Report for dataset 3prime_v4_GEX_hPBMCs_10k_Rep1-Cell1 (all samples)

Dataset 2841a41a-4df9-48c5-9ab4-878ca3a92ccb

Summary

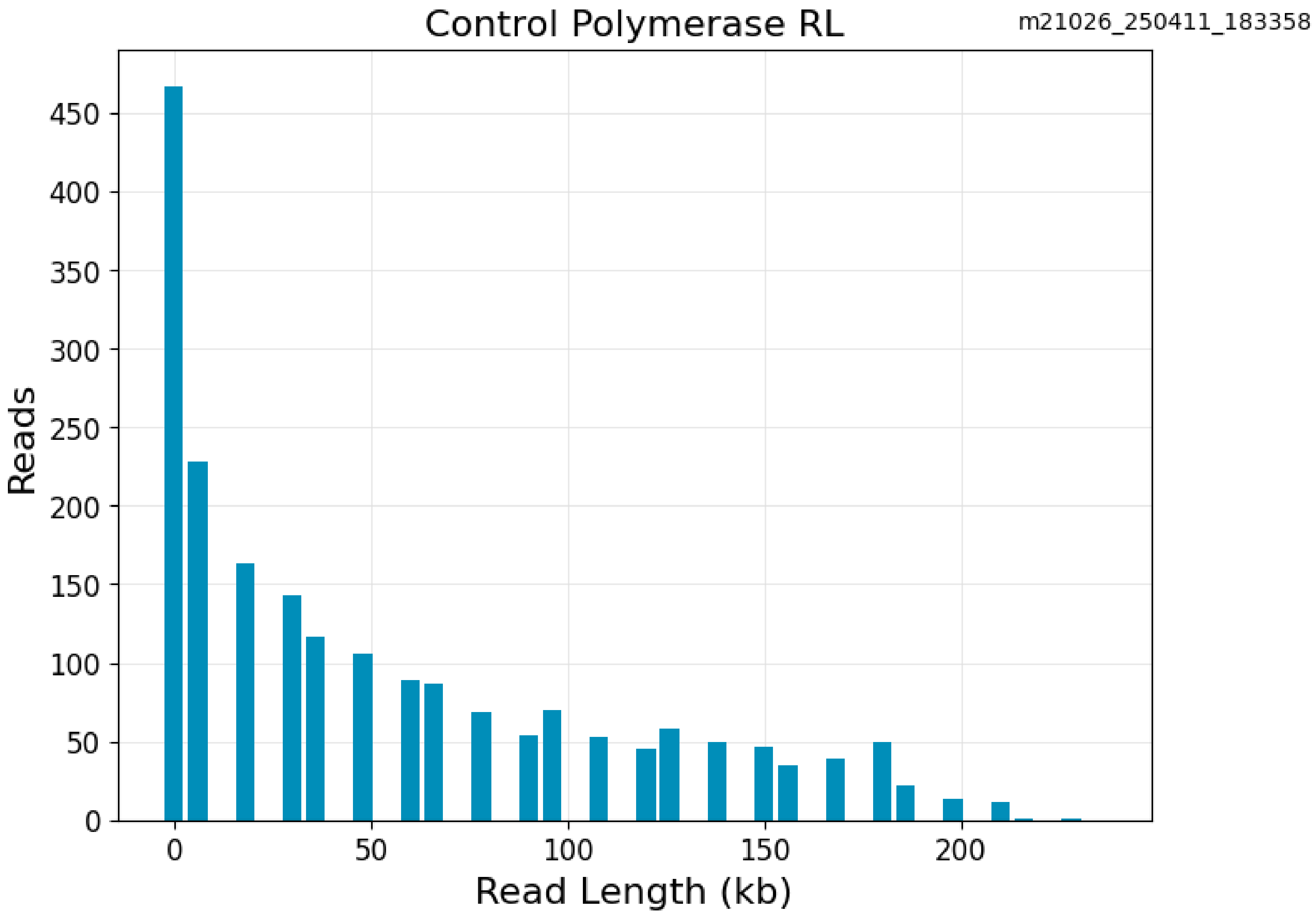
Name	3prime_v4_GEX_hPBMCs_10k_Rep1-Cell1 (all samples)
Created At	2025-04-13 08:49:49.545
Number of Records	4,372,029
Total Length	64,857,376,652
Movie Name	m21026_250411_183358
ICS Version	1.1.0.46.43
Well Sample	3prime_v4_GEX_hPBMCs_10k_Rep1
Biological Sample	3prime_v4_GEX_hPBMCs_10k_Rep1
Barcode Name	bcM0001--bcM0001

Control Report

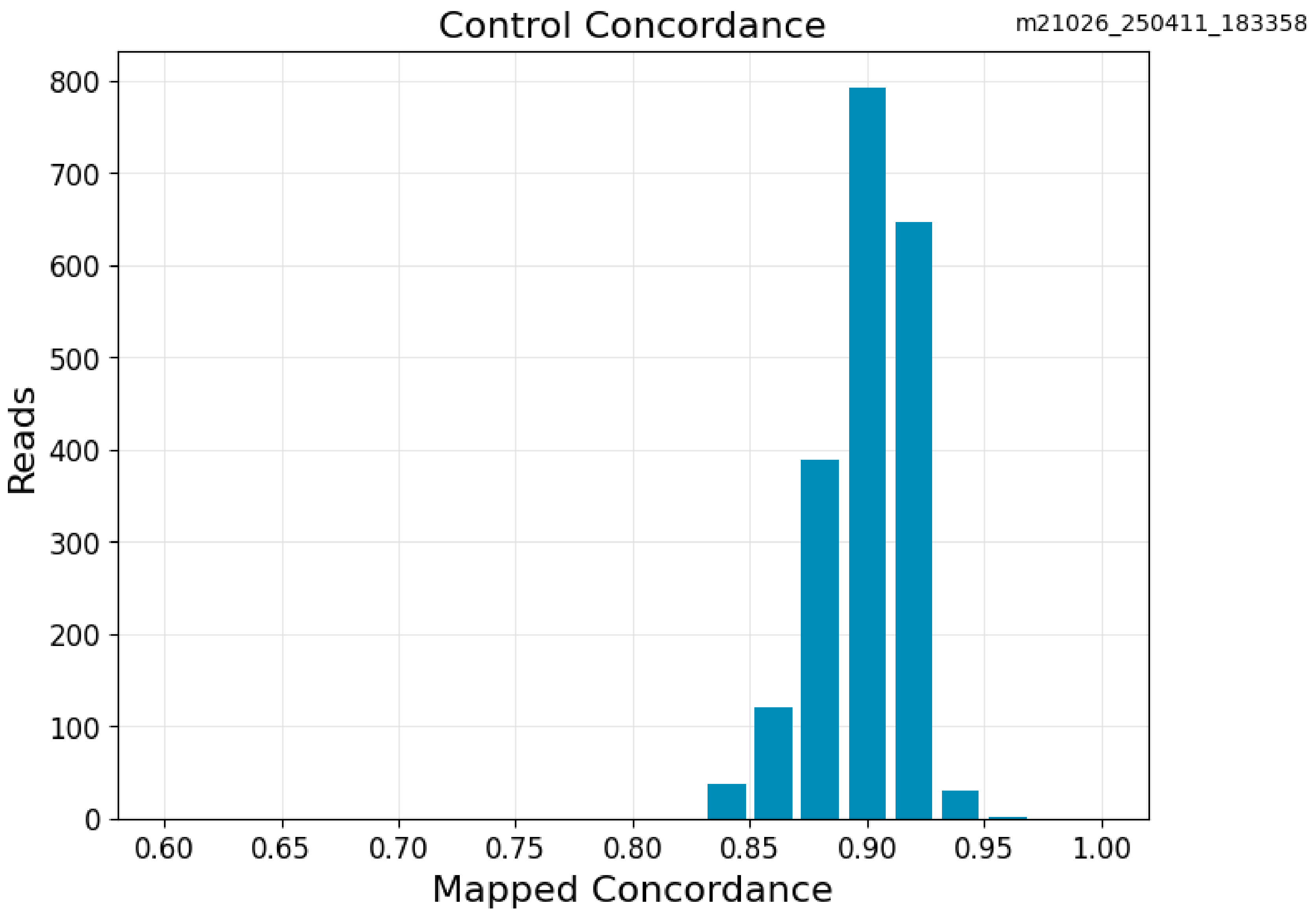
Summary

Number of Control Reads	2,016
Control Read Length Mean	55,958
Control Read Concordance Mean	0.91
Control Read Concordance Mode	0.91

Control Polymerase RL



Control Concordance



Barcodes

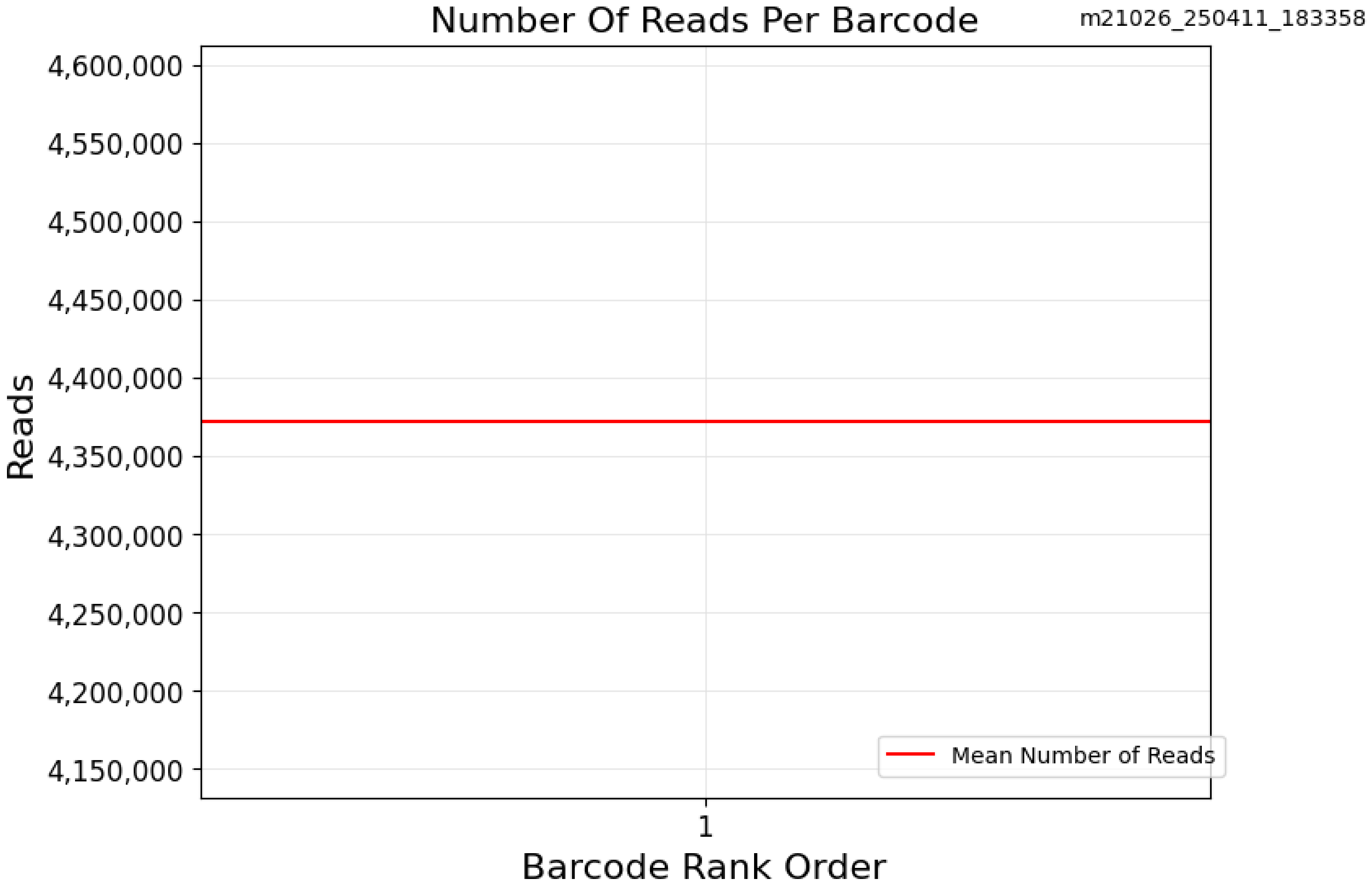
Summary

Unique Barcodes	1
Barcoded HiFi Reads	4,372,029
Unbarcoded HiFi Reads	8,327
Barcoded HiFi Reads (%)	99.81 %
Barcoded HiFi yield (Gb)	64.86 Gb
Unbarcoded HiFi yield (Gb)	0.09 Gb
Barcoded HiFi Yield (%)	99.86 %
Mean HiFi Reads per Barcode	4,372,029
Max. HiFi Reads per Barcode	4,372,029
Min. HiFi Reads per Barcode	4,372,029
Barcoded HiFi read length (mean, kb)	14.83 kb
Unbarcoded HiFi read length (mean, kb)	11.07 kb

