Supplemental Information for:

Decoding the Protein Composition of Whole Nucleosomes with Nuc-MS

Luis F. Schachner[1,2], Kevin Jooß[1,3], Marc A. Morgan[4,5], Andrea Piunti[4,5], Matthew J. Meiners[6], Jared O. Kafader[1,3], Alexander S. Lee[2,3,5], Marta Iwanaszko[4,5], Marcus A. Cheek[6], Jonathan M. Burg[6], Sarah A. Howard[6], Michael-Christopher Keogh[6], Ali Shilatifard[4,5], and Neil L. Kelleher[1,2,3,4,5]*

[1]Department of Chemistry, Northwestern University, Evanston, Illinois, USA

[2]The Chemistry of Life Processes Institute, Northwestern University, Evanston, Illinois, USA

[3]The Proteomics Center of Excellence, Northwestern University, Evanston, Illinois, USA

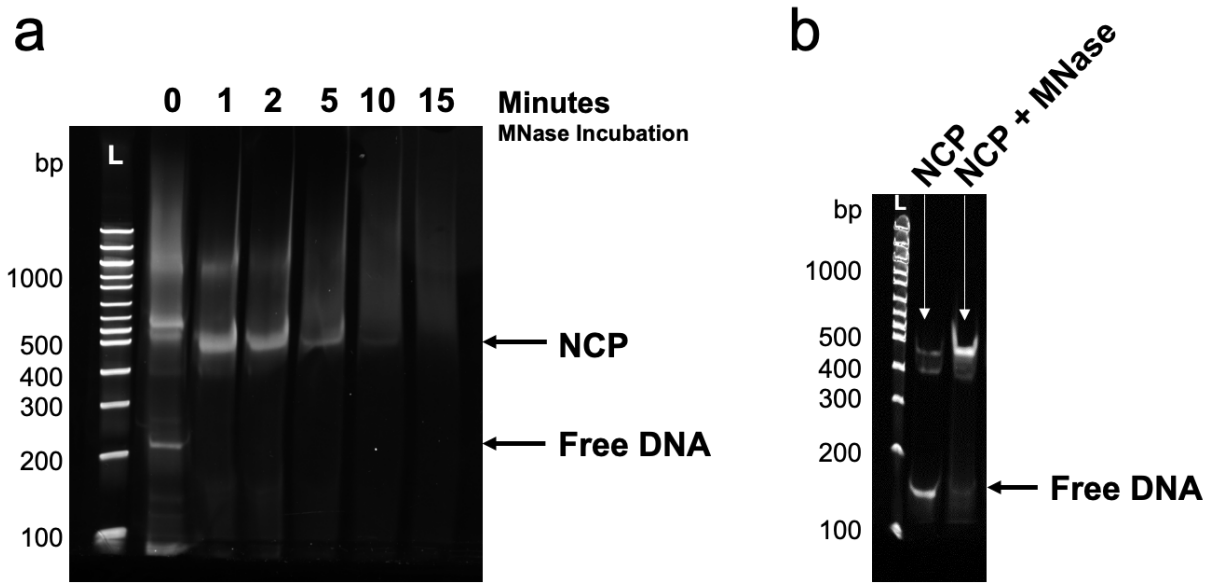
[4]Simpson Querrey Institute for Epigenetics, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

[5]Department of Biochemistry and Molecular Genetics, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

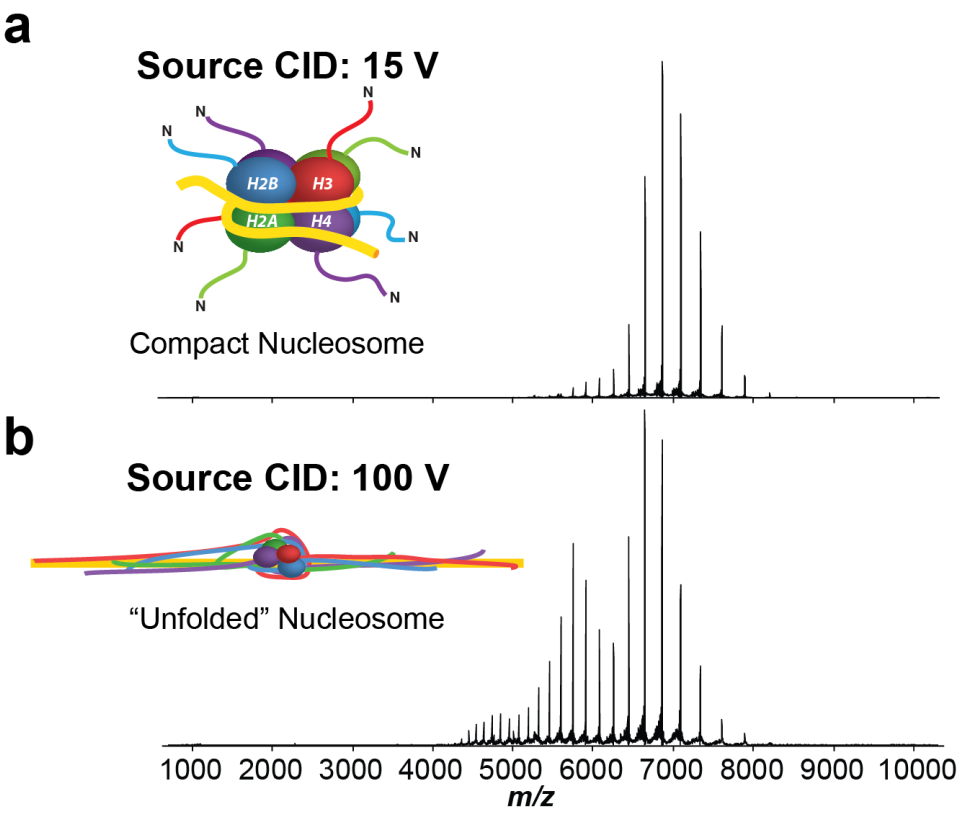
[6]EpiCypher, Inc., Research Triangle Park, North Carolina, USA

*Address reprint requests and communications to:

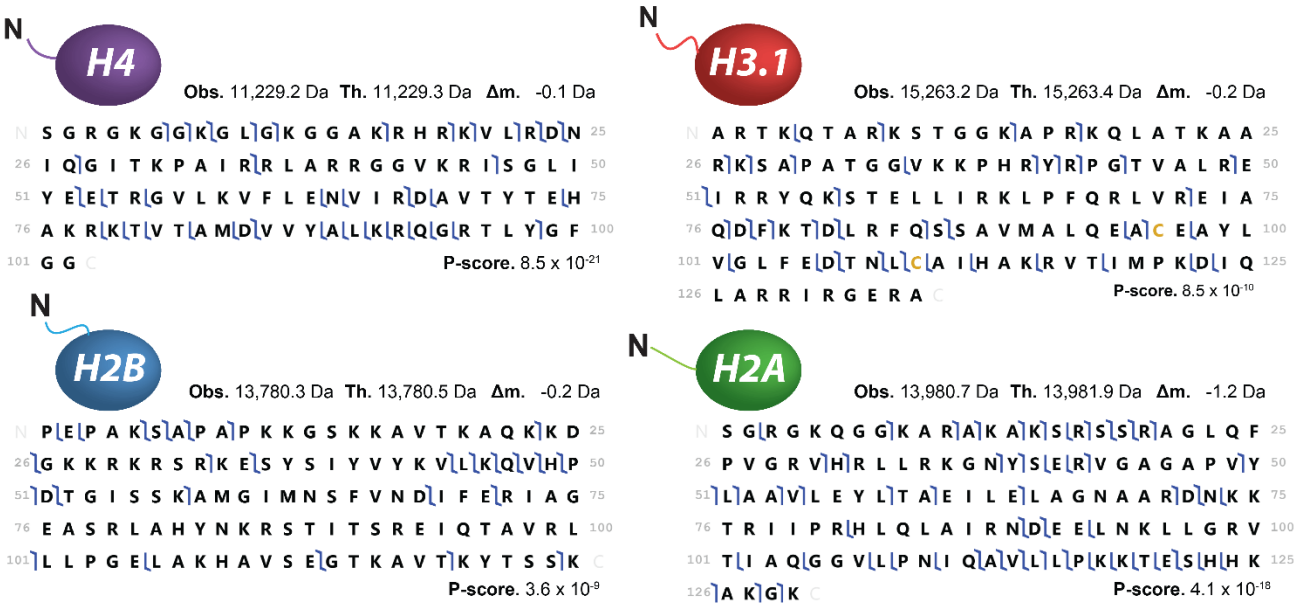
Neil L. Kelleher
2170 Tech Dr., Silverman Hall, Northwestern University, Evanston, Illinois, 60208, United States
Ph. 847-491-3731
Fax. 847-467-1566
n-kelleher@northwestern.edu



Supplemental Figure 1. Synthetic nucleosomes were assembled by salt dialysis and MNase treated to eliminate free DNA prior to analysis by Nuc-MS (employing native electrospray).[1] **(a)** Nucleosome core particle (NCP) samples with 208 bp DNA were treated with 20U of MNase for 0-15 minutes, and quenched with 25 mM EDTA. Samples are resolved by TBE native PAGE and stained with GelRed (*Biotium*). **(b)** NCP with 147 bp DNA treated with MNase for 1 minute. The bp ladder lane is indicated by the letter “L”.



Supplemental Figure 2. "Unfolding" of the nucleosome in the Electrospray source. MS^1 of nucleosome at low **(a, 15 Volt)** and high **(b, 100 Volt)** source activation energy (source-induced dissociation, source CID). An increase in collisional energy results in a concomitant increase in the average charge state and width of the charge state distribution in the native mass spectrum; spray solution was pH ~7, 150 mM ammonium acetate. This observation is consistent with a current model of nucleosome unfolding,[2] where the DNA in the nucleosome core particle first linearizes upon activation, thereby extending the strongly-bound histones across the length of the strand, and potentially increasing exposure of additional sites to protonation during electrospray ionization.



