Cell Reports Methods, Volume *4*

Supplemental information

Single chromatin fiber profiling and nucleosome position mapping in the human brain

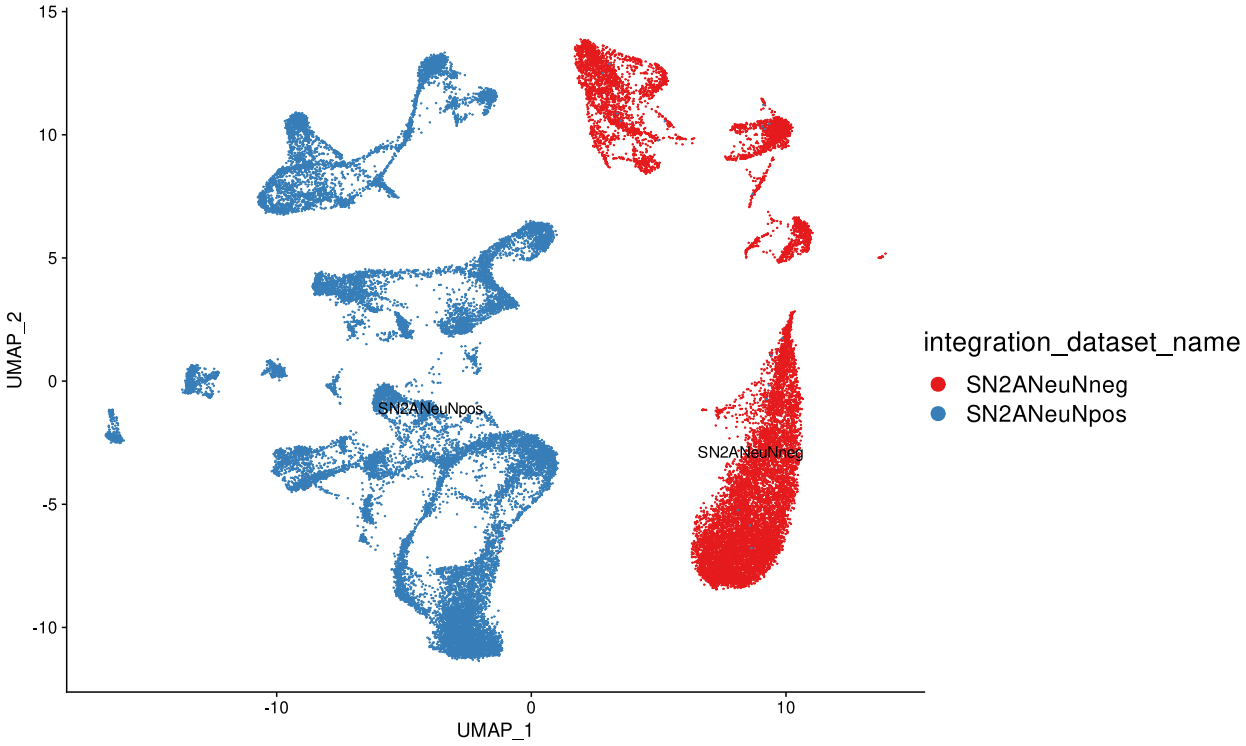
Cyril J. Peter, Aman Agarwal, Risa Watanabe, Bibi S. Kassim, Xuedi Wang, Tova Y. Lambert, Behnam Javidfar, Viviana Evans, Travis Dawson, Maya Fridrikh, Kiran Girdhar, Panos Roussos, Sathiji K. Nageshwaran, Nadejda M. Tsankova, Robert P. Sebra, Mitchell R. Vollger, Andrew B. Stergachis, Dan Hasson, and Schahram Akbarian

Supplementary figures S1-S9, legends and figures.

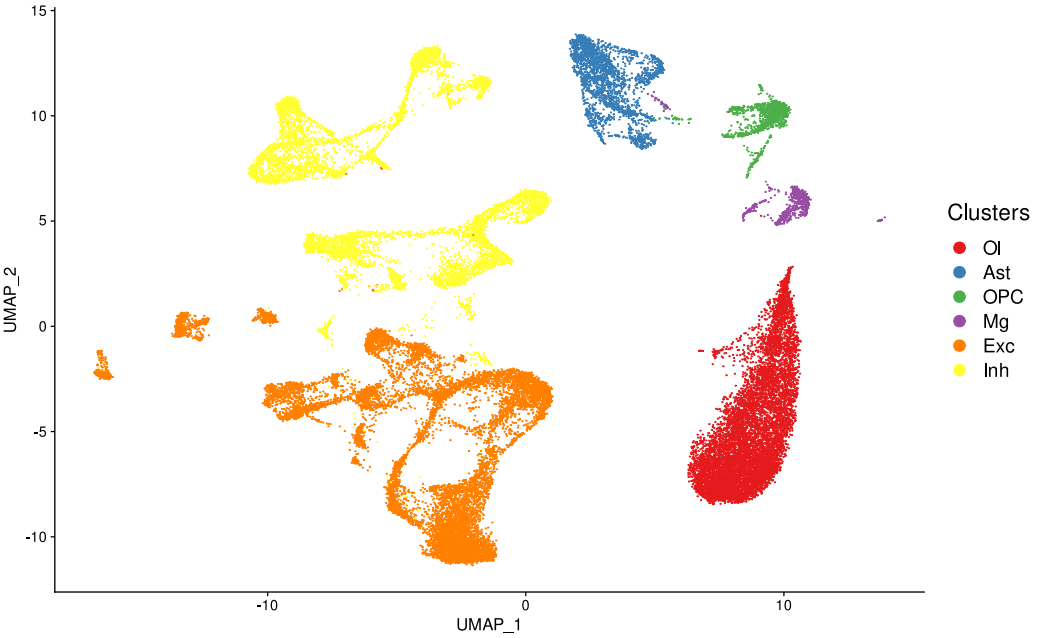
Supplemental Figure S1: Single nucleus transcriptomes confirm complete separation of neuronal and glial population after sorting with NeuN antibody (related to Figure 1).

(A, B) Uniform Manifold Approximation and Projection (UMAP) plots generated after combined analysis of single nucleus-RNAseq of 22,844 NeuN+ and 13,977 NeuN- nuclei sorted from frontal cortex tissue block of a 24y/o male donor (Brain ID SN2a, Table S1), with median genes expressed/ nucleus NeuN+, 2,859 and NeuN-, 2,675 , respectively. Notice (**A**) complete separation of NeuN+ and NeuN- nuclei populations, with (**B**) NeuN+ comprised of excitatory (Exc) and inhibitory (Inh) neuronal population, and NeuN- comprised of oligodendrocytes (Ol) and oligodendrocyte precursors (OPC), astrocytes (Ast) and microglia (Mg), as indicated. (**C**) Single nucleus expression levels of RBFOX3 as the NeuN encoding transcript confirms overwhelming proportion of RBFOX3+ nuclei locating to excitatory and inhibitory neuronal subtype.