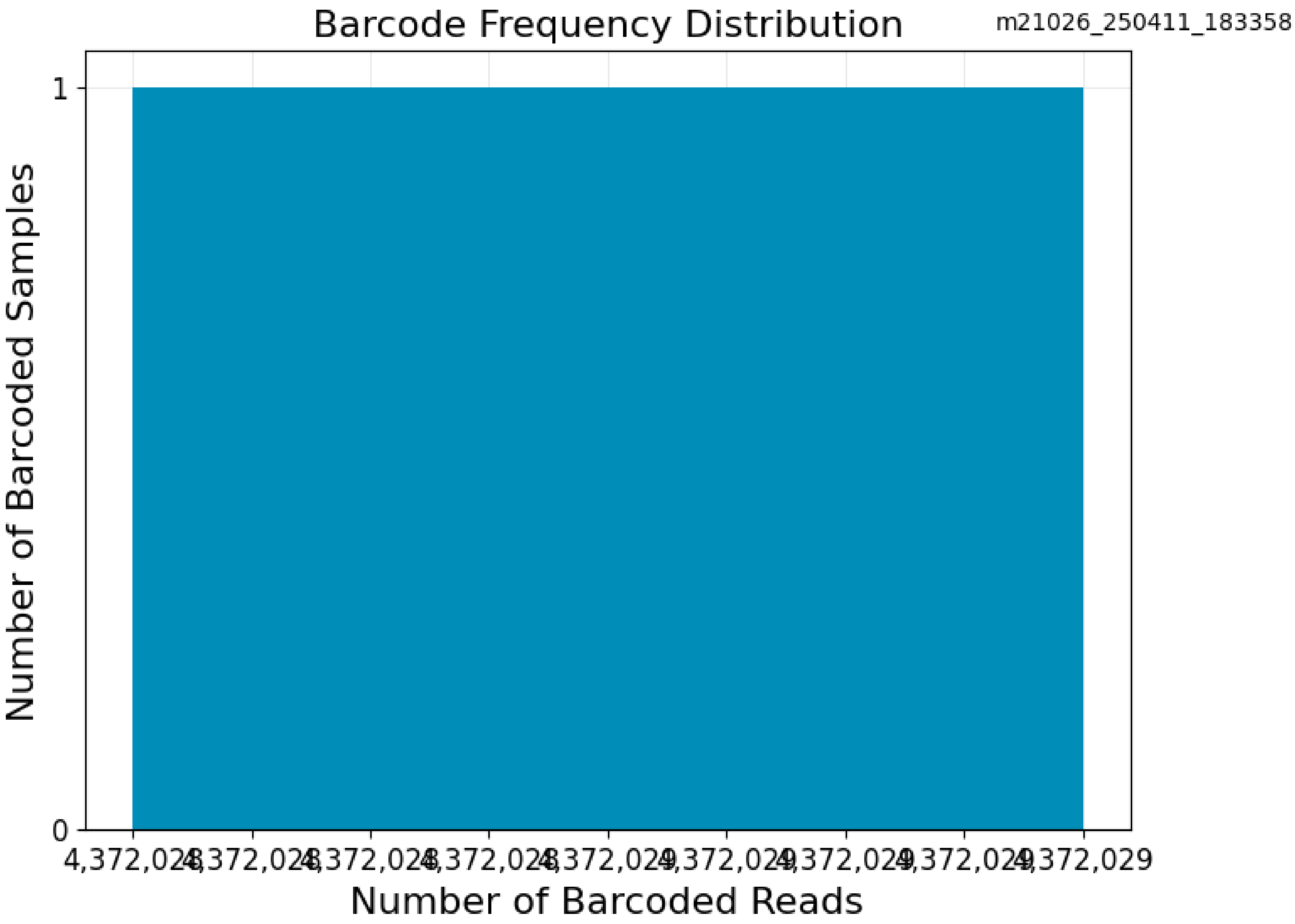
Barcode Data

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	97.4	4372029	14834	Q29	64857376652	146709
No Name	Not Barcoded	0.0	8327	11074	Q28	92214793	128422