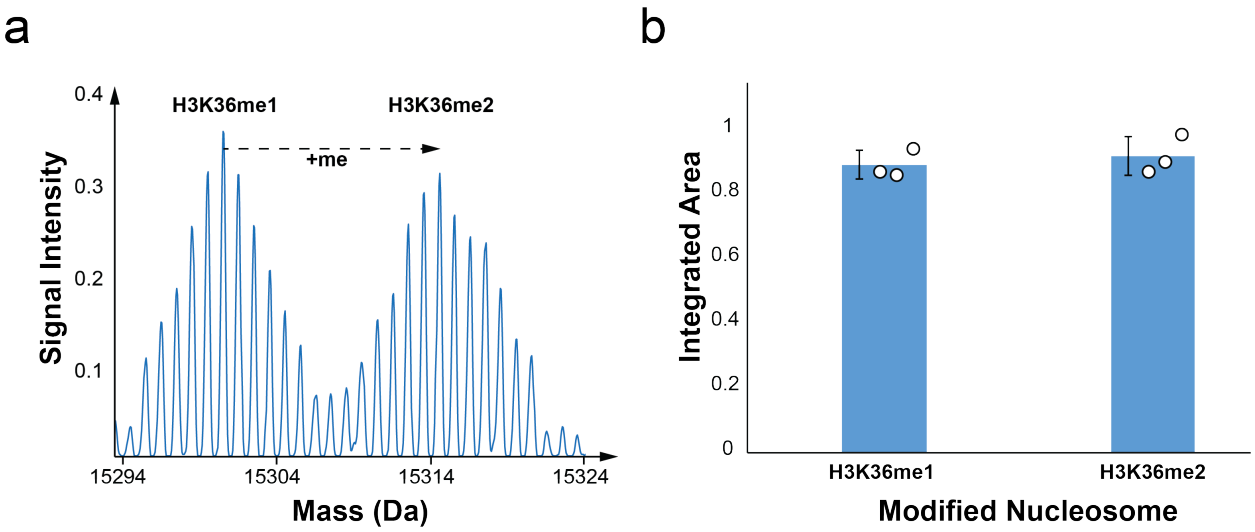
Supplemental Figure 3. Fragment maps for all four histones ejected from unmodified recombinant nucleosomes (**Fig. 1b).** Fragments were manually validated using TDValidator.[3] P-scores were calculated using ProSight Lite.[4]

*Supplemental Discussion for Nuc-MS of unmodified recombinant nucleosomes (**Fig. 1 and Supplemental Figs. 2-3**).* After measuring the intact nucleosome mass(es), and then liberating core histones and their sequence ions, Nuc-MS provides isoform and variant-specific identification, with PTM localization (i.e. the complete proteoform; **Fig. 1b**, and **Supplemental Figs. 1** and **2**). The theoretical average mass of a recombinant nucleosome, comprising eight histones (108,576.6 Da) and 147 bp of biotinylated 601 DNA nucleosome positioning sequence (91,294.5 Da), was calculated to be 199,871.1 Da.[1] Deconvolution of the charge states in the MS^1 spectrum yielded an observed average mass of 199,867.9 ±12.5 Da, a -3.1 Da difference (Δm) from the theoretical value (**Fig. 1b**, MS^1) and within the standard error range for such large ions.[5]

The detection of all histone types in the same spectrum with minimal overlap is a major benefit of native electrospray (nESI). Namely, nESI keeps nucleosomes intact and lowers the

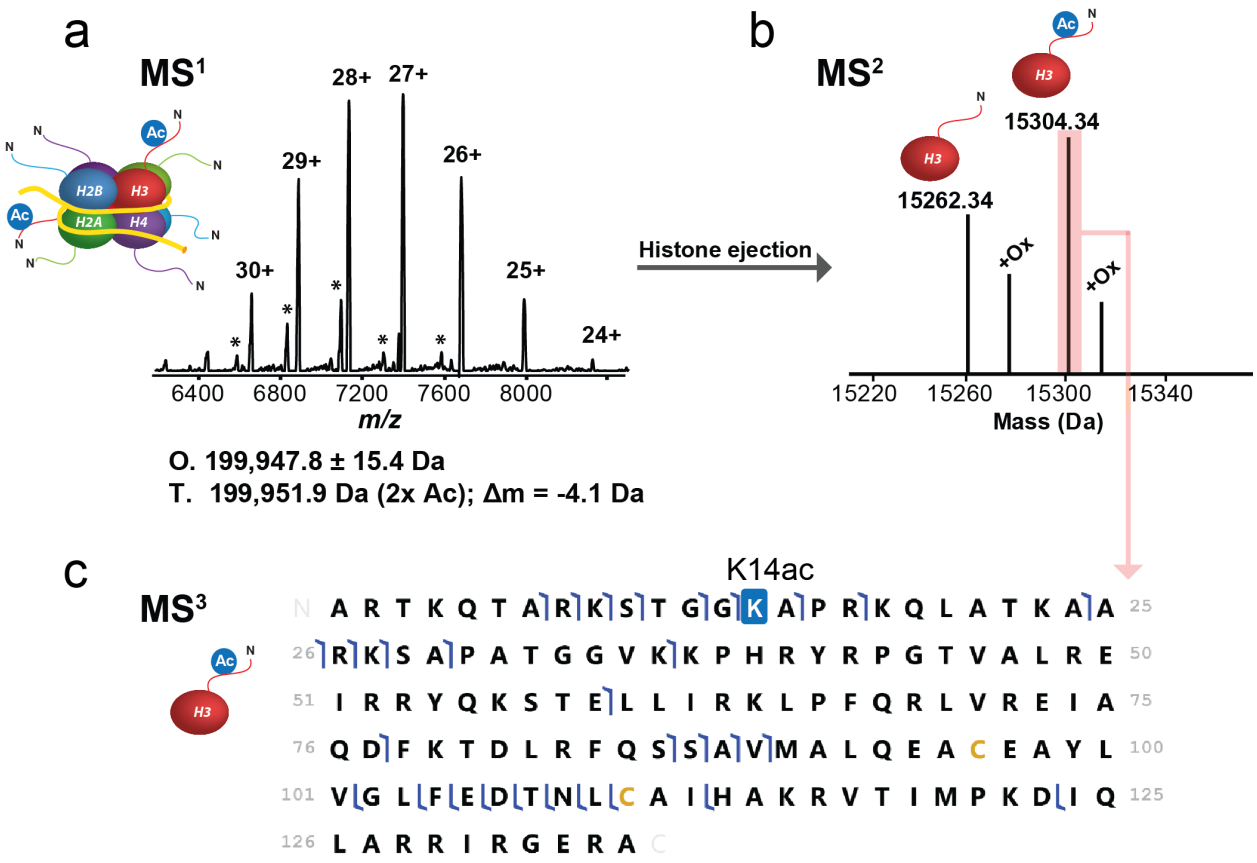
charge state of histones, thereby increasing peak capacity and allowing detection of all ejected histones simultaneously in the same spectrum without any upfront separation (i.e., this contributes strongly to the low-bias nature of the method). In contrast, denaturing ESI relies on successful separation of histones prior to MS to avoid signal overlap that greatly complicates detection and characterization of histone proteoforms.

Upon ejection of histone proteoforms, they were isolated and further fragmented. Additional HCD fragmentation causes protein ions to fragment along the backbone in predictable ion types (i.e., *b* and *y* ions, containing the N- and C-termini, respectively). These fragment ions were mapped onto the protein primary sequence for PTM localization on specific histones (**Fig. 1b**, **MS^3** for H3.1; **Supplemental Fig. 3** for H4, H2A, and H2B).[6]



Supplemental Figure 4. Nuc-MS analysis of equimolar H3K36me1 and H3K36me2 synthetic nucleosomes reveals (**a**) equivalent signal intensities for ejected histones in the deconvoluted mass spectrum showing their isotopic distributions, and (**b**) equivalent mean integrated areas upon quantitation. The bar chart denotes the mean area for three measurement replicates. Individual measurements were overlaid as a scatter plot and are ordered left to right according to data file of origin. Specifically, the ratio of the integrated areas of ejected H3K36me1 and H3K36me2 proteoforms was calculated to be 49.2 ±2.5% and 50.8 ±3.3%, respectively. The error bars denote ±1 standard deviation.

Supplemental Discussion on the Quantitative Nature of the Nuc-MS Assay. The finding in **Supp. Fig. 4** shows that the endogenous proteoform distributions of ejected histones faithfully reflect their relative abundances in the population of nucleosomes presented. In accordance with this observation, relative quantitation is only performed for histones of the same type to minimize concerns about differences in ionization and ejection efficiency.[7] **Fig. 3** and **Supp. Fig. 9** outline some biological insights possible through relative quantitation in Nuc-MS.

