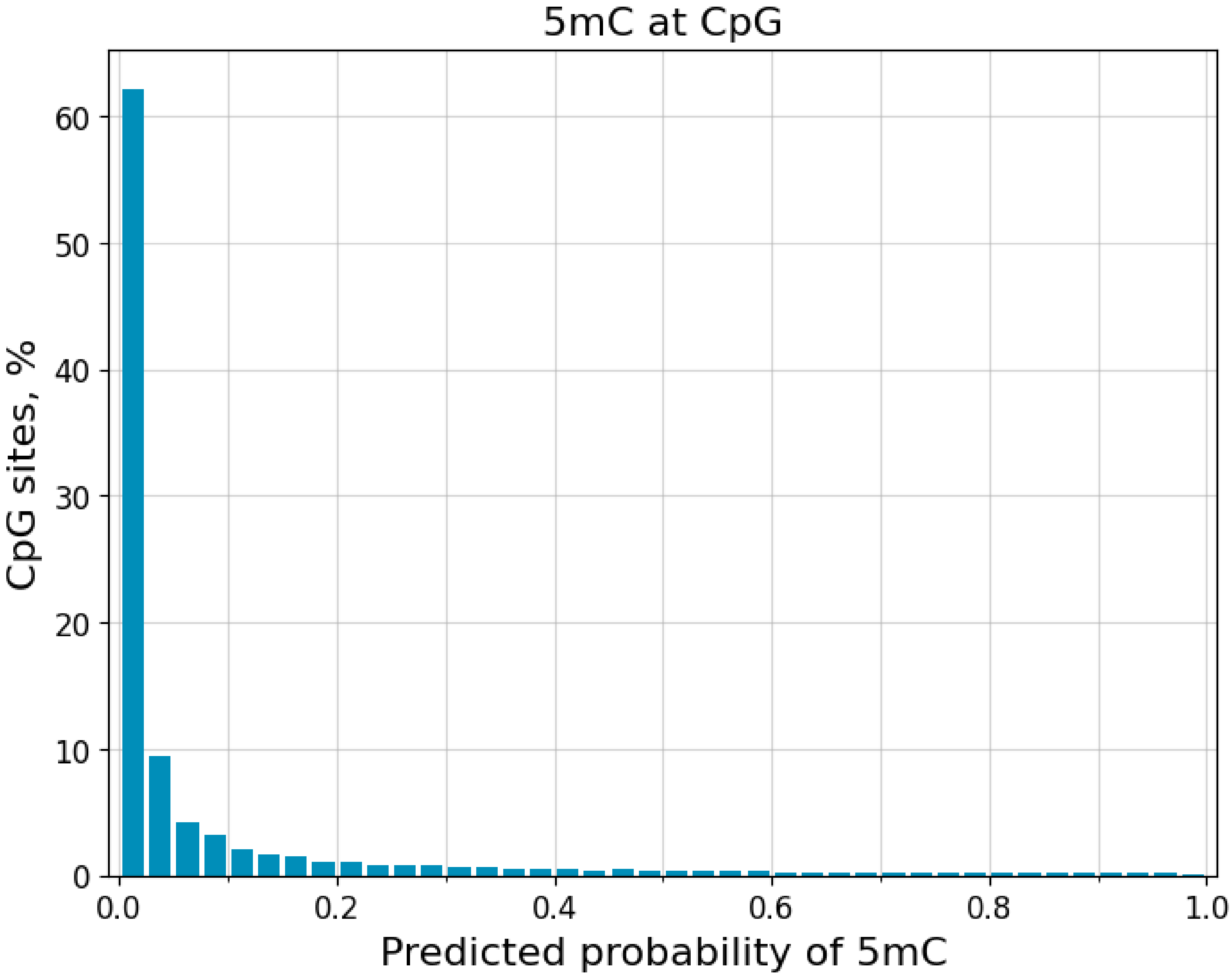


Methylation

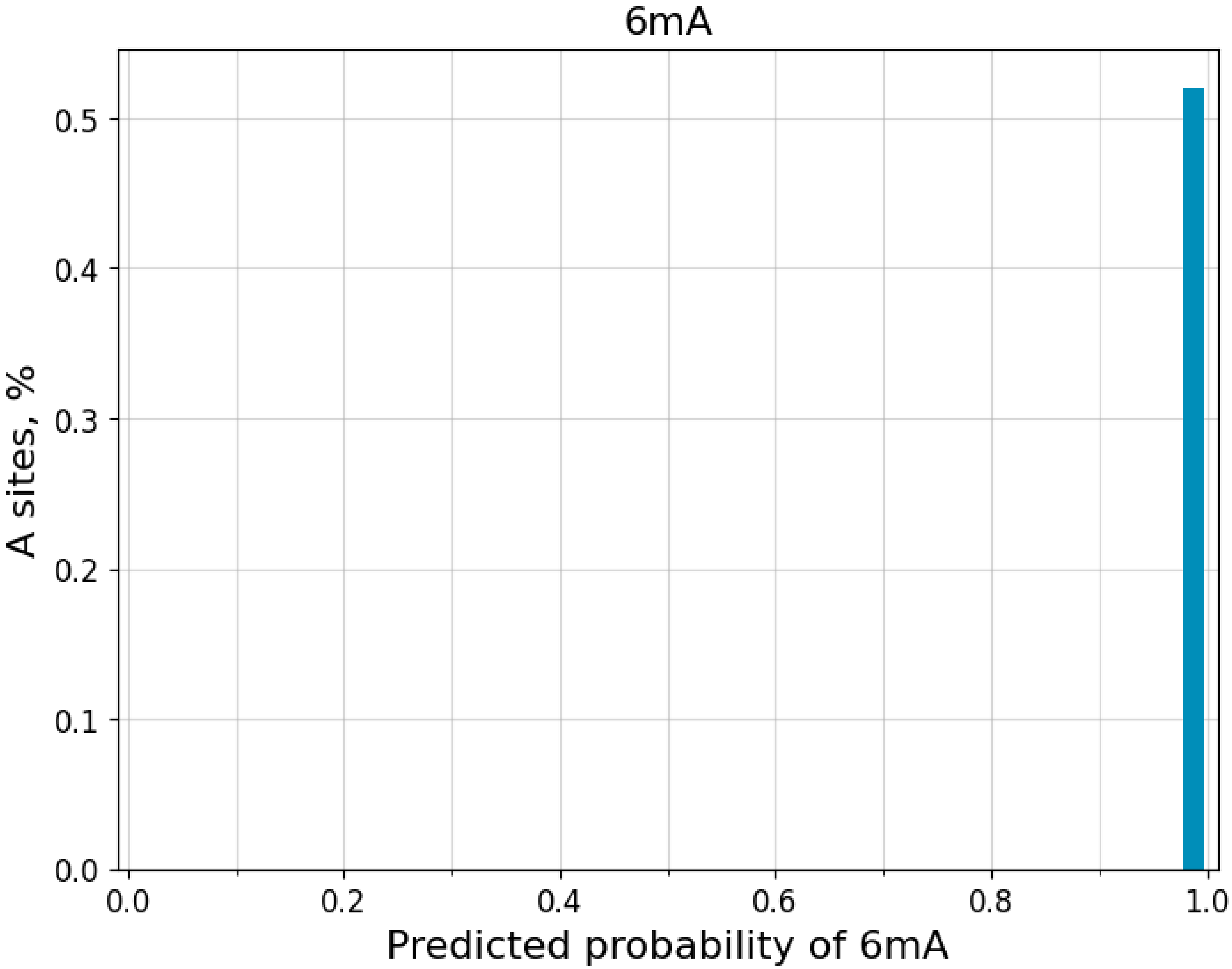
Summary

Modification	Motif	Scored sites	Modified sites (Pr > 0.5)
5mC	CpG	98.3%	5.2%
6mA	A	0.5%	0.5%