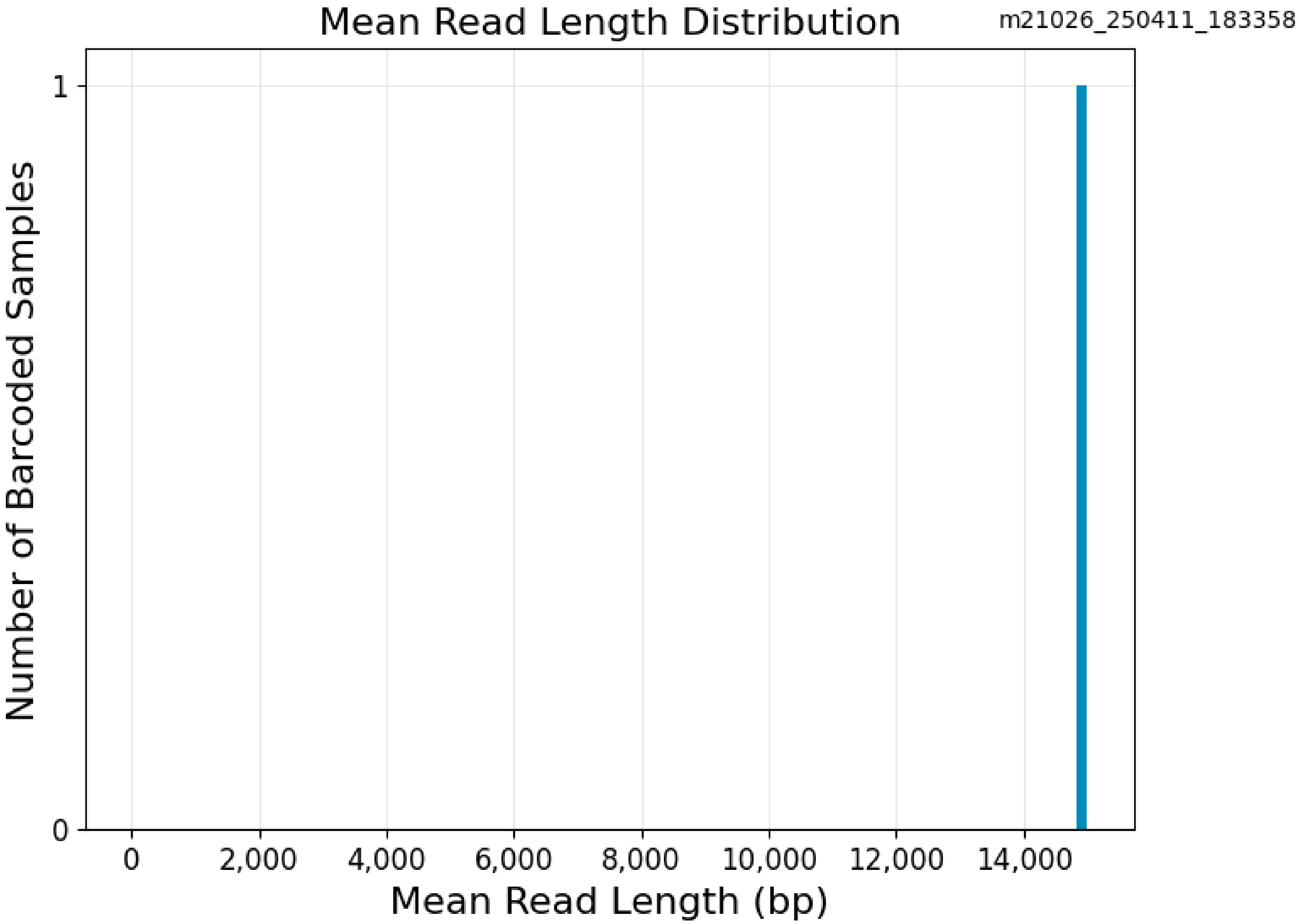
Number Of Reads Per Barcode



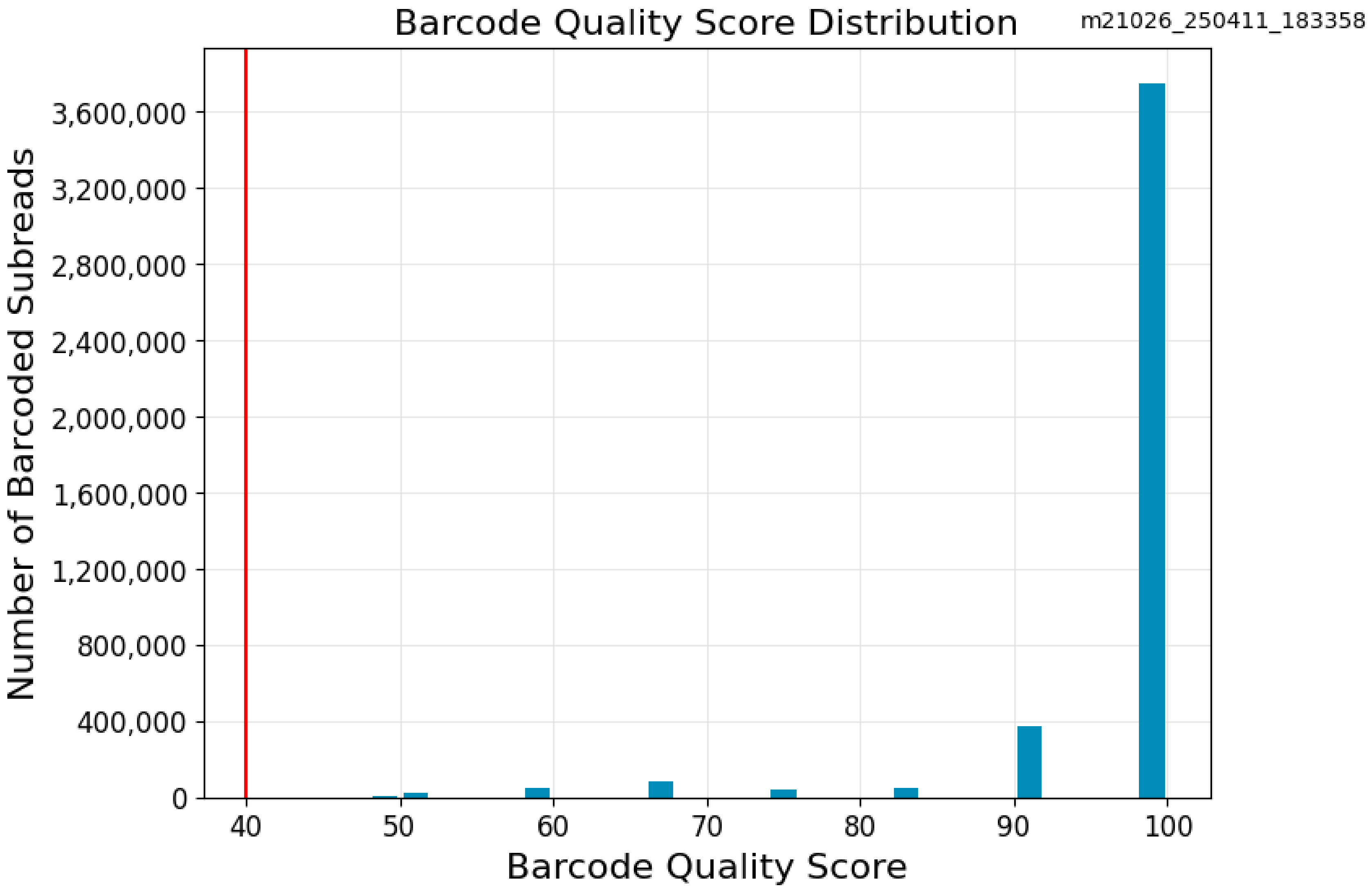
Barcode Frequency Distribution



Mean Read Length Distribution



Barcode Quality Score Distribution

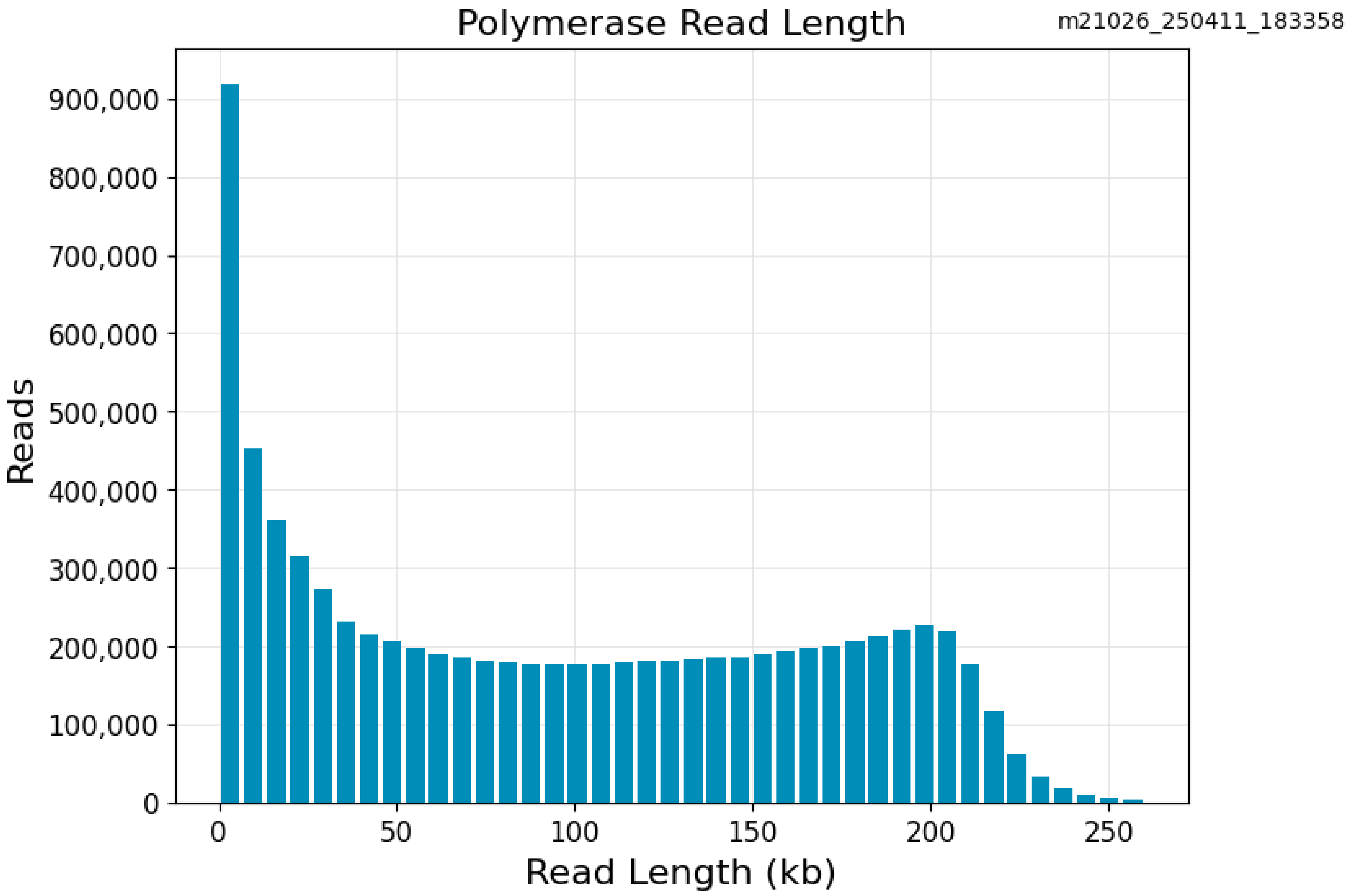


Raw Data Report

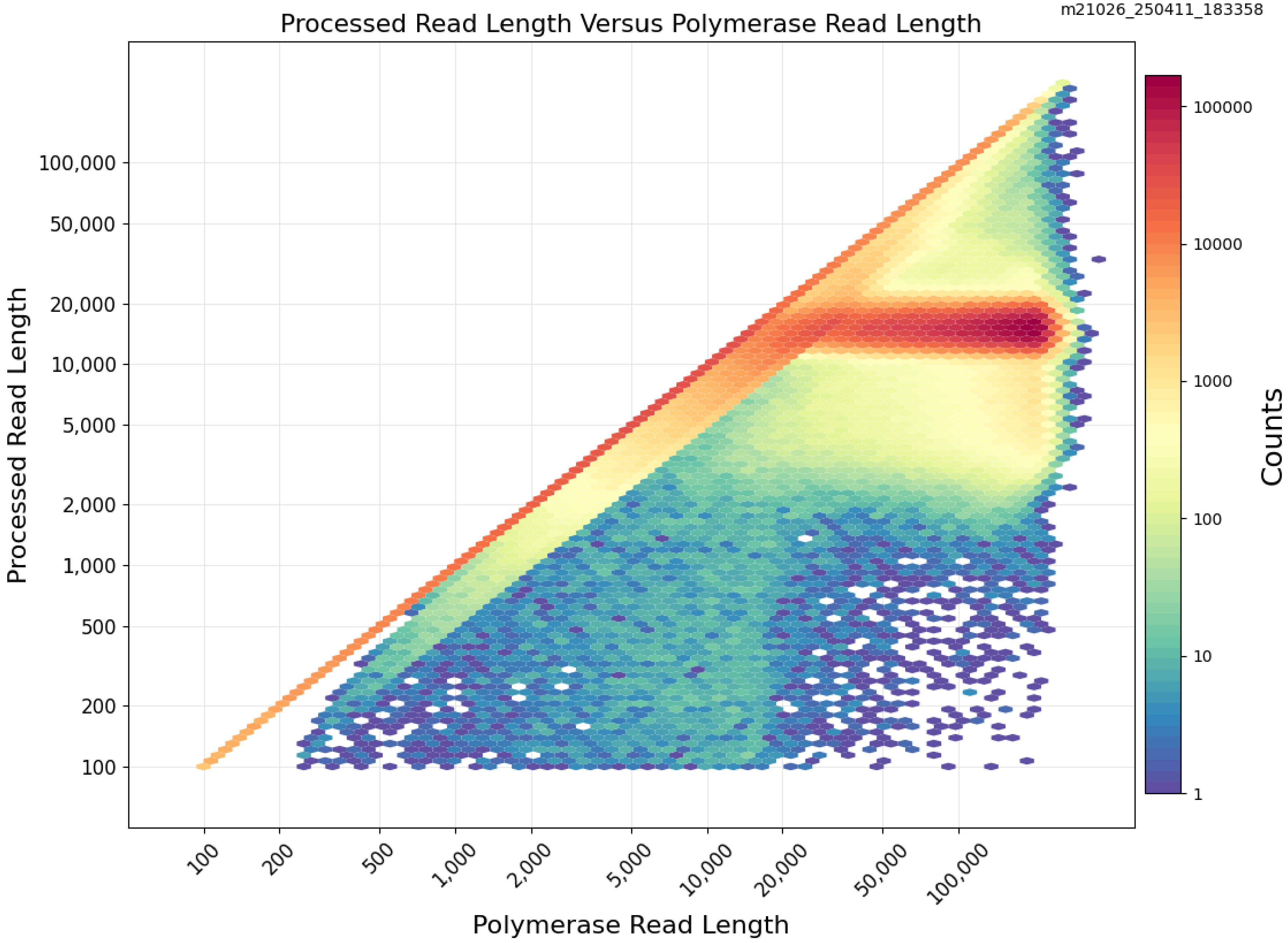
Summary

Polymerase Read Bases	753,056,768,242
Polymerase Reads	8,004,498
Polymerase Read Length (mean)	94.08 kb
Polymerase read length (N50)	160.25 kb
Polymerase read length longest subread length (mean)	16.69 kb
Polymerase read length longest subread length (N50)	16.75 kb
Unique Molecular Yield	124,712,714,240
Local Base Rate	2.67

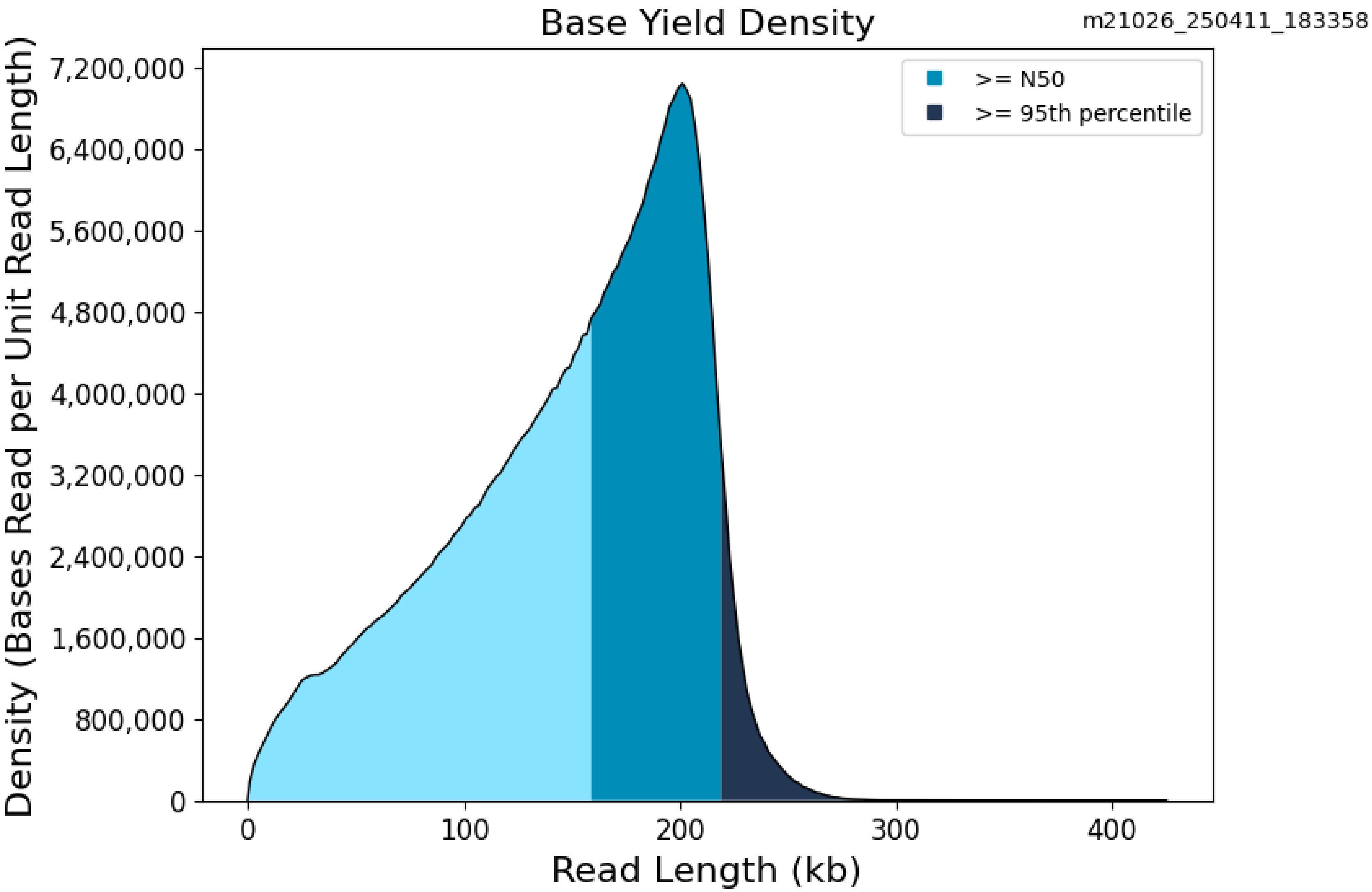
Polymerase Read Length



Longest Subread Length Versus Polymerase Read Length



Base Yield Density



CCS Analysis Report

Summary

HiFi reads	4.4 M
HiFi reads yield	64.95 Gb
HiFi reads length (mean)	14.83 kb
HiFi reads length (median, bp)	14,858
HiFi Read Length N50 (bp)	15,115
HiFi Read Quality (median)	Q29
HiFi Read Quality (median)	29
Base Quality Q30 (%)	95.51%
HiFi Number of Passes (mean)	9
Missing adapters (%)	3.09%

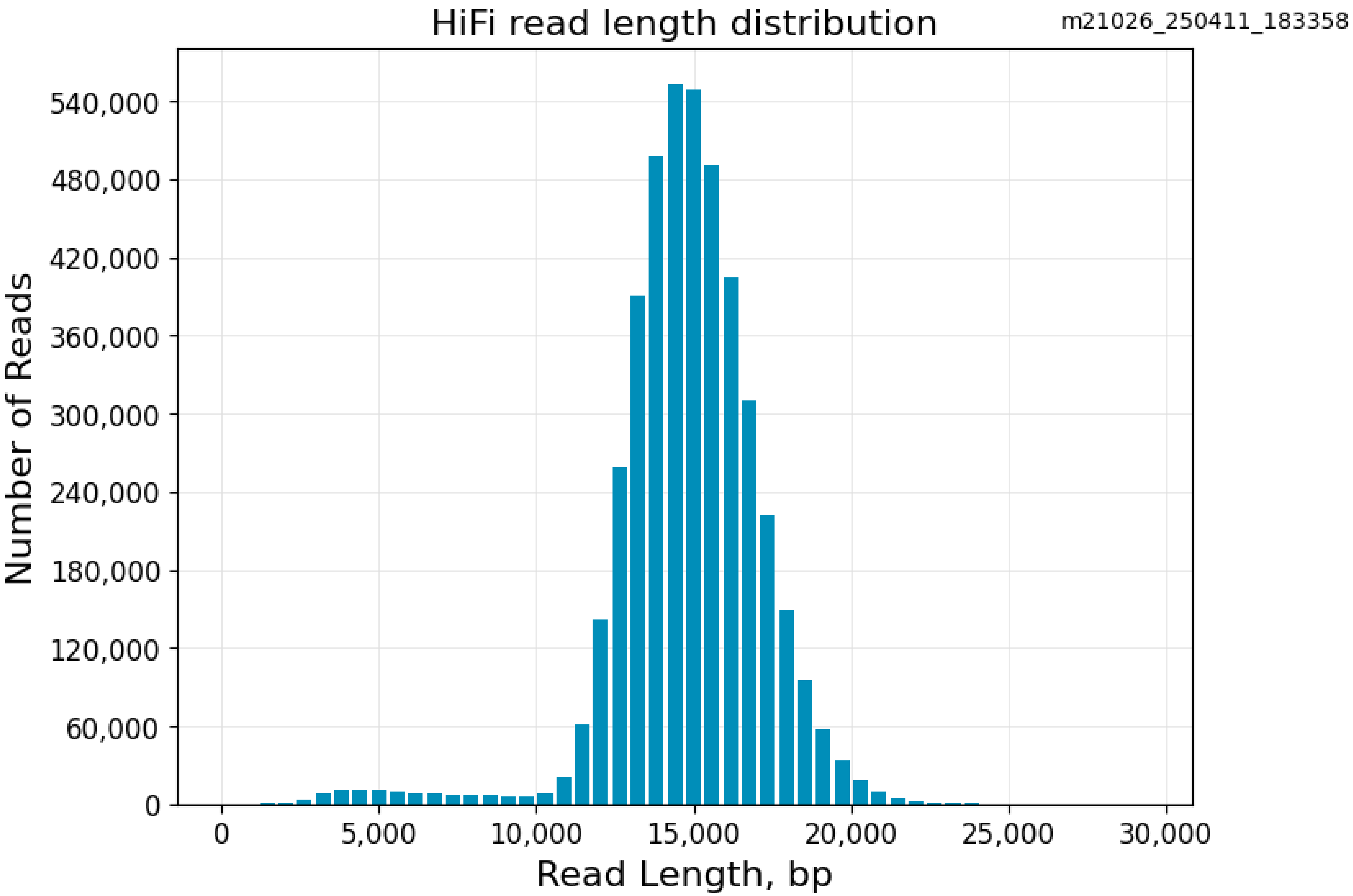
HiFi Read Length Summary

Read Length (kb)	Reads	Reads (%)	Yield (Gb)	Yield (%)
0	4380356	100	64.95 Gb	100
5,000	4339030	99	64.79 Gb	100
10,000	4273785	98	64.32 Gb	99
15,000	2057674	47	34.08 Gb	52
20,000	37970	1	0.79 Gb	1
25,000	76	0	0.00 Gb	0
30,000	6	0	0.00 Gb	0
35,000	0	0	0.00 Gb	0
40,000	0	0	0.00 Gb	0