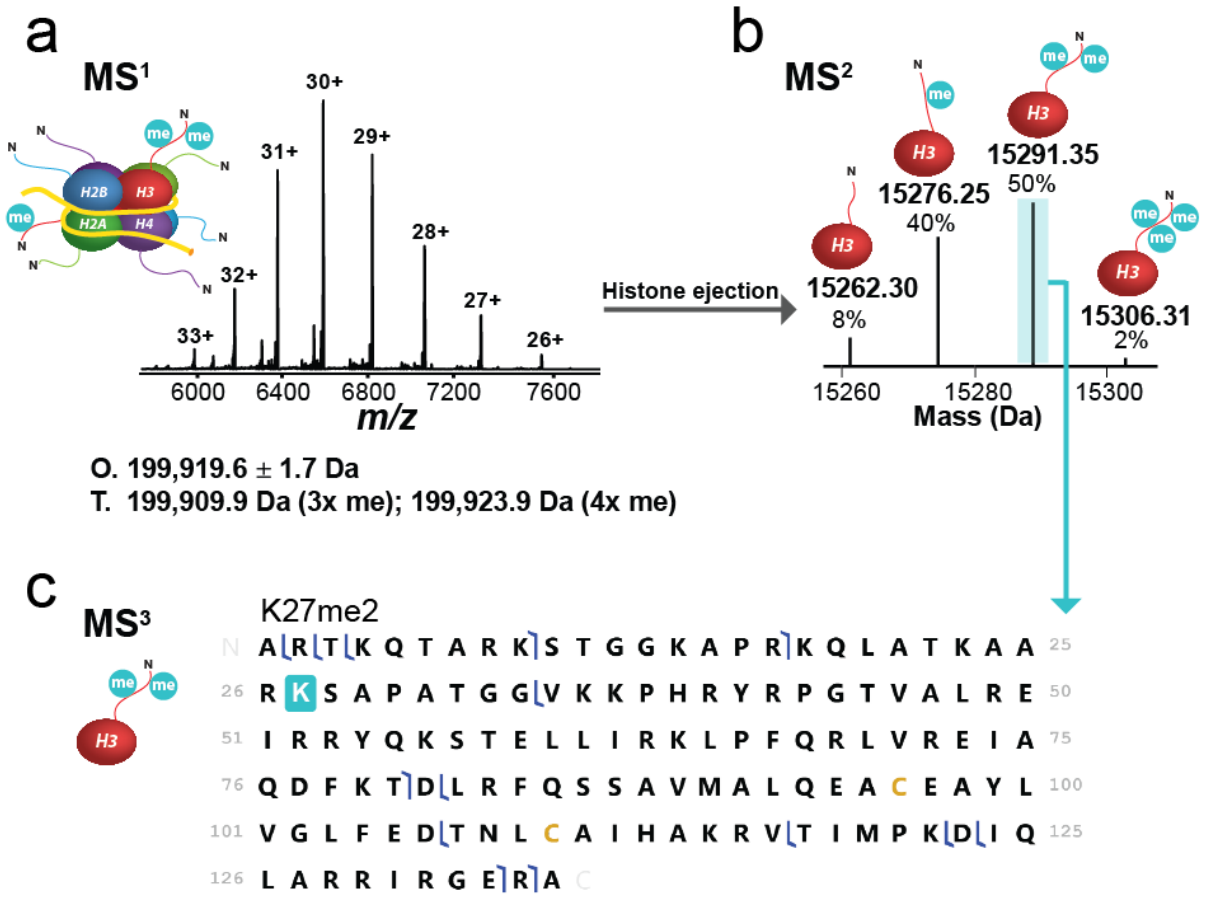


Supplemental Figure 5. Nuc-MS of synthetic nucleosomes after treatment with PCAF acetyltransferase. (a) MS[1]: the full charge state distribution of a synthetic nucleosome acetylated with PCAF (492-658) *in vitro* (O, observed average mass and precision at 1σ; T, theoretical mass for doubly-acetylated nucleosome; Δ*m*, error). * Peaks denote a nucleosome species with a shorter DNA strand. **(b)** MS[2]: spectral region reporting monoisotopic mass values for histone H3 proteoforms ejected from acetylated nucleosomes in panel **a**. **(c)** MS[3]: graphical fragment map of acetylated H3, with complete localization of K14ac, consistent with the activity of PCAF (see supplemental discussion on this figure).[8] Fragments were asserted and manually verified using TDValidator.

Supplemental Discussion of nucleosome modification with PCAF in vitro. A synthetic unmodified nucleosome was assembled and treated with the catalytic domain of PCAF/KAT2B (aa 492-658) for five minutes at room temperature. MS[1] data on the resulting products yielded an

observed average mass of 199,947.8 ± 15.4 Da, which is -4.1 Da from the theoretical mass of a doubly-acetylated nucleosome (**Supplemental Fig. 5a**). This mass error likely indicates an unresolved mixture of acetylated nucleosomes. MS^2 thus helps to elucidate the nature of these acetylation events (**Supplemental Fig. 5b**). These data show masses corresponding to a mixture of 45% unmodified and 55% mono-acetylated H3, and also reveal significant oxidation. Thus, if we consider a mass addition from histone oxidation, the MS^1 and MS^2 data together suggest that the mixture consists of approximately equal amounts of doubly- and singly-acetylated nucleosomes, and up to 20% of unmodified nucleosomes.

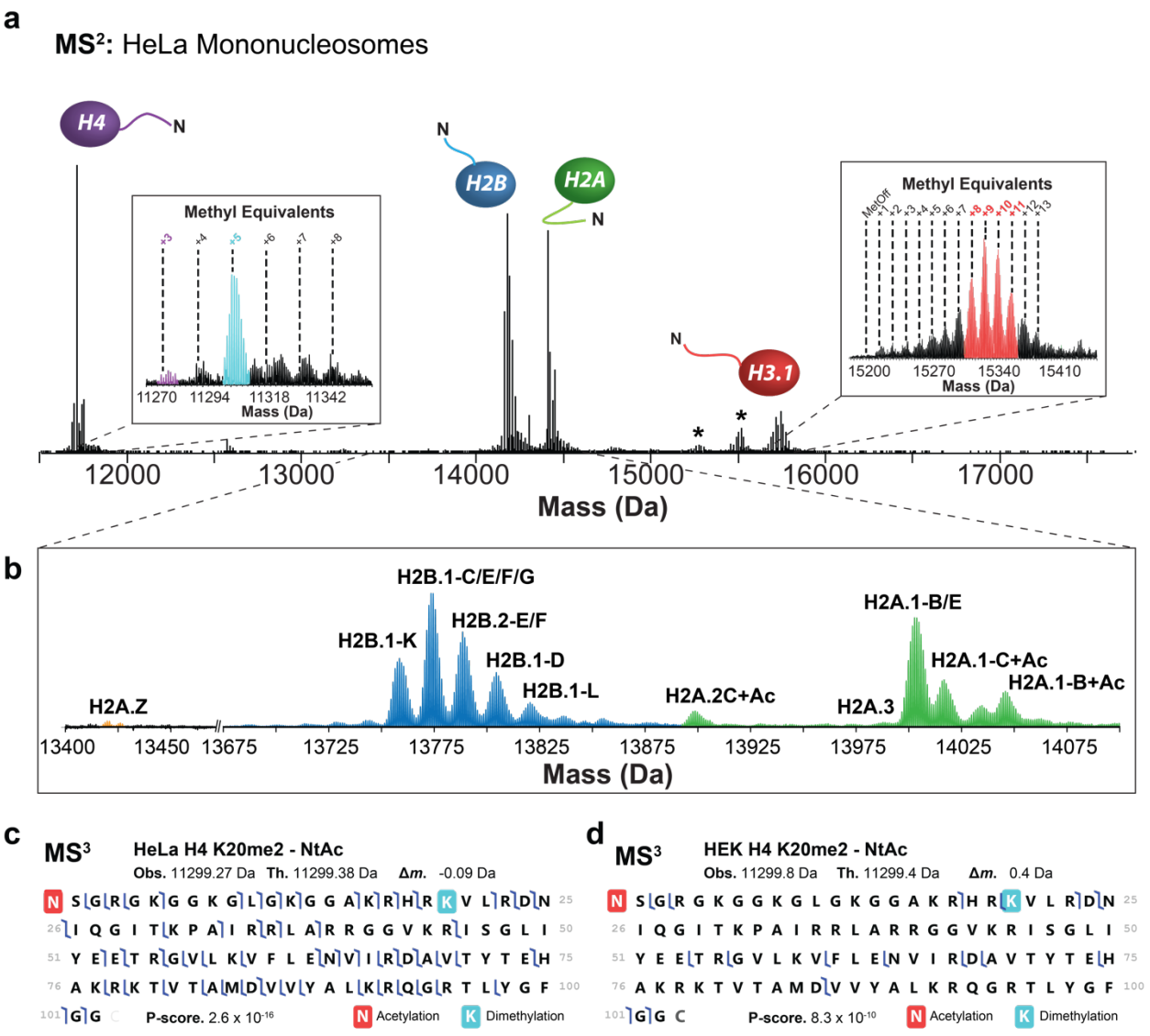
Finally, the MS^3 data can be used to localize the site of acetylations within H3. Hence, these data were converted into a fragmentation map revealing acetylation at K14, a proteoform consistent with PCAF activity (**Supplemental Fig. 5c**).[8] Inspection of the data using TDValidator and mMass confirmed the K14ac site, and that no other positional acetylation isomer was present down to ~5% relative abundance (when nucleosomes are *in vitro* acetylated by PCAF for 5 minutes or up to 2 hours). Previous studies of PCAF report acetylation at K9 and other sites beyond K14 when using H3 peptide substrates,[9] but almost exclusive K14 acetylation when using nucleosomal substrate,[8, 10] suggesting that substrate context yields differing results.



Supplemental Figure 6. Nuc-MS data of synthetic nucleosomes upon treatment by PRC2 methyltransferase. (a) MS[1]: full charge state distribution of a synthetic nucleosome methylated with PRC2 *in vitro* (O, observed average mass and precision at 1σ; T, theoretical mass for tri- and tetra-methylated nucleosome). **(b)** MS[2]: spectral region reporting monoisotopic neutral masses for the methylated histone H3 proteoforms ejected from PRC2-trated nucleosomes. **(c)** MS[3]: graphical fragment map of di-methylated H3, consistent with H3.1K27me2.[11] Fragments were asserted and manually verified using TDValidator.

