Figure S1

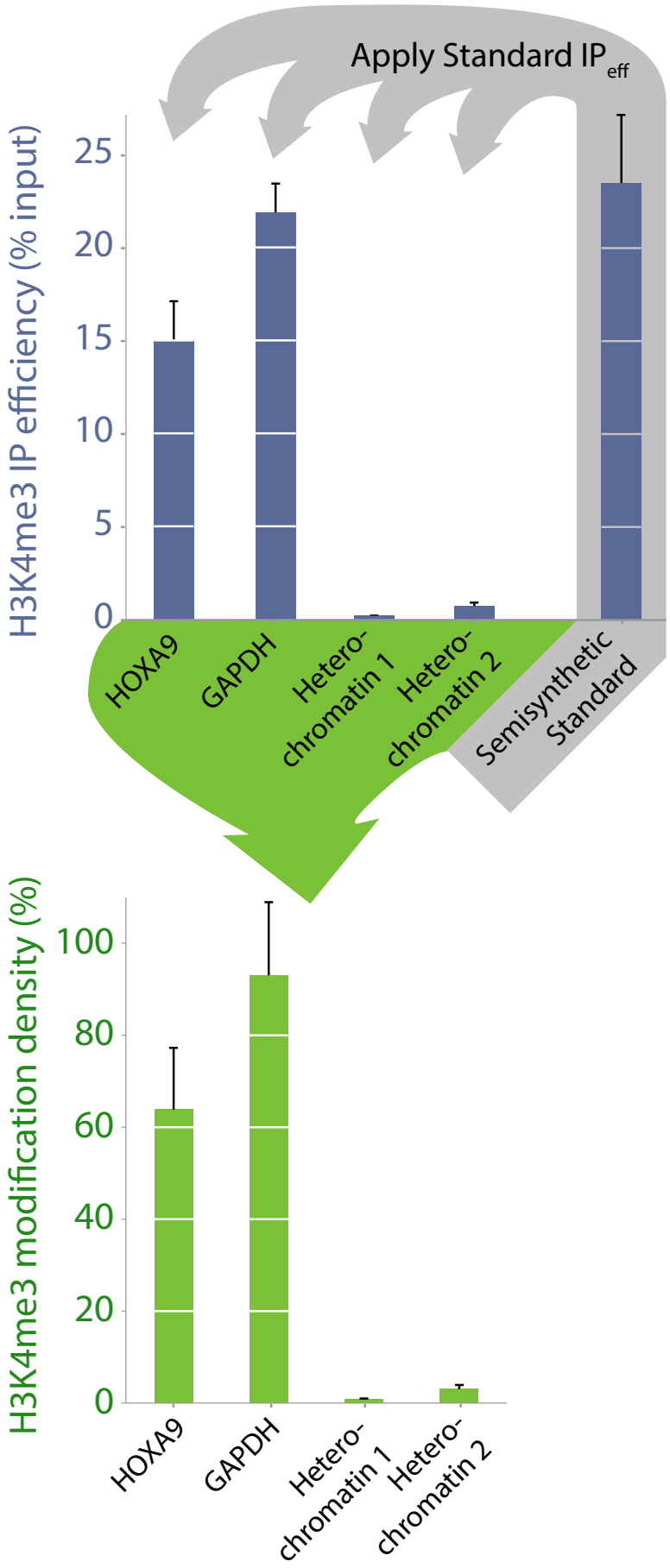
Apply Standard IP eff
H3K4me3 IP efficiency (% input)
25
20
15
10
5
0
HOXA9
GAPDH
Hetero-chromatin 1
Hetero-chromatin 2
Semisynthetic Standard
H3K4me3 modification density (%)
100
80
60
40
20
0
HOXA9
GAPDH
Hetero-chromatin 1
Hetero-chromatin 2

Figure S2

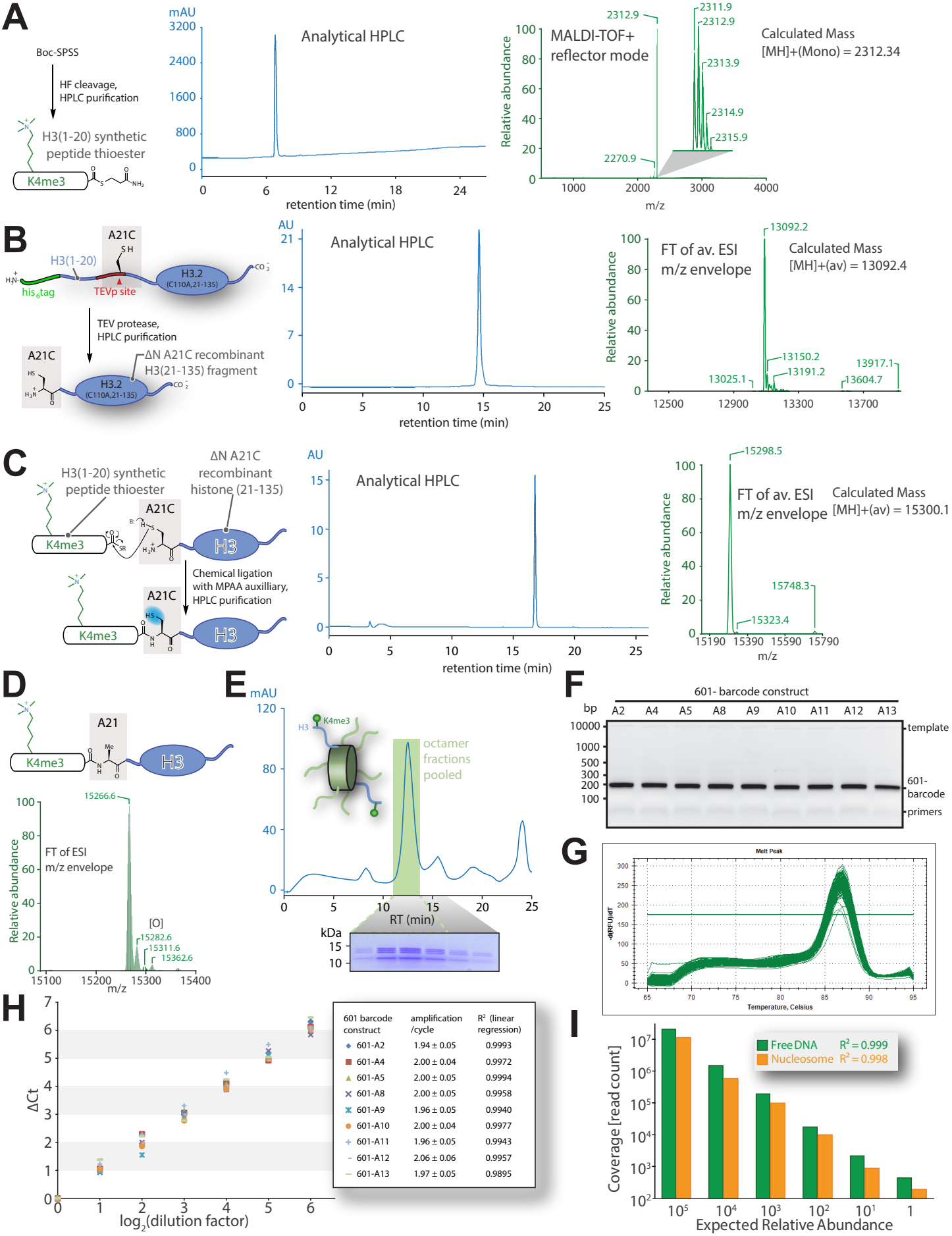
A
Boc-SPSS
HF cleavage,
HPLC purification
H3(1-20) synthetic
peptide thioester
K4me3
mAU
Analytical HPLC
retention time (min)
MALDI-TOF+
reflector mode
Relative abundance
m/z
2312.9
2311.9
2313.9
2314.9
2315.9
2270.9
Calculated Mass
[MH]+(Mono) = 2312.34
B
A21C
H3(1-20)
his6tag
TEVp site
H3.2
(C110A,21-135)
TEV protease,
HPLC purification
ΔN A21C recombinant
H3(21-135) fragment
AU
Analytical HPLC
retention time (min)
FT of av. ESI
m/z envelope
13092.2
13150.2
13191.2
13025.1
13917.1
13604.7
Calculated Mass
[MH]+(av) = 13092.4
C
H3(1-20) synthetic
peptide thioester
ΔN A21C
recombinant
histone (21-135)
H3
Chemical ligation
with MPAA auxilliary,
HPLC purification
AU
Analytical HPLC
retention time (min)
FT of av. ESI
m/z envelope
15298.5
15748.3
15323.4
Calculated Mass
[MH]+(av) = 15300.1
D
A21
FT of ESI
m/z envelope
15266.6
[O]
15282.6
15311.6
15362.6
E
mAU
K4me3
octamer
fractions
pooled
RT (min)
kDa
15
10
F
601- barcode construct
bp
A2
A4
A5
A8
A9
A10
A11
A12
A13
10000
1000
500
300
200
100
template
601-
barcode
primers
G
Melt Peak
-d(RFU)/dT
Temperature, Celsius
H
ΔCt
log2(dilution factor)
601 barcode construct | amplification /cycle | R² (linear regression)
601-A2 | 1.94 ± 0.05 | 0.9993
601-A4 | 2.00 ± 0.04 | 0.9972
601-A5 | 2.00 ± 0.05 | 0.9994
601-A8 | 2.00 ± 0.05 | 0.9958
601-A9 | 1.96 ± 0.05 | 0.9940
601-A10 | 2.00 ± 0.04 | 0.9977
601-A11 | 1.96 ± 0.05 | 0.9943
601-A12 | 2.06 ± 0.06 | 0.9957
601-A13 | 1.97 ± 0.05 | 0.9895
I
Free DNA R² = 0.999
Nucleosome R² = 0.998
Coverage [read count]
Expected Relative Abundance