5mC methylation at CpG in reads


5mC at CpG
CpG sites, %
60
50
40
30
20
10
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability of 5mC

6mA methylation in reads


6mA
A sites, %
0.0
0.1
0.2
0.3
0.4
0.5
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability of 6mA

No Sample Setup found

No instrument run found

Parent jobs (1)

Job 124482

Summary

Job Type	import-dataset
Name	import-dataset
Comments	Description for job Import PacBio DataSet
Created At	2025-08-01 13:55:11.074
SMRT Link Version	25.3.0.273777

Child jobs (1)

Job 124625

Summary

Job Type	analysis
Pipeline	cromwell.workflows.pb_segment_reads_and_sc_isoseq
Name	KinnexRelease-10x-3pGEX-PBMC-20k-rep1-2RevioCell-largeMem
Comments	Description for job Run Analysis Application
Created At	2025-08-02 13:23:49.637
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_Rep1-Cell3 (all samples)	85868c24-0602-4d91-8160-6a2ba131c641	9,605,565

Output files

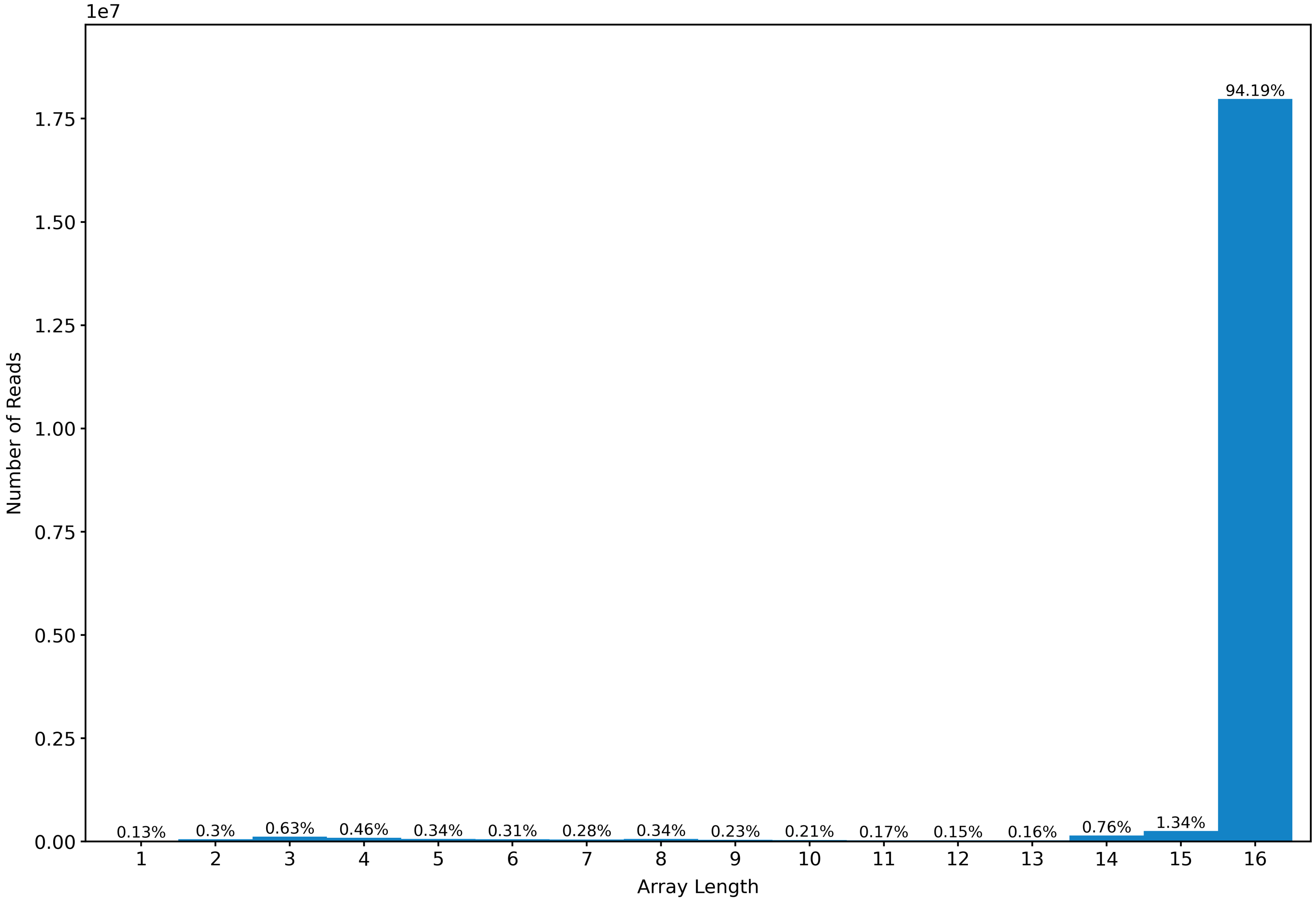
File name	Size	Type
KinnexRelease-10x-3pGEX-PBMC-20k-rep1-2RevioCell-largeMem Segmented Reads	47,593	ConsensusReadSet
Deduplicated reads after cell barcode correction, unmapped, BAM	65,514,118,740	bam
Single-cell isoform and gene matrix, tar-gzipped	6,068,666,618	tgz
Segmented Reads, passing, unaligned	95,625,358,339	bam
Segmented Reads, passing, unaligned	96,879,374,686	bam
Non-passing reads, unaligned	2,946,550,795	bam
Non-passing reads, unaligned	2,928,151,183	bam
Unique mapped transcripts, GFF	3,750,131,236	gff
Unique mapped transcripts, filtered, GFF	1,514,726,065	gff
Deduplicated reads after cell barcode correction, mapped, BAM	36,293,406,032	bam
Deduplicated reads after cell barcode correction, mapped, BAM index	46,833,832	bam_bai
Unique mapped transcripts, junctions TXT	2,687,522,465	txt
Unique mapped transcripts, filtered, junctions TXT	1,381,362,938	txt
Deduplicated transcripts	48,668,057,448	Fasta
Collapsed transcript groups	3,768,489,358	txt
Unique mapped transcripts, classification TXT	2,803,368,347	txt
Unique mapped transcripts, filtered, classification TXT	921,652,951	txt
bcstats_report.tsv.gz	17,291,534	tsv
Analysis Log	98,371	log
737K_august_2016.txt.gz	2,238,609	txt
3M_3pgex_may_2023.REVCOMP.txt.gz	19,830,613	txt
Analysis Log	4,092	log
SMRT Link Log	20,111	log