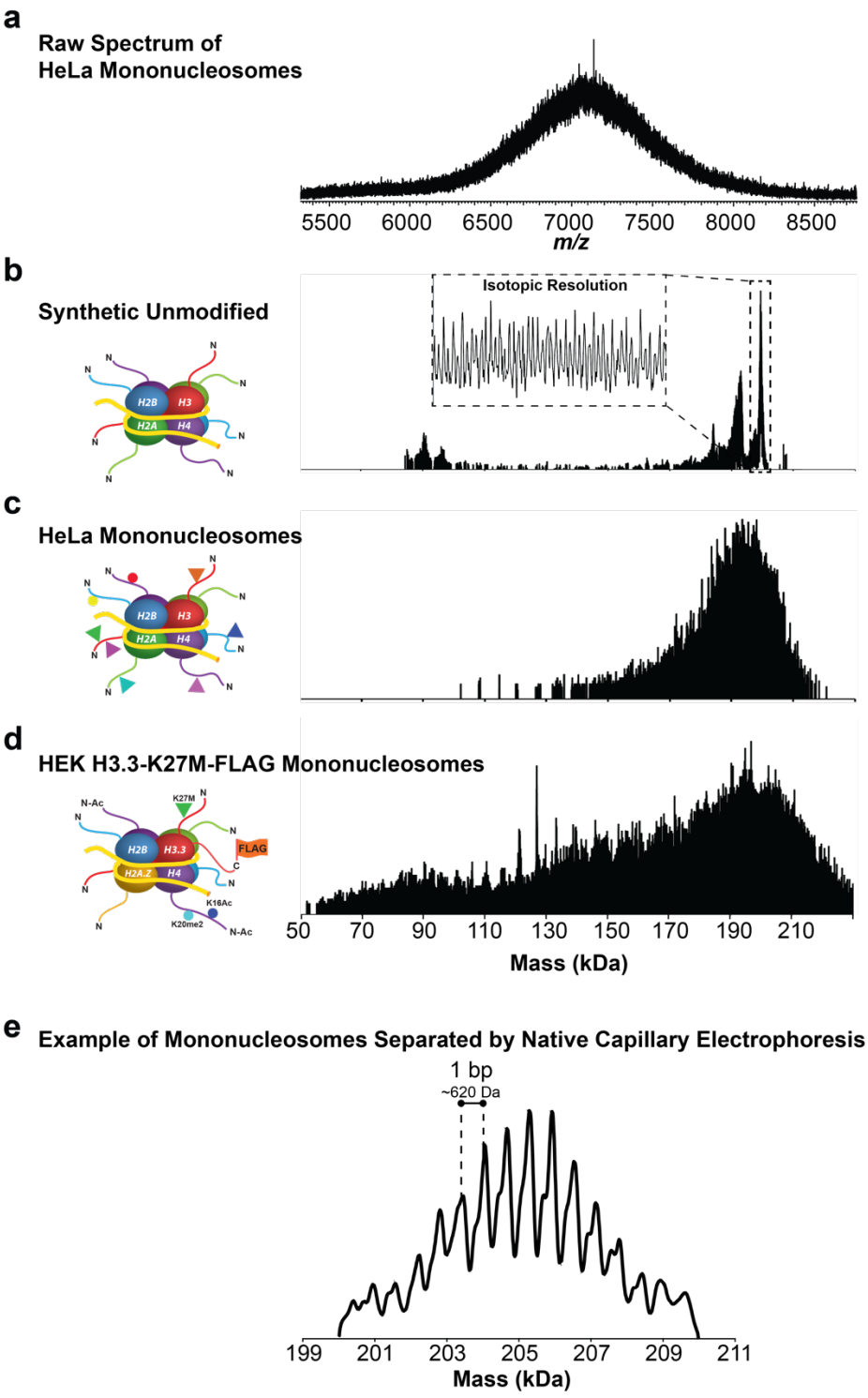
Supplemental Discussion of nucleosome modification with PRC2 in vitro. We analyzed synthetic nucleosomes incubated for 18 hours with the Polycomb repressive complex 2 (PRC2), comprising full-length EZH2, SUZ12, EED, and RbAp46/48. The MS[1] spectrum (**Supplemental Fig. 6a**) reports an observed average mass of 199,919.6 ±1.7 Da after PRC2 treatment, indicating a mixture of mostly tri- and tetra-methylated nucleosomes (199,909.9 Da and 199,923.9 Da,

respectively). The ejected proteoforms in the MS^2 (**Supplemental Fig. 6b**) reveal 8% unmodified, 40% mono-, 50% di-, and 2% tri-methylated H3. The low abundance of tri-methylation supports previous results that PRC2 is most efficient at catalyzing mono- and di-methylation.[12, 13] Moreover, the MS^2 data is consistent with a nucleosome mixture bearing on average three to four methylations. The fragmentation map created from the MS^3 data of the di-methylated H3.1 proteoform (**Supplemental Fig. 6c**) is consistent with H3K27me2 installed by the SET domain of EZH2 within the PRC2 complex.



Supplemental Figure 7. Nuc-MS analysis of endogenous mononucleosomes from HeLa cells. (a) MS^2 spectrum of ejected histones, demonstrating detection of all core histones and their proteoform distributions in the same spectrum. Asterisks denote proteoform distributions consistent with N-terminal loss of the first two (AR) and four (ARTK) amino acids from H3.1. Insets display spectral regions in the mass domain containing the isotopic distributions for selected proteoforms of histones H3.1 and H4. **(b)** Spectral region showing ejected proteoforms of H2A and H2B, including detection of H2A.Z (gold). **(c)** The H4 proteoform highlighted in cyan in (a) was isolated and fragmented to produce the fragmentation map shown, thus characterizing the proteoform as N-terminally acetylated H4K20me2. **(d)** The same N-terminally acetylated H4K20me2 proteoform in HEK cells was also fragmented and characterized.

Supplemental Table 1 contains a list of >115 histone masses ejected from HeLa mononucleosomes to illustrate the information on proteoforms and their PTMs (histone marks) that can be generally detected in a Nuc-MS experiment and supports some central claims in this work.



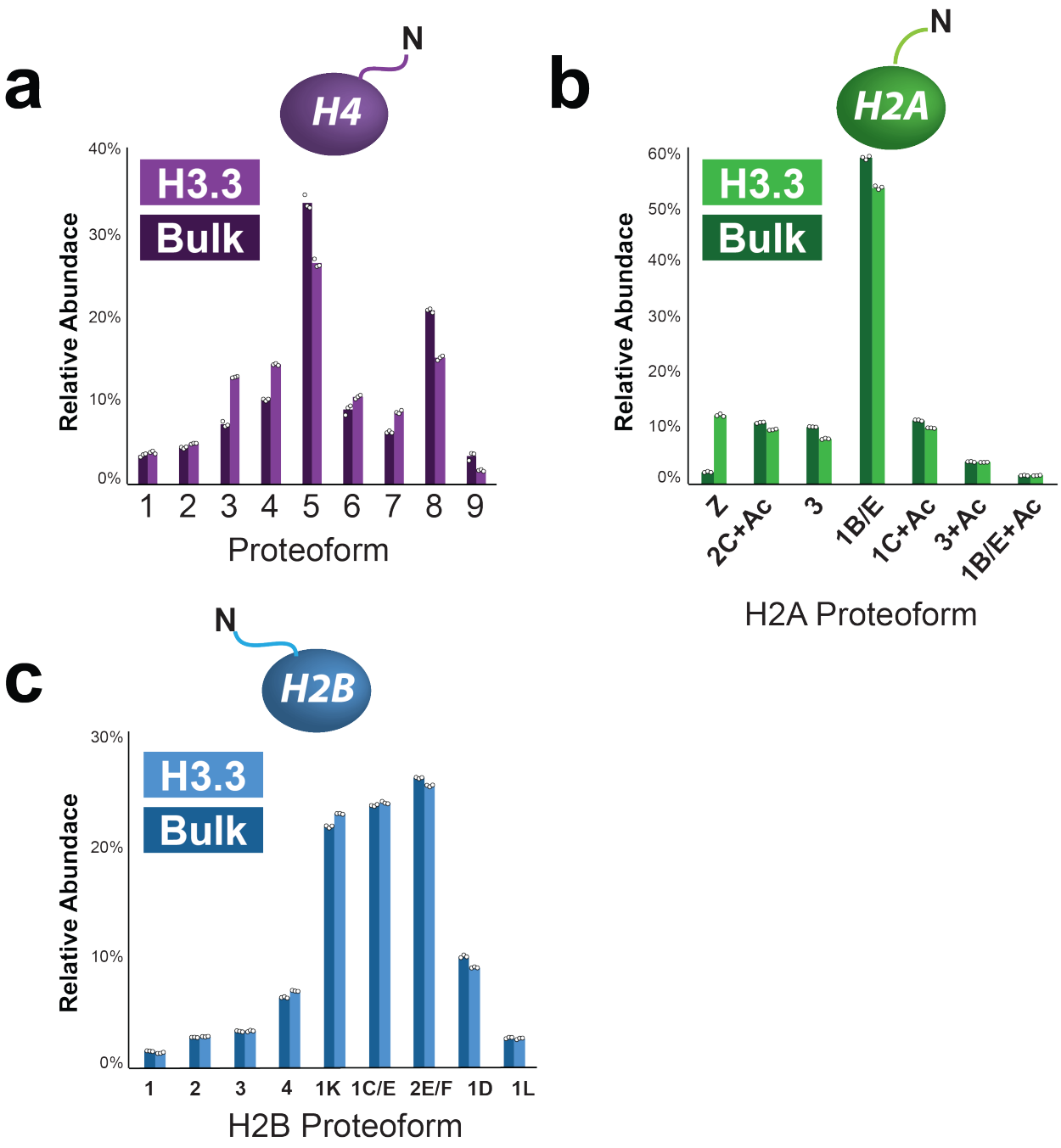
Supplemental Figure 8. (a) Intact native mass spectrum of endogenous mononucleosomes from HeLa after MNase treatment. Given the chemical complexity of modifications and DNA masses present on the population of ~100M nucleosomes contained in diploid cells, the MS^1 spectrum did not yield a successful