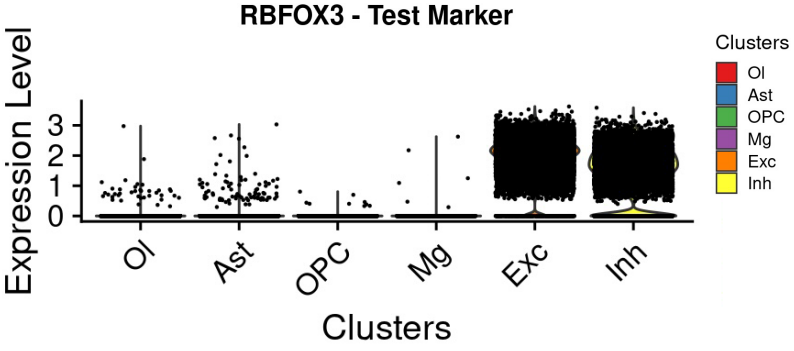
A


UMAP_2
UMAP_1
SN2ANeuNpos
SN2ANeuNneg
integration_dataset_name
SN2ANeuNneg
SN2ANeuNpos

B

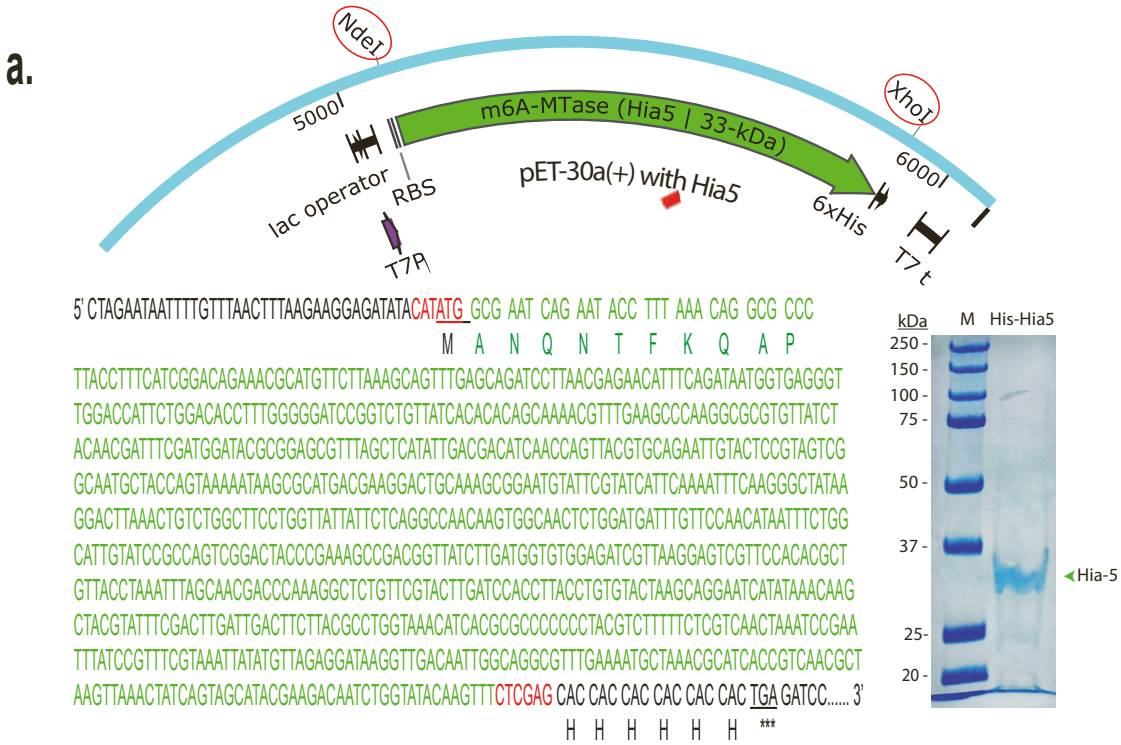

UMAP_2
UMAP_1
Clusters
Ol
Ast
OPC
Mg
Exc
Inh

C


RBFOX3 - Test Marker
Expression Level
Clusters
Ol
Ast
OPC
Mg
Exc
Inh

Supplemental Figure S2: Production of the Hia5 Enzyme, ELISA Detection of m6A, and Quality Control Checks for the m6A Background Signal in Brain (related to STAR Methods).

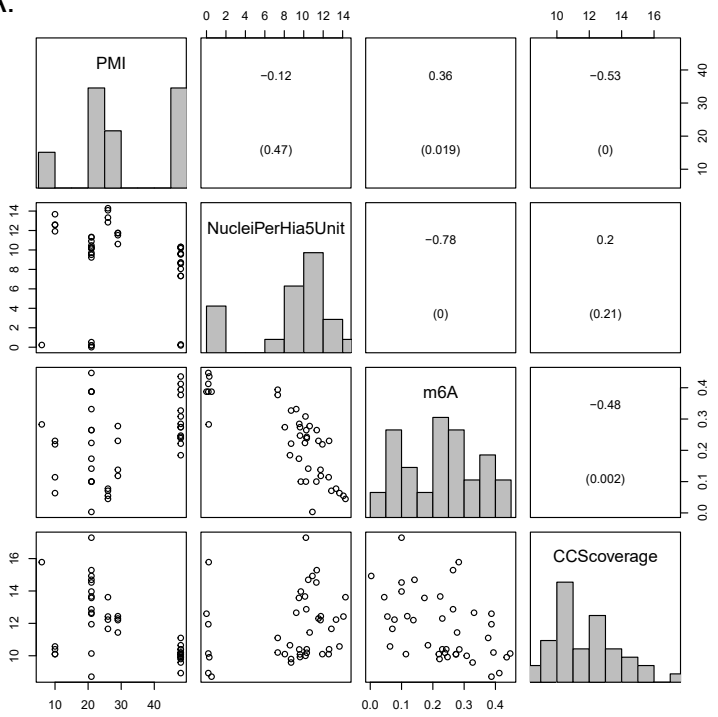
(a) Plasmid construction (top) and DNA sequence (bottom left) of the *Hemophilus influenza* Hia5 m6A adenylmethyltransferase to be expressed in E. coli. (bottom right) SDS-PAGE gel of the recombinant produced HiA5 purified protein. (**b**) Synthetic 1026 bp DNA construct with adenine locations marked, as indicated. (bottom left) Polyacrylamide gel electrophoresis (PAGE). Unmethylated synthetic DNA (Crtl) is not cleaved after treatment with the restriction enzyme DpnI. Synthetic DNA treated with in-house generated Hia5, and commercially acquired EcoGII enzyme shows expected cleavage product after DpnI digest. (bottom right) slot-blot incubated with anti-m6A antibody. Synthetic DNA substrate (top to bottom) untreated (no enzyme control), HIa5 treated, EcoGII treated. (**c**) (left) Graphical representation of ELISA assay to quantify activity of recombinant Hia5 produced in-house. Each step of the reaction is given its own box (green box, add material to the reaction; blue box depict wash steps where the substrate remains unchanged; red box shows the blocking step. Reagents marked in bold font; black font marks volume, blue incubation time, red temperature (RT: room temperature). (right) The line plot and well-plate show increasing ELISA fluorescence detected with increasing amounts of methyltransferase (blue, red dots), but no fluorescence when enzyme is not added (yellow, blue dots). (**d**) Graphical depiction of *in situ* Hia5 methyltransferase treatment of FACS sorted NeuN+ and NeuN- nuclei as indicated. (**e**) Slot blot-assay with anti-m6A antibody as quality control check for (test mouse M1 and M2) DNA from FACS-sorted brain nuclei after being processed as libraries for sequencing. (**f**) Histograms showing the per-read detected adenine methylation for Hia5 treated (red, left) and untreated (blue, right) human postmortem cerebral cortex. Treated samples showed a median proportion of adenines methylated at 0.03, while untreated samples had a mean of 0.