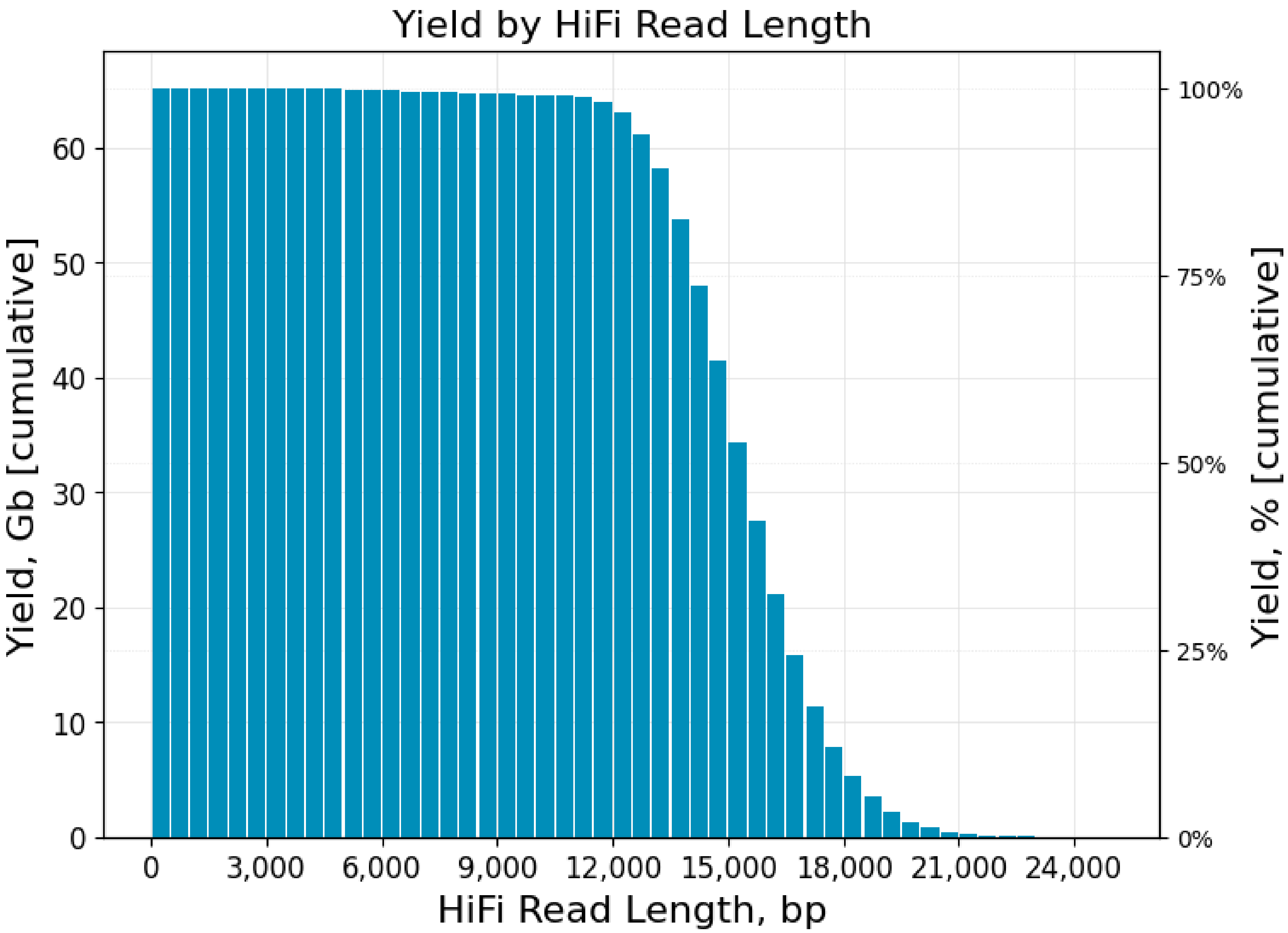
HiFi Read Quality Summary

Read Quality (Phred)	Reads	Reads (%)	Yield (Gb)	Yield (%)
Q20	4380356	100	64.95 Gb	100
Q30	1682624	38	24.58 Gb	38
Q40	19779	0	0.11 Gb	0
Q50	5234	0	0.02 Gb	0