mass deconvolution. This motivated the use of the new approach of individual ion MS (I^2MS), as described in the methods.[14] This novel analytical technique relies on the direct assignment of charge to each individual mononucleosome ion (mainly found to be 26-28+ for the ~1 million individual ions sampled to generate these spectra). **(b)** Mass-domain spectrum of synthetic, unmodified nucleosomes showing isotopic resolution for the predominant mass, centered at 200 kDa. The smaller species correspond to nucleosomes with smaller lengths of DNA that likely result from PCR error during DNA synthesis. The low abundance peaks present <170 kDa are artifacts from charge state mis-assignment. **(c)** Mass-domain spectrum of HeLa mononucleosomes, showing a mass distribution centered between 190-200 kDa. The apparent heterogeneity – a broad mass range between 170-210 kDa – likely stems from variability in MNase digestion products, resulting in mononucleosomes with varying lengths of DNA. **(d)** Mass-domain spectrum of H3.3K27M-FLAG mononucleosomes showing a broad distribution of masses, with the most abundant species centered around 200 kDa. It is possible that the H3.3-containing mononucleosome sample includes partial nucleosome particles and other sub-species. However, the ejection experiment in Nuc-MS isolates species between 6500-8000 m/z (i.e., species ~200 kDa in size) and thus ensures that only whole mononucleosomes are dissociated and characterized. **(e)** Deconvoluted mass spectrum of an exemplary section from a Capillary Electrophoresis separation of endogenous nucleosomes, showing individual "nucleoforms" differing in mass by a base pair or single nucleotide.[15] See *Supplemental Discussion on Accurate Mass Measurements of Intact Nucleosomes*.

Supplemental Discussion on Accurate Mass Measurements of Intact Nucleosomes. The mass measurement of the various mononucleosomes shown in **Supplemental Figure 8** demonstrate the immense complexity presented by chromatin. The current instrumentation is capable of accurately measuring the masses of nucleosomes, as shown by the isotopically-resolved synthetic nucleosome. Nonetheless, the endogenous heterogeneity challenges precise nucleosome characterization and precludes tandem MS on individual, isobaric and molecularly distinct nucleosomes, or '*nucleoforms'*. For this reason, the platform cannot yet rigorously connect a single '*nucleoform'* to the histone proteoforms present in it (note that isobaric and isomeric forms can exist at all levels: MS^1, MS^2 and MS^3).

We thus focused on the value, relative simplicity and quantitative data provided in the MS^2 step: simultaneous detection of all histone proteoforms present in a population of nucleosomes above a ~1% abundance threshold. Much information about nucleosome composition can now be accessed and reassembled from the quantitative Nuc-MS readout, namely, a survey of the chromatin landscape. In the H3.3K27M system (**Fig. 3**), for example, we describe the average mono-nucleosome composition by aggregating the quantitative MS^2 and MS^3 results, enabling us to provide comparative and reliable results from which functional insights can be inferred.

With the goal of better understanding nucleosome heterogeneity, we separated endogenous mononucleosomes based on their charge density using capillary electrophoresis (CE) coupled to Nuc-MS.[15] This Nuc-CE-MS platform was optimized with synthetic Nucs to reduce sample requirements by >100-fold relative to direct infusion with a low attomole limit of detection. Using this setup, we resolved 60 molecularly distinct '*nucleoforms'*, at base pair resolution (~620 Da) and discovered that the predominant nucleosome heterogeneity at our MS^1 level comes from differences in DNA-strand length. Since DNA is negatively charged, the CE sufficiently separated the nucleosomes such that Nuc-MS could report distinct nucleosome species with an overall mass range of 192,332 - 228,854 Da (an example section of the separation is shown in **Supplemental Fig. 8e**). We are reasoning that these ~60 *nucleoforms* are generated by inconsistent activity of Micrococcal nuclease (MNase), the enzyme used to digest native chromatin in standard biochemistry. As a result, the specific time and concentration of MNase used in the preparation created a distribution of mononucleosome sizes that differ by a base pair or single nucleotide. This artificial 620 Da shift obscures the histone PTMs of interest to biologists (i.e., methylation, +14 Da; acetylation, +42 Da; and phosphorylation, +80).

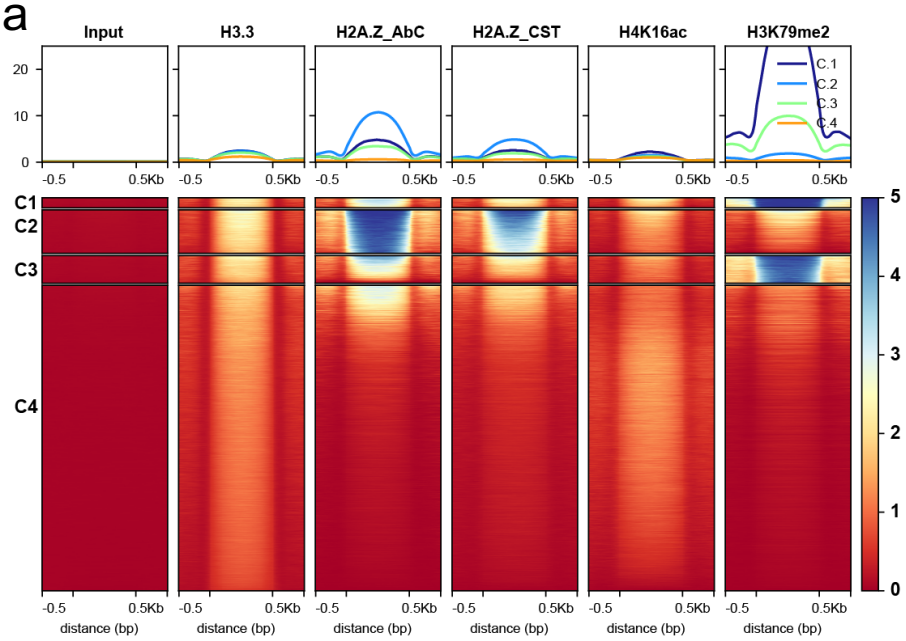


Supplemental Figure 9. Quantitation of **(a)** H4, **(b)** H2A, and (**c**) H2B proteoforms ejected from HEK bulk nucleosomes (Bulk) compared to those enriched for H3.3-FLAG (H3.3). Data points from three measurement replicates (n = 3) are displayed as a scatter plot and ordered left to right according to data file of origin. The mean integrated peak area for each histone proteoform is represented with the histogram. The sum of the abundances of all proteoforms in each histogram equals 100%. H4

proteoforms are displayed in order of methyl equivalence and are listed from low to high mass as proteoforms 1-9 (as in **Fig. 3c**). Additional proteoform assignments can be found in **Supplemental Table 1**.

Supplemental Discussion on Quantitation of Histones ejected from HEK bulk chromatin. Nuc-MS detected an elevation of unmodified H4 (3), H4K20me1 (4), and acetylated H4K20me1 (7) proteoforms in H3.3 nucleosomes by 80 ±3.4%, 40 ±0.7%, and 37.5 ±1.2%, respectively (with p = $7.5x10^{-6}$, $2.3x10^{-6}$, and $6.8x10^{-5}$, and numbers representing aforementioned H4 proteoform order in the histogram). Overall, H4 proteoforms have a higher abundance of me0 and me1 states in H3.3 nucleosomes compared to bulk, suggesting incorporation of newly expressed H4 and thus higher H3-H4 turnover.[16, 17]