a.
NdeI
5000l
m6A-MTase (Hia5 | 33-kDa)
XhoI
6000l
lac operator
RBS
pET-30a(+) with Hia5
6xHis
T7P
T7 t
5' CTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATG GCG AAT CAG AAT ACC TTT AAA CAG GCG CCC
M A N Q N T F K Q A P
TTACCTTTCATCGGACAGAAACGCATGTTCTTAAAGCAGTTTGAGCAGATCCTTAACGAGAACATTTCAGATAATGGTGAGGGT
TGGACCATTCTGGACACCTTTGGGGGATCCGGTCTGTTATCACACACAGCAAAACGTTTGAAGCCCAAGGCGCGTGTTATCT
ACAACGATTTCGATGGATACGCGGAGCGTTTAGCTCATATTGACGACATCAACCAGTTACGTGCAGAATTGTACTCCGTAGTCG
GCAATGCTACCAGTAAAAATAAGCGCATGACGAAGGACTGCAAAGCGGAATGTATTCGTATCATTCAAAATTTCAAGGGCTATAA
GGACTTAAACTGTCTGGCTTCCTGGTTATTATTCTCAGGCCAACAAGTGGCAACTCTGGATGATTTGTTCCAACATAATTTCTGG
CATTGTATCCGCCAGTCGGACTACCCGAAAGCCGACGGTTATCTTGATGGTGTGGAGATCGTTAAGGAGTCGTTCCACACGCT
GTTACCTAAATTTAGCAACGACCCAAAGGCTCTGTTCGTACTTGATCCACCTTACCTGTGTACTAAGCAGGAATCATATAAACAAG
CTACGTATTTCGACTTGATTGACTTCTTACGCCTGGTAAACATCACGCGCCCCCCCTACGTCTTTTTCTCGTCAACTAAATCCGAA
TTTATCCGTTTCGTAAATTATATGTTAGAGGATAAGGTTGACAATTGGCAGGCGTTTGAAAATGCTAAACGCATCACCGTCAACGCT
AAGTTAAACTATCAGTAGCATACGAAGACAATCTGGTATACAAGTTTCTCGAG CAC CAC CAC CAC CAC CAC TGA GATCC..... 3'
H H H H H H ***
kDa
M
His-Hia5
250
150
100
75
50
37
25
20
Hia-5
b.
DpnI
1
459
645
669
960
1026
458-bp
186
24
291
66
M
-Ctrl
-EcoGII
-Hia-5
1000-
500-
300-
200-
100-
-1026 (S)
-291
-186
-66
DpnI detects Methylation
Assay: 90-min (PAGE)
Substrate
DNA (1-µg)
No Enzyme
m6A-MTase
Hia-5
EcoGII
Anti-m6A detects Methylation
Assay: Slot-Blot
c.
Add DNA Substrate
Substrate DNA
100 µl
60-min
RT
Wash
TBS-T
150 µl
5-min
RT
Block
5% milk in TBS-T
100 µl
30-min
RT
Wash
TBS-T
150 µl
5-min
RT
Wash
Activation Buffer
150 µl
5-min
RT
Adenine Methylation
SAM
SAM
Hia5
Activation Buffer
100 µl
60-min
37°C
Wash (X3)
TBS-T
150 µl
5-min
RT
Primary Antibody
Rabbit Anti-m6A Ab
100 µl
30-min
RT
Wash (X3)
SPB-T
150 µl
10-min
RT
Secondary Antibody
Ab-HRP
100 µl
30-min
RT
Wash (X3)
SPB-T
150 µl
10-min
RT
Detection
TMBL HRP substrate
450 nm
TMBL Substrate
15-min
RT
Plate Reader
Streptavidin
Biotin
DNA Substrate
Methyltransferase
Methyl-source
m6A
m6A rpAb
m6A HRP pAb
HRP
Hia5
No Enzyme
A450 (m6A production)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Batch-1: 22 u/µl
1 Unit=1 nmol substrate per min
A450
100,000
80,000
60,000
40,000
20,000
10,000
Hia5 dilutions
d.
NeuN-
Hia5
SDS
m6A Methylation
37°C
20-min
1500-rpm
Hia5 Inactivation
1% SDS; 37°C
5-min
1500-rpm
NeuN+
e.
HiFi Library 1-4 (unsorted)
HiFi library-5-8 (FACS sorted)
Hia-5
+
-
-
+
Hia-5
+
+
+
+
NeuN-
NeuN+
NeuN-
NeuN+
M1
M2
f.
Hia5 Treated Human PFC Nuclei
0.03
0.1%>0.3
Reads
Proportion of adenines with m6A per read
(for reads with 8+ CCS coverage)
Hia5 Un-Treated Human PFC Nuclei
0
0%>0.3
Proportion of adenines with m6A per read
(for reads with 8+ CCS coverage)

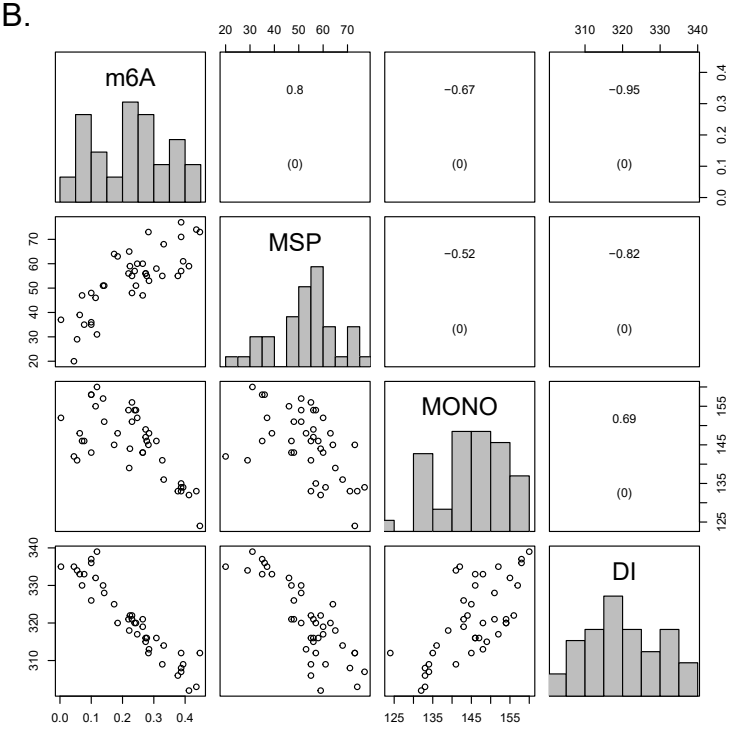
Supplemental Figure S3: Impact of postmortem interval, nuclei to HIa5 ratio and CCS on methyl-adenine labeling. Impact of genome-level adenine methylation on Fiber-seq computational outcomes (related to Figure 1 and STAR Methods).