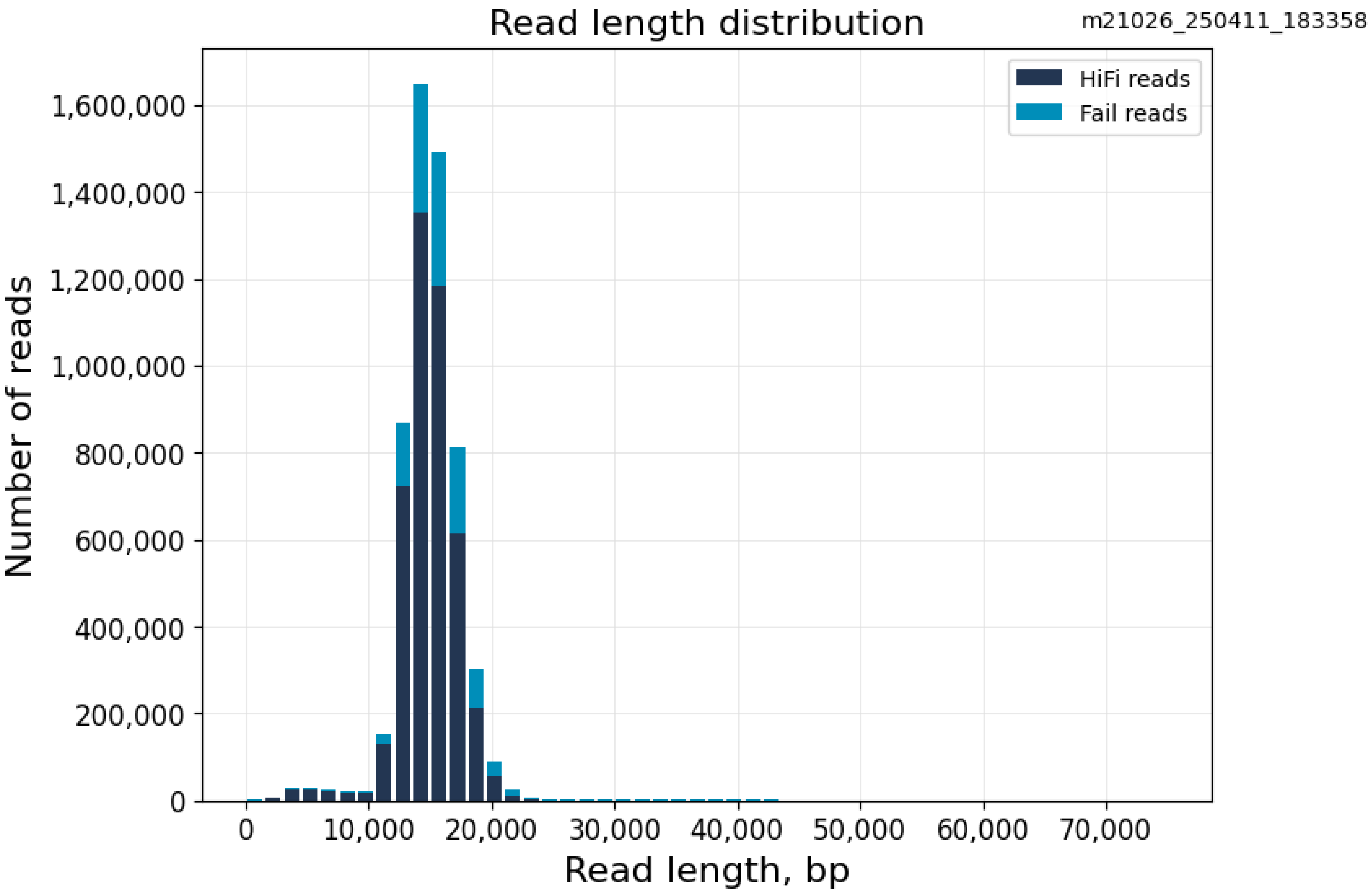
HiFi read length distribution



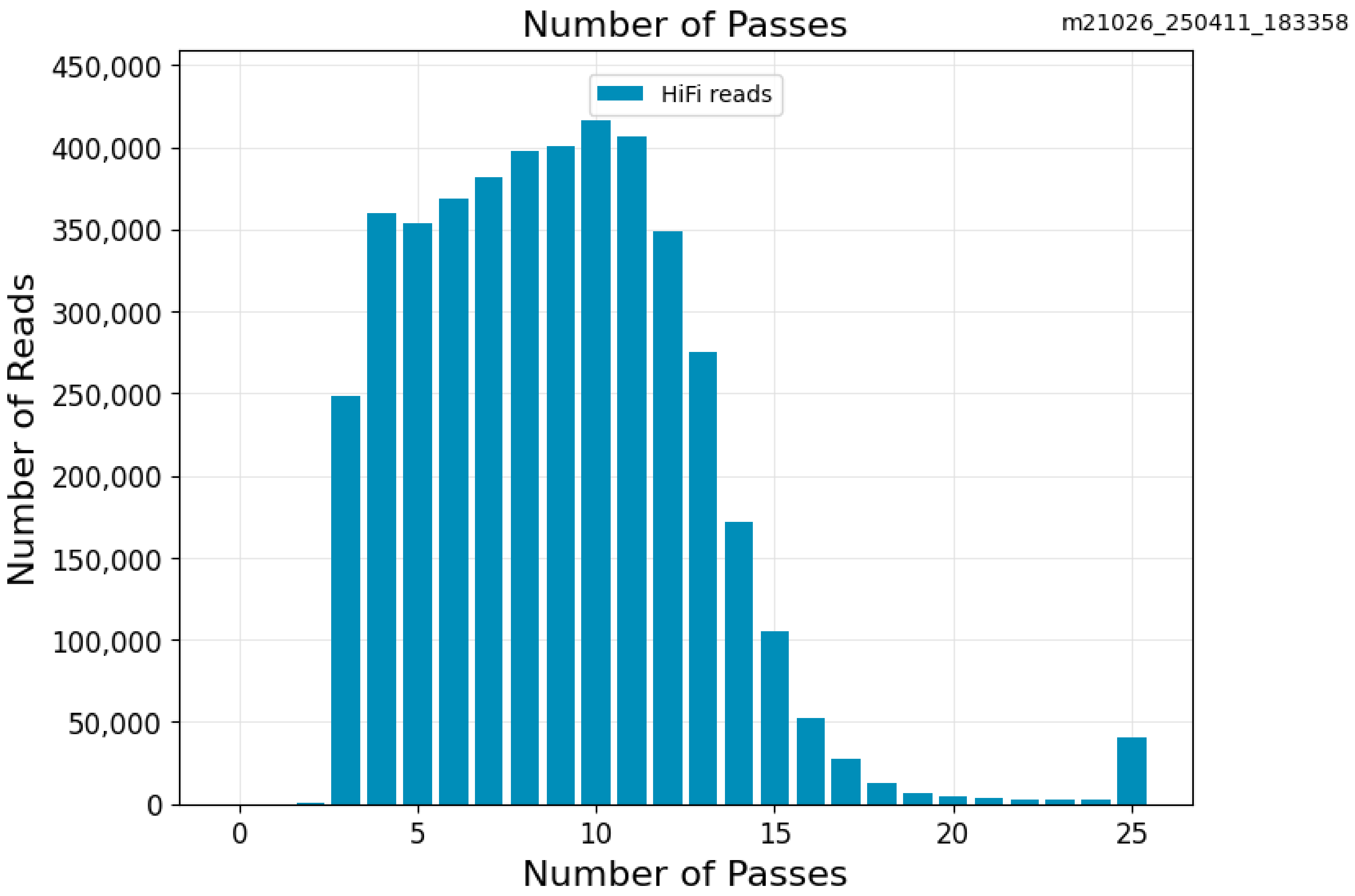
Yield by HiFi Read Length



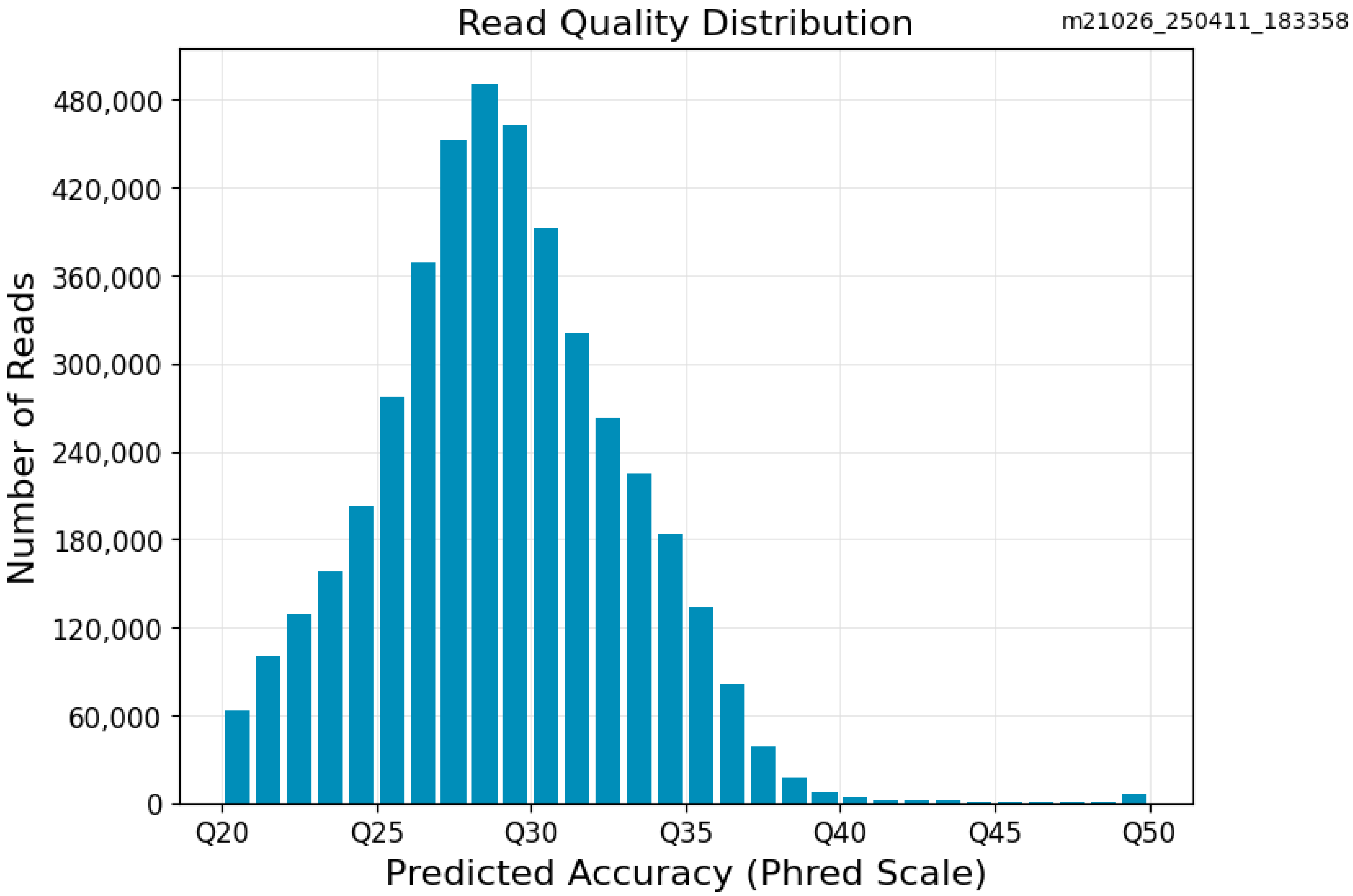
Read length distribution



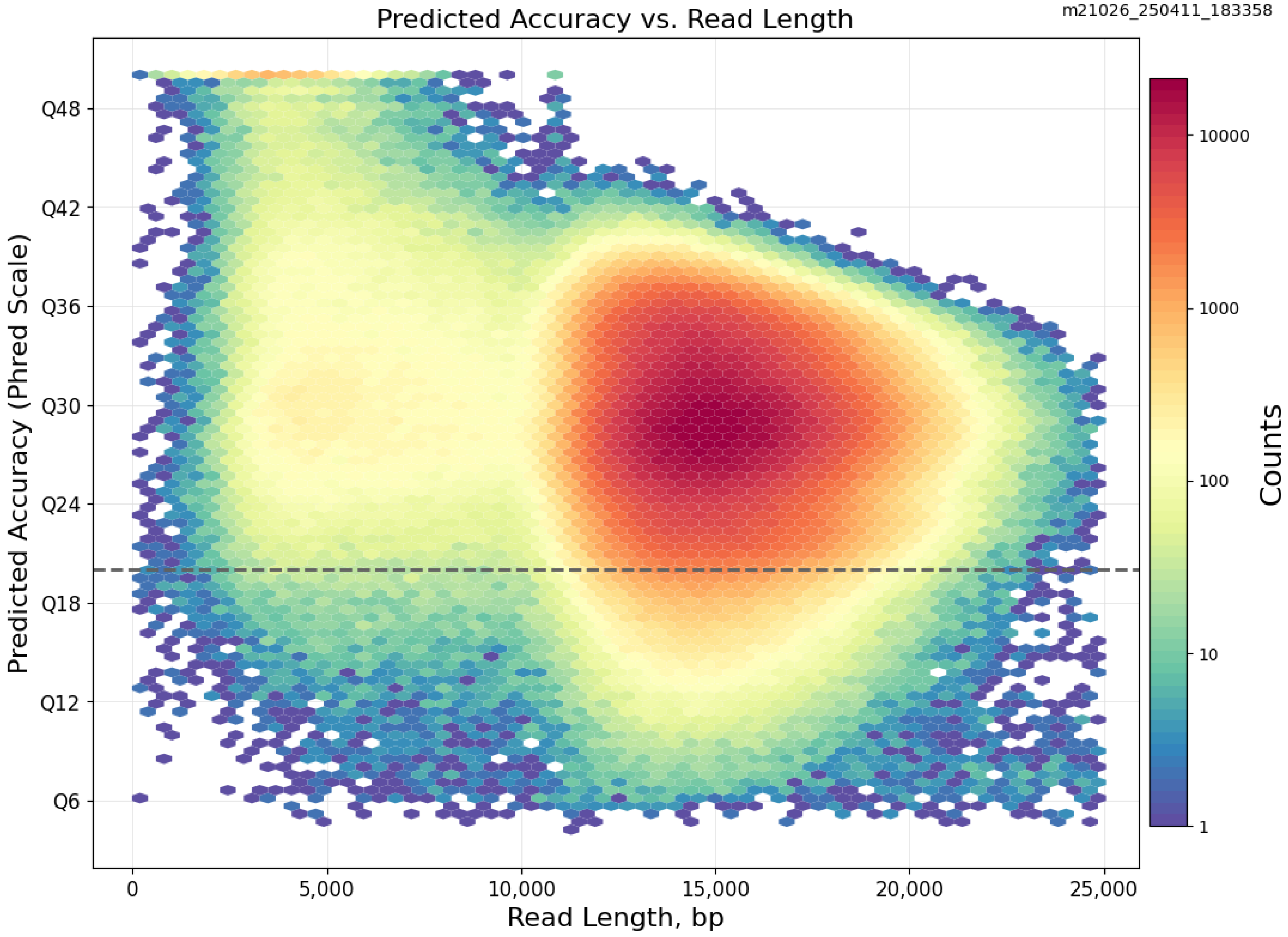
Number of Passes



Read Quality Distribution



Predicted Accuracy vs. Read Length



Loading Report

Summary

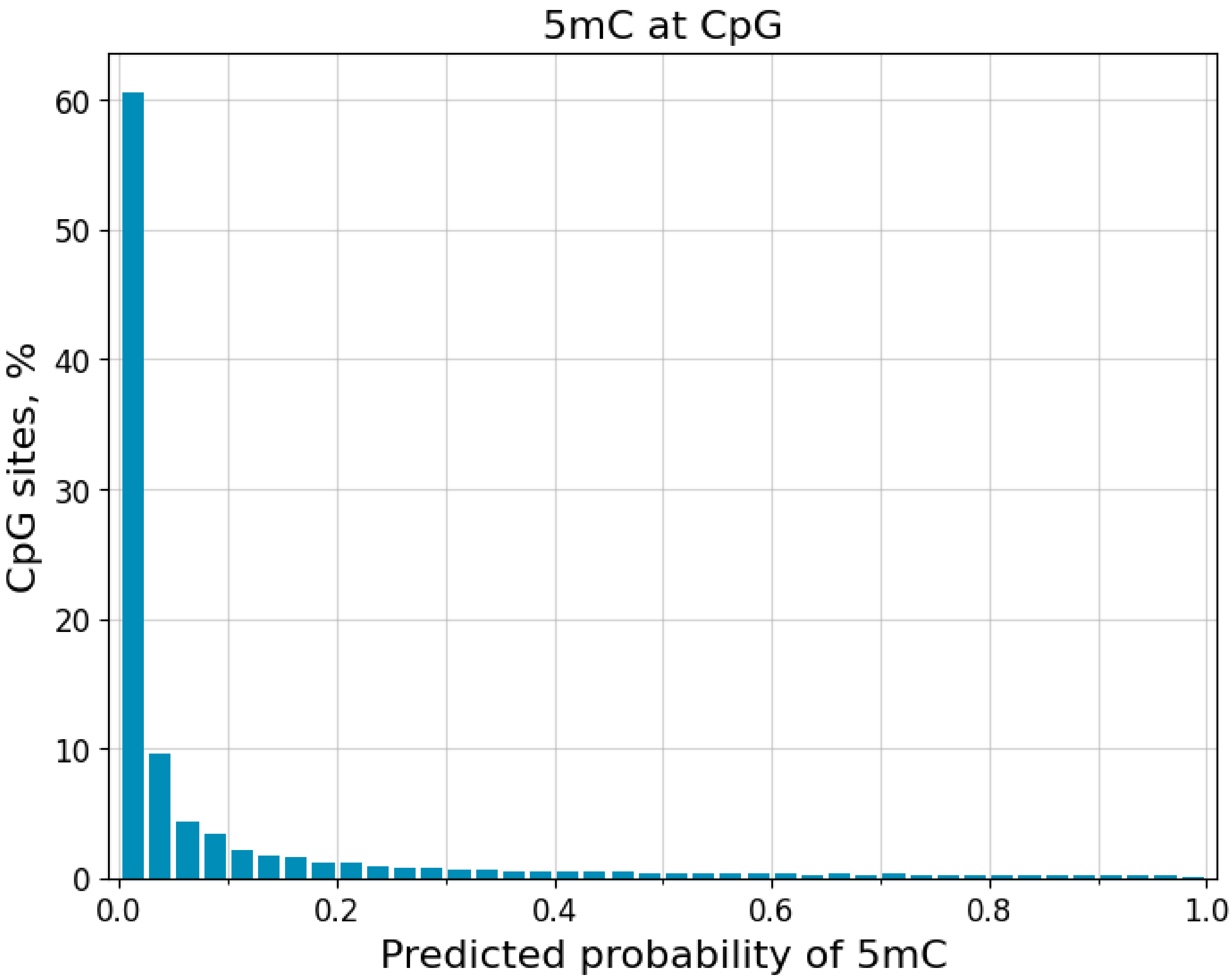
Loading Level (%)	71.05 %

Methylation

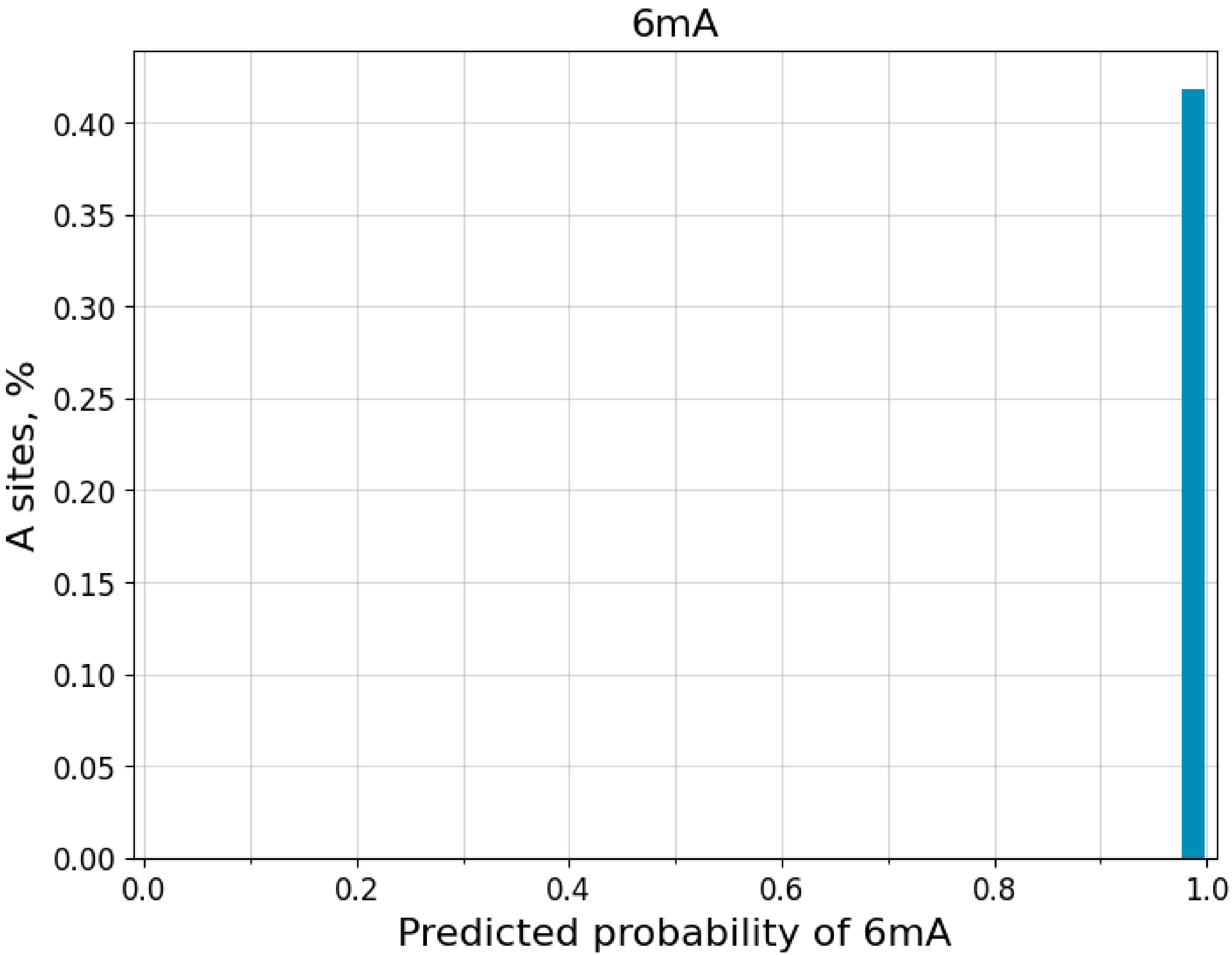
Summary

Modification	Motif	Scored sites	Modified sites (Pr > 0.5)
5mC	CpG	98.4%	5.6%
6mA	A	0.4%	0.4%

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



No Sample Setup found

No instrument run found

Parent jobs (1)

Job 124486

Summary

Job Type	import-dataset
Name	import-dataset
Comments	Description for job Import PacBio DataSet
Created At	2025-08-01 13:58:33.016
SMRT Link Version	25.3.0.273777

Child jobs (1)

Job 124496

Summary

Job Type	analysis
Pipeline	cromwell.workflows.pb_segment_reads_and_sc_isoseq
Name	KinnexRelease-10x-3pGEX-PBMC-10k-Vega
Comments	Description for job Run Analysis Application
Created At	2025-08-01 14:07:17.591
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