Figure S3

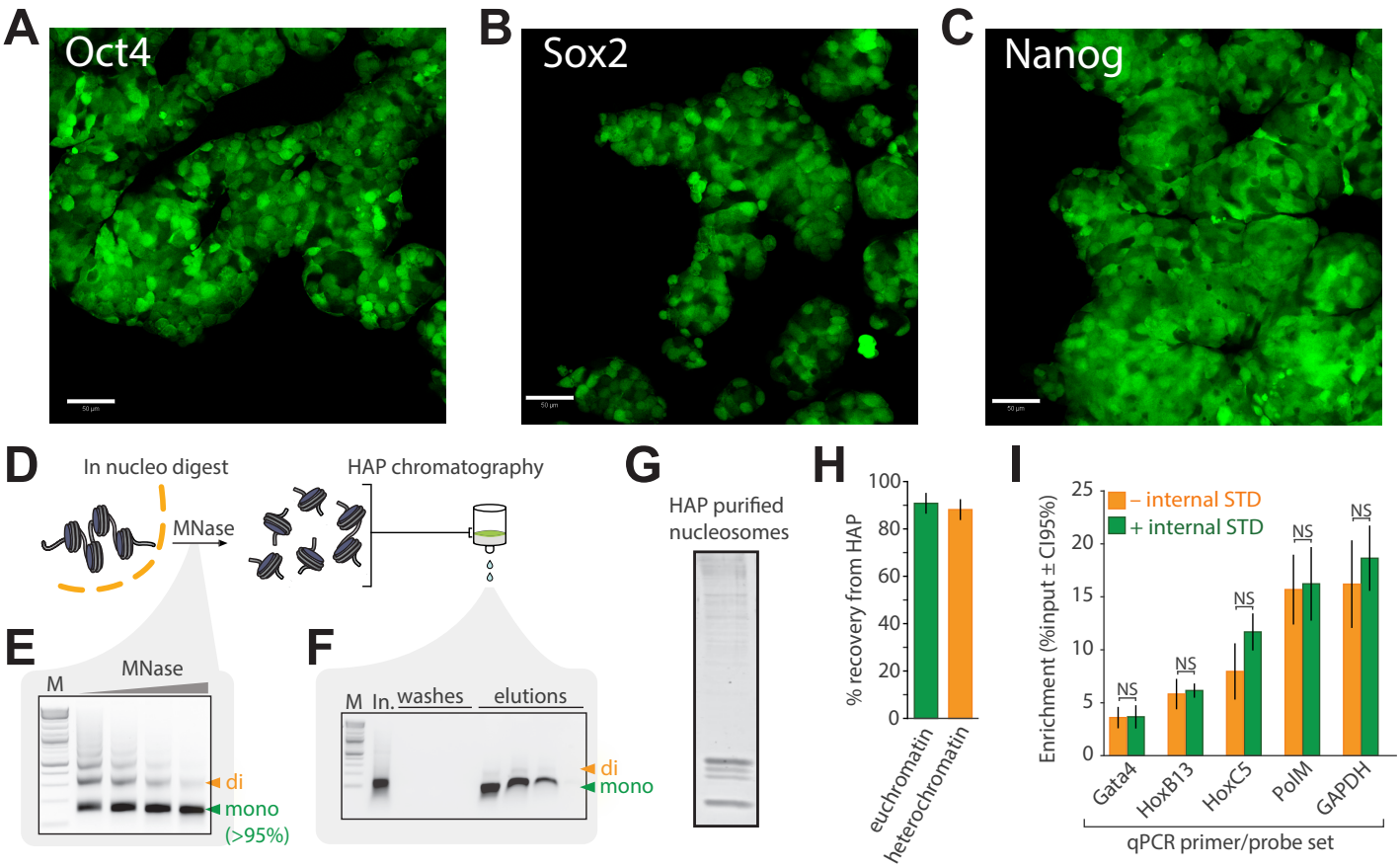
A
Oct4
B
Sox2
C
Nanog
D
In nucleo digest
MNase
HAP chromatography
E
M
MNase
di
mono
(>95%)
F
M In. washes elutions
di
mono
G
HAP purified
nucleosomes
H
% recovery from HAP
100
80
60
40
20
0
euchromatin
heterochromatin
I
Enrichment (%input ± CI95%)
25
20
15
10
5
0
– internal STD
+ internal STD
NS
Gata4
HoxB13
HoxC5
PolM
GAPDH
qPCR primer/probe set

Figure S4

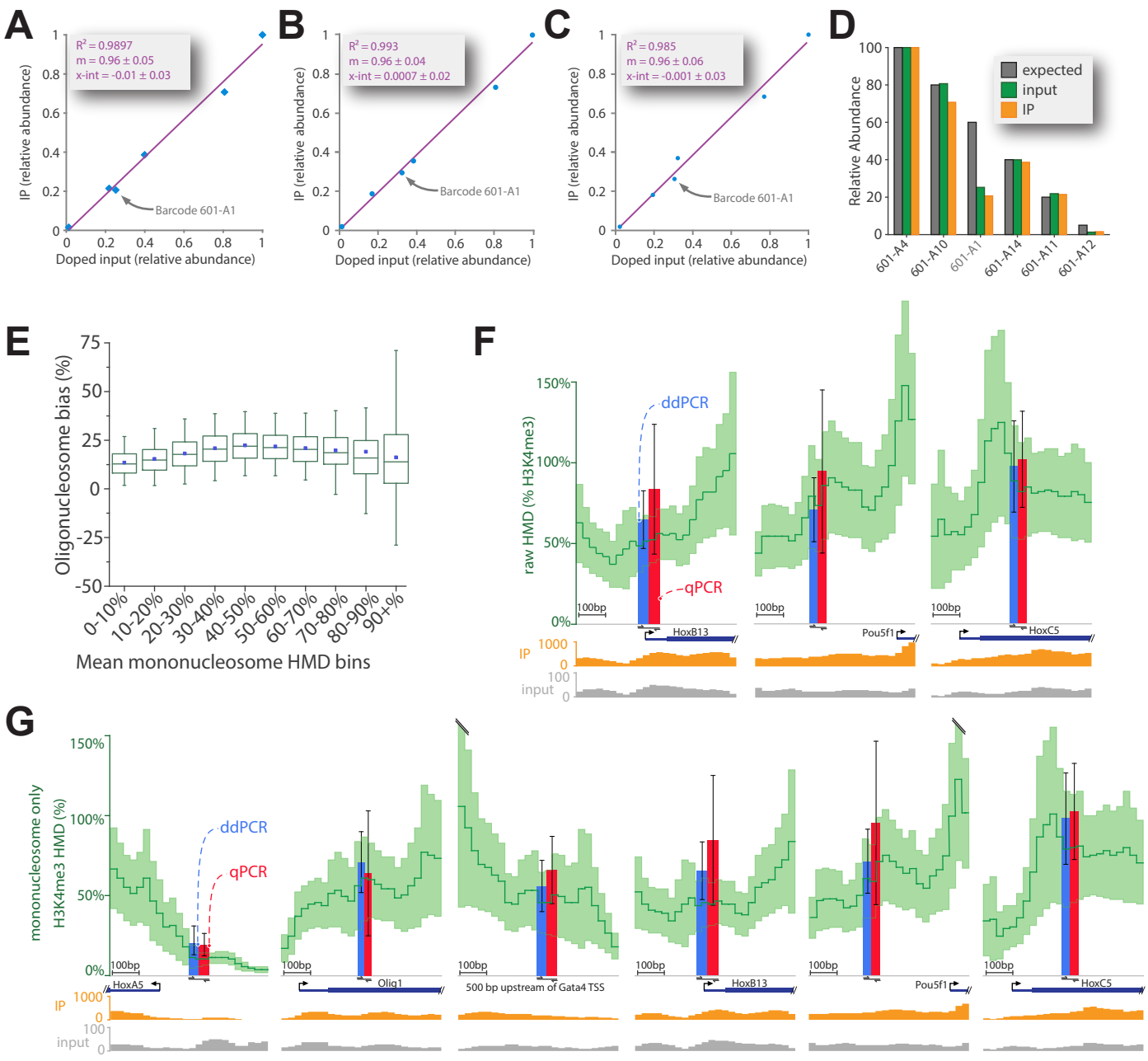
A
R² = 0.9897
m = 0.96 ± 0.05
x-int = -0.01 ± 0.03
IP (relative abundance)
Doped input (relative abundance)
Barcode 601-A1
B
R² = 0.993
m = 0.96 ± 0.04
x-int = 0.0007 ± 0.02
IP (relative abundance)
Doped input (relative abundance)
Barcode 601-A1
C
R² = 0.985
m = 0.96 ± 0.06
x-int = -0.001 ± 0.03
IP (relative abundance)
Doped input (relative abundance)
Barcode 601-A1
D
Relative Abundance
expected
input
IP
601-A4
601-A10
601-A1
601-A14
601-A11
601-A12
E
Oligonucleosome bias (%)
0-10%
10-20%
20-30%
30-40%
40-50%
50-60%
60-70%
70-80%
80-90%
90+%
Mean mononucleosome HMD bins
F
raw HMD (% H3K4me3)
ddPCR
qPCR
100bp
HoxB13
Pou5f1
HoxC5
IP
input
G
mononucleosome only H3K4me3 HMD (%)
ddPCR
qPCR
100bp
HoxA5
Olig1
500 bp upstream of Gata4 TSS
HoxB13
Pou5f1
HoxC5
IP
input

Figure S5

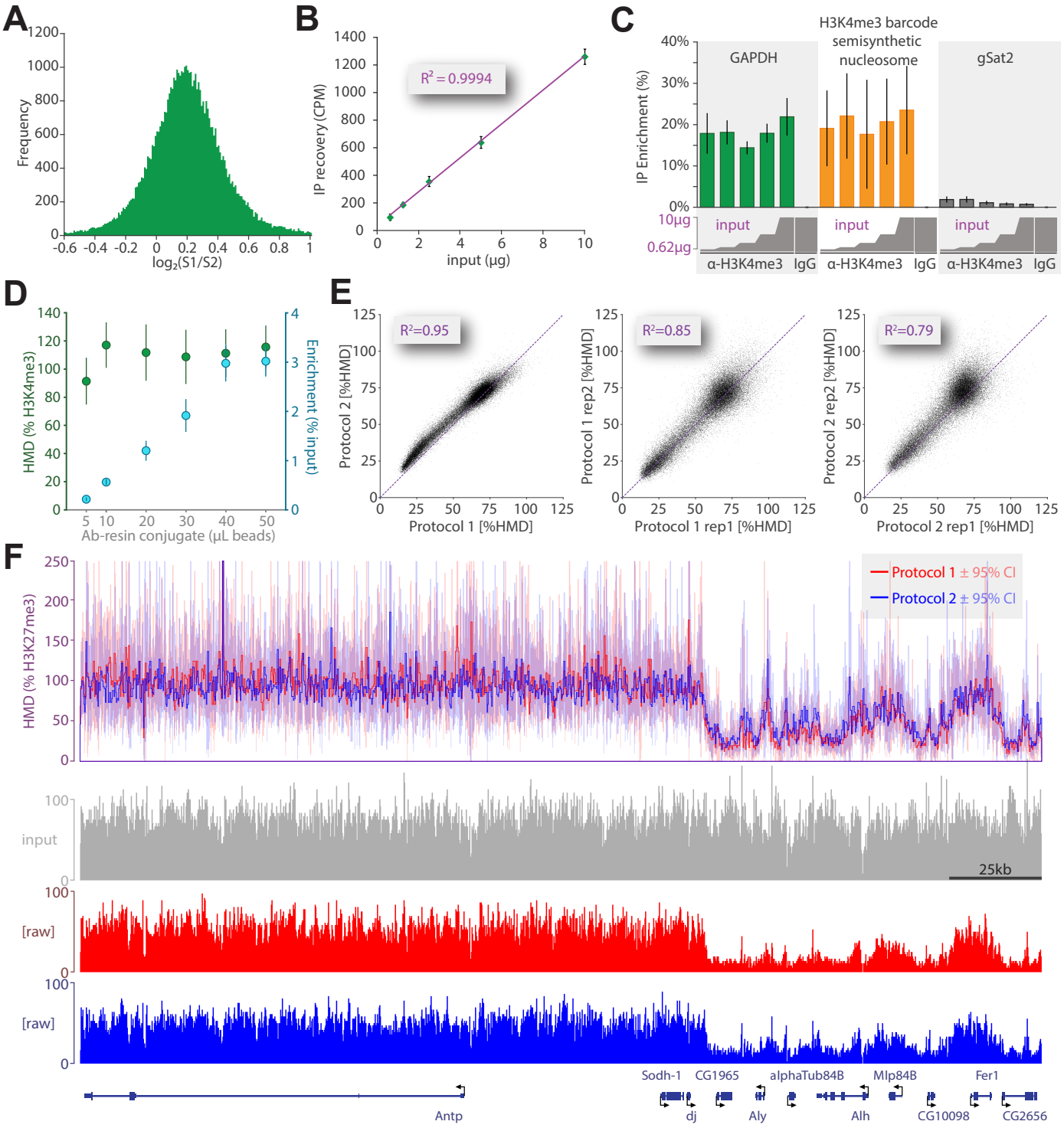
A
Frequency
log2(S1/S2)
B
IP recovery (CPM)
input (µg)
R² = 0.9994
C
H3K4me3 barcode semisynthetic nucleosome
GAPDH
gSat2
IP Enrichment (%)
10µg
0.62µg
input
α-H3K4me3
IgG
D
HMD (% H3K4me3)
Enrichment (% input)
Ab-resin conjugate (µL beads)
E
R²=0.95
R²=0.85
R²=0.79
Protocol 2 [%HMD]
Protocol 1 [%HMD]
Protocol 1 rep2 [%HMD]
Protocol 1 rep1 [%HMD]
Protocol 2 rep2 [%HMD]
Protocol 2 rep1 [%HMD]
F
HMD (% H3K27me3)
Protocol 1 ± 95% CI
Protocol 2 ± 95% CI
input
25kb
[raw]
Sodh-1
CG1965
alphaTub84B
Mlp84B
Fer1
Antp
dj
Aly
Alh
CG10098
CG2656