a

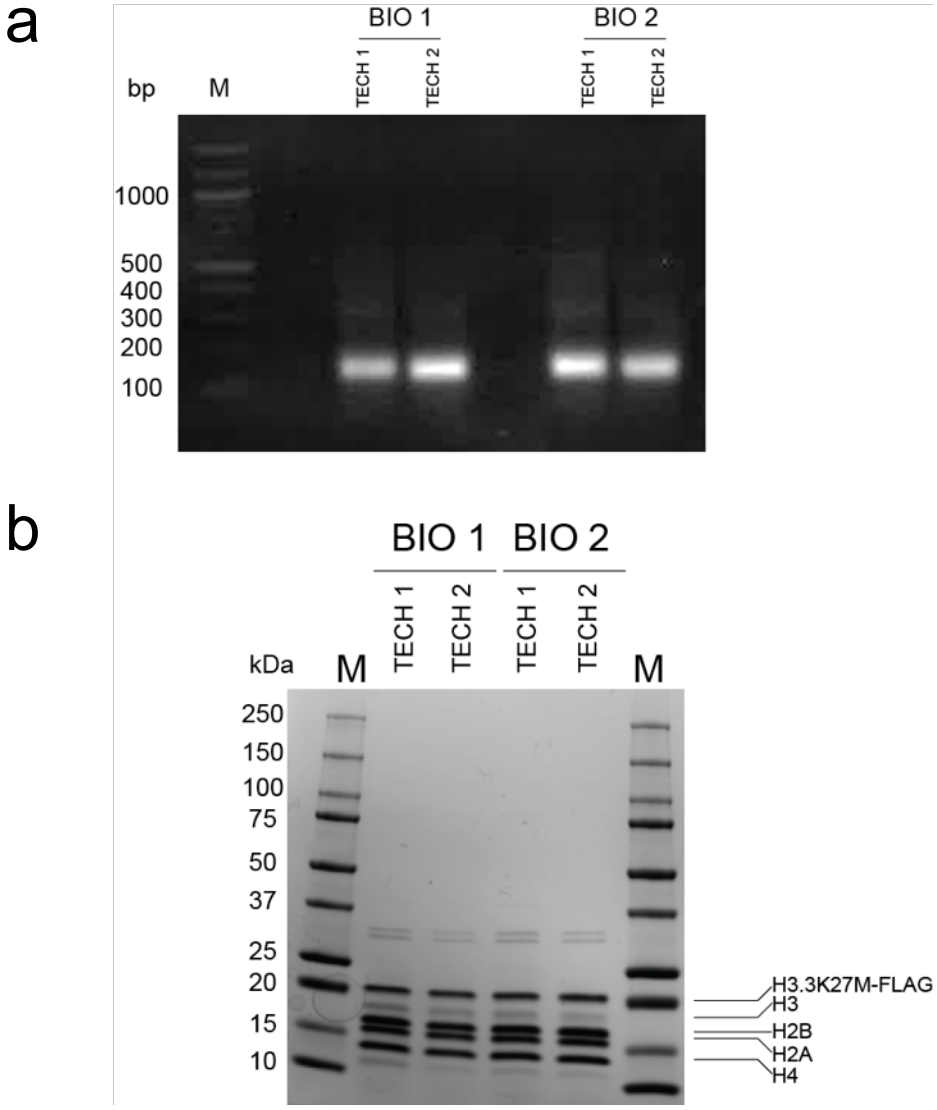


b

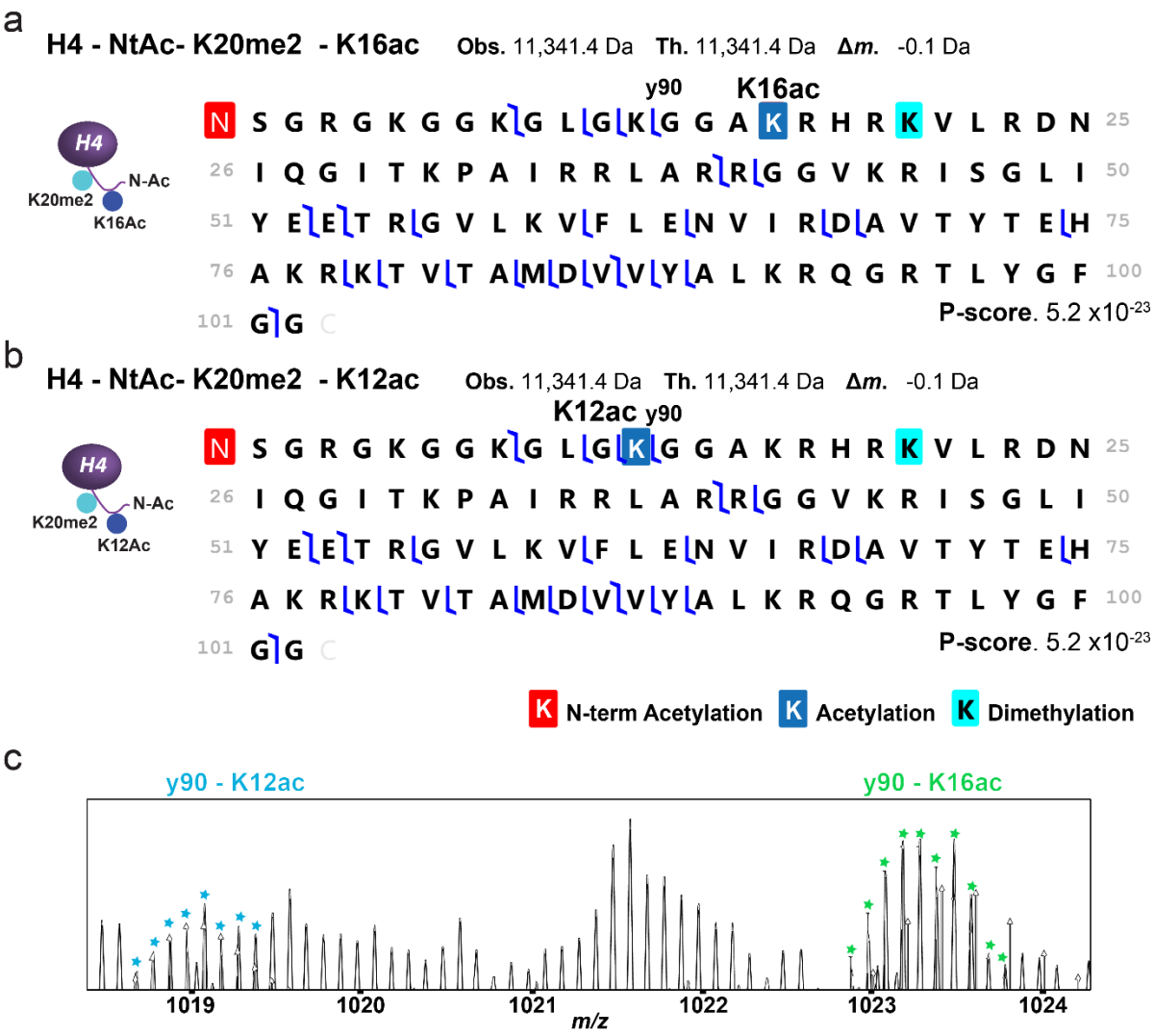
	Cluster 1		Cluster 2		Cluster 3		Cluster 4	
Annotation	**# of peaks**	**%Total**	**# of peaks**	**%Total**	**# of peaks**	**%Total**	**# of peaks**	**%Total**
3UTR	10	0.5%	33	0.4%	26	0.5%	1827	3.2%
miRNA	0	0.0%	0	0.0%	0	0.0%	7	0.0%
ncRNA	17	0.9%	98	1.2%	37	0.7%	434	0.8%
TTS	54	2.9%	124	1.5%	109	2.1%	2653	4.6%
pseudo	0	0.0%	10	0.1%	5	0.1%	76	0.1%
Exon	237	12.6%	675	8.3%	635	12.3%	7889	13.8%
Intron	**889**	**47.2%**	**1827**	**22.5%**	**2346**	**45.3%**	**25530**	**44.7%**
Intergenic	24	1.3%	1127	13.9%	140	2.7%	**12903**	22.6%
Promoter	**521**	**27.7%**	**3718**	**45.8%**	**1587**	**30.7%**	**5320**	**9.3%**
5UTR	132	7.0%	505	6.2%	291	5.6%	471	0.8%
snoRNA	0	0.0%	0	0.0%	0	0.0%	0	0.0%
scRNA	0	0.0%	0	0.0%	0	0.0%	0	0.0%
rRNA	0	0.0%	0	0.0%	0	0.0%	0	0.0%
Total	1884		8117		5176		57110	

Supplemental Figure 10. (a) Heatmap centered on H3.3 peaks ± 0.5 kb showing the correlation of *ChIP-seq* signal for input, H3.3, H3K79me2, H2A.Z (all antibodies in Methods), and H4K16ac, and extending 0.5 kb in each direction. **(b)** Table summarizing genetic composition of each cluster listed in **(a)**.

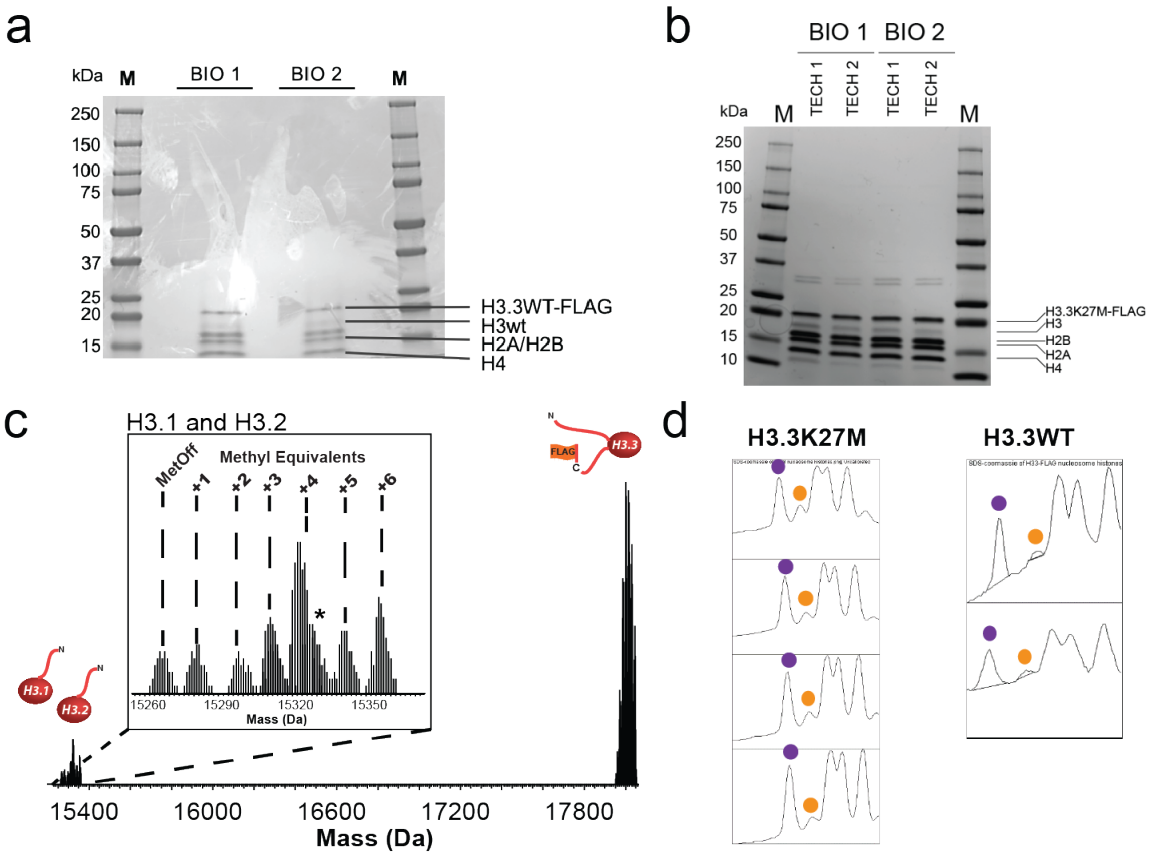
Supplemental Discussion on ChIP-seq results. H4K16ac shows the strongest Pearson correlation to H3.3 (0.48, **Fig. 3f**) with signal intensity across the whole genome and no cluster-specific enrichment (**Fig. 3h**). In contrast, the heatmap reveals that H3K79me2 signal is mostly localized to promoters and introns (Cluster 2, **Supplemental Figs. 10, 19**). Close inspection of a representative gene in **Supplemental Fig. 18** shows that H3K79me2 displays a prominent promoter-associated peak that co-localizes with H3.3 and H2A.Z;[18] however, unlike these variants, H3K79me2 signal extends beyond the transcription start site into the gene body.