Read Segmentation

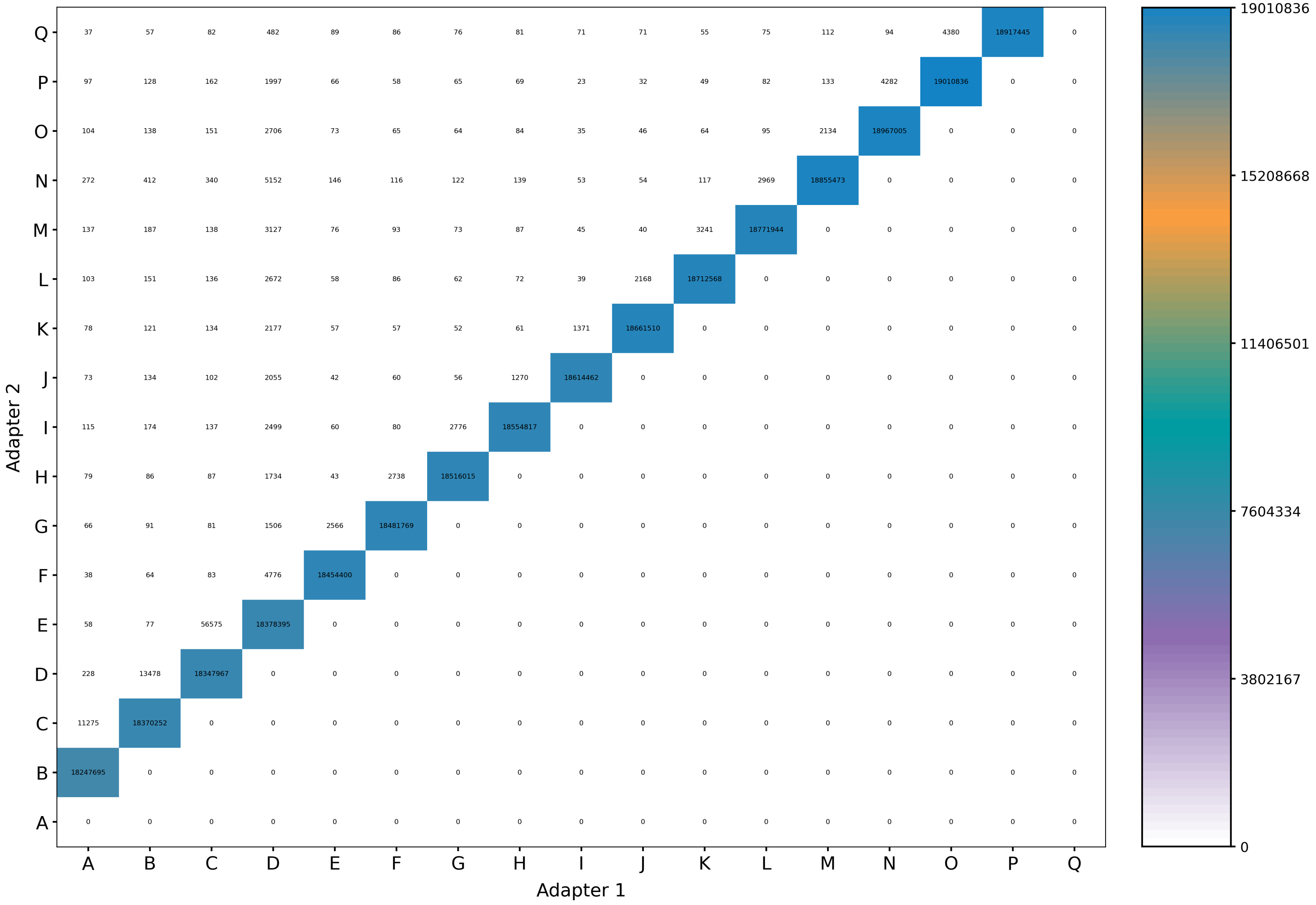
Summary

Reads	19,094,849
Segmented reads (S-reads)	297,862,553
Mean length of S-reads	963
Percent of reads with full arrays	94.17 %
Mean array size (concatenation factor)	15.60