Figure S6

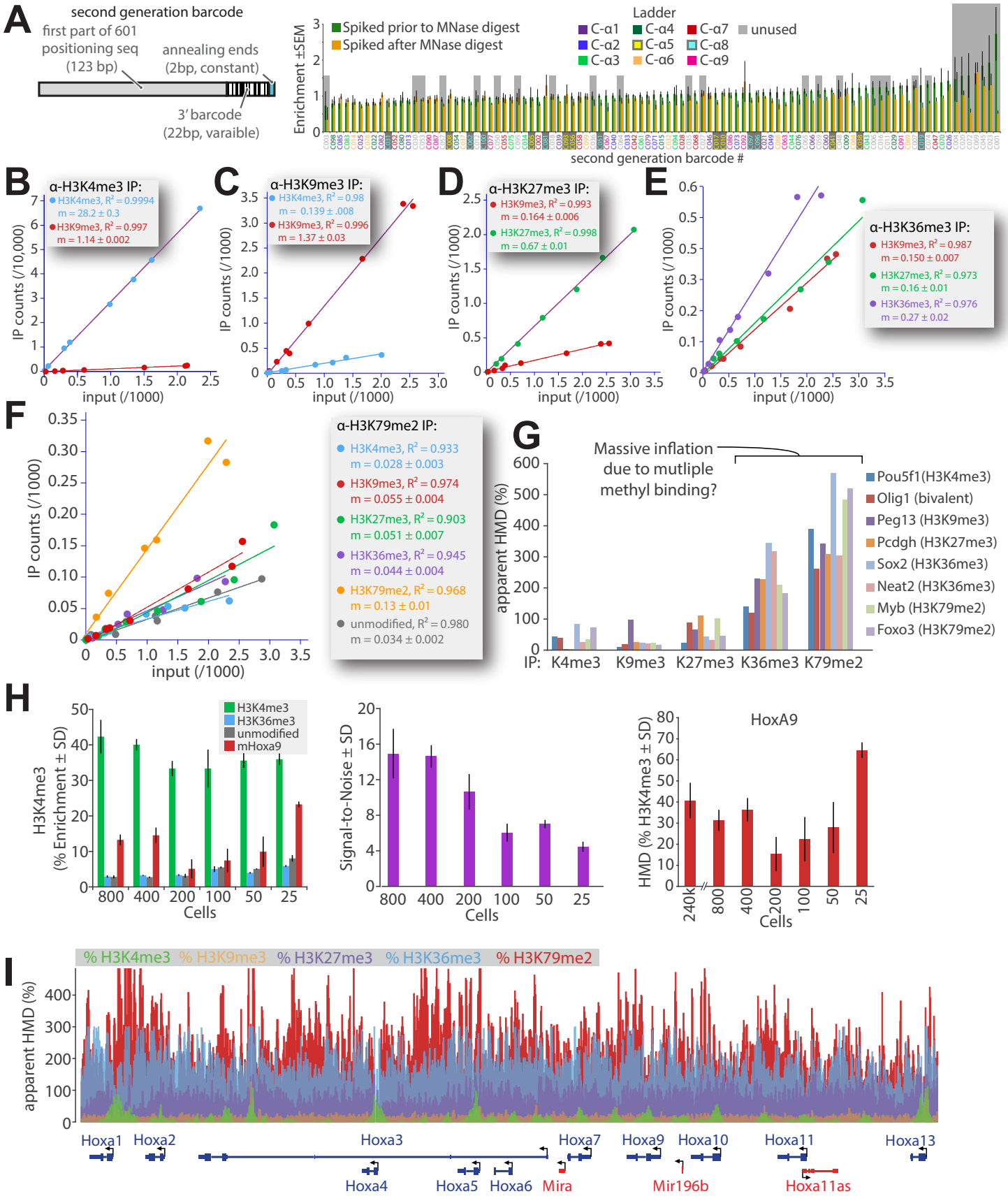
A
second generation barcode
first part of 601 positioning seq (123 bp)
annealing ends (2bp, constant)
3' barcode (22bp, varaible)
Spiked prior to MNase digest
Spiked after MNase digest
Ladder
C-α1
C-α2
C-α3
C-α4
C-α5
C-α6
C-α7
C-α8
C-α9
unused
Enrichment ±SEM
second generation barcode #
B
α-H3K4me3 IP:
H3K4me3, R² = 0.9994
m = 28.2 ± 0.3
H3K9me3, R² = 0.997
m = 1.14 ± 0.002
IP counts (/10,000)
input (/1000)
C
α-H3K9me3 IP:
H3K4me3, R² = 0.98
m = 0.139 ± .008
H3K9me3, R² = 0.996
m = 1.37 ± 0.03
IP counts (/1000)
input (/1000)
D
α-H3K27me3 IP:
H3K9me3, R² = 0.993
m = 0.164 ± 0.006
H3K27me3, R² = 0.998
m = 0.67 ± 0.01
IP counts (/1000)
input (/1000)
E
α-H3K36me3 IP:
H3K9me3, R² = 0.987
m = 0.150 ± 0.007
H3K27me3, R² = 0.973
m = 0.16 ± 0.01
H3K36me3, R² = 0.976
m = 0.27 ± 0.02
IP counts (/1000)
input (/1000)
F
α-H3K79me2 IP:
H3K4me3, R² = 0.933
m = 0.028 ± 0.003
H3K9me3, R² = 0.974
m = 0.055 ± 0.004
H3K27me3, R² = 0.903
m = 0.051 ± 0.007
H3K36me3, R² = 0.945
m = 0.044 ± 0.004
H3K79me2, R² = 0.968
m = 0.13 ± 0.01
unmodified, R² = 0.980
m = 0.034 ± 0.002
IP counts (/1000)
input (/1000)
G
Massive inflation due to mutliple methyl binding?
apparent HMD (%)
IP: K4me3 K9me3 K27me3 K36me3 K79me2
Pou5f1(H3K4me3)
Olig1 (bivalent)
Peg13 (H3K9me3)
Pcdgh (H3K27me3)
Sox2 (H3K36me3)
Neat2 (H3K36me3)
Myb (H3K79me2)
Foxo3 (H3K79me2)
H
H3K4me3
H3K36me3
unmodified
mHoxa9
H3K4me3 (% Enrichment ± SD)
Cells
Signal-to-Noise ± SD
HoxA9
HMD (% H3K4me3 ± SD)
I
% H3K4me3 % H3K9me3 % H3K27me3 % H3K36me3 % H3K79me2
apparent HMD (%)
Hoxa1 Hoxa2 Hoxa3 Hoxa4 Hoxa5 Hoxa6 Mira Hoxa7 Hoxa9 Hoxa10 Mir196b Hoxa11 Hoxa11as Hoxa13