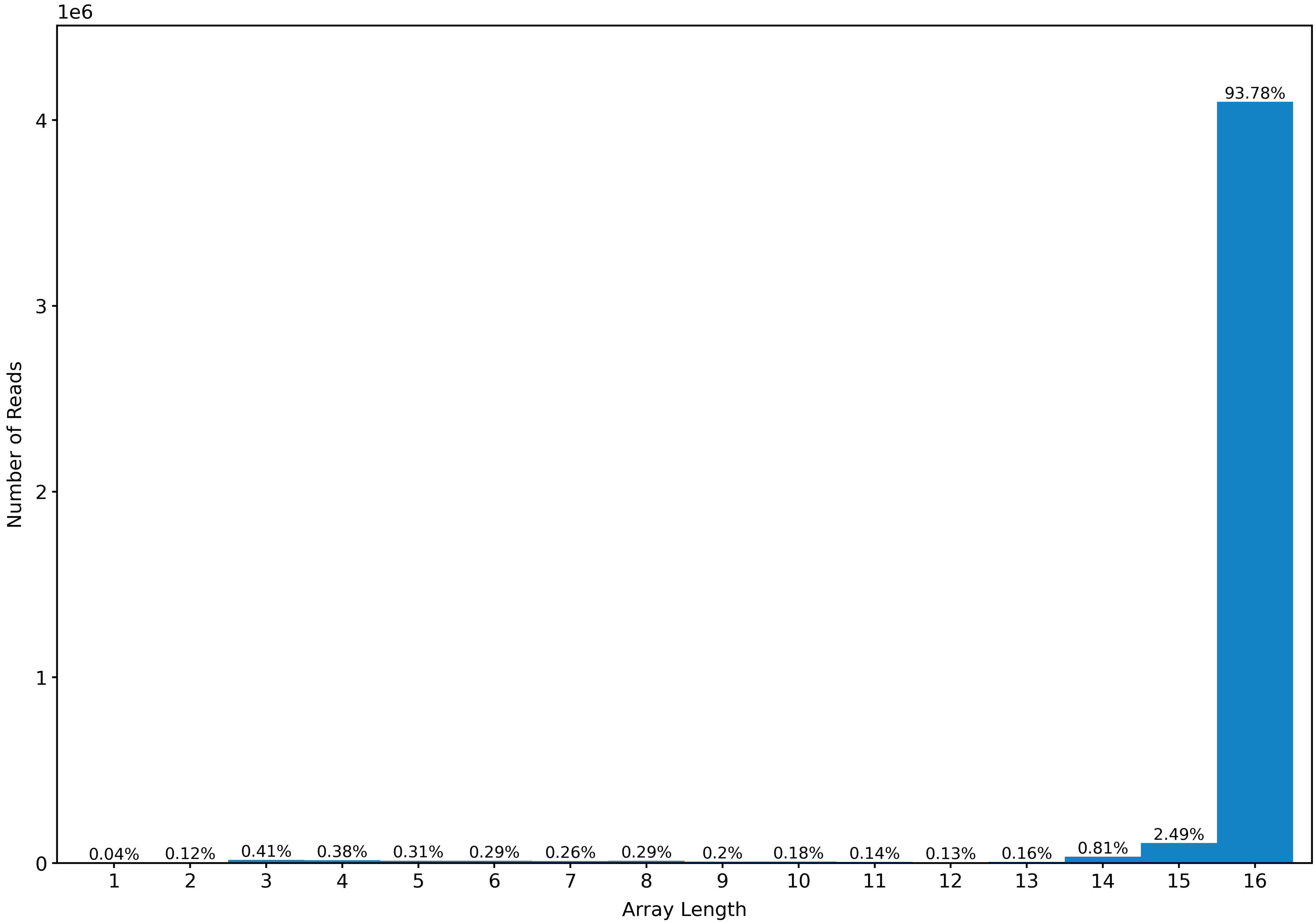
File name	Size	Type
3prime_v4_GEX_hPBMCs_10k_Rep1-Cell1 (all samples) Segmented Reads	21,946	ConsensusReadSet
Deduplicated reads after cell barcode correction, unmapped, BAM	15,275,338,243	bam
Single-cell isoform and gene matrix, tar-gzipped	1,634,116,422	tgz
Segmented Reads, passing, unaligned	42,311,221,838	bam
Non-passing reads, unaligned	1,631,594,569	bam
Unique mapped transcripts, GFF	1,269,699,595	gff
Unique mapped transcripts, filtered, GFF	557,879,859	gff
Deduplicated reads after cell barcode correction, mapped, BAM	8,128,612,166	bam
Deduplicated reads after cell barcode correction, mapped, BAM index	11,276,448	bam_bai
Unique mapped transcripts, junctions TXT	866,505,761	txt
Unique mapped transcripts, filtered, junctions TXT	511,964,681	txt
Deduplicated transcripts	11,335,279,081	Fasta
Collapsed transcript groups	941,214,084	txt
Unique mapped transcripts, classification TXT	918,909,658	txt
Unique mapped transcripts, filtered, classification TXT	304,456,884	txt
bcstats_report.tsv.gz	7,547,479	tsv
Analysis Log	69,916	log
737K_august_2016.txt.gz	2,238,609	txt
3M_3pgex_may_2023.REVCOMP.txt.gz	19,830,613	txt
SMRT Link Log	14,587	log

Read Segmentation

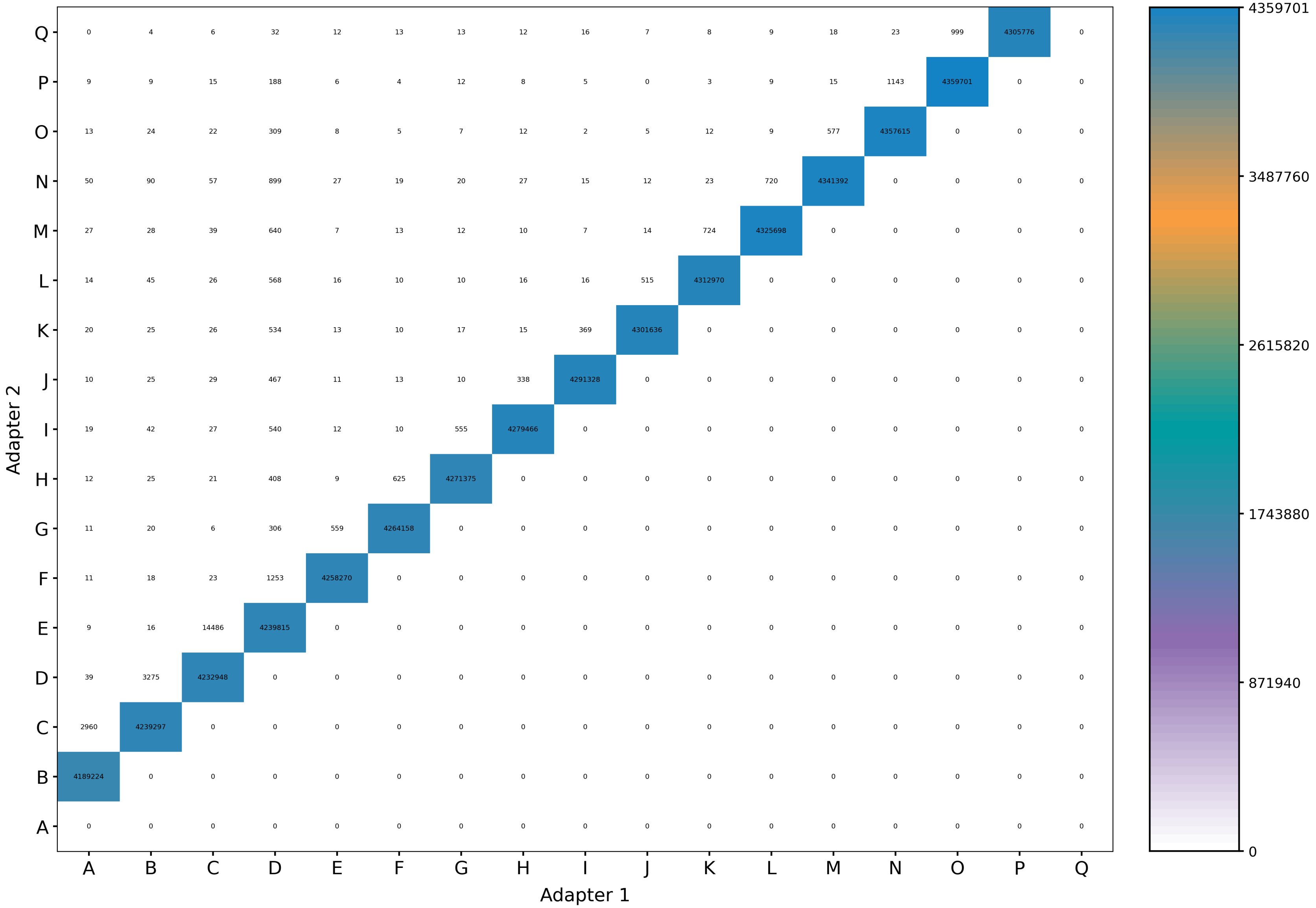
Summary

Reads	4,372,029
Segmented reads (S-reads)	68,570,669
Mean length of S-reads	924
Percent of reads with full arrays	93.78 %
Mean array size (concatenation factor)	15.68

Histogram of concatemers

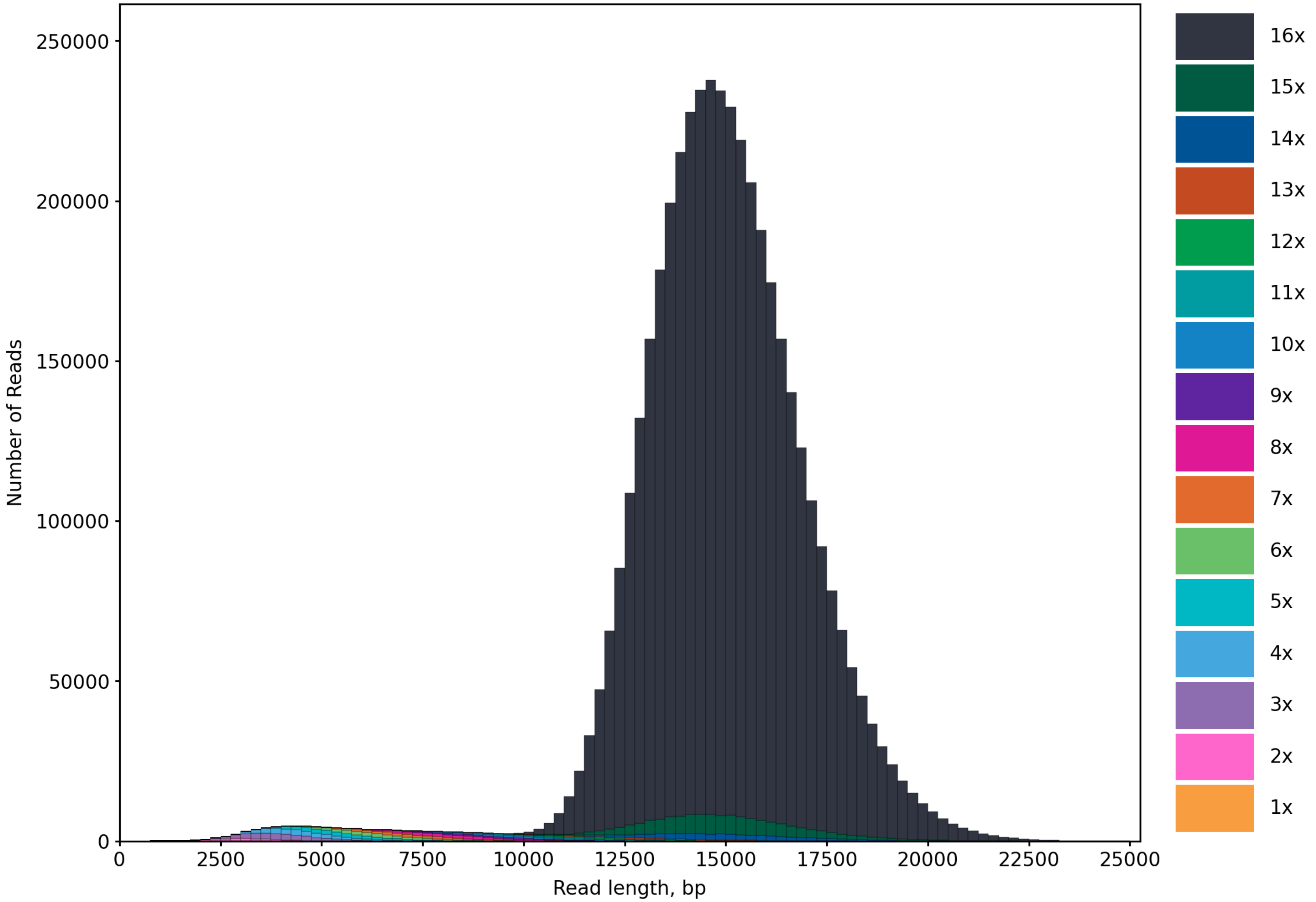


Heatmap of concatemer ligations

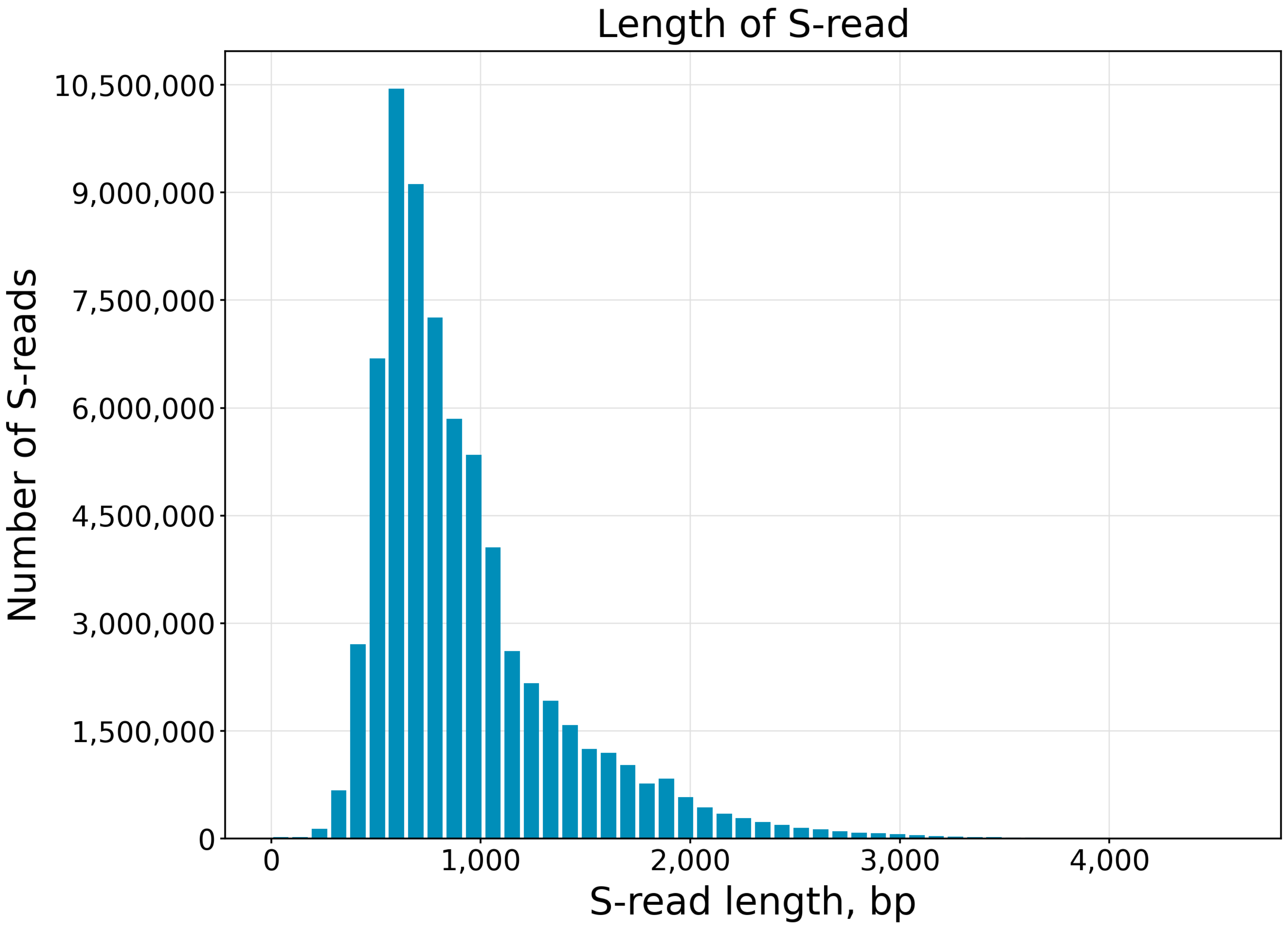


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Q	0	4	6	32	12	13	13	12	16	7	8	9	18	23	999	4305776	0
P	9	9	15	188	6	4	12	8	5	0	3	9	15	1143	4359701	0	0
O	13	24	22	309	8	5	7	12	2	5	12	9	577	4357615	0	0	0
N	50	90	57	899	27	19	20	27	15	12	23	720	4341392	0	0	0	0
M	27	28	39	640	7	13	12	10	7	14	724	4325698	0	0	0	0	0
L	14	45	26	568	16	10	10	16	16	515	4312970	0	0	0	0	0	0
K	20	25	26	534	13	10	17	15	369	4301636	0	0	0	0	0	0	0
J	10	25	29	467	11	13	10	338	4291328	0	0	0	0	0	0	0	0
I	19	42	27	540	12	10	555	4279466	0	0	0	0	0	0	0	0	0
H	12	25	21	408	9	625	4271375	0	0	0	0	0	0	0	0	0	0
G	11	20	6	306	559	4264158	0	0	0	0	0	0	0	0	0	0	0
F	11	18	23	1253	4258270	0	0	0	0	0	0	0	0	0	0	0	0
E	9	16	14486	4239815	0	0	0	0	0	0	0	0	0	0	0	0	0
D	39	3275	4232948	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	2960	4239297	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	4189224	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