(A) Correlation plots to assess impact of postmortem interval (PMI), circular consensus sequence (CCS), N nuclei / Hia5 unit on methyl-adenine m6A/A proportion, for 41 PFC Fiber-seq assays. Data are shown after log-transformation and square root-transformation of *N* nuclei per Hia5 Uni,t and of m6A proportiosn, respectively. Lower panels, diagonal panels, and the upper panels show scatter plot, histogram, and pearson's correlations coefficient with p-values, respectively. See also results of linear mixed model effects, ***Table S2A***. (**B**) Data distribution and the correlation of m6A and MTase-sensitive patches (MSP) and computational nucleosome sizes. (**C**) Data distribution and the correlation of m6A, sequenced reads, and Fiber-seq Inferred Regulatory Elements (FIREs). To ensure normality, m6A and FIRE were transformed using the square-root function, and the number of reads were transformed using log. For the samples that failed to call FIRE, a value 0.00001 was given before the transformation. Lower panels, diagonal panels, and the upper panels show scatter plot, histogram, and pearson's correlations with p-values, respectively. The m6A, the proportion of methylated adenines. MSP, the average size (bp) of MTase-sensitive patches. MONO, average mono-nucleosome size (bp). DI, average di-nucleosome size (bp). FIRE, the number of Fiber-seq Inferred Regulatory Elements that were called. Reads, the number of sequenced reads. See also ***Table S2B.***

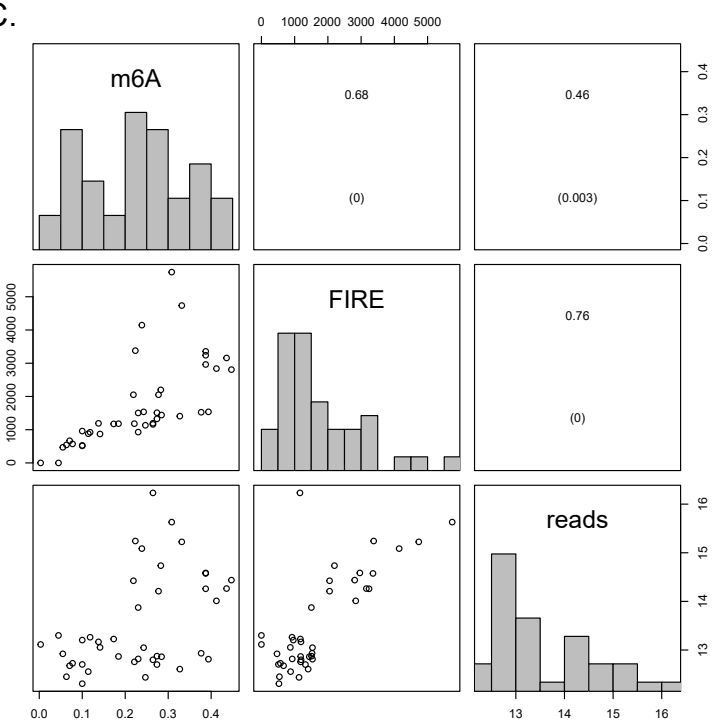
A.


PMI
NucleiPerHia5Unit
m6A
CCScoverage
−0.12
(0.47)
0.36
(0.019)
−0.53
(0)
−0.78
(0)
0.2
(0.21)
−0.48
(0.002)

B.

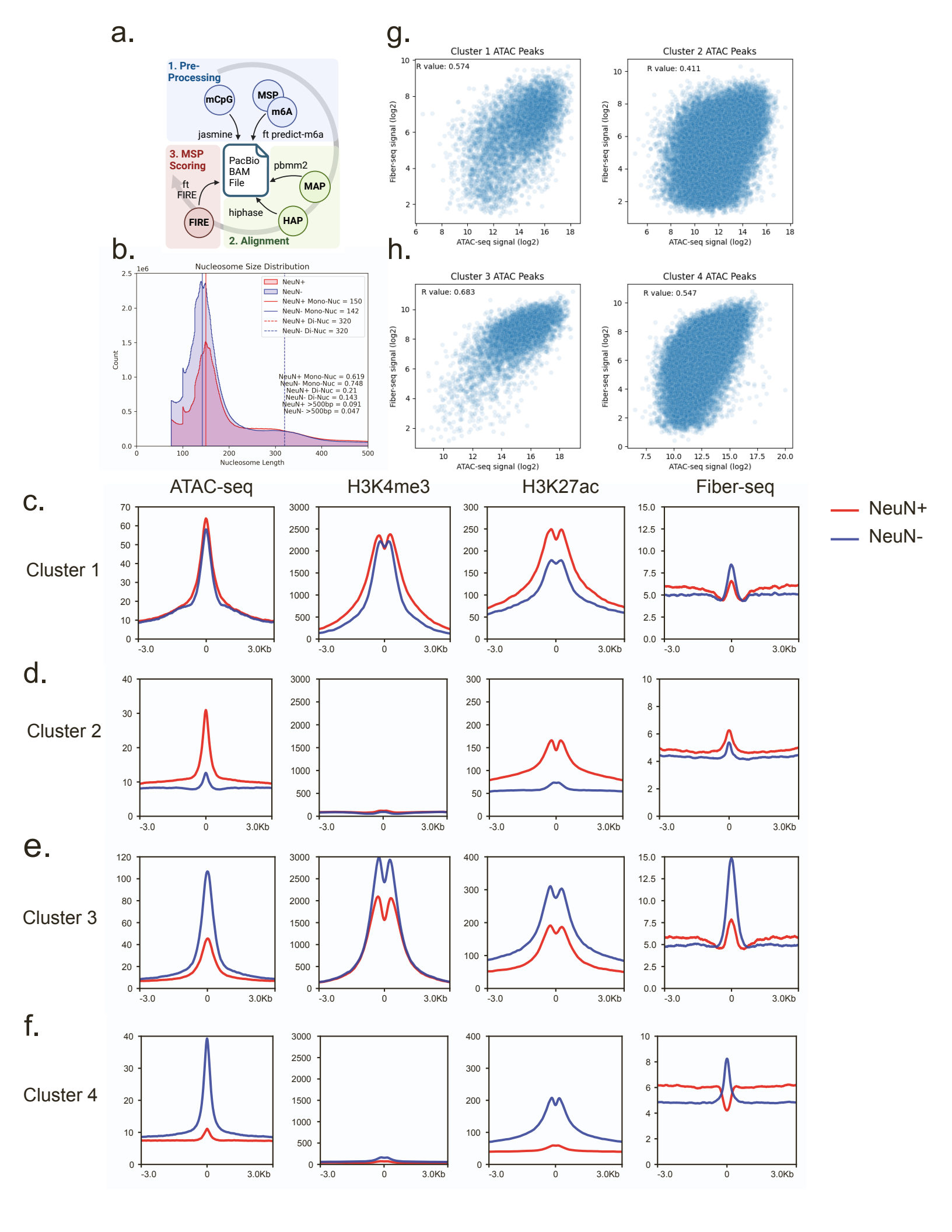

m6A
MSP
MONO
DI
0.8
(0)
−0.67
(0)
−0.95
(0)
−0.52
(0)
−0.82
(0)
0.69
(0)

C.


m6A
FIRE
reads
0.68
(0)
0.46
(0.003)
0.76
(0)

Supplemental Figure S4: Cell-type Specific Fiber-seq Data Processing and Correlation with ATAC-seq (related to Figure 2).