Figure S7

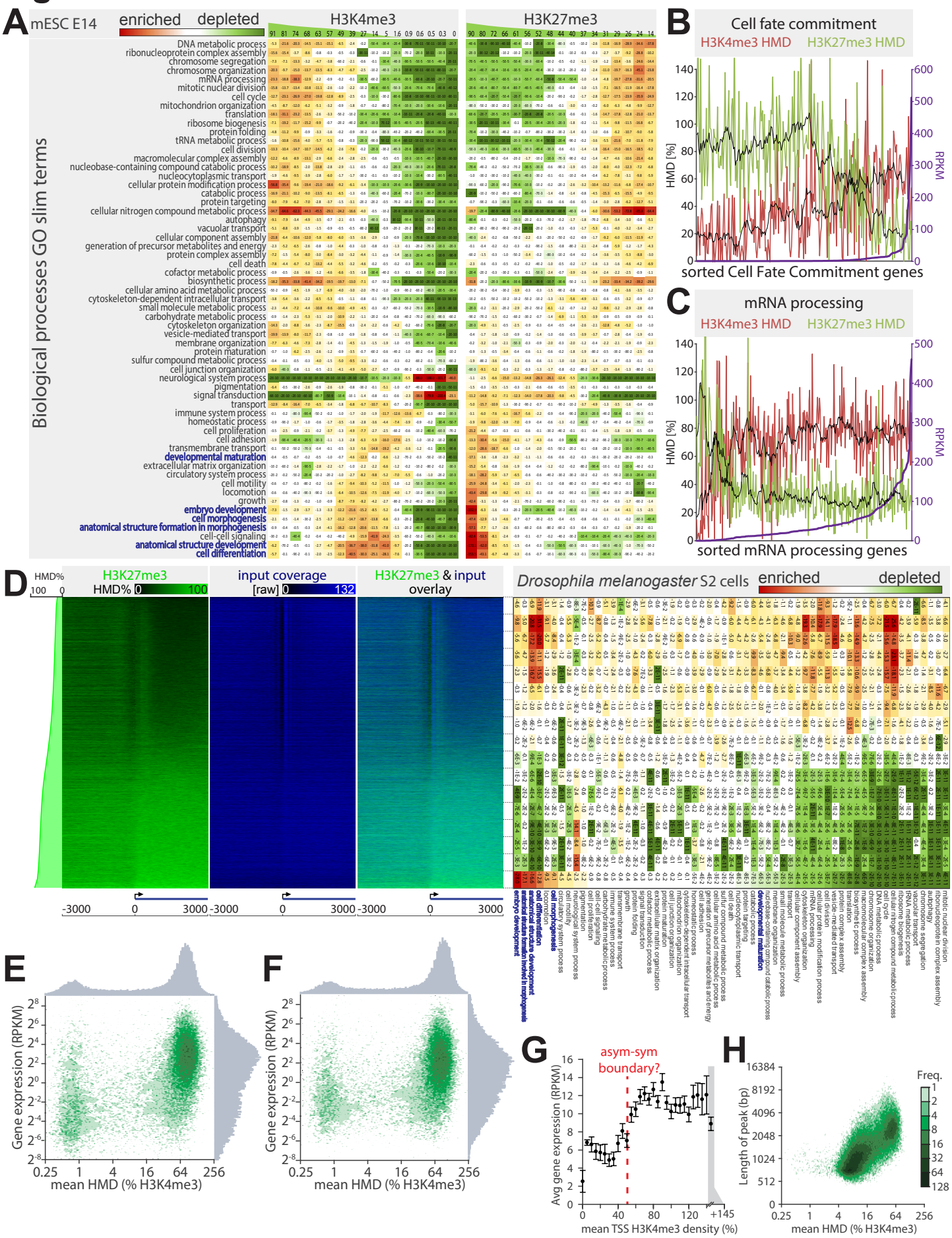
A
mESC E14
enriched
depleted
H3K4me3
H3K27me3
Biological processes GO slim terms
DNA metabolic process
ribonucleoprotein complex assembly
chromosome segregation
chromosome organization
mRNA processing
mitotic nuclear division
cell cycle
mitochondrion organization
translation
ribosome biogenesis
protein folding
tRNA metabolic process
cell division
macromolecular complex assembly
nucleobase-containing compound catabolic process
nucleocytoplasmic transport
cellular protein modification process
catabolic process
protein targeting
cellular nitrogen compound metabolic process
autophagy
vacuolar transport
cellular component assembly
generation of precursor metabolites and energy
protein complex assembly
cell death
cofactor metabolic process
biosynthetic process
cellular amino acid metabolic process
cytoskeleton-dependent intracellular transport
small molecule metabolic process
carbohydrate metabolic process
cytoskeleton organization
vesicle-mediated transport
membrane organization
protein maturation
sulfur compound metabolic process
cell junction organization
neurological system process
pigmentation
signal transduction
transport
immune system process
homeostatic process
cell proliferation
cell adhesion
transmembrane transport
developmental maturation
extracellular matrix organization
circulatory system process
cell motility
locomotion
growth
embryo development
cell morphogenesis
anatomical structure formation in morphogenesis
cell-cell signaling
anatomical structure development
cell differentiation
B
Cell fate commitment
H3K4me3 HMD
H3K27me3 HMD
HMD [%]
RPKM
sorted Cell Fate Commitment genes
C
mRNA processing
H3K4me3 HMD
H3K27me3 HMD
HMD [%]
RPKM
sorted mRNA processing genes
D
HMD%
H3K27me3
HMD%
input coverage
[raw]
H3K27me3 & input
overlay
-3000
0
3000
Drosophila melanogaster S2 cells
enriched
depleted
E
Gene expression (RPKM)
mean HMD (% H3K4me3)
F
Gene expression (RPKM)
mean HMD (% H3K4me3)
G
asym-sym
boundary?
Avg gene expression (RPKM)
mean TSS H3K4me3 density (%)
H
Length of peak (bp)
Freq.
mean HMD (% H3K4me3)

Figure S1. Pilot ICeChIP experiment using qPCR, related to Figure 1 Schematic of the ICeChIP internal standard approach applied to ChIP with qPCR readout probed with primer sets for the HOXA9, GAPDH and two heterochromatin regions. Briefly, the mean IP efficiency of the of internal standards spiked into the native ChIP experiment is taken to represent maximal modification so that the ratio of IP efficiencies at a genomic locus versus the internal standard, converted to a percentage is used to compute H3K4me3 modification density over the span of amplicons.

Figure S2. Preparation of barcoded semisynthetic nucleosomes, related to Figure 2 (A) Characterization of the 2-mercaptopropionamide peptide thioester by analytical HPLC (middle panel) and MALDI-TOF+ in reflector mode (right panel) of the purified product. **(B)** Characterization of the ΔN A21C recombinant H3(21-135) fragment by analytical HPLC (middle panel) and ESI MS (right panel) of the purified product. **(C)** Characterization of the H3K4me3 A21C species after the native chemical ligation by analytical HPLC (middle panel) and ESI MS (right panel) of the purified product. **(D)** Characterization of semisynthetic H3K4me3 after radical initiated desulfurization by Fourier transform (FT) of the average mass to charge ion envelope of the ESI+ MS of the purified product reveals approximate average mass with an oxidation, most likely a DTT-reversible adduct to methionine, as a minor impurity. **(E)** Chromatogram of the octamer purification on a gel filtration column (Superdex 200 10/300GL), with Coomassie stained SDS-PAGE of the octamer and proximal fractions. **(F)** Gel analysis of the set of barcoded constructs selected for consistent single-species amplification was observed. **(G)** Product melt curve and **(H)** plot of quantitative PCR (qPCR) data via SYBR-green fluorescence of a dilution series for each of the library members quantitatively demonstrating even 2-fold amplification per PCR cycle in triplicate. The linearity of fits to these curves (expressed as R^2 for the linear regression) and calculated per-cycle amplification efficiencies are given in the accompanying table. Figure 1G is based on the end-product melting of these barcoded constructs. **(I)** A mixture of barcoded nucleosome positioning sequences, where each barcode was added in relative amount ranging between 1 to 100000 with 10 fold increment between barcodes was used to prepare an Illumina sequencing library and subjected to SE50 sequencing. The same mixture was used in a single tube to prepare a ladder of H3K4me3 semisynthetic nucleosomes, which upon purification was then subjected to DNA isolation, Illumina library preparation, and sequencing. In this experiment, we also confirm that the Illumina 2500 HiSeq sequencing platform has at least 5 orders of magnitude of linear range.

Figure S3. Histone Modification Density of mouse and human HOXA gene cluster, related to Figure 3 Immunocytochemistry staining of mESCs (E14 cell line) with canonical markers of ES cells for **(A)** Oct4, **(B)** Sox2, and **(C)** Nanog, white scale bar is

50 µm. **(D)** Schematic representation of *in nucleo* digestion and hydroxyapatite chromatography. **(E)** In nucleo digestion was optimized to yield 95% pure mononucleosomes, where 80% of total chromatin is in soluble fraction. Purified DNA resolved with 2% agarose gel electrophoresis stained with ethidium bromide, M is NEB 2-log DNA ladder. **(F)** Nucleosomes were quantitatively recovered from HAP chromatography column. Purified DNA resolved with 2% agarose gel electrophoresis stained with ethidium bromide, M is NEB 2-log DNA Ladder. **(G)** Representative purity of HAP purified nucleosomes used in ICeChIP resolved by SDS-PAGE then stained with SYPRO Ruby. **(H)** No significant difference in recovery from HAP was observed between euchromatin and heterochromatin loci. Measured with SYBR qPCR with primer set HOXA9 for euchromatin and primer set for a heterochromatin region of human chromosome 8, 26.6 kb downstream 3' flanking region of the *myc* gene. Error bars are 95%CI. **(I)** The addition of H3K4me3 internal standard ladder representing ~0.0001% of the total chromatin input is indistinguishable from the corresponding undoped “classical” ChIP-qPCR experiment. A comparison of H3K4me3 enrichment in ChIP-qPCR for spiked and non-spiked samples performed in triplicate with three technical replicates is presented ±95%CI.

Figure S4. A Critical examination of ICeChIP, related to Figure 4 (A) The relative abundance of barcode tags normalized to the most abundant ladder member measured in IP and input from HEK293 H3K4me3 ICeChIP-seq. **(B, C)** The relative abundance of barcode tags normalized to the most abundant ladder member measured in IP and input from E14 ICeChIP-seq, samples 1 and 2, respectively. **(D)** The observed abundance of each barcode closely followed expected value, except for barcode 601-A1, which displayed nearly equivalent IP enrichment, but repeatedly failed to amplify and is removed from all other analyses. **(E)** Oligonucleosome bias for binned mean HMD averaged over peaks does not show significant difference between bins. Measured with H3K4me3 (AM39159) ICeChIP-seq. The box represents 25^{th} to 75^{th} percentile, whiskers represent 5^{th} to 95^{th} percentile, horizontal line indicates the median, and a small blue square indicates the mean. **(F)** ICeChIP-seq compared to ddPCR and qPCR: the dark green line represents uncorrected H3K4me3 Histone Modification Density (HMD) ± 95%CI (light green) in the mESC E14 cell line as a function of chromosomal coordinate for three additional 600bp windows. signal was averaged over 25bp non-overlapping bins. Blue and red bars represent H3K4me3 measured by ddPCR and qPCR respectively on the same HMD scale (error bars are 95%CI), positioned over the indicated amplicon. **(G)** The same six 600 bp windows of the genome as Figure 4B and S4F, but with mononucleosome only derived data (correcting for oligonucleosome avidity bias).

Figure S5. The reproducibility and robustness of ICeChIP, related to Figure 5 (A) Frequency plot of the two replicates over MACS2-called peaks ($p<10^{-20}$) on the two biological replicates (S1 and S2) of ICeChIP-seq from E14 cells. This analysis suggests a systematic variation between sample 1 and 2 to be a modest 14.7% on average. **(B)** Amount of IP recovery of H3K4me3 (AM39159) pull-down has changed linearly with increasing amount of input (0.625-10 µg). IP recovery was measured with increasing amounts of radiolabeled native nucleosomes as input and amount of total IP elution was measured by scintillation counting. Error bars are 95%CI. **(C)** Enrichment of H3K4me3 pull-down is independent of input in the examined range (0.625-10 µg, 2 fold concentration series) as measured by qPCR with primer sets for hGAPDH (green, euchromatin), a synthetic DNA sequence common to the H3K4me3 ladder nucleosomes doped into the native pool of nucleosomes (orange), and gamma satellite repeat 2 (grey, heterochromatin), IP enrichment is calculated as percent of loaded input. This is the raw data used to compute figure 5D. **(D)** Plot of HMD with a HOXA9 primer set computed from ICeChIP-qPCR for a variety of experimental conditions. Note the ladder IP-enrichment is highly variable, whereas the HMD measurement is relatively constant. (**E)** Comparison of two H3K27me3-directed ICeChIP experiments from Drosophila S2 cells, staged from the same input but with great variation in IP and washes. Sample 1 data was generated using our standard ICeChIP conditions (15 minute incubation of resin-Ab conjugate with input, followed by five washes over 50 minutes) whereas the Sample 2 IP was performed with a shorter incubation and flow washes of the resin with the same volumes applied over the span of one minute. Each data point corresponds to mean H3K27me3 averaged over 3000bp non-overlapping window (N=41158); windows with insufficient input depth were excluded from analysis (cut-off>5). Leftmost panel was performed in technical triplicate for each protocol (independent IPs and measurements). **(F)** The superposition of HMD derived for two abovementioned protocols (blue and red) of the H3K27me3 ICeChIP-seq in S2 *Drosophila melanogaster* cell line displayed at the Antennapedia complex. Raw tag count for input and each IP is shown below, each track was measured in triplicate. Error is 95%CI. Signal was averaged over 25bp non-overlapping bins.