Histogram of concatemers

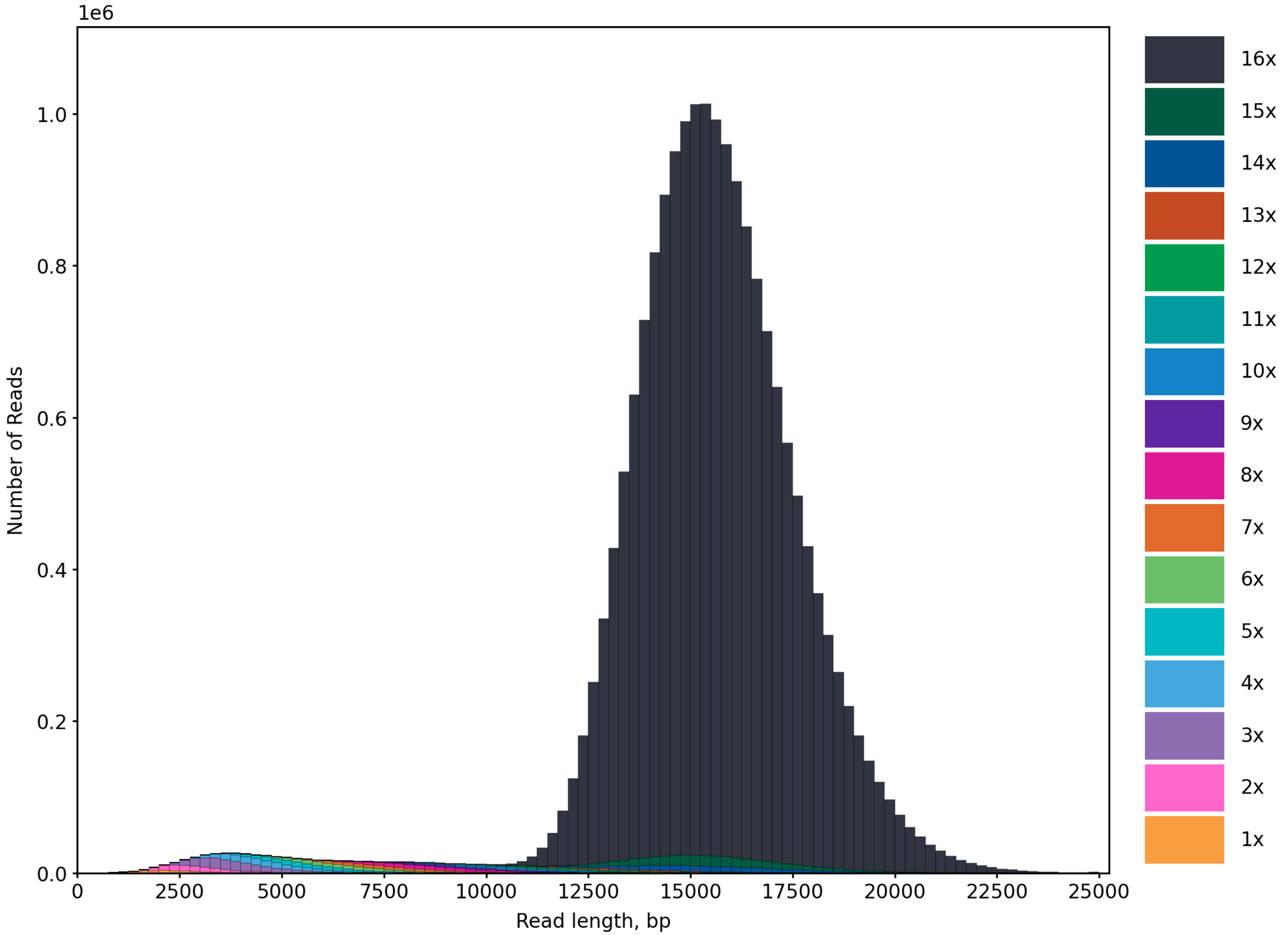


Heatmap of concatemer ligations

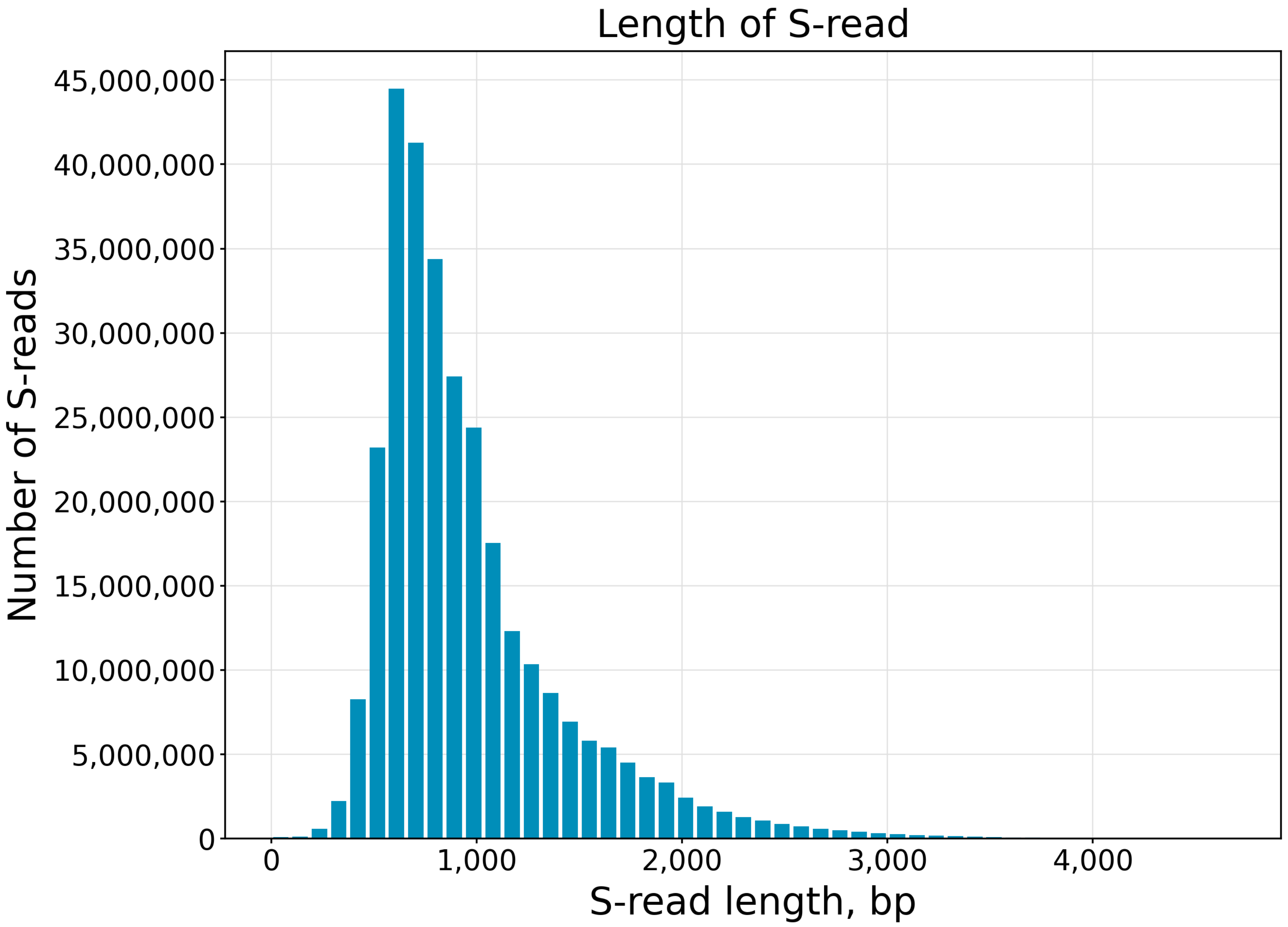


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Q	37	57	82	482	89	86	76	81	71	71	55	75	112	94	4380	18917445	0
P	97	128	162	1997	66	58	65	69	23	32	49	82	133	4282	19010836	0	0
O	104	138	151	2706	73	65	64	84	35	46	64	95	2134	18967005	0	0	0
N	272	412	340	5152	146	116	122	139	53	54	117	2969	18855473	0	0	0	0
M	137	187	138	3127	76	93	73	87	45	40	3241	18771944	0	0	0	0	0
L	103	151	136	2672	58	86	62	72	39	2168	18712568	0	0	0	0	0	0
K	78	121	134	2177	57	57	52	61	1371	18661510	0	0	0	0	0	0	0
J	73	134	102	2055	42	60	56	1270	18614462	0	0	0	0	0	0	0	0
I	115	174	137	2499	60	80	2776	18554817	0	0	0	0	0	0	0	0	0
H	79	86	87	1734	43	2738	18516015	0	0	0	0	0	0	0	0	0	0
G	66	91	81	1506	2566	18481769	0	0	0	0	0	0	0	0	0	0	0
F	38	64	83	4776	18454400	0	0	0	0	0	0	0	0	0	0	0	0
E	58	77	56575	18378395	0	0	0	0	0	0	0	0	0	0	0	0	0
D	228	13478	18347967	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	11275	18370252	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	18247695	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Length of Reads