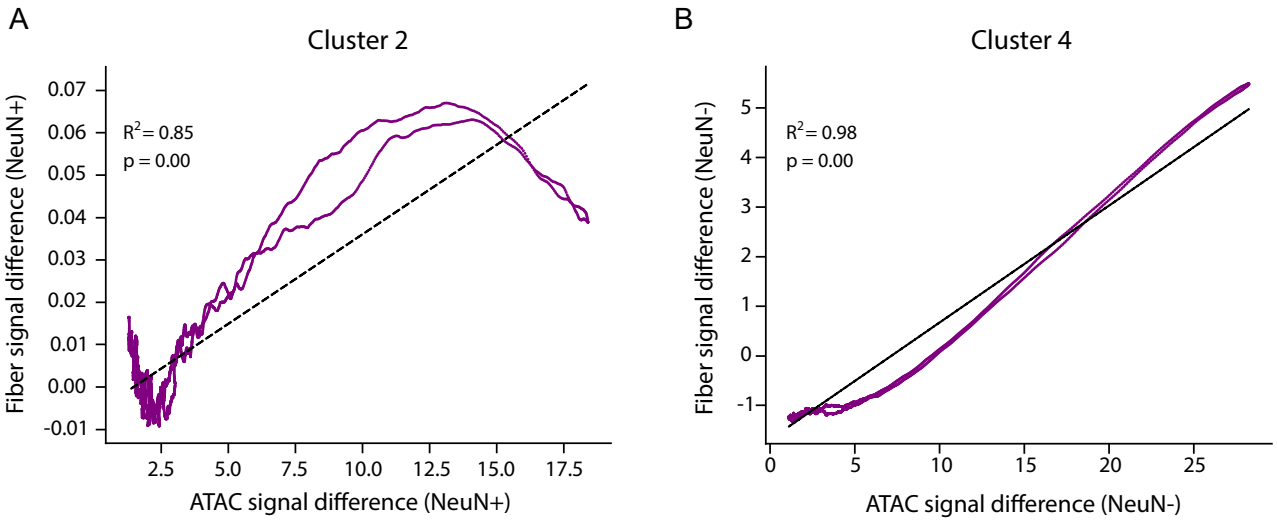
(a) Graphic depicting the standard Fiber-seq data processing workflow for PacBio Sequel II sequenced BAM files. See methods for more details on processing BAM files after sequencing. (**b**) Histogram showing the distribution of sizes of predicted nucleosomes in two deeply sequenced NeuN+ (red) and NeuN- (blue) Fiber-seq reference libraries. Median mono-nucleosome length is 150bp for NeuN+ and 142bp for NeuN-; median di-nucleosome or larger length is 320bp for both NeuN+ and NeuN-. (**c-f**) Profile plots showing the average signal enrichment for ATAC-seq, H3K4me3 and H3K27ac ChIP-seq, and Fiber-seq, from left to right for ATAC-seq peaks clusters 1 (c), 2 (d), 3 (e), and 4 (f). (**g-h**) Correlation scatter plots comparing log2 total Fiber-seq pseudobulk accessibility signal versus log2 total ATAC-seq signal within NeuN+ (g) and NeuN- (h) ATAC-seq peaks. (left) Promoter (right) enhancer peaks.


a.
1. Pre-Processing
mCpG
MSP
m6A
jasmine
ft predict-m6a
3. MSP Scoring
ft FIRE
FIRE
PacBio BAM File
pbmm2
MAP
hiphase
HAP
2. Alignment
b.
Nucleosome Size Distribution
1e6
NeuN+
NeuN-
NeuN+ Mono-Nuc = 150
NeuN- Mono-Nuc = 142
NeuN+ Di-Nuc = 320
NeuN- Di-Nuc = 320
NeuN+ Mono-Nuc = 0.619
NeuN- Mono-Nuc = 0.748
NeuN+ Di-Nuc = 0.21
NeuN- Di-Nuc = 0.143
NeuN+ >500bp = 0.091
NeuN- >500bp = 0.047
Count
Nucleosome Length
g.
Cluster 1 ATAC Peaks
R value: 0.574
Cluster 2 ATAC Peaks
R value: 0.411
Fiber-seq signal (log2)
ATAC-seq signal (log2)
h.
Cluster 3 ATAC Peaks
R value: 0.683
Cluster 4 ATAC Peaks
R value: 0.547
ATAC-seq
H3K4me3
H3K27ac
Fiber-seq
c.
Cluster 1
d.
Cluster 2
e.
Cluster 3
f.
Cluster 4
-3.0
0
3.0Kb
NeuN+
NeuN-

Supplementary Figure S5: NeuN+ versus NeuN- Fiber-seq peak differentials are tracking corresponding cell type specific difference in ATAC-seq dataset (related to Figure 2).

a. (NeuN+ enhancers) This is a scatter plot that shows the average difference in fiber-seq percent accessibility vs the average difference ATAC-seq counts at each position ±3kb of the ATAC-seq peak center from cluster 2 in main **Figure 2E**. NeuN+ enhancers shown in purple. The line of best fit (black) is calculated using a linear regression and has an R-squared value of 0.85 with a p-value of 0.

b. (NeuN- enhancers) This is a scatter plot that shows the average difference in fiber-seq percent accessibility vs the average difference ATAC-seq counts at each position ±3kb of the ATAC-seq peak center from cluster 4 in in main **Figure 2F**. The NeuN- enhancers are shown in purple. The line of best fit (black) is calculated using a linear regression and has an R-squared value of 0.98 with a p-value of 0.


A
Cluster 2
Fiber signal difference (NeuN+)
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
-0.01
R2 = 0.85
p = 0.00
2.5
5.0
7.5
10.0
12.5
15.0
17.5
ATAC signal difference (NeuN+)
B
Cluster 4
Fiber signal difference (NeuN-)
5
4
3
2
1
0
-1
R2 = 0.98
p = 0.00
0
5
10
15
20
25
ATAC signal difference (NeuN-)

Supplemental Figure S6: Fiber-seq Peak Detection Limits (related to Figure 3).

(**a-b**) Line plot showing the number of peaks called in Fiber-seq data with an FDR < 0.05 as a function of the proportion of accessible single fibers at the peak for our (**a**) NeuN- and (**b**) NeuN+ reference sets. Barplot at the top shows the cumulative peak count as a function of the percent accessibility, broken into 5% bins. (**c**) Heatmap showing the ATAC-seq signal at cluster 4 FIRE peaks in NeuN+ (left) and NeuN- (right). Scale has been adjusted to show signal at background levels. (**d**) Fiber-seq cluster-specific proportional representation for repeat- and non-repeat sequences (clusters 1-4 in Figure 3A). (**e**) Regulatory CTCF motif footprints with Fiber cluster-specific binding scores as indicated (***, P< 1e-16, Kruskal-Wallis), note binding sites sharply marked at NDR center. (**f**) Plots depicting the average proportion of adenines that are methylated in fibers centered at CTCF motif that are not in an MSP with no footprint (n=18,997 fibers). Notice minimal nucleosomal positioning effects, in sharp contrast to fibers with MSP over the CTCF motif (see main ***Figure 3G***).