Figure S6. ICeChIP with multiple internal standards, related to Figure 6 (A) Second generation barcoded DNA templates—designed to be largely MNase resistant—were reconstituted into H3K4me3 nucleosomes, doped into a mESC native chromatin, prior to or after MNase digestion, and subjected to ChIP with H3K4me3 antibody (AM39159). The recovered DNA was subjected to qPCR and each barcode compared in the input and IP under the two conditions. From this set, we composed nine ladders (8 members each), and discarded the subset of sequences that behaved poorly in ChIP or were especially susceptible to MNase digestion. **(B)** Calculation of IP-enrichment in multi-standard ICeChIP experiments from mESCs presented as raw ladder read counts in the

IP versus input for the on-target mark, as well as the highest off-target background ladders for H3K4me3 (Active Motif AM39159), **(C)** H3K9me3 (M309M3-A (Hattori et al., 2013)), **(D)** H3K27me3 (Millipore 07-449), **(E)** H3K36me3 (Abcam AB9050), **(F)** H3K79me2 (Abcam AB3594). **(G)** qPCR analysis of loci implicated as strongly enriched for one or more of these marks from multi-standard ICeChIPs from mESCs. Note the massive inflation of apparent HMD at nearly all loci that results from off-target and non-specific binding by the H3K36me3 and H3K79me2 antibodies as compared to physically plausible values for the other three antibodies - converging arrows in Figure 6D denote qPCR amplicon. **(H)** Chromatin input titration for a small scale ICeChIP experiment as presented in Figure 6A. The method works well down to the chromatin equivalent of 400 cells. **(I)** Overlaid ICeChIP tracks demonstrating the comparative power of using the same % HMD scale, as well as the danger of poor antibodies inflating these values (for H3K36me3 and H3K79me2); The HMD peaks are not corrected for off-target recognition as they are in Figure 6C-D. Tracks represent 1bp non-overlapping bins.

Figure S7. ICeChIP affords insight histone mark distributions and function, related to Figure 7 (A) Generic biological processes GO slim enrichment analysis of mESC E14 cell line shows enrichment and depletion of genes classes correlated with certain H3K4me3 HMD and H3K27me3 HMD abundances. Enrichment analysis was performed on equal-sized fractions of H3K4me3 HMD averaged over 400bp downstream TSS and H3K27me3 averaged over gene bodies, Go terms were sorted by H3K27me3 HMD enrichment. Numbers in each cell are log10 of p-value. GO terms that are developmental processes are highlighted in purple. Go slim categories were developed by GO Consortium (Ashburner et al., 2000). Correlation between H3K4me3 HMD <400bp downstream TSS>, H3K27me3 HMD <400bp downstream TSS> and RPKM for all genes presented in Figure F7C belonging to cell fate commitment **(B),** and mRNA processing **(C)**. **(D)** Heatmaps of H3K27me3 HMD and input depth calculated for 3000bp around mappable TSS (N = 16319, where input depth ≥ 5) of *Drosophila melanogaster* S2 cell line. Plot on the left shows average H3K27me3 HMD over gene bodies, tracks were sorted by H3K27me3 HMD. To the right of heatmaps is generic biological processes GO slim enrichment analysis of classes correlated with H3K27me3 HMD abundances. Enrichment analysis was performed on equal-sized fractions of H3K27me3 averaged over gene bodies; GO terms were sorted by H3K27me3 HMD enrichment in the highest H3K27me3 HMD abundance class. Numbers in each cell are $\log_{10}$ of p-value. GO terms that are developmental processes are highlighted in purple. GO slim categories were developed by GO Consortium (Ashburner et al., 2000). The black trend line is a moving average window over 10 genes; genes are sorted by gene expression. **(E)** Heatmap of RNA-seq from E14 mESCs (Xiao et al., 2012) as a function of the mean H3K4me3 HMD averaged over TSS ± 200bp for the corresponding genes. **(F)** Heatmap of gene expression of the mESCs E14 cell line using expression data from

(Xu et al., 2011) as the function of the mean H3K4me3 HMD averaged over TSS ± 200bp, reveals very similar distribution to Figure S7F. **(G)** Binned average gene expression RPKM ± SEM for mESC E14 cell line using Xu et al. expression dataset as the function of binned mean H3K4me3 HMD averaged over TSS ± 200bp. **(H)** Heatmap of length of MACS peak ($p < 1e-20$) as the function of mean H3K4me3 HMD averaged over MACS peak performed for mESC E14 cell line.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Semi-synthetic Histone Preparation

Human histone H3.2(C110A)K4me3 was made by semi-synthesis (He et al., 2003; Ruthenburg et al., 2011; Shogren-Knaak et al., 2003), but distinct in one critical respect—the ligation junction is scarless following a desulfurization step (Chiang et al., 2009; Wan and Danishefsky, 2007)—the resulting histone is identical to the native modified histone save for the C110A mutation that is frequently made to ease of handling in recombinant histone. The sequence corresponding to residues 1-20 of histone 3, bearing the K4me3 modification was synthesized as peptide thioester by Boc-chemistry SPPS on S-trityl-β-mercaptopronionyl-p-methyl-benzhydrylamine resin (Nova Biochem) (Alewood et al., 1997). Resin was swelled for 1 hour with DMF and subsequently deprotected by washing it three times for three minutes with 95% TFA, 2.5% triisopropylsilane, and 2.5% H_2O. All amino acid couplings were performed with 4 molar equivalents of Boc-protected amino acid, 3.9 molar equivalents of HBTU and 6 molar equivalents of DIPEA incubated with resin for 10 minutes under nitrogen agitation. Following coupling, the resin was washed three times with DMF (with exception of glutamine where DCM was used instead), and Boc deprotection effected with three washes of TFA, where the first one is a flow wash. After last amino acid deprotection, the resin was washed sequentially with DMF, DCM and methanol. HF cleavage, ether precipitation, and lyophilization was performed at Midwest bio-tech (Fishers, IN).

Boc-Lys(me3)-OH was prepared by treating Boc-Lys-OH with 16 molar equivalents of methyl Iodide in 250 molar equivalents of methanol in the presence of 10 molar equivalents of $KHCO_3$ at the room temperature for 72 hrs under constant stirring (Chen and Benoiton, 1976). The solid residue was removed by filtration over Whatman filter paper and the liquid phase was evaporated on a rotary film evaporator. The crude mixture was resuspended in 50 mL of water and was slowly acidified to pH 1.0 with 4M aqueous HCl. This material was degassed and the desired product was purified from residual starting material and methyl ester byproduct by preparative HPLC (column YMC C18 pro pack, 150 mm x 20 mm, 5 µm, 12 nm).

Truncated histone H3.2Δ20 (C110A) was expressed recombinantly with His_6-tag at N-terminus and a TEV protease cleavage site (ENLYFQ^C) inserted after position H3.2L20, replacing A21, so that upon TEV protease cleavage, an N-terminal cysteine is released (Supplementary Figure 2B). The C-terminal peptidyl thioester described above was ligated to the recombinant histone H3.2Δ20-A21C fragment through native chemical ligation (Dawson et al., 1994), using MPAA (4-mercaptophenlacetic acid) as a ligation catalyst (Johnson and Kent, 2006). Briefly, equimolar amounts of peptidyl 3-mercaptopropionamide thioester and truncated histone were mixed at 2mM final concentration in NCL buffer (6M Guanidinium chloride, 200 mM phosphate pH 7.0) in the presence of 30 mM MPAA and 20 mM TCEP. The pH was adjusted to 7.0-7.2 and the reaction was incubated for 12-16 hrs at room temperature. Subsequently, the completion of reaction was validated with MALDI-TOF MS and product was purified by semipreparative HPLC (column YMC pack C8, 250 mm*10 mm, 5 µm, 30 nm). The native alanine at position 21 was restored by radical-mediated desulfurization of cysteine (Chiang et al., 2009; Wan and Danishefsky, 2007). Briefly, to 2-10 mg of the fully ligated H3 is added 300 µl of NCL buffer (6M GdmCl, 200 mM sodium phosphate pH 7), 300 ul of 0.5 M TCEP in NCL buffer (pH adjusted to 7), 20 µl 10% β-me (v/v in NCL buffer), and 20 µl of 200 mM VA-044 in MilliQ. The resulting solution is shaked at 45°C in a thermoshaker for 40 minutes and then subjected to HPLC purification. The amount of β-me and reaction temperature is very important. Temperature above 50°C leads to side reactions. Excess of β-me or reaction temperature below 42°C lead to incomplete reaction. The completion of reaction was validated by ESI MS. The masses of histones are determined on a LTQ-FT mass spectrometer (Thermo Scientific) with a short analytical HPLC run before the mass-spectrometry. 20 µl of filtered histone solution in milliQ water with concentration around 0.1 mg/ml is injected on the analytical HPLC column (Agilent, poroshell 120 SB-C18 2.7 µm) each time. The method runs at 0.5 ml/min with the gradient ramping from 5% B to 100 % B in 10 min (buffer A, 0.1% formic acid in water; buffer B, 0.1% formic acid, 90% ACN in water). After desulfurization, histones were purified by semipreparative HPLC (YMC pack C8, 250 mm*10 mm, 5 µm, 30 nm) and subsequently lyophilized.

All the intermediates and products are characterized by analytical HPLC and MALDI-TOF or ESI-MS (Figure 2 and Figure S2). The theoretical masses are calculated using UCSF protein prospector (http://prospector.ucsf.edu/prospector/mshome.htm). For the final desulfurized H3K4me3 (C110A), the theoretical monoisotopic mass is 15258.548 Da and the observed monoisotopic mass is 15258.586 Da (calculated from the z = 22 ion): identical within 0.038Da (2.5 ppm), Figure 2D displays the average charge envelope of a direct infusion experiment, with the mass computed from the monoisotopic mass of z=22 charge state (grey inset). This product is >97.5% pure by HPLC (Figure 2C-D). H3K36me3 histone was prepared as in (Chen et al., 2014).

Cell Culture

HEK293 cells were grown in 37°C, 5% CO_2, 95% humidity in high glucose DMEM (Invitrogen) supplemented with 10% (v/v) HyClone FBS Characterized U.S. and 1X Penicillin/Streptomycin (Gibco). HEK293 cultures were grown in 15 cm polystyrene dishes, passaged 1:20 every 5 days, Cells used in ChIP experiments were harvested when 80-90% confluent with media changed at least 3 hours prior harvesting.