Supplemental Figure 11. DNA and protein gels showing masses consistent with FLAG-tagged H3.3K27M mononucleosomes **(a)** and its constituent histone proteoforms **(b)**.

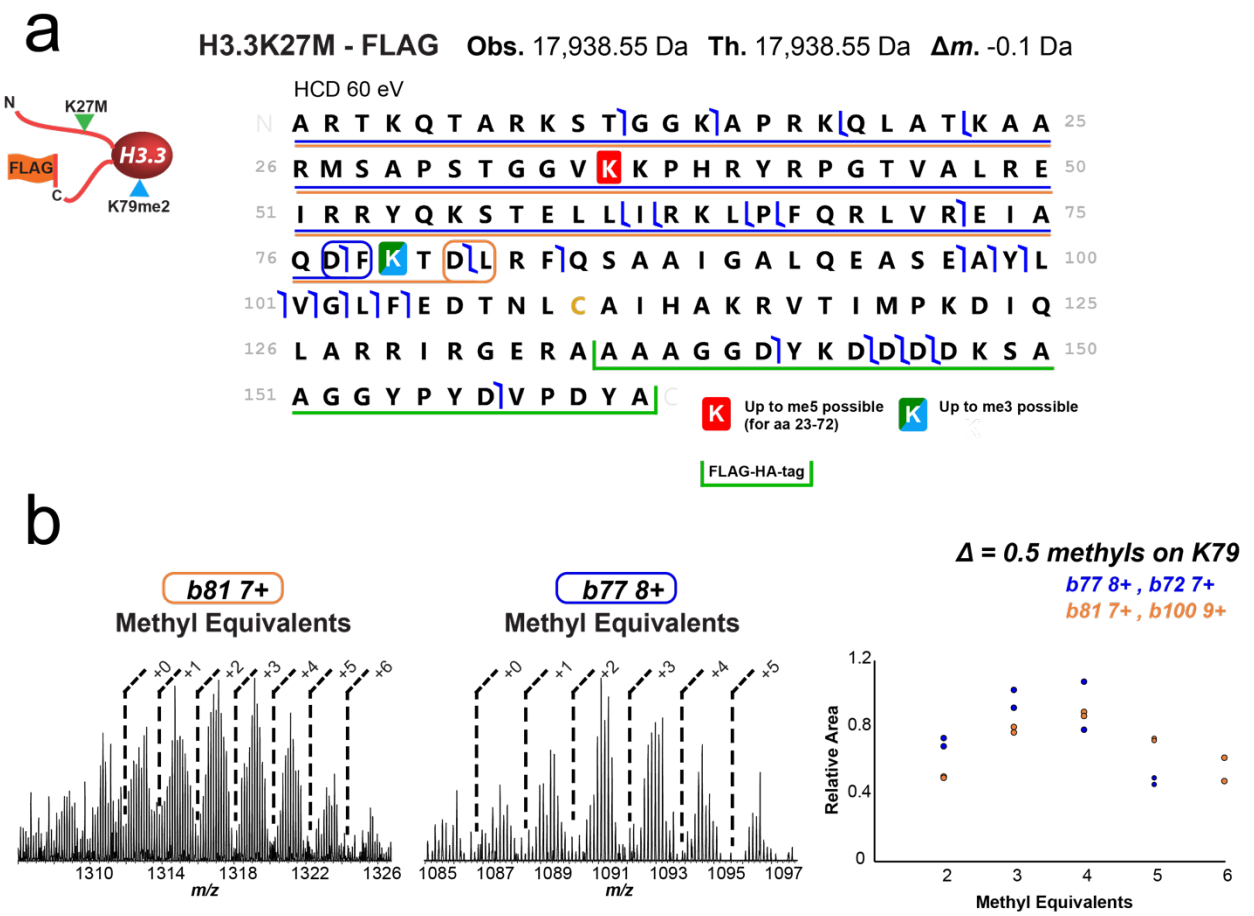


Supplemental Figure 12. Fragmentation maps of H4 proteoforms detected from H3.3K27M mononucleosomes. Fragmentation data support H4K20me2 co-existing with K16 and K12 acetylation (**a** and **b**, respectively). Moreover, diagnostic ion y90 reveals that H4K16ac is the predominant proteoform, as shown by the isotopic distributions in the TDValidator output[3] of panel **(c)**.



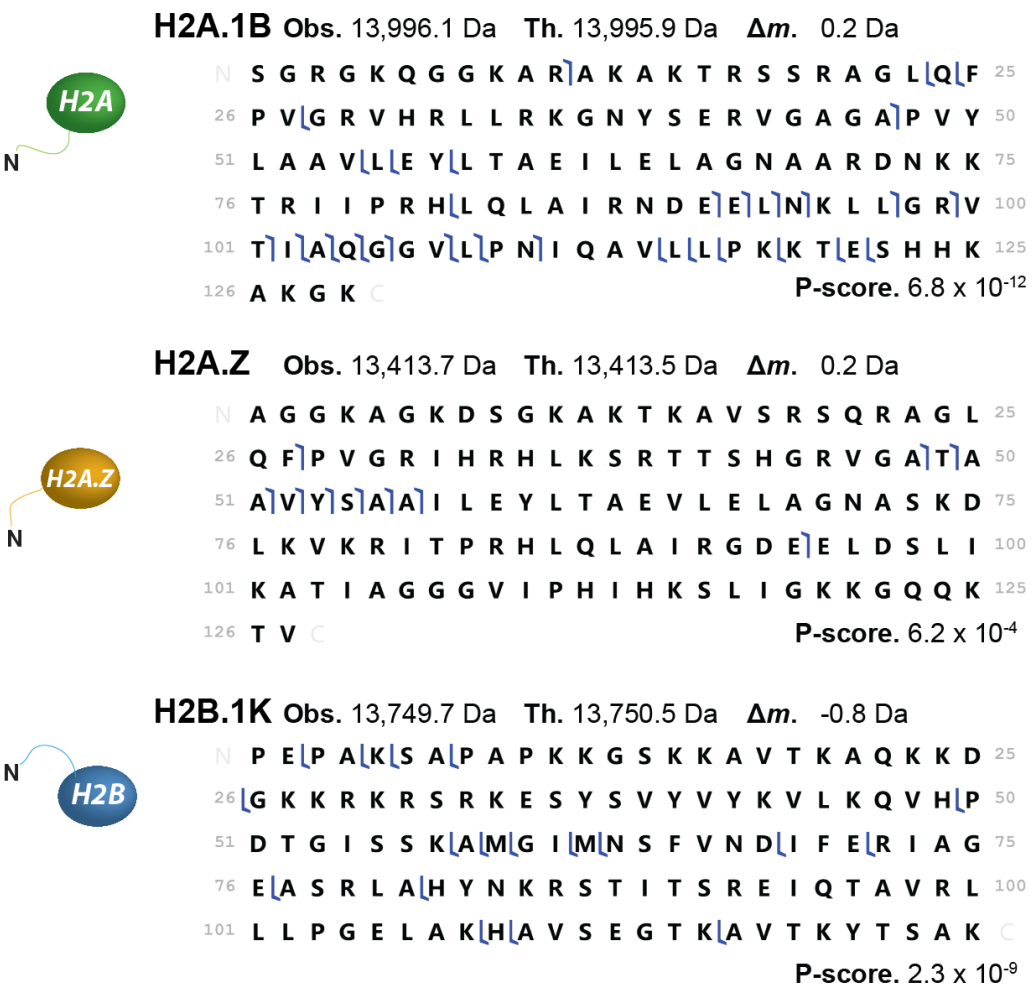
Supplemental Figure 13. Coomassie-stained SDS-PAGE of histones enriched after FLAG IP of (**a**) H3.3WT (**b**) H3.3K27M nucleosomes, showing significantly lower abundance of H3 WT relative to the FLAG-tagged counterpart. (**c**) Nuc-MS detects low abundance distributions for wild-type H3.1 and H3.2 proteoforms ejected from H3.3WT-FLAG nucleosomes. The inset displays the methylation series for H3.2 (9.9% of total H3), and a minor series for H3.1 (1.7% of total H3, H3.1me4 proteoform denoted with an asterisk). H3.3WT-FLAG makes up 88.4% of total H3 signal. Only the H3.2me4 proteoform was detected in H3.3K27M-FLAG nucleosomes, suggesting that the full distribution of proteoforms may have been below the limit of detection. (**d**) ImageJ-generated traces of the intensities of the Coomassie-stained bands in the gels of **a** and **b**. The areas underneath the purple-labeled peaks (H3.3K27M-FLAG or H3.3WT-FLAG) and orange-labeled peaks (H3WT) were integrated to determine the relative ratios of these proteoforms. These were found to be 88% H3.3K27M-FLAG compared to 12% H3WT (n = 4), and 90% H3.3WT-FLAG compared to 10% H3WT (n = 2).