Length of S-reads

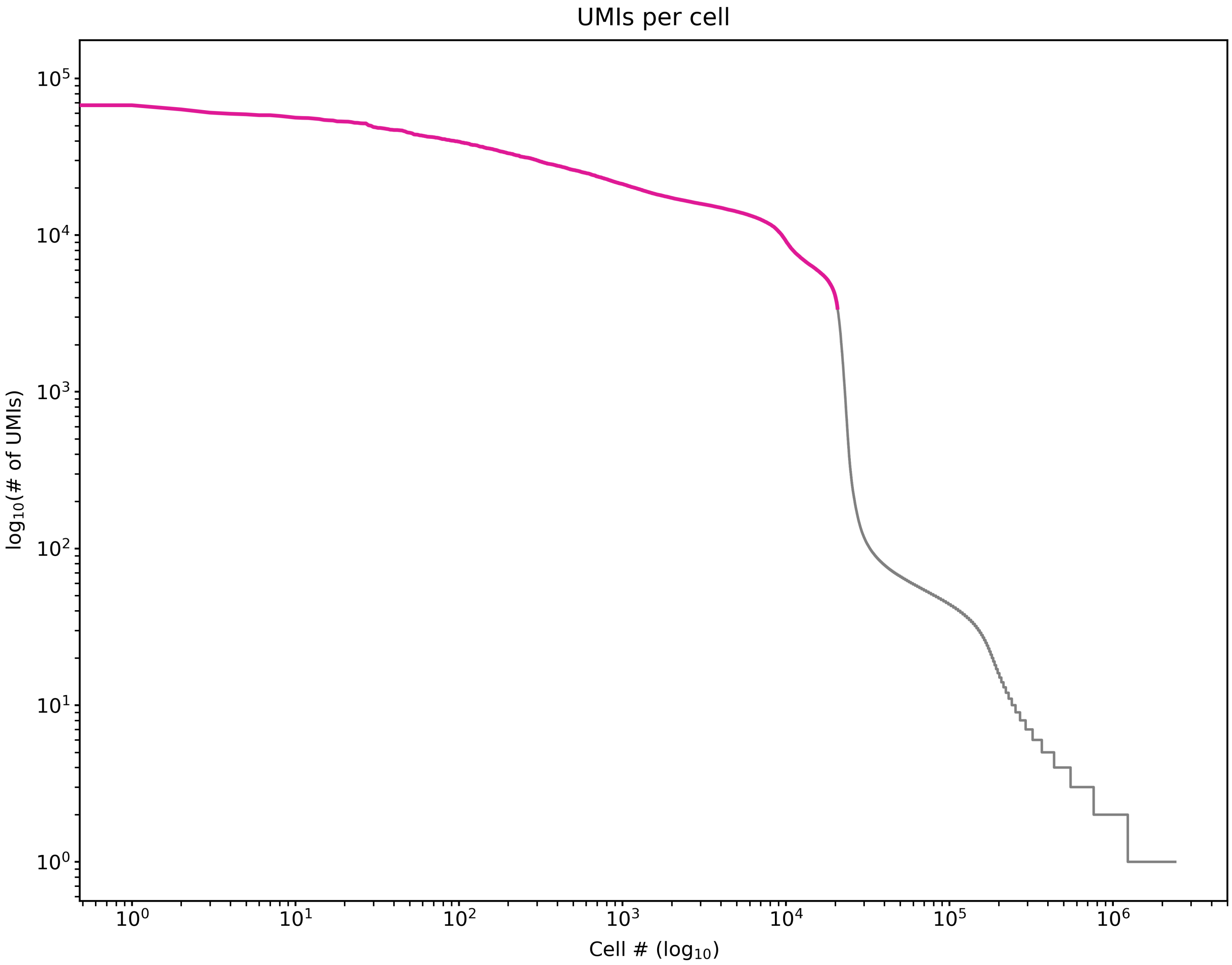


Cell Statistics

Summary

Estimated Number of Cells	20,426
Reads in Cells	91.77%
Mean Reads per Cell	12,964
Median UMIs per Cell	8,882

Barcode Rank Plot



Read Statistics

Summary

Reads	297,862,553
Read Type	SEGMENT
Reads with 5' and 3' Primers with extracted UMIs and Barcodes	295,436,944
Non-Concatamer Reads with 5' and 3' Primers and Poly-A Tail (FLNC reads)	288,569,391
FLNC Reads with Valid Barcodes	282,849,411
FLNC Reads with Valid Barcodes, corrected	287,921,294
Reads after Barcode Correction and UMI Deduplication	218,678,519

Transcript Statistics

Summary

FLNC Reads Mapped Confidently to Genome	233,818,089
FLNC Reads Mapped Confidently to Transcriptome	141,208,760
Total Unique Genes	3,404,460
Total Unique Genes, filtered	226,088
Total Unique Genes, known genes only	43,407
Total Unique Genes, filtered, known genes only	33,141
Total Unique Transcripts	8,203,385
Total Unique Transcripts, filtered	1,841,952
Total Unique Transcripts, known transcripts only	117,310
Total Unique Transcripts, filtered, known transcripts only	107,850

Transcript Summary

Value	Analysis Metric
3,557	Median Genes per Cell
2,504	Median Genes per Cell, known genes only
4,725	Median Transcripts per Cell
2,024	Median Transcripts per Cell, known transcripts only
3,404,460	Total Unique Genes
43,407	Total Unique Genes, known genes only
8,203,385	Total Unique Transcripts
117,310	Total Unique Transcripts, known transcripts only

Transcript Summary, filtered

Value	Analysis Metric
1,806	Median Genes per Cell
1,764	Median Genes per Cell, known genes only
2,351	Median Transcripts per Cell
1,742	Median Transcripts per Cell, known transcripts only
226,088	Total Unique Genes
33,141	Total Unique Genes, known genes only
1,841,952	Total Unique Transcripts
107,850	Total Unique Transcripts, known transcripts only