Mouse Embryonic Stem Cells (mESC) E14 cell line (129/Ola background) (Hooper et al., 1987) were grown at 37°C, 5% CO_2, 95% humidity in ES media high glucose DMEM (Invitrogen) supplemented with 15% (v/v) FBS (Gibco), 2mM L-glutamine (Gibco), 1% (v/v) non-essential amino acids (Gibco), 1X Penicillin/Streptomycin (Gibco), 0.1 mM 2-mercaptoethanol (Gibco), 1000U/mL LIF (ESG1107, Millipore), 3 μM CHIR99021 (04-0004, Stemgent), 1 μM PD0325901 (04-0006, Stemgent), sterile filtered 0.1 μm membrane, stored in 4°C for up to 1 week. E14 cells were cultured on dishes coated with 0.1% gelatin (Sigma) without feeder cells, and passaged daily in 1:3 ratio (media was changed 3 hours prior passage). Cells were harvested at 80-90% confluence, with media changed at least 3 hours prior.

The stem-cell state of the E14 cells was verified by immunofluorescent staining for classical markers of pluripotency Oct4, Sox2, Nanog (Figure S2E-G). Immunostaining of mESCs was performed as follows: E14 mESCs were grown on acid-washed glass coverslips coated with 0.1% gelatin in the ES media. Before fixation, cells were washed with PBS. Cells were fixed by 10 minute incubation with 0.4% freshly cracked formaldehyde in PBS at room temperature. If not stated differently, all subsequent washes, were done by three subsequent 5 minutes incubations in PBS. Fixed cells were washed and subsequently cell membranes were permeabilized via 100% methanol -20°C incubation for 10 minutes, then rinsed with PBS for 5 minutes. After blocking with blocking buffer (1x PBS, 0.3% Triton X-100, 10% Normal Donkey Serum) for 30 min, cells were stained by 1 hr incubation with primary antibody diluted in binding buffer (1% BSA, 1XPBS, 0.3% Triton X-100). The following primary antibody dilutions were used: 1:200 Rabbit α-Nanog (ab80892), 1:500 Rabbit α-Sox2 (ab97959), 1:250 Rabbit α-Oct4 (ab19857). Cells were washed and incubated for 1 hr in darkness with secondary antibody 1:2000 Alexafluor 488 donkey α-rabbit IgG (Invitrogen A21206) diluted in binding buffer. Cells were washed, mounted on glass slide and imaged with confocal microscope.

Octamer and Nucleosome Reconstitution

Octamers were prepared on 250-500 μg scale as previously described (Dyer et al., 2003; Luger et al., 1999), using human histones expressed in *E. coli* (Ruthenburg et al., 2011). Briefly, equimolar core histones were mixed in unfolding buffer (50 mM Tris-HCl pH 8, 6.3 M Guanidine-HCl, 10 mM 2-mercaptoethanol, 4 mM EDTA) to final

concentration of total histone ≥ 1 mg/mL, and dialyzed against two changes of 500 volumes of refolding buffer (20 mM Tris-HCl pH 7.8, 2M NaCl, 1mM EDTA, 5 mM DTT) over 16 hours in 3500 MWCO SnakeSkin dialysis tubing (Pierce) at 4°C. The soluble fraction of crude octamer was subjected to gel filtration chromatography (Superdex 200 10/300 GL, GE Healthcare) resolved in refolding buffer. Fractions containing pure octamer were pooled and concentrated with Amicon Ultra-4 centrifugal filters (10k MWCO, Millipore) to a final concentration of 5-15 μM (measured spectroscopically, ε_{280nm}=44700 $M^{-1}cm^{-1}$, blanked with concentrator flow-through).

DNA for nucleosome reconstitution is based on 601 nucleosome positioning sequence (Lowary and Widom, 1998). To each end of 601 sequence we have appended 22 bp barcode sequences-- each composed of two catenated 11 bp "unwords", sequences absent in human and mouse genome (Herold et al., 2008), flanked by constant 6 bp of linker DNA (see Figure 2F and Table S1). To preclude potential misalignment of our 22-mer barcodes, we computationally filtered potential sequences to exclude candidates that had fewer than two mismatches or indels relative to the assembled mouse or human genomes (Table S1). We thereby ensure that even poor quality reads from our library are not likely to be inappropriately aligned to the genome, and vice-versa. We have created a range of barcoded nucleosome positioning sequences with thus construction, and from this set we have selected sequences which when amplified by PCR yield a single product and amplify with an efficiency equal to 2. The barcode sequences are longer than necessary for specific recognition by sequencing because they were designed to also work well in qPCR and ddPCR. Assessment of a single amplified species was made by running 100 ng of PCR product on 6% native PAGE in 1xTBE(89 mM tris-base, 89 mM boric acid, 2 mM EDTA) with post-staining in 1x SYBR® Gold (Invitrogen™). Amplification efficiency was measured by running qPCR reaction with SYBR green chemistry for serially diluted barcoded nucleosome positioning sequences template. Amplification efficiency was calculated from the slope of C_t([template]) with the following equation $\eta = DF^{(-1/slope)}$, where DF is dilution factor for serial dilution.

Nucleosomes were reconstituted by mixing equimolar histone octamer and DNA to final concentrations of 1 μM and dialyzing this solution in dialysis buttons (Hampton Research) against a non-linear gradient starting with 2M NaCl and ending at 200 mM NaCl over the course of 12-16 hours in a buffer containing 20 mM Tris-HCl pH 7.5, 1 mM EDTA, 10 mM 2-mercaptoetanol (Ruthenburg et al., 2011). Subsequent to dialysis, semi-synthetic nucleosomes were diluted 1:1 with 2x storage buffer (20mM Na•Cacodylate pH 7.5, 10% v/v glycerol, 1mM EDTA, 1X RL Protease Inhibitor Cocktail [1 mM PMSF, 1mM ABESF, 0.8 μM aprotinin, 20 μM leupeptin, 15 μM pepstatin A, 40 μM bestatin, 15 μM E-64], 200 μM PMSF) and kept at 4°C. The concentration of nucleosomes was measured in triplicate by stripping DNA with 2M NaCl and measuring concentration of DNA by densitometry of ethidium bromide-stained agarose gels

calibrated in situ with the Thermo Scientific MassRuler Low Range DNA Ladder. As we reconstitute the internal standard nucleosome ladder from modified histone octamer and a concentration series of different barcoded DNAs in a single tube, we sought examine whether there was any over- or under-representation of the barcoded nucleosomes relative to their corresponding DNA species. Illumina sequencing of free DNA versus nucleosome ladder revealed that our ladder reconstitutions accurately reflect the relative concentrations of DNA onto which they are reconstituted (Figure S2I). The dynamic range of sequencing is much larger than it is necessary for ICeChIP-seq, hence we do not consider it a limiting factor for our experiments. Working concentrations of semi-synthetic nucleosomes were prepared by dilution to the desired concentrations in long-term storage buffer (10 mM Na•Cacodylate pH 7.5, 100 mM NaCl, 50% Glycerol, 1 mM EDTA, 1X RL Protease Inhibitor Cocktail, 200 µM PMSF) and stored at -20°C.

ICeChIP

Our ICeChIP protocol is based on input preparation largely following the Dilworth lab native ChIP protocol (Brand et al., 2008). Plate-adhered cells (~10^7 cells) were washed twice with 10 mL of PBS, and released by 5 mL Accutase (Millipore) for 5 minutes in 37°C, quenched with 2 mL of complete media, and collected by centrifugation (500 x g, for 5 minutes at 4°C). All subsequent steps were performed on ice with ice-cold buffers. Cells were washed twice with 10 mL PBS, and twice with 5 mL of Buffer N (15 mM Tris pH 7.5, 15 mM NaCl, 60 mM KCl, 8.5% (w/v) Sucrose, 5 mM $MgCl_2$, 1 mM $CaCl_2$ 1 mM DTT, 200 µM PMSF, 1x RL Protease Inhibitor Cocktail). Cells were resuspended in 2 PCVs (packed cell volumes) of Buffer N and lysed by adding 2 PCV of 2x Lysis Buffer (Buffer N supplemented with 0.6% NP-40 substitute (Sigma)) for 10 minutes at 4°C. Nuclei were collected by centrifugation (500 x g for 5 minutes at 4°C) and were resuspended in 6 PCVs of Buffer N. To remove cell debris, resuspended nuclei were overlaid on the surface of 7.5 mL of sucrose cushion (10 mM HEPES pH 7.9, 30%(w/v) sucrose, 1.5 mM $MgCl_2$) in a 50 mL centrifuge tube centrifuged (1300 x g, Sorvall Legend XTR swinging bucket rotor for 12 minutes at 4°C). Most cell debris remained in upper layer while nuclei sedimented through the sucrose cushion and pelleted on the bottom of the tube. The supernatant was discarded and nuclei were resuspended in 2 PCVs of Buffer N. To measure apparent concentration of chromatin, 2 µL of resuspended nuclei were diluted in 98 µL of 2M NaCl in triplicate, total nucleic acid absorbance was measured at 260 nm by Nanodrop (Thermo Scientific), and the conversion factor assuming $1A_{260}$ = 50 ng/µL of chromatin employed. Based on these measurements, apparent concentration of chromatin was adjusted to 1 µg/µL with Buffer N. The quantity and quality of nuclei were also assessed using hemocytometer.

Following exhaustive MNase digests we experimented with sucrose gradient fractionation (O'Neill and Turner, 2003) and hydroxyapatite chromatography (Brand et

al., 2008) (Figure S3A-C), and found that the latter affords better recovery of native mononucleosomes with superior purity, at the expense of discrimination against di-nucleosomes. Paired-end sequencing circumvents this avidity bias, as one can simply exclude fragments larger than 220 bp from the analysis. This number was chosen to be significantly larger than the average human and mouse nucleosome repeat lengths (van Holde, 2011; Widom, 1992). At this stage, a ladder of semisynthetic nucleosomes was doped into the pool of native nucleosomes. The amount of spiked ladder was comparable to estimated amount of genome copies in the pool based on the nuclei counting times the average DNA content per cell (~2.5 copies of genome per cell).

To remove debris coming from nuclei lysis and MNase digestion as well as strip chromatin bound factors, the pool of nucleosomes was subjected to hydroxyapatite chromatography purification (Brand et al., 2008). Fragmented chromatin with internal standard ladders were split into 100 µg total nucleic acid fractions and each fraction was mixed with 66 mg of hydroxyapatite (HAP) resin (Bio-Rad Macro-Prep® Ceramic Hydroxyapatite Type I 20 µm) rehydrated with 200 µL of HAP buffer 1(3.42 mM Na_2HPO_4 and 1.58 mM NaH_2PO_4 final pH 7.2, 600 mM NaCl, 1 mM EDTA, 200 µM PMSF), incubated for 10 minutes at 4°C on rotator and subsequently was applied to the centrifugal filter unit (Millipore Ultrafree MC–HV Centrifugal Filter 0.45 µm). The chromatin-loaded resin in the column was drained and then washed four times with 200 µL HAP buffer 1 and four times with 200 µL of HAP buffer 2 (3.42 mM Na_2HPO_4 and 1.58 mM NaH_2PO_4 final pH 7.2, 100 mM NaCl, 1 mM EDTA, 200 µM PMSF) by centrifugation (600 x g, 1 minute at 4°C in fixed angle rotor. Nucleosomes were eluted from the HAP column with three 100 µL washes of HAP elution buffer (342 mM Na_2HPO_4 and 158 mM NaH_2PO_4 final pH 7.2, 100 mM NaCl, 1 mM EDTA, 200 µM PMSF). To measure apparent concentration of HAP purified chromatin fragments, 10 µL of HAP elution was diluted in 40 µL of 2M NaCl in triplicate, and absorbance measured at 260 nm averaged and adjusted ($1A_{260}$ = 50 ng/µL of chromatin). Apparent concentration of chromatin was adjusted to 20 µg/mL with ChIP Buffer 1(25 mM Tris pH 7.5, 5 mM $MgCl_2$, 100 mM KCl, 10% (v/v) glycerol, 0.1% (v/v) NP-40 substitute).