Supplemental Discussion on Nucleosome Symmetry. Given the low expression levels of H3.3-FLAG constructs relative to total H3 upon doxycycline induction,[19] the predominance of homotypic (2x) FLAG-tagged nucleosomes may have resulted from avidity bias during immunoprecipitation. Nonetheless, the ejection of wild-type H3.1 and H3.2 (12% of total H3) from H3.3WT-FLAG reveals that up to 24% of nucleosomes could be heterotypic (i.e. contain one copy of H3.3WT-FLAG and one copy of wild-type H3.1 or H3.2). Moreover, the detection of six methylated H3.1 and H3.2 proteoforms in H3.3WT-FLAG nucleosomes demonstrate that Nuc-MS can quantitatively interrogate the homotypic vs. heterotypic nature of nucleosomes.



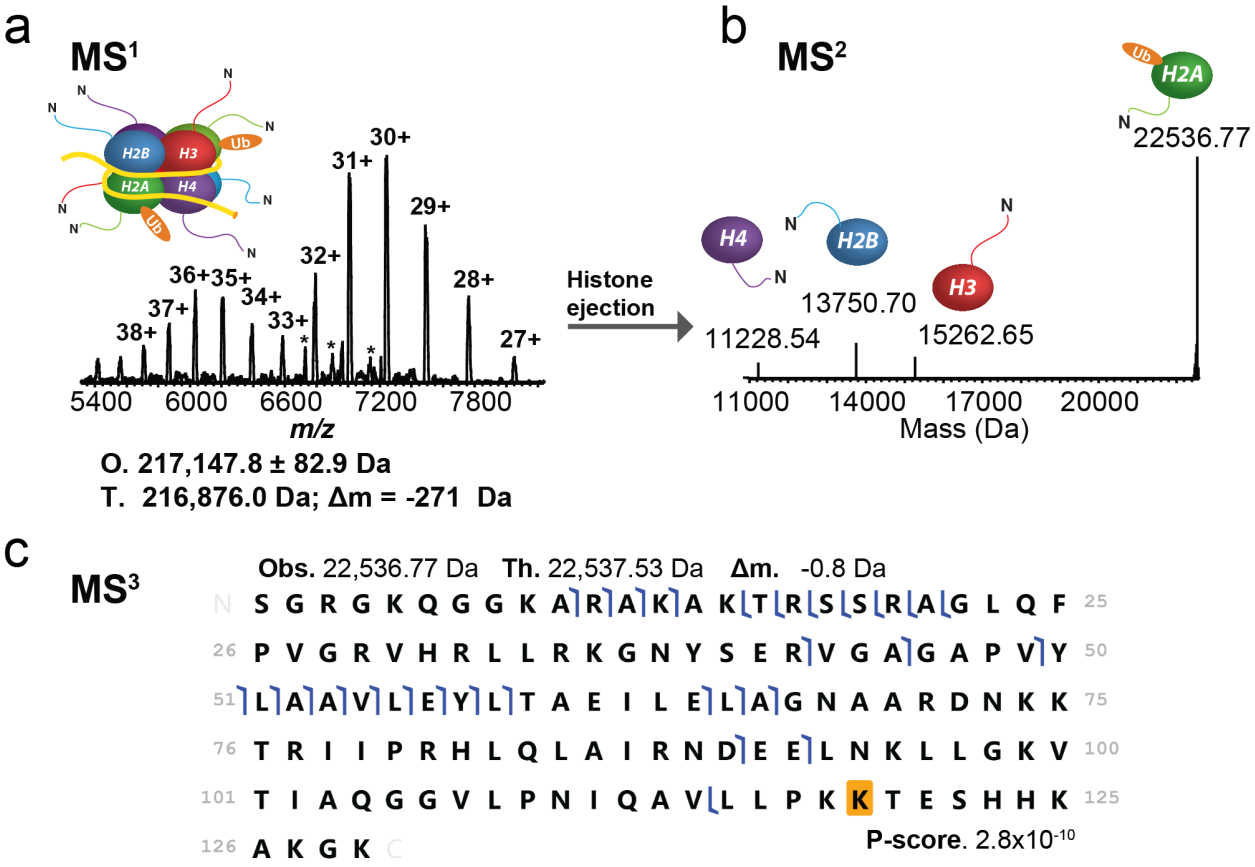
Supplemental Figure 14. Fragmentation of H3.3K27M by MS^3 at HCD of 60 eV. (a) Analysis of backbone fragment ions of H3.3K27M proteoform enables partial localization of methylation. First, the data are consistent with two to three methyl equivalents partially localized to residues 23-72. This is most likely an elevation of K36me2/3, which has been detected previously with the K27M mutation but only asserted to co-occur in the same nucleosomes after extensive *ChIP-Seq*.[20] **(b)** Subtracting the average methyl equivalence for diagnostic ions *b81* and b100 (orange) from that of *b77* and *b72* (blue), results in an average decrease of 0.5 methyl equivalents when moving from the *b81* to *b77* fragment ion pair (n = 4). This indicates a 50% probability of mono-methylation and 25% probability of di-methylation between amino acids 78-81 (FKTD); K79 being the most likely residue to contain this PTM. A 25% probability of K79me2 constitutes at least 15-fold enrichment relative to bulk H3 (1-2%). Additional fragment ions (*y100-101* 8+, *y96* 8+ & 9+, and *y97* 9+) further localize the assignment of K79me2 being co-occurring with H3.3K27M and part of the reprogrammed code on these mutant nucleosomes. The data do not support PTMs N-

terminal to H3K23, indicating that these are low abundance and below the limit of detection. The fragmentation result aggregates two measurement replicates.



Supplemental Figure 15. Fragmentation maps showing identification of H2A, H2A.Z, and H2B proteoforms after ejection from intact H3.3K27M mononucleosomes by Nuc-MS. Fragment ions were manually validated using TDValidator. P-scores were calculated using ProSight Lite.

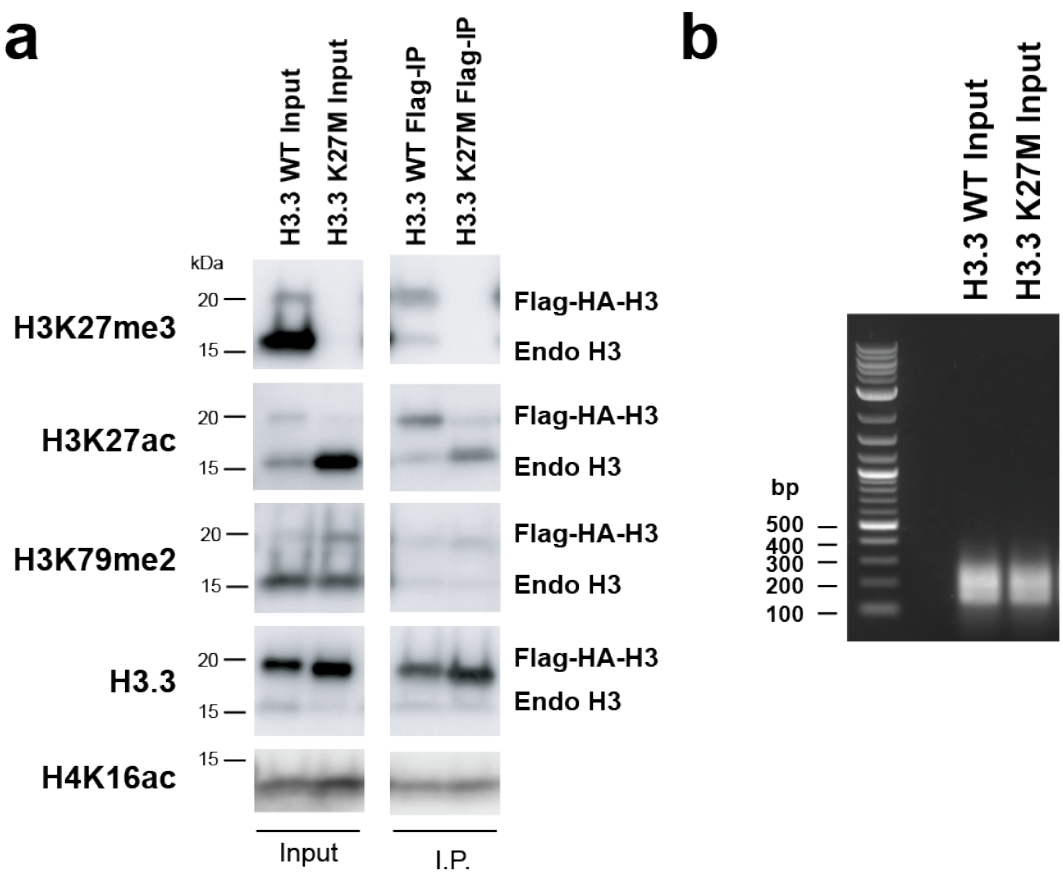
As mentioned in the main text (**Fig. 3e)**, examination of H2A proteoforms exposed elevated acetylation of H2A.1C, and a three-fold decrease in H2A.Z abundance in H3.3K27M-FLAG nucleosomes compared to H3.3WT-FLAG (13.7 ±0.03% in H3.3WT vs. 5.1 ±0.2% in H3.3K27M; $p = 1.7 \times 10^{-4}$). Moreover, H2A.Z acetylation at 3.3 ±0.4% in H3.3WT nucleosomes was undetectable in H3.3K27M nucleosomes.