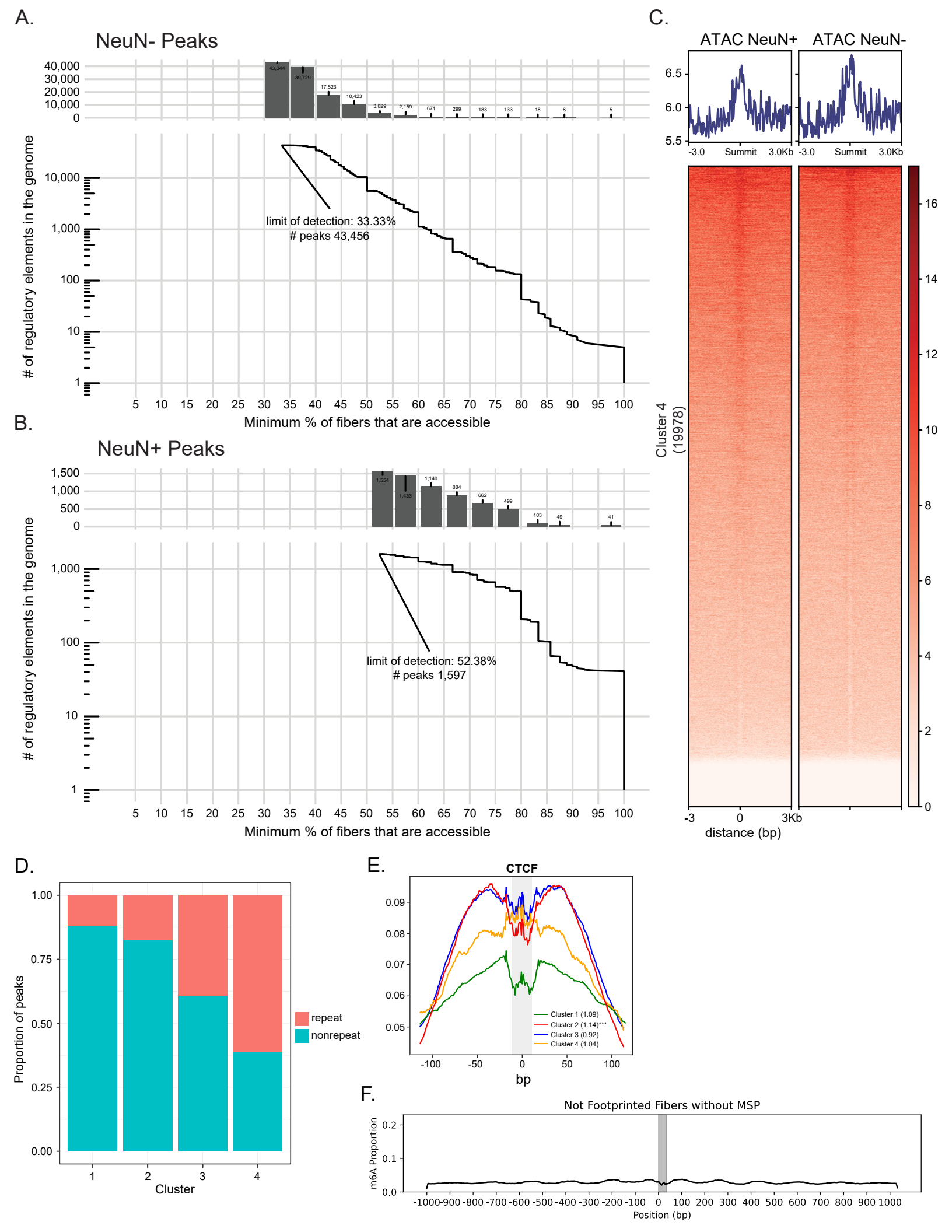
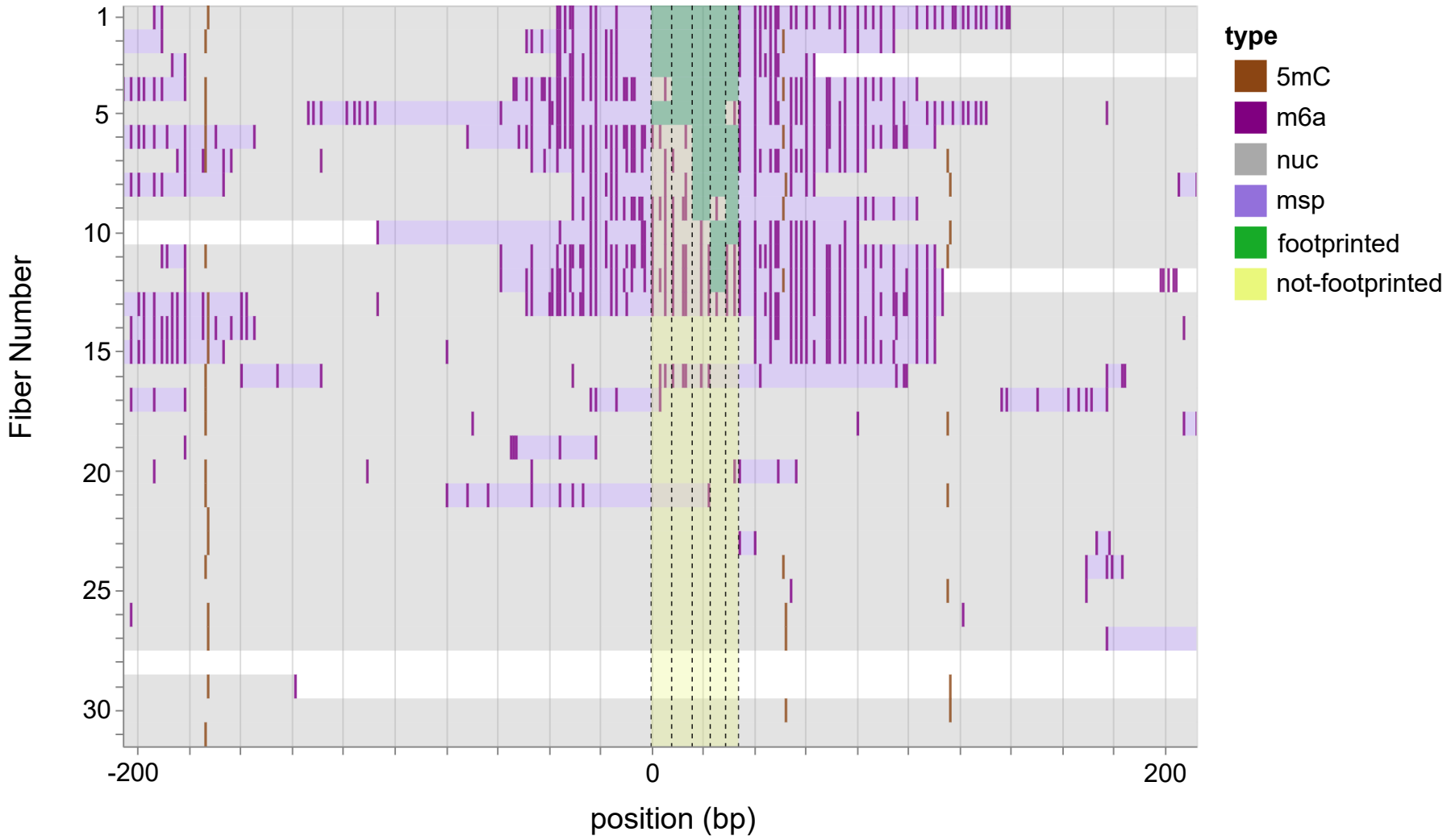
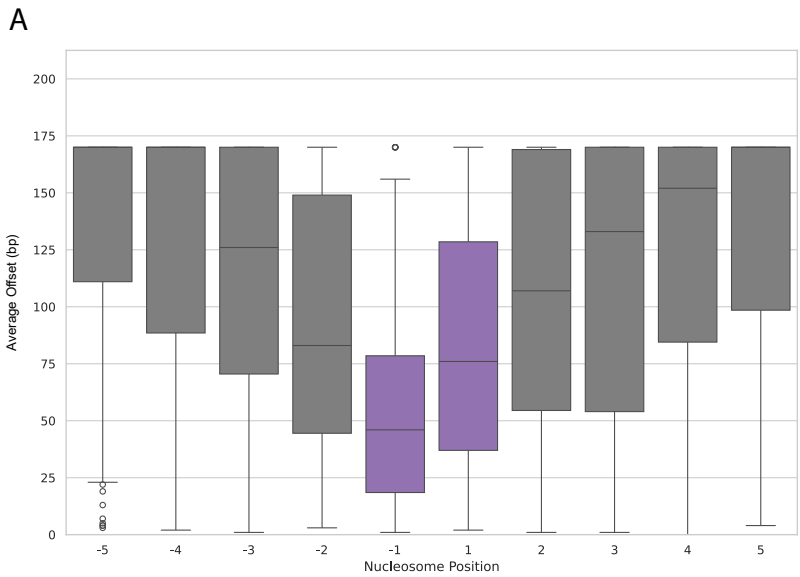