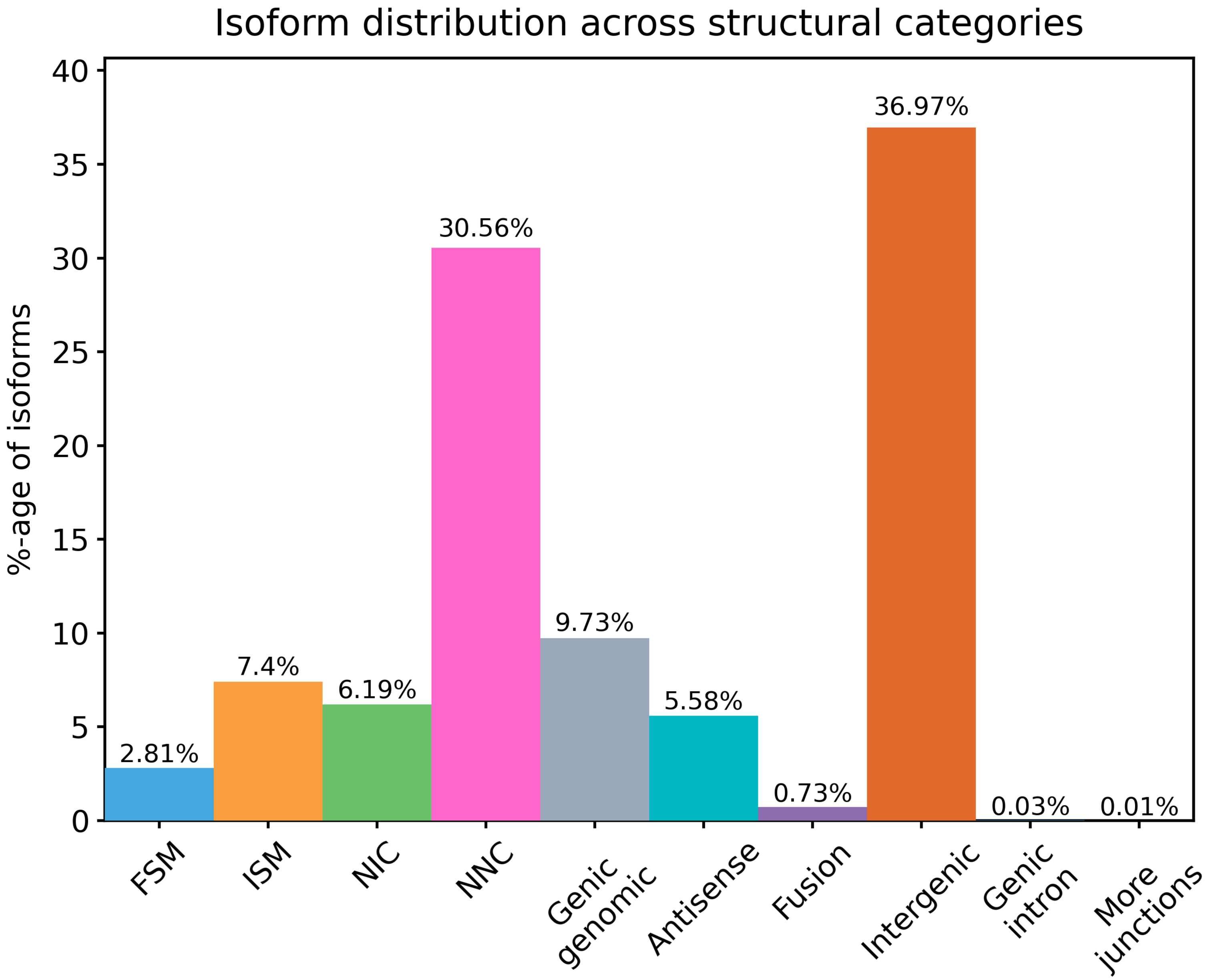
Transcript Classification

Category	Count	CAGE Detected	CAGE Detected (%)	polyA Motif Detected	polyA Motif Detected (%)
FSM	252662	142005	56.20%	72161	28.56%
ISM	666164	267482	40.15%	191129	28.69%
NIC	557760	357644	64.12%	173342	31.08%
NNC	2751523	1694810	61.60%	979309	35.59%
Antisense	502622	17401	3.46%	130438	25.95%
Fusion	66138	37016	55.97%	24539	37.10%
More junctions	930	423	45.48%	269	28.92%
Genic intron	2437	8	0.33%	684	28.07%
Genic genomic	875737	213170	24.34%	254462	29.06%
Intergenic	3328928	31057	0.93%	1087968	32.68%

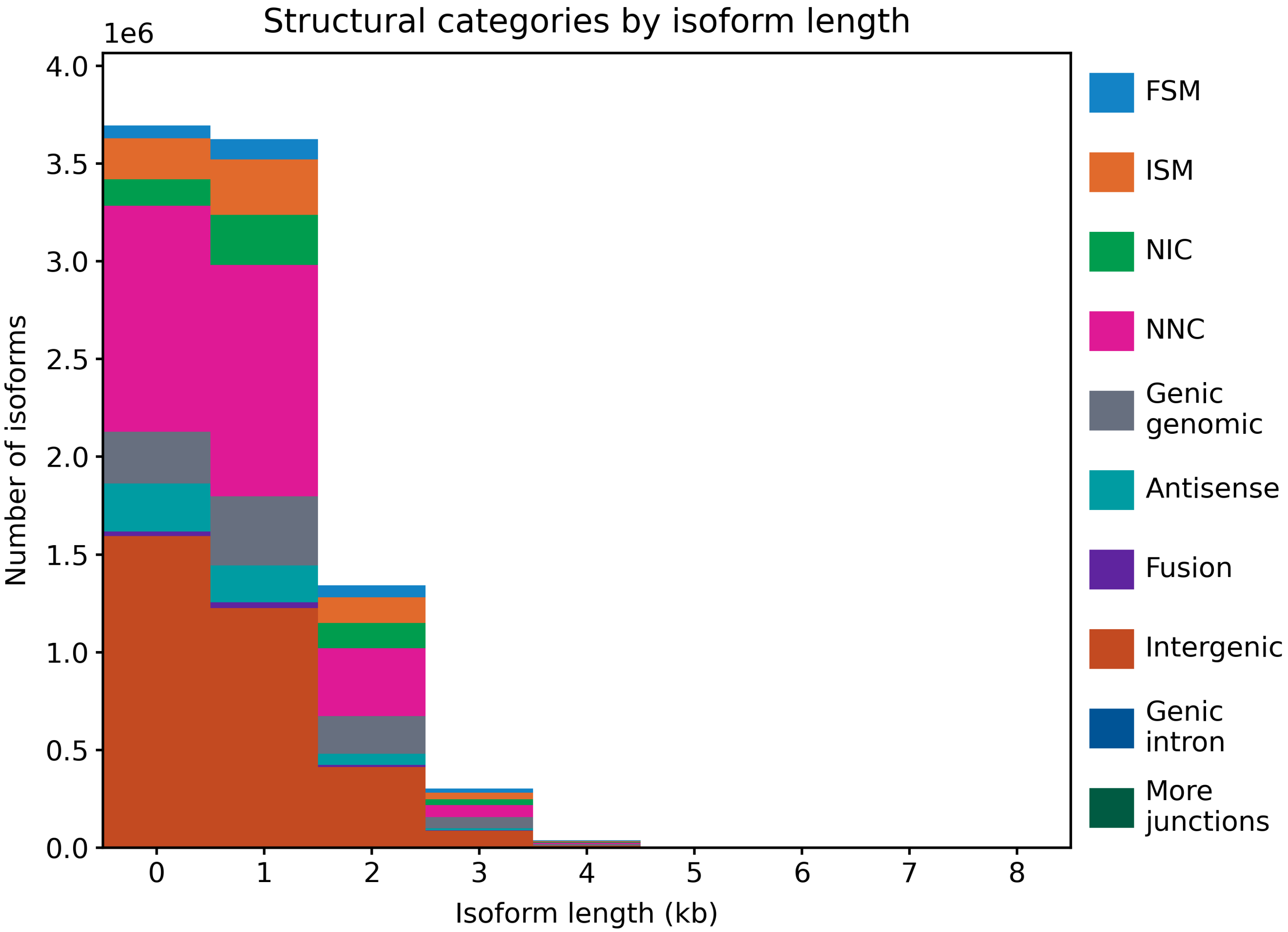
Transcript Classification, filtered

Category	Count	CAGE Detected	CAGE Detected, (%)	polyA Detected	polyA Detected, (%)
FSM	215414	124039	57.58%	72161	33.50%
ISM	479344	211772	44.18%	169828	35.43%
NIC	449764	290143	64.51%	172018	38.25%
NNC	984611	622549	63.23%	435717	44.25%
Antisense	41574	3747	9.01%	23960	57.63%
Fusion	31312	18086	57.76%	14264	45.55%
More junctions	566	261	46.11%	207	36.57%
Genic intron	2	0	0.00%	1	50.00%
Genic genomic	42998	19202	44.66%	19239	44.74%
Intergenic	183275	4338	2.37%	122457	66.82%