H3K4me3 ChIP was performed with 10 µg of chromatin and 15 µL of AM39159 antibody, H3 and H4 ChIP was performed with 1 µg of chromatin and 15 µL of AM61277 and AM61299 antibody, respectively (Active Motif). 10% of initial chromatin for each IP was set aside to serve as ChIP input. Each IP experiment used 50 µL of Protein A Dynabeads (Invitrogen) that were washed twice with 1mL of ChIP buffer 1 with 1 min collection on magnetic rack after each wash. To prepare the resin, 15 µL of antibody and 85 µL of ChIP buffer 1 was added to Protein A Dynabeads and incubated for 10 minutes at room temperature on a rotator, then washed twice with 1mL of ChIP Buffer 1. Chromatin (10 µg unless otherwise indicated) in 500 µL of ChIP buffer 1 was then added to magnetic beads and incubated for 15 minutes at room temperature on rotator. Beads were washed 3 times with 1 mL of ChIP Buffer 2 (mM Tris pH 7.5, 5 mM

$MgCl_2$, 300 mM KCl, 10% (v/v) glycerol, 0.1% (v/v) NP-40 substitute), then twice with ChIP buffer 3 (10 mM Tris pH 7.5, 250 mM LiCl, 1mM EDTA, 0.5% Na•Deoxycholate, 0.5%(v/v) NP-40 substitute), each wash consisting of a 10 minute rotating incubation and 1 minute collection on magnetic rack at 4ºC. During the course of washing, at least two tube changes reduced non-specific background. Beads were then rinsed with 1 mL of ChIP Buffer 1 and 1 mL of TE buffer, followed with two 200 µL ChIP elution buffer steps (50 mM Tris pH 7.5, 1 mM EDTA, 1% w/v SDS). Each elution step consisted of 10 minute incubation at 65ºC in a Thermoshaker (Eppendorf) at 900 rpm. Elutions were combined and ChIP elution buffer was added to inputs to match volume of ChIP elutions. After adjusting the buffer to 200mM NaCl, 100 ng of RNase A was added into the mixture and incubated at 65ºC for 45 minutes in Thermoshaker at 800 rpm, and terminated with 10 mM EDTA. Next, protein digestion was accomplished with 20 ug of proteinase K (Roche) for 2 hrs at 42ºC in the Thermoshaker at 800 rpm. DNA was recovered and purified with Qiaquick columns (Qiagen): 6 volumes of PB buffer were added to the digestion and this solution applied to the column (17900 x g, 30s) followed by 3x 750 µL of PE buffer washes (17900 x g, 30s) with an extra 1 minute spin to remove residual ethanol. DNA was eluted by applying two times 25 µL of TE buffer at 50ºC and centrifuging (17900 x g, 1 min).

Illumina library preparation

For library preparation 10 ng of DNA isolated from IP or input was used. In cases in which the total amount of DNA was below 10 ng, all available DNA was subjected to library preparation. Ends of DNA were blunted using the End-it™DNA End-Repair Kit (Epicentre) (7µL 10x End-It buffer, 7µL 2.5 mM dNTP Mix, 7µL 10 mM ATP, 1.4µL of End-Repair Enzyme Mix and 47.6µL of DNA in TE buffer, incubated for 45 minutes at room temperature. DNA was purified with 126µL (1.8 volume) of Ampure XP Beads (Beckman Coulter). Beads were mixed with End repair mixture by pipetting 10 times up and down followed by 5 minutes incubation at room temperature. Magnetic beads were collected on side of the tube by magnet and two 30 sec 250 µL 80% EtOH washes on magnet were performed. Tubes were removed from the magnetic rack and 34 µL of TE buffer was added to beads and pipetted 10 times up and down. Magnetic beads were not removed from elution and remained in the tube during A-tailing. Addition of single adenosine to 3' ends of DNA was accomplished by adding to 5 µL NEB buffer 2, 10 µL 1 mM dATP, 1 µL Klenow fragment (3'→5' exo-, NEB) to the End-repaired DNA, and with incubation at 37ºC for 30 minutes. To purify DNA, 110 µL (2.2 volume) of SPRI Buffer (20% PEG6000, 2.5M NaCl) was added to the reaction and was pipetted 10 times up and down followed by 5 minutes incubation at room temperature. Magnetic beads were collected on side of the tube with magnet and two 30 sec 200 µL 80% EtOH washes on magnet were performed. Tubes were then taken out of magnetic rack and 13 µL of TE buffer was added to beads and mixed by pipette. Magnetic beads were not

removed from elution and remained in the tube during adaptor ligation. To ligate adaptors, the following mixture was prepared: 2x Quick DNA ligase buffer, 2 µL 2 µM of adaptor duplex, 1 µL of Quick DNA ligase (NEB) and added to 13 µL of A-tailed DNA. The reaction was incubated for 15 minutes at room temperature. To purify DNA, 21 µL (0.7 volume) of SPRI Buffer was added to the reaction and was pipetted 10 times up and down followed by 5 minutes incubation at room temperature. Magnetic beads were collected via magnet and washed twice with 30 sec 200 µL 80% EtOH incubations, and eluted with 46 µL of TE buffer. The supernatant was transferred to the new siliconized tube.

Quantitative-PCR was run to estimate minimal number of PCR cycles to amplify DNA library. 7.15 µL of H2O, 1 µL of 10xAccuPrime PCR buffer II, 0.25 µL of 20 x EvaGreen® dye (Biotum) to final 0.5x dilution, 1µL of DNA library, 0.2 µL of 25 µM MP_PCR_Primer1, 0.2 µL of 25 µM MP_PCR_Primer2, and 0.2 µL AccuPrime Taq DNA Polymerase (Invitrogen #12339-016). Bio-Rad CFX384 qPCR machine program was set to: 1- 95ºC for 5 min, 2- 95ºC for 80s, 3- 65ºC for 90s – read at the end, 4- go back to step 2 for 24 times. Based on the readings, cycle number to amplify library was set to C_t + 3 cycles. If the observed C_t value was below 7 cycles, the template was diluted 10 fold and procedure would be repeated.

DNA library was amplified by mixing: 40 µL of DNA library, 5 µL 10xAccuPrime PCR buffer II, 1µL 25 µM MP_PCR_Primer1, 1 µL 25 µM MP_PCR_Primer2_INDEX, 1 µL AccuPrime Taq DNA Polymerase, and 2 µL of H_2O, followed by thermal cycling in a C1000 (Bio-Rad). The machine was set to: 1- 95ºC for 5min, 2- 95ºC for 80s, 3- 65ºC for 90s, 4- go back to step 2 for number of cycles determined with qPCR (C_t+3 cycles). Detailed primer information can be found in supplementary Table S2. Amplified DNA was purified with 90µL (1.8 volume) of Agencourt Ampure XP Beads. Beads were mixed with PCR mixture by pipetting 10 times up and down, followed by 5 minutes incubation at room temperature. Magnetic beads were collected on the side of the tube via magnet and two 30 s 250 µL 80% EtOH washes on magnet were performed. The tube was removed from the magnetic rack and 25 µL of TE buffer was added to beads and pipetted 10 times up and down. Magnetic beads were collected on the side of the tube and supernatant was moved to new siliconized tube. Size distribution and concentration of amplified library was assessed with Agilent Technologies 2100 Bioanalyzer.

Sequencing and Data Analysis

Cluster generation and sequencing was performed using the standard Illumina protocol for Illumina HiSeq 2500 by the University of Chicago Functional Genomics Core facility. Data analysis was performed with Galaxy (Blankenberg et al., 2010; Giardine et al., 2005; Goecks et al., 2010). Raw reads in FastQ fromat were first submitted to FastQ Groomer. Reads were mapped with Bowtie2 (Langmead et al., 2009) (sensitive preset

option, end-to-end alignment), depending on organism of origin, to human (hg19) or mouse (mm10) reference genomes with sequences of barcodes catenated at the end (each barcode with its own entry) (For barcoded nucleosome templates, see Table S1). Resulting SAM files were then filtered using SAMtools (Li et al., 2009). Reads that were unmapped, unpaired (distance >1000 bp) and paired in wrong pair were removed from the set by this data analysis pipeline. To remove noise coming from low quality reads and contaminants as well to mask repeatable genomic sequences, reads with mapping quality lower than 20 were removed. To avoid signal artifacts and not distort Poisson sampling statistics paired reads were merged together into single entries (overlapping fragments were flattened and gaps were filled). To avoid oligonucleosome avidity bias, reads longer than 220 bps were removed, except where explicitly stated otherwise. BEDTools (Quinlan and Hall, 2010) were used to create genome coverage bedgraphs. When indicated data was averaged over 25bp non-overlapping windows.

In order to get high precision we have aimed to achieve IP coverage ranging between 1000 and 10 reads of depth and average depth of input to be at least ~20. However, the deeper input sequencing the better, as it is in most cases a limiting factor for precision. In order to compute barcode IP enrichment, we calculated the ratio of integrated coverage over the whole sequence of each baracode in IP over the input, (Table S3).

$$Barcode\ IP\ enrichment = \frac{\sum_1^N IP}{\sum_1^N input}$$

where, *n* is the length of barcoded construct, in this case it is 203 bp, *IP* and *input* represent integrated tag counts.

To increase accuracy, we have averaged the barcode IP enrichment values for multiple barcodes. To calculate Histone Modification Density (HMD) we have applied the following equation to genome coverage information for IP and input:

$$HMD\ (per/bp) = \frac{\left(\frac{IP}{input}\right) * 100\%}{Barcode\ IP\ enrichment}$$

To estimate HMDs confidence intervals 95% we have applied following equation:

$$CI^{95\%}_{HMD(\frac{per}{bp})} \cong 1.96 * \sqrt{\left(\frac{100\% * \sqrt{IP}}{input * Barcode\ IP\ enrichment}\right)^2 + \left(-\frac{100\% * IP * \sqrt{input}}{input^2 * Barcode\ IP\ enrichment}\right)^2}$$

Here we assume that standard deviation of Barcode IP enrichment is negligible. To calculate total HMD content genome-wide (percent of all nucleosomes in all genomic loci bearing modification), HMD signal was integrated and subsequently divided by total number of bps for which we had genome coverage or alternatively by reported total genome size. Figures of HMD/bp over selected genomic intervals were prepared using

IGV (http://www.broadinstitute.org/igv), and gene structure is derived from the relevant Refseq annotation.