A.
NeuN- Peaks
43,344
39,729
17,523
10,423
3,829
2,159
671
299
183
133
18
8
5
# of regulatory elements in the genome
limit of detection: 33.33%
peaks 43,456
Minimum % of fibers that are accessible
B.
NeuN+ Peaks
1,554
1,433
1,140
884
662
499
103
49
41
# of regulatory elements in the genome
limit of detection: 52.38%
peaks 1,597
Minimum % of fibers that are accessible
C.
ATAC NeuN+
ATAC NeuN-
-3.0
Summit
3.0Kb
Cluster 4
(19978)
-3
0
3Kb
distance (bp)
D.
Proportion of peaks
repeat
nonrepeat
Cluster
E.
CTCF
Cluster 1 (1.09)
Cluster 2 (1.14)***
Cluster 3 (0.92)
Cluster 4 (1.04)
bp
F.
Not Footprinted Fibers without MSP
m6A Proportion
Position (bp)

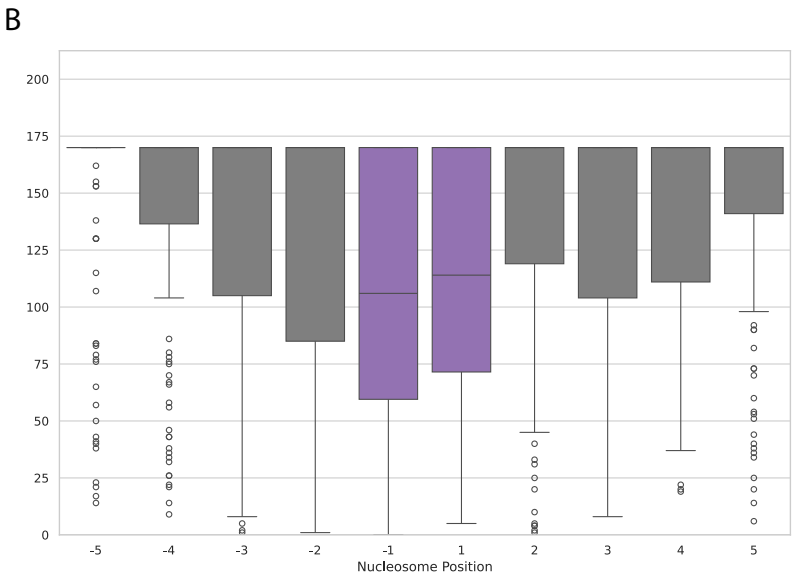
Figure S7: Fiber-specific resolution of a CTCF footprint at a single locus (related to Figure 3). This visualization shows the single fibers (n=31) that overlap a CTCF motif site found within a cluster 2 Fiber-seq peak, located at chr1:110064114-11006448, and centered at the start of the motif. Each row indicates a fiber that overlaps the locus, with the fibers numbered on the left vertical axis, and the genomic position relative to the motif start on the lower horizontal axis. On each fiber, the positions of m6A (dark purple) and mCpG (brown) are indicated with vertical lines. The fiber is divided into nucleosome (gray) and MSP (light purple) patches. The motif region, indicated by dotted lines, where each division represents a binding domain in the motif, is denoted as footprinted (green) or not footprinted (yellow), based on the presence of m6A marks in the motif region in that fiber.

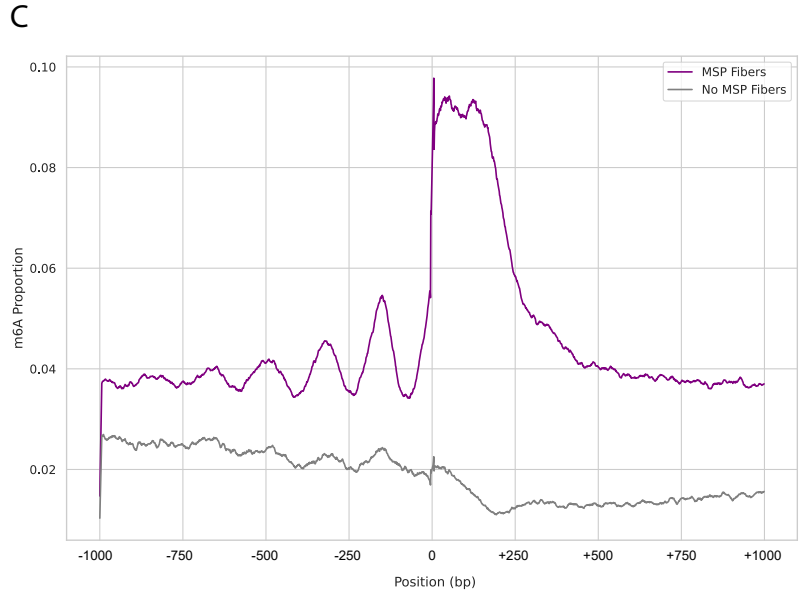

Fiber Number
1
5
10
15
20
25
30
-200
0
200
position (bp)
type
5mC
m6a
nuc
msp
footprinted
not-footprinted