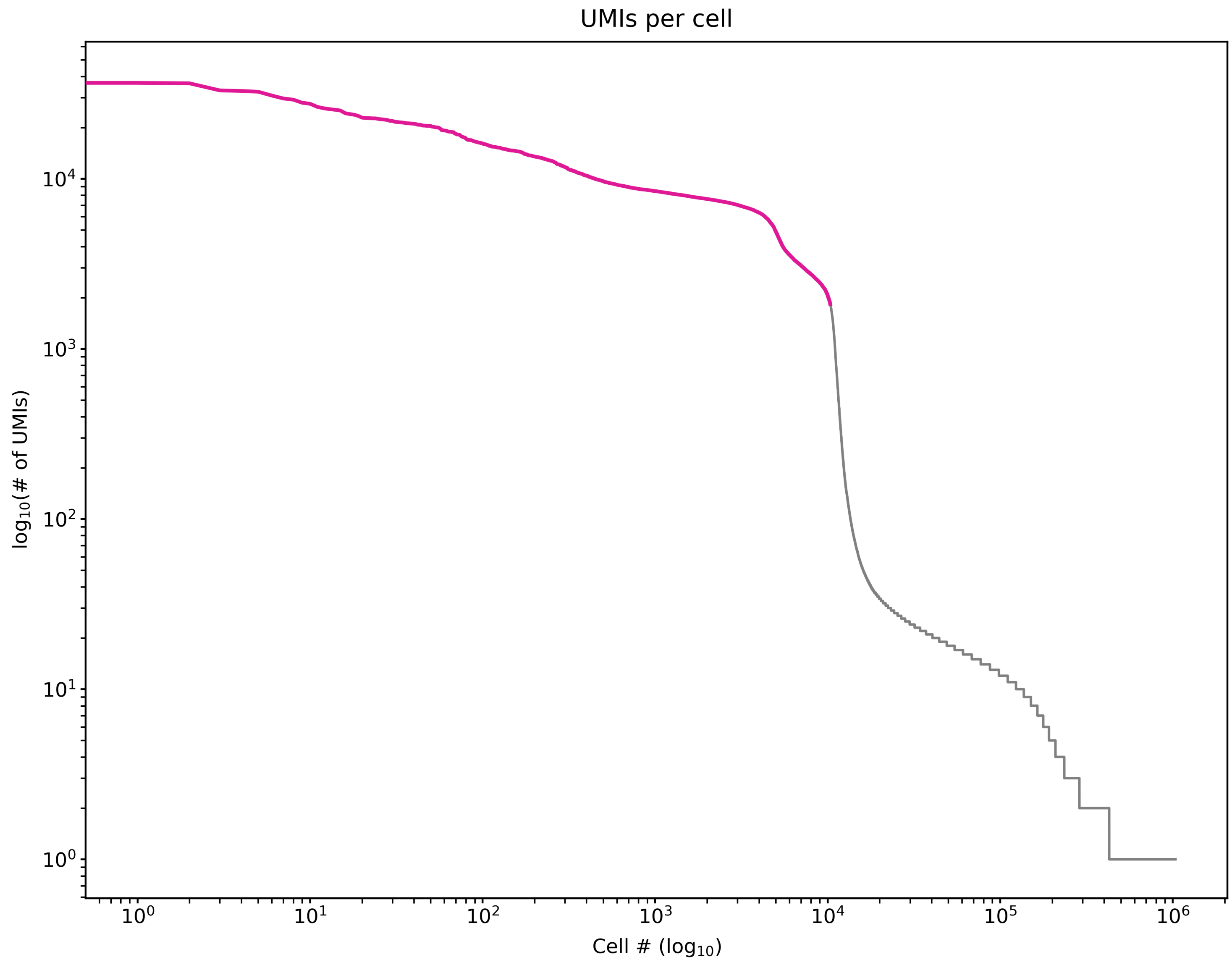


Cell Statistics

Summary

Estimated Number of Cells	10,329
Reads in Cells	90.70%
Mean Reads per Cell	5,818
Median UMIs per Cell	4,561

Barcode Rank Plot



Read Statistics

Summary

Reads	68,570,669
Read Type	SEGMENT
Reads with 5' and 3' Primers with extracted UMIs and Barcodes	68,007,729
Non-Concatamer Reads with 5' and 3' Primers and Poly-A Tail (FLNC reads)	66,266,566
FLNC Reads with Valid Barcodes	64,948,820
FLNC Reads with Valid Barcodes, corrected	66,113,774
Reads after Barcode Correction and UMI Deduplication	54,980,296

Transcript Statistics

Summary

FLNC Reads Mapped Confidently to Genome	54,152,125
FLNC Reads Mapped Confidently to Transcriptome	35,486,110
Total Unique Genes	1,310,725
Total Unique Genes, filtered	62,925
Total Unique Genes, known genes only	33,435
Total Unique Genes, filtered, known genes only	25,832
Total Unique Transcripts	2,819,731
Total Unique Transcripts, filtered	604,684
Total Unique Transcripts, known transcripts only	92,696
Total Unique Transcripts, filtered, known transcripts only	82,608

Transcript Summary

Value	Analysis Metric
2,002	Median Genes per Cell
1,510	Median Genes per Cell, known genes only
2,466	Median Transcripts per Cell
1,253	Median Transcripts per Cell, known transcripts only
1,310,725	Total Unique Genes
33,435	Total Unique Genes, known genes only
2,819,731	Total Unique Transcripts
92,696	Total Unique Transcripts, known transcripts only

Transcript Summary, filtered

Value	Analysis Metric
1,108	Median Genes per Cell
1,095	Median Genes per Cell, known genes only
1,348	Median Transcripts per Cell
1,086	Median Transcripts per Cell, known transcripts only
62,925	Total Unique Genes
25,832	Total Unique Genes, known genes only
604,684	Total Unique Transcripts
82,608	Total Unique Transcripts, known transcripts only

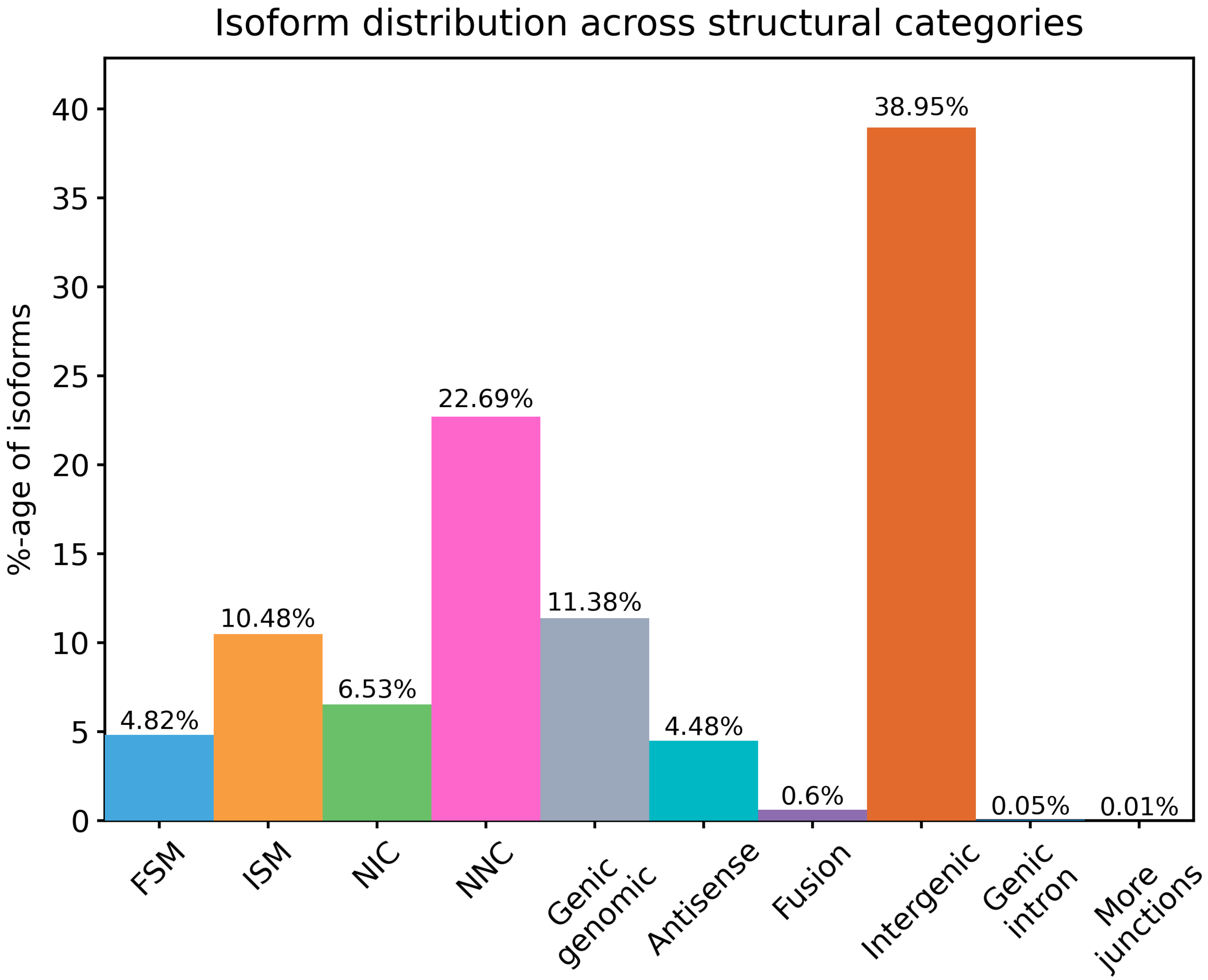
Transcript Classification

Category	Count	CAGE Detected	CAGE Detected (%)	polyA Motif Detected	polyA Motif Detected (%)
FSM	155050	97462	62.86%	47102	30.38%
ISM	337549	155273	46.00%	98048	29.05%
NIC	210362	147978	70.34%	69196	32.89%
NNC	730606	489006	66.93%	272795	37.34%
Antisense	144203	7403	5.13%	39078	27.10%
Fusion	19435	11682	60.11%	7867	40.48%
More junctions	239	122	51.05%	73	30.54%
Genic intron	1694	12	0.71%	550	32.47%
Genic genomic	366435	97488	26.60%	110782	30.23%
Intergenic	1254061	15628	1.25%	424965	33.89%

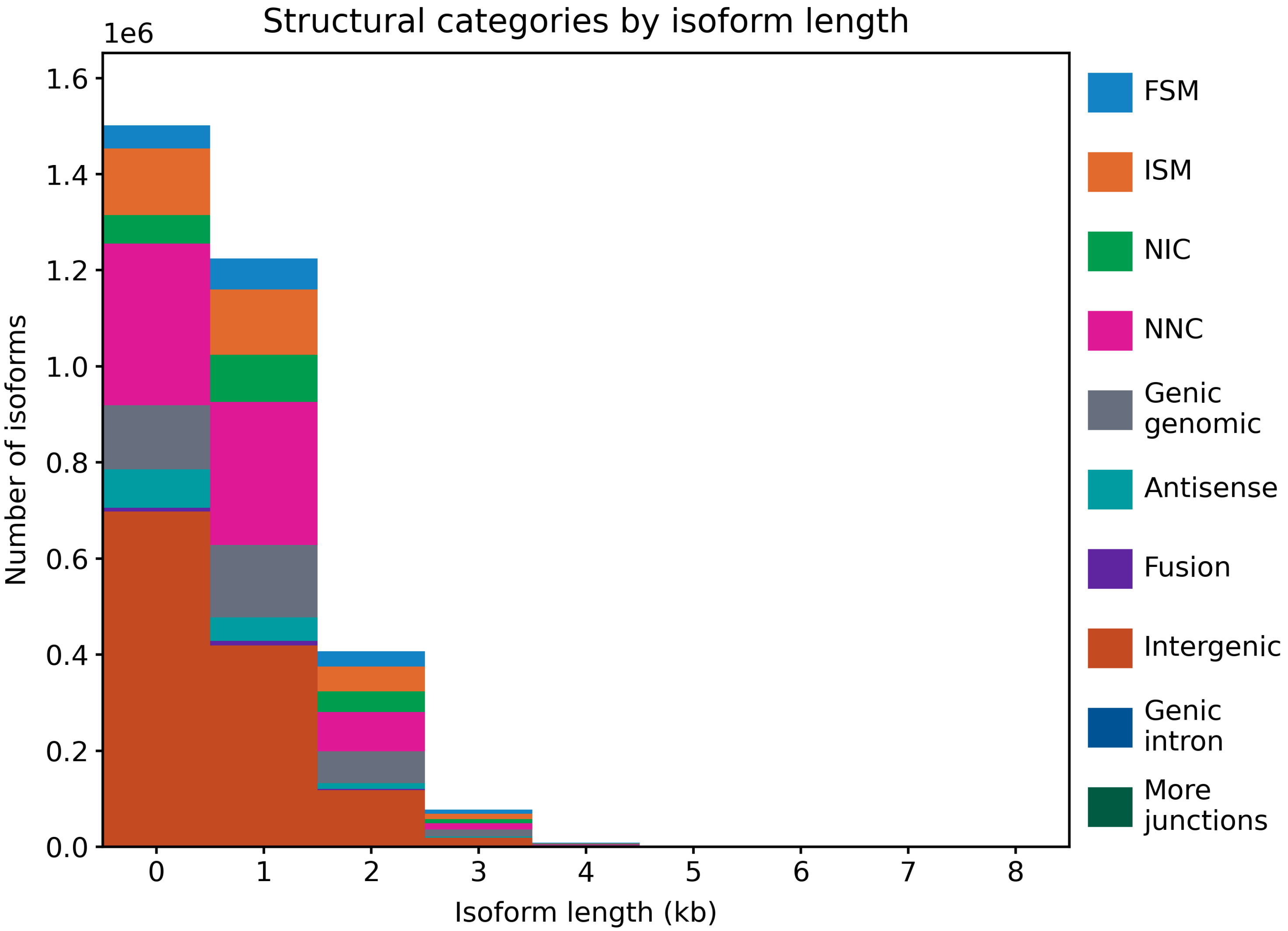
Transcript Classification, filtered

Category	Count	CAGE Detected	CAGE Detected, (%)	polyA Detected	polyA Detected, (%)
FSM	131927	84716	64.21%	47102	35.70%
ISM	227511	120138	52.81%	80142	35.23%
NIC	169708	120081	70.76%	68508	40.37%
NNC	289868	196647	67.84%	131232	45.27%
Antisense	9729	1259	12.94%	5429	55.80%
Fusion	9621	5876	61.07%	4633	48.16%
More junctions	167	83	49.70%	64	38.32%
Genic intron	0	0	0.00%	0	0.00%
Genic genomic	10112	5517	54.56%	4269	42.22%
Intergenic	32871	1412	4.30%	22540	68.57%