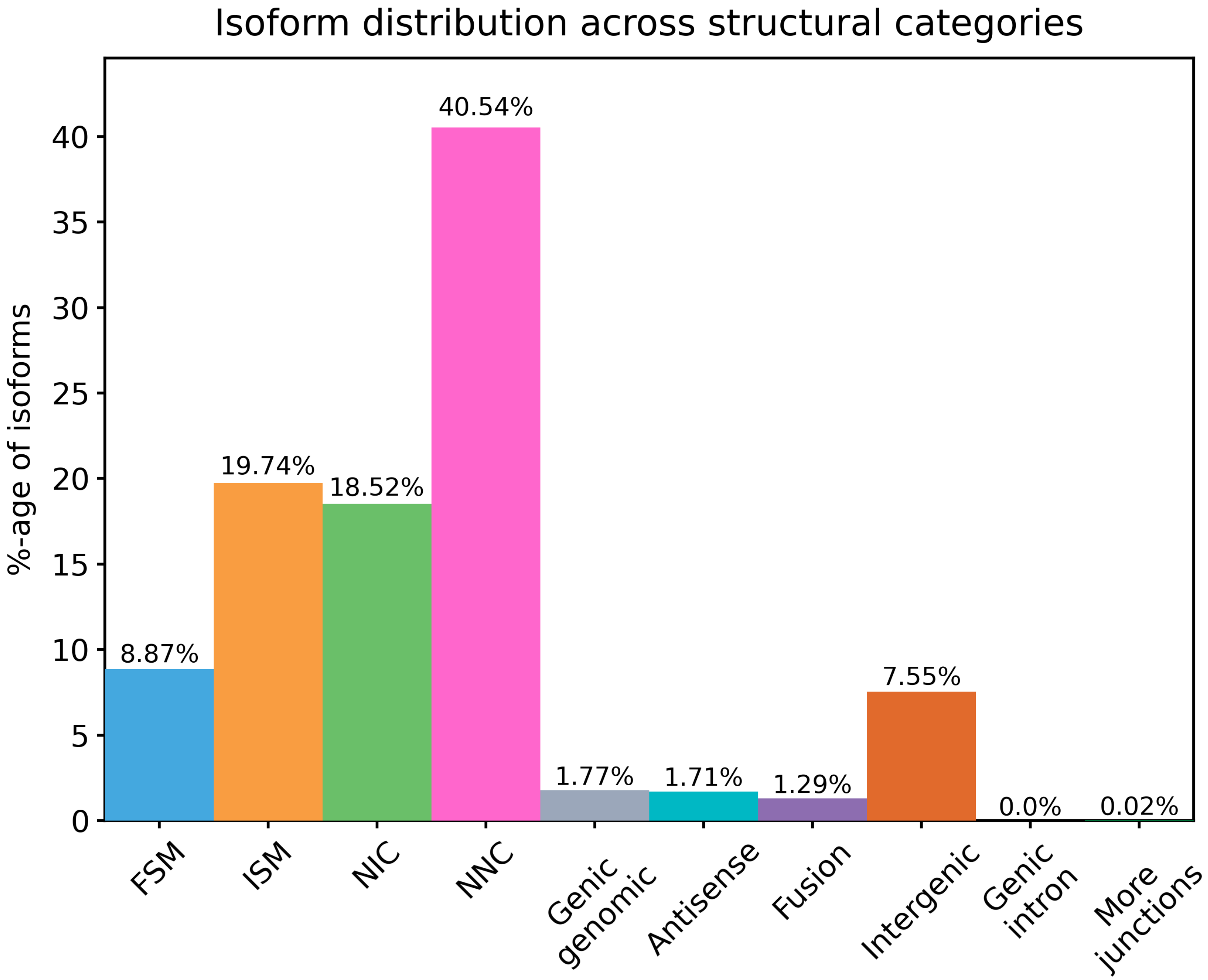
Isoform Distribution Across Categories


Isoform distribution across structural categories
%-age of isoforms
0
5
10
15
20
25
30
35
40
2.81%
7.4%
6.19%
30.56%
9.73%
5.58%
0.73%
36.97%
0.03%
0.01%
FSM
ISM
NIC
NNC
Genic genomic
Antisense
Fusion
Intergenic
Genic intron
More junctions

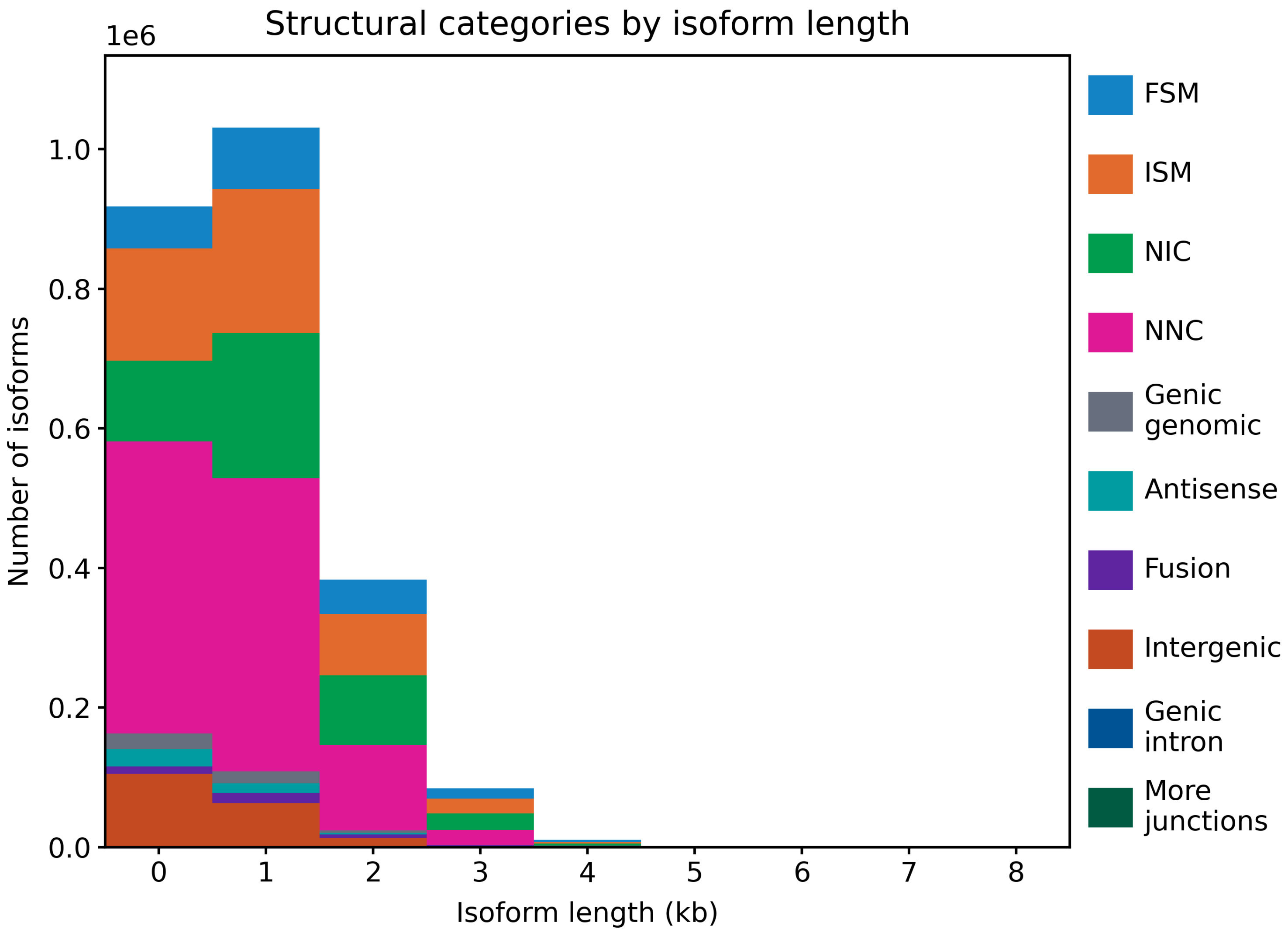
Structural Categories by Transcript Length