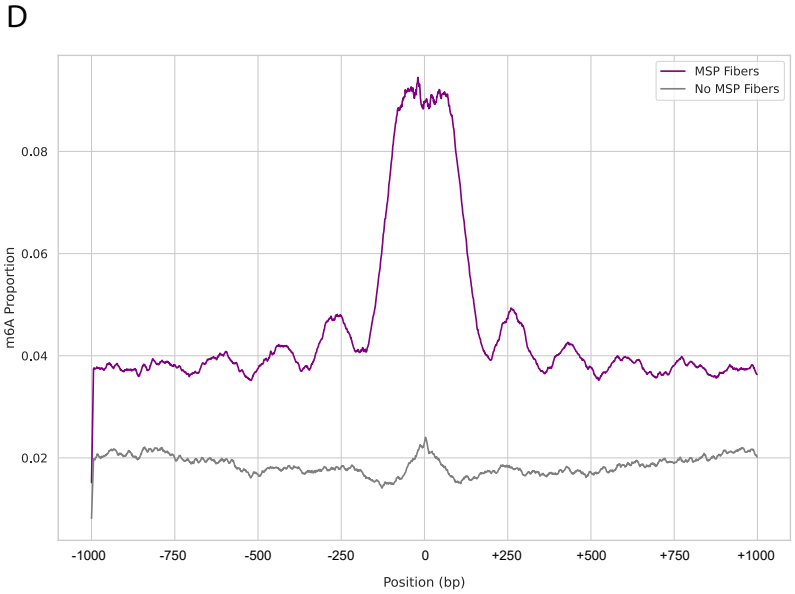
Supplemental Figure S8: **Transcription factor footprinting and nucleosome positioning at transcription start sites** (related to Figure 4).

(A) This series of box plots depicts the average nucleosome offset centered at the TSS for genes that have a FIRE peak overlapping their promoter (n=119) and have at least 20 fibers and 6 fibers per category. Only fibers that have an MSP overlapping the TSS were used. The -1 and +1 nucleosomes have a significantly different offset when compared to all other nucleosomes with p-values of 1.34×10^{-6} and 2.79×10^{-4} respectively. The -2 nucleosome shows significance before multiple-test correction (p=0.038) but does not pass the threshold post correction. (**B**) This series of box plots depicts the average nucleosome offset centered at the TSS for genes that have a FIRE peak overlapping their promoter. Only fibers that have do not have an MSP overlapping the TSS were used. The -1 and +1 nucleosomes do not a have significant different offset when compared to all other nucleosomes. (**C,D**) These plot depict the average proportion of adenines that are methylated (m6a prop) in fibers overlapping the TSS of genes that have a FIRE peak in their promoter, either (**C**) centered at the TSS or (**D**) centered at the FIRE peak. The fibers are split into those with an MSP overlapping the TSS (purple) and those without one (gray), and the average m6a prop at each position 1kb upstream and downstream is calculated for each class of fibers. The x-axis shows the position in base pairs, and the y-axis shows the average proportion of methylated adenines.


A
Average Offset (bp)
Nucleosome Position


B
Nucleosome Position

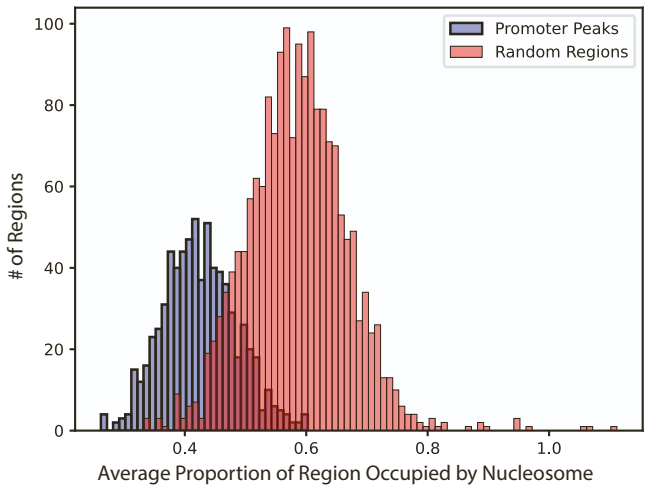

C
MSP Fibers
No MSP Fibers
m6A Proportion
Position (bp)


D
MSP Fibers
No MSP Fibers
m6A Proportion
Position (bp)

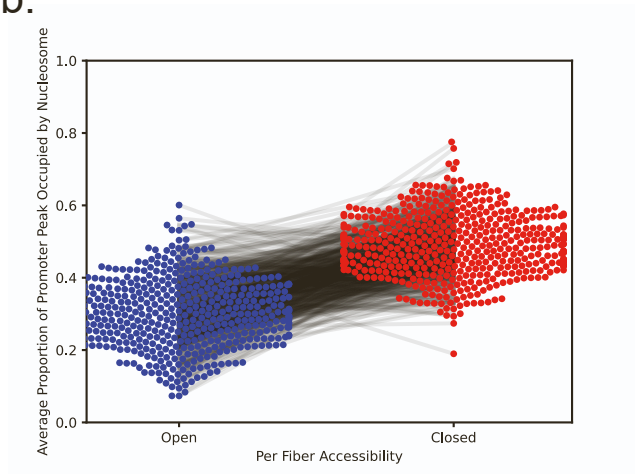
Supplemental Figure 9: Nucleosome sliding in single fibers and observed versus expected proportion of co-actuated peaks (related to Figure 4).

(**a**) Histogram showing the average single-fiber nucleosome encroachment into promoter FIRE peaks (blue) versus randomly selected internucleosomal linker regions (red) ($P < 1e\text{-}17$, Kruskal-Wallis Test). (**b**) Paired swarm plot showing the average single-fiber nucleosome encroachment for open (red) versus closed (blue) fibers at promoter FIRE peaks ($P < 1e\text{-}12$, Kruskal Wallis Test). **(c,d)** Plots shows the actual proportion of co-actuated fibers (green) compared to that peak's expected co-actuation proportion (orange) for (**c**) significantly co-actuated peaks ($p < 0.05$, Fisher Test) and (**d**) for non-significant co-actuated peaks ($p \geq 0.05$, Fisher Test). The x-axis denotes the two metrics that are computed for each peak and the y-axis depicts the proportion. The slope of the blue line indicates the average difference in the actual vs expected proportion for these peaks (**c**, m=0.149; **d,** m=0.036). P-value shown at figure bottom ($P= 1.64 \times 10^{-227}$) indicates significance of one-tailed t-test between the average differences of significant and non-significant peaks.

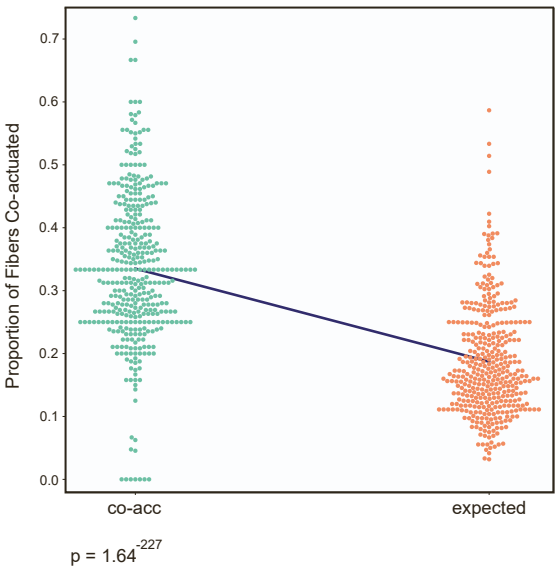
a.



b.



c.



$p = 1.64^{-227}$

d.