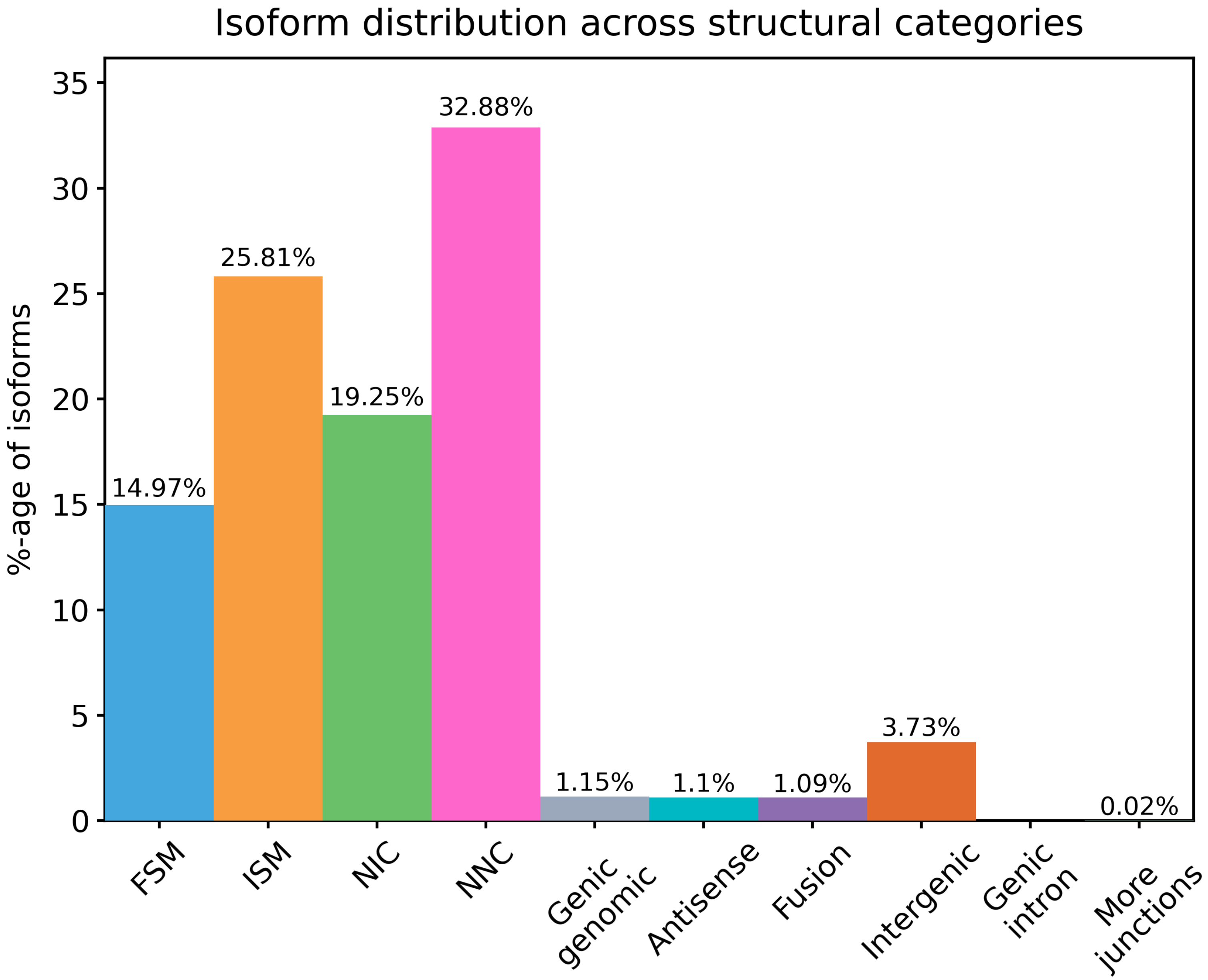
Isoform Distribution Across Categories


Isoform distribution across structural categories
%-age of isoforms
0
5
10
15
20
25
30
35
40
4.82%
10.48%
6.53%
22.69%
11.38%
4.48%
0.6%
38.95%
0.05%
0.01%
FSM
ISM
NIC
NNC
Genic genomic
Antisense
Fusion
Intergenic
Genic intron
More junctions

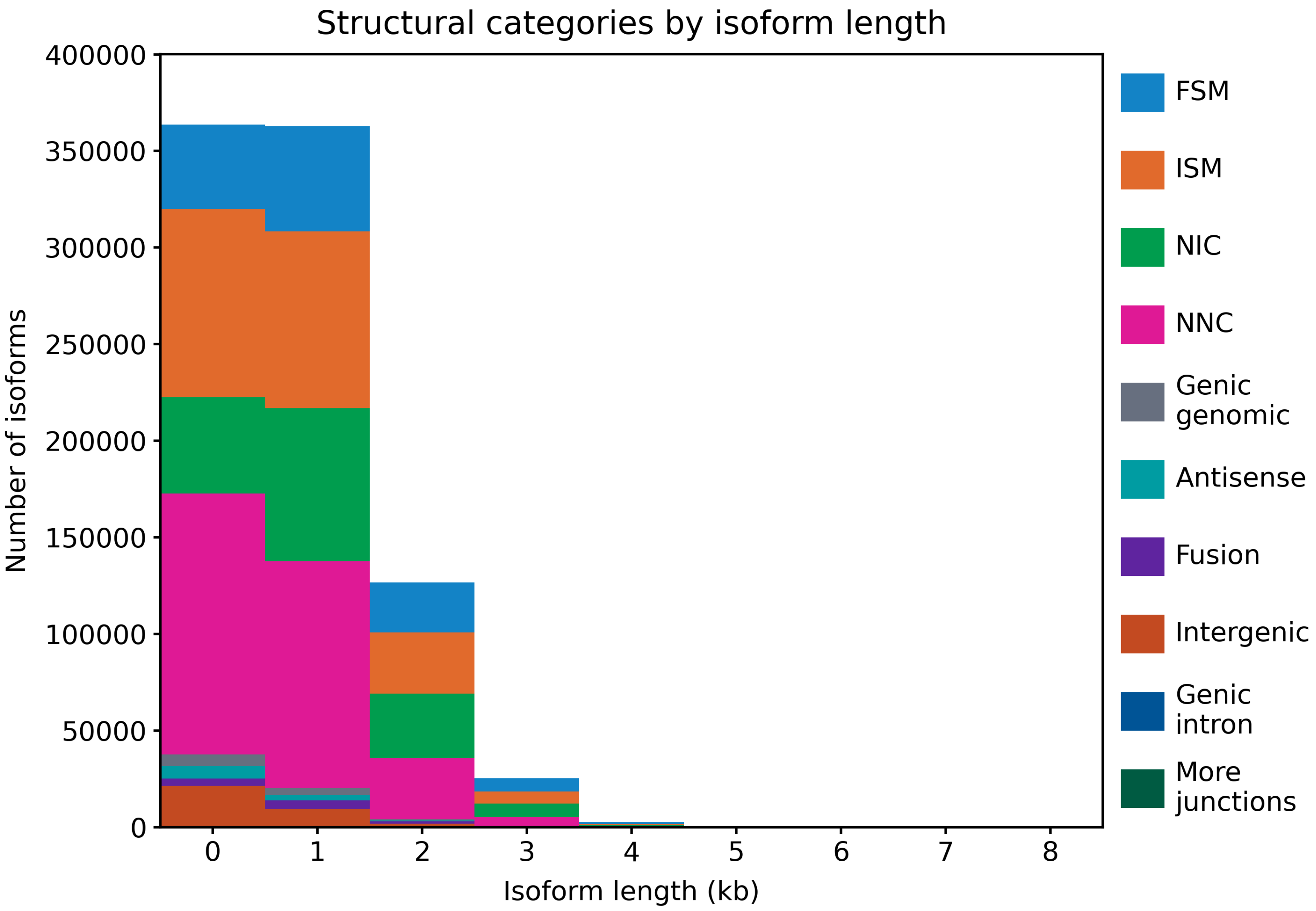
Structural Categories by Transcript Length



Isoform Distribution Across Categories


Isoform distribution across structural categories
%-age of isoforms
0
5
10
15
20
25
30
35
14.97%
25.81%
19.25%
32.88%
1.15%
1.1%
1.09%
3.73%
0.02%
FSM
ISM
NIC
NNC
Genic genomic
Antisense
Fusion
Intergenic
Genic intron
More junctions

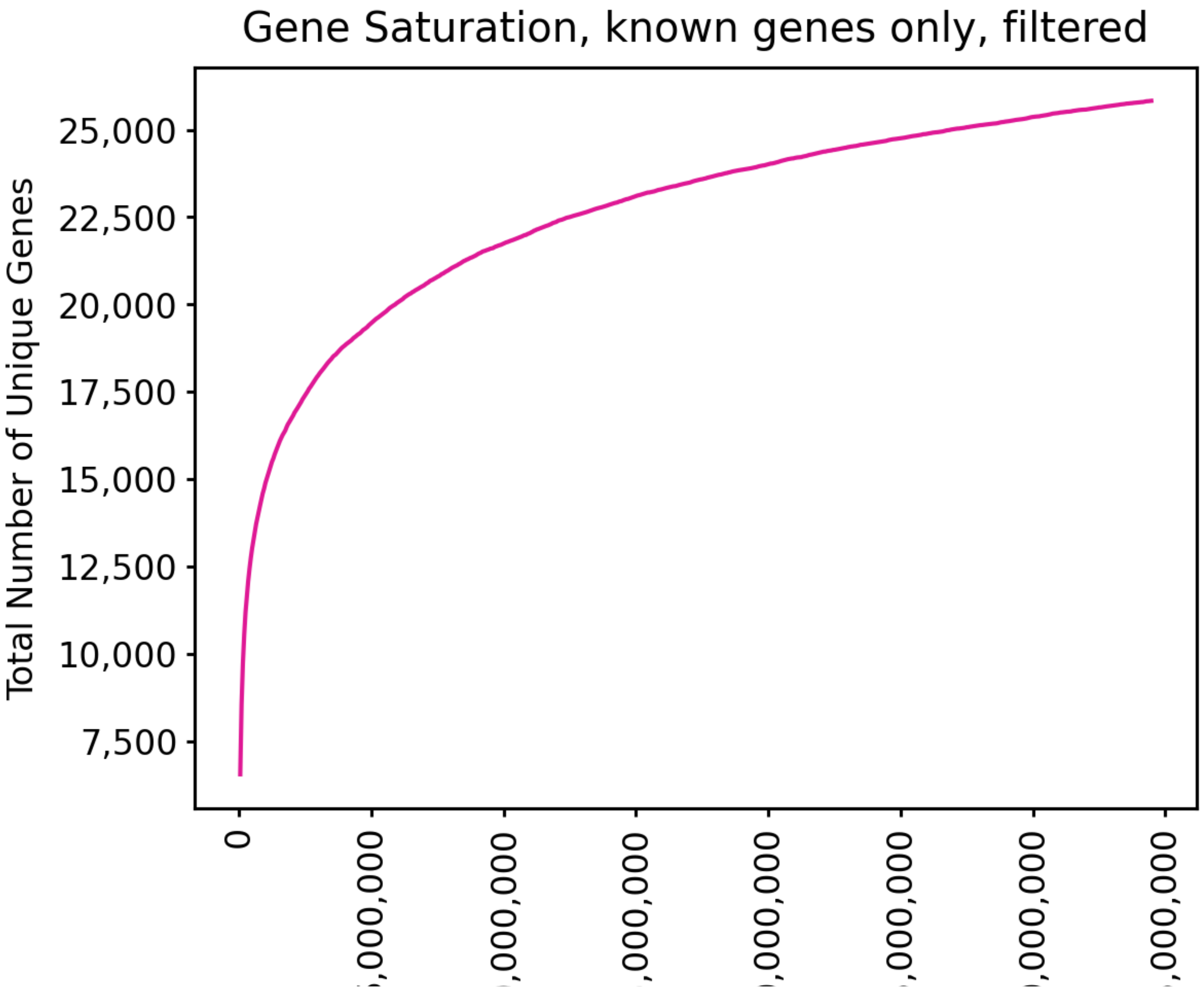
Structural Categories by Transcript Length



sc_isoseq_transcript_statistics.gene_saturation_plotgroup.gene_saturation_plot_0

5
10
15
20
25
30
35
Total Reads

sc_isoseq_transcript_statistics.gene_saturation_plotgroup.gene_saturation_plot_1




5
10
15
20
25
30
35
Total Reads