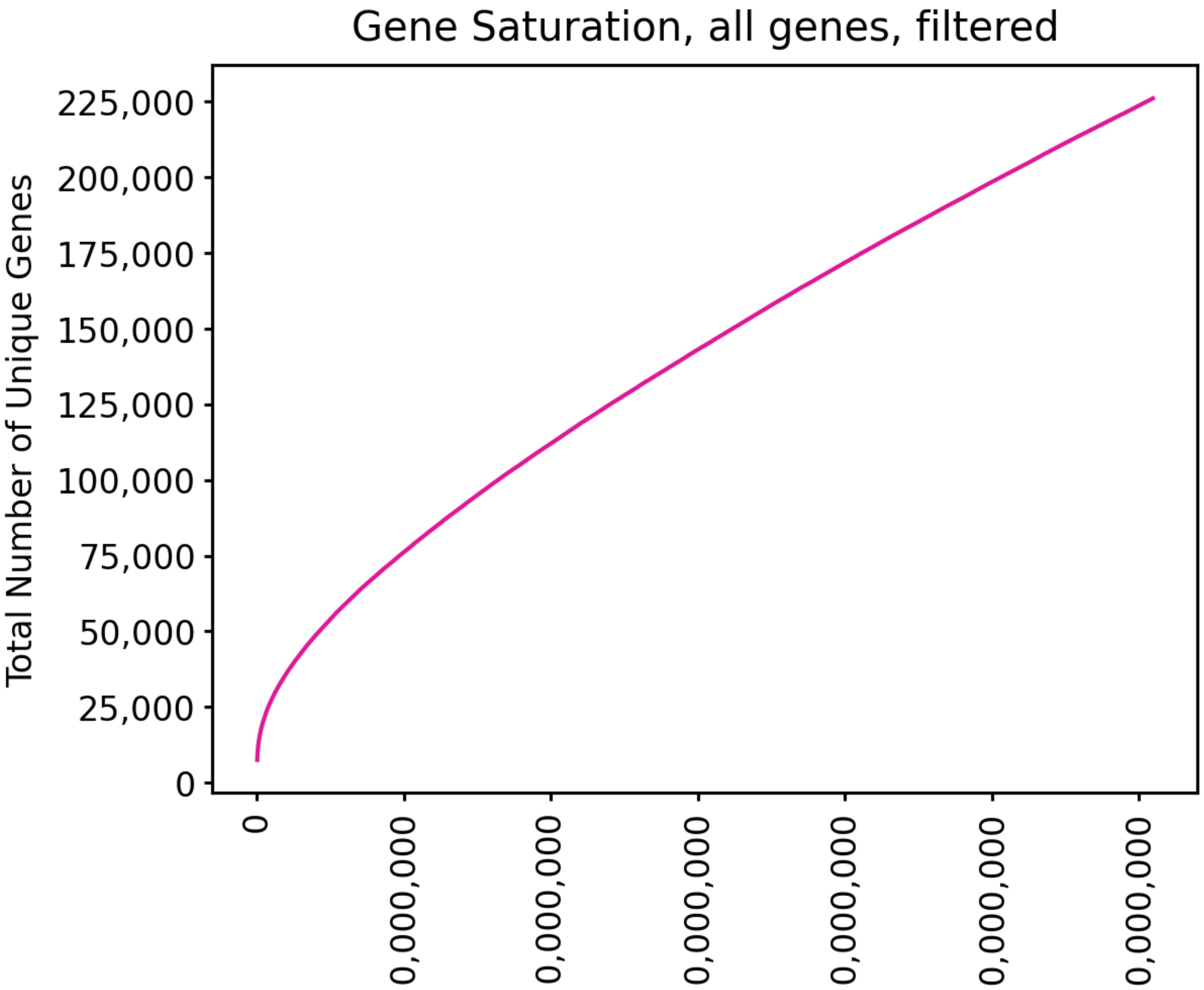
Isoform Distribution Across Categories

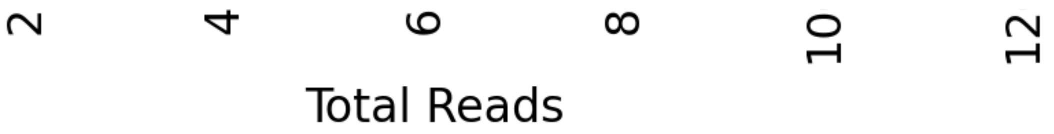

Isoform distribution across structural categories
%-age of isoforms
0
5
10
15
20
25
30
35
40
8.87%
19.74%
18.52%
40.54%
1.77%
1.71%
1.29%
7.55%
0.0%
0.02%
FSM
ISM
NIC
NNC
Genic genomic
Antisense
Fusion
Intergenic
Genic intron
More junctions

Structural Categories by Transcript Length



sc_isoseq_transcript_statistics.gene_saturation_plotgroup.gene_saturation_plot_0




2
4
6
8
10
12
Total Reads

sc_isoseq_transcript_statistics.gene_saturation_plotgroup.gene_saturation_plot_1


20
40
60
80
100
120
Total Reads