Heatmaps and GO analysis were made using Homer software(Heinz et al., 2010).

ICeChIP correction for off-target specificity

For indicated datasets, ICeChIP-seq tracks were corrected for off-specificity by solving the following matrix equation: $A * x = b$,

where, x is a matrix of corrected HMD scores, A is a matrix of correction factors and b is a matrix of non-corrected HMD scores,

$$A = \begin{vmatrix} t_a^a & \cdots & t_z^a \\ \vdots & \ddots & \vdots \\ t_a^z & \dots & t_z^z \end{vmatrix}, b = \begin{vmatrix} HMD_1^a & \cdots & HMD_n^a \\ \vdots & \ddots & \vdots \\ HMD_1^z & \cdots & HMD_n^z \end{vmatrix}, x = \begin{vmatrix} HMD(Cor)_1^a & \cdots & HMD(Cor)_n^a \\ \vdots & \ddots & \vdots \\ HMD(Cor)_1^z & \cdots & HMD(Cor)_n^z \end{vmatrix}$$

where, *t* is correction factor for specificity toward histone marks from the set of 'a' to 'z' histone marks (subscript), in the immunoprecipitation using antibody toward a histone mark from the set of 'a' to 'z' histone marks (superscript); *HMD* is histone modification density for a given histone mark ('a' to 'z') from the 1st to the nth locus; *HMD(Cor)* is corrected histone modification density for a given histone mark from the 1st to the nth locus,

$$t_z^a = \frac{\frac{\sum_1^N IP_z^a}{\sum_1^N input_z}}{\frac{\sum_1^N IP_a^a}{\sum_1^N input_a}}$$

where, $\sum_1^N IP$ and $\sum_1^N input$ refer to abundance of the given barcode in the IP or in the input, superscript refers to histone mark toward which antibody was raised, while subscript refers to mark on the semisynthetic nucleosome that was pulled-down.

Whenever indicated the abovementioned correction was applied. H3K4me3, H3K9me3 and H3K27me3 HMD were corrected using raw HMD for H3K4me3, H3K9me3 and H3K27me3 but not H3K36me3 and H3K79me2. H3K36me3 and H3K79me2 were excluded because it is apparent that the signal is not coming from accountable sources and hence cannot be correctly deconvoluted. An attempt was made to deconvolute H3K36me3 and H3K79me2 using all available tracks, but as mentioned above the signals for these two marks is convoluted with a noise of unknown to us source.

ChIP with radiolabeled nucleosomes

ChIP with radiolabeled nucleosomes was performed the same way as IceChIP with following differences: 1.) After Hydroxyapatite chromatography, digested fragments of chromatin were subjected to dialysis in Thermo Scientific SnakeSkin 10k MWCO to the

T4 PNK buffer without Mg^{2+} (70 mM Tris pH 7.5, 50 mM NaCl, 5 mM DTT, 25 ug/mL NEB BSA, 200 uM PMSF, 1X RL Protease Inhibitor Cocktail). After 8 hours, the nucleosomes were recovered from dialysis tubing and $MgCl_2$ was added to final 10 mM. Radiolabeling was performed as follows: 1 pmol of 5' DNA ends, 10 µL 10x T4 PNK buffer, 1 µL γ-^{32}P ATP (10 µCi/µL), 20 U of T4 PNK and H_2O up to 100 µL was mixed and incubated for 1 hr in 37°C. The same operation was done to radiolabel semi-synthetic nucleosomes. Radiolabeled standard nucleosomes were subsequently doped into cold digested fragments of chromatin. Further steps were performed as in ICeChIP. A LS6000IC liquid scintillation counter (Beckman) with triplicate samples prepared in 4 mL of Ecoscint A scintillation fluid (National Diagnostics) was used to readout amount of recovered radiolabeled chromatin and/or radiolabeled barcodes.

Alternate modalities for calibrating ChIP with semisynthetic nucleosomes

There are two alternate ways we could envision to calibrate ChIP: global histone modification density calibration using an external standard and direct internal standard calibration. Like the relative internal standard approach that was predominantly employed in this work, these two can yield results expressed in "histone modification density" units, which are equal to apparent ratio of probed epitope to all other epitopes available in the given locus.

Global histone modification density calibration relies on a measurement of the total ratio of modification relative to the amount of histone, for example, knowing the percentage of all H3 that is K4 trimethylated. This global histone modification density, derived from either mass spectrometry or quantitative western blot measurements can be then redistributed among all IP peaks corrected for input depth in any given locus. The drawback of this method is vulnerability to the errors in making the global abundance measurement (for example, MS accuracy plus the ambiguity of perhaps not observing all potential forms of the modification). We have tried using quantitative western blot for global histone modification density calibration, however this was problematic. Beside the fact that quantitative western blot still requires standards, it also suffers from very short linear range, low sensitivity, low saturation point and quantitative western blot seems to be sensitive to differences in complexity of sample and standard, perhaps due to differences in mass transport during binding and washing in these two instances, which results in a significant shifting standard error of the measurement. Such external measurements by orthogonal methodologies need to be made from the same nucleosomal sample used in the ChIP, and differential sample handling losses in both techniques are likely to be a considerable source of error.

Direct internal standard calibration measures the tag count of a spiked-in barcoded nucleosome standard through the ChIP process, knowing the precise molar concentrations of each internal standard ladder member in the input to extrapolate

absolute molar abundance of probed epitope in the original sample. This sort of calibration is limited by the accuracy of counting the number of nuclei subjected to the micrococcal nuclease digest and biased loses that mount on the way from this well quantified number to exhaustively fragmented chromatin isolate. As we recover little more than 80% of the total nucleic acid from digested nuclei under highly optimized digest and isolation conditions, this method would suffer from some systematic error due to biased genome recovery (Henikoff et al., 2009). Given the added experimental complexity and the additional sources of error endemic to direct calibration, we have chosen to avoid this approach, although in principle, it could work.

SUPPLEMENTAL REFERENCES

Alewood, P., Alewood, D., Miranda, L., Love, S., Meutermans, W., and Wilson, D. (1997). Rapid in situ neutralization protocols for Boc and Fmoc solid-phase chemistries. Meth Enzymol *289*, 14–29.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *25*, 25–29.

Blankenberg, D., Kuster, Von, G., Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A., and Taylor, J. (2010). Galaxy: a web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol *Chapter 19*, Unit19.10.1–Unit19.10.21.

Brand, M., Rampalli, S., Chaturvedi, C.-P., and Dilworth, F.J. (2008). Analysis of epigenetic modifications of chromatin at specific gene loci by native chromatin immunoprecipitation of nucleosomes isolated using hydroxyapatite chromatography. Nature Protocols *3*, 398–409.

Chen, F.C., and Benoiton, N.L. (1976). A new method of quaternizing amines and its use in amino acid and peptide chemistry. Canadian Journal of Chemistry *54*, 3310–3311.

Chiang, K.P., Jensen, M.S., McGinty, R.K., and Muir, T.W. (2009). A semisynthetic strategy to generate phosphorylated and acetylated histone H2B. Chembiochem *10*, 2182–2187.

Dawson, P.E., Muir, T.W., Clark-Lewis, I., and Kent, S.B. (1994). Synthesis of proteins by native chemical ligation. Science *266*, 776–779.

Dyer, P.N., Edayathumangalam, R.S., White, C.L., Bao, Y., Chakravarthy, S., Muthurajan, U.M., and Luger, K. (2003). Reconstitution of nucleosome core particles from recombinant histones and DNA. Meth Enzymol *375*, 23–44.

Giardine, B., Riemer, C., Hardison, R.C., Burhans, R., Elnitski, L., Shah, P., Zhang, Y.,

Blankenberg, D., Albert, I., Taylor, J., et al. (2005). Galaxy: a platform for interactive large-scale genome analysis. Genome Res *15*, 1451–1455.

Goecks, J., Nekrutenko, A., Taylor, J., Galaxy Team (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol *11*, R86.

Hattori, T., Taft, J.M., Swist, K.M., Luo, H., Witt, H., Slattery, M., Koide, A., Ruthenburg, A.J., Krajewski, K., Strahl, B.D., et al. (2013). Recombinant antibodies to histone post-translational modifications. Nat Methods *10*, 992–995.

He, S., Bauman, D., Davis, J.S., Loyola, A., Nishioka, K., Gronlund, J.L., Reinberg, D., Meng, F., Kelleher, N., and McCafferty, D.G. (2003). Facile synthesis of site-specifically acetylated and methylated histone proteins: reagents for evaluation of the histone code hypothesis. Proc Natl Acad Sci USA *100*, 12033–12038.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576–589.

Henikoff, S., Henikoff, J.G., Sakai, A., Loeb, G.B., and Ahmad, K. (2009). Genome-wide profiling of salt fractions maps physical properties of chromatin. Genome Res *19*, 460–469.

Herold, J., Kurtz, S., and Giegerich, R. (2008). Efficient computation of absent words in genomic sequences. BMC Bioinformatics *9*, 167.

Hooper, M., Hardy, K., Handyside, A., Hunter, S., and Monk, M. (1987). HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. Nature *326*, 292–295.

Johnson, E.C.B., and Kent, S.B.H. (2006). Insights into the mechanism and catalysis of the native chemical ligation reaction. J Am Chem Soc *128*, 6640–6646.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J Mol Biol *276*, 19–42.

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Preparation of nucleosome

core particle from recombinant histones. Meth Enzymol *304*, 3–19.

O'Neill, L.P., and Turner, B.M. (2003). Immunoprecipitation of native chromatin: NChIP. Methods *31*, 76–82.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Rozowsky, J., Euskirchen, G., Auerbach, R.K., Zhang, Z.D., Gibson, T., Bjornson, R., Carriero, N., Snyder, M., and Gerstein, M.B. (2009). PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls. Nat Biotechnol *27*, 66–75.

Ruthenburg, A.J., Li, H., Milne, T.A., Dewell, S., McGinty, R.K., Yuen, M., Ueberheide, B., Dou, Y., Muir, T.W., Patel, D.J., et al. (2011). Recognition of a Mononucleosomal Histone Modification Pattern by BPTF via Multivalent Interactions. Cell *145*, 692–706.

Shogren-Knaak, M.A., Fry, C.J., and Peterson, C.L. (2003). A native peptide ligation strategy for deciphering nucleosomal histone modifications. J Biol Chem *278*, 15744–15748.

van Holde, K.E. (2011). Chromatin (Springer New York).

Wan, Q., and Danishefsky, S.J. (2007). Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides. Angew. Chem. Int. Ed. Engl. *46*, 9248–9252.

Widom, J. (1992). A relationship between the helical twist of DNA and the ordered positioning of nucleosomes in all eukaryotic cells. Proc Natl Acad Sci USA *89*, 1095–1099.

Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A.J., Zheng, L., Zhang, H., Huang, S., et al. (2011). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. Mol Cell *42*, 451–464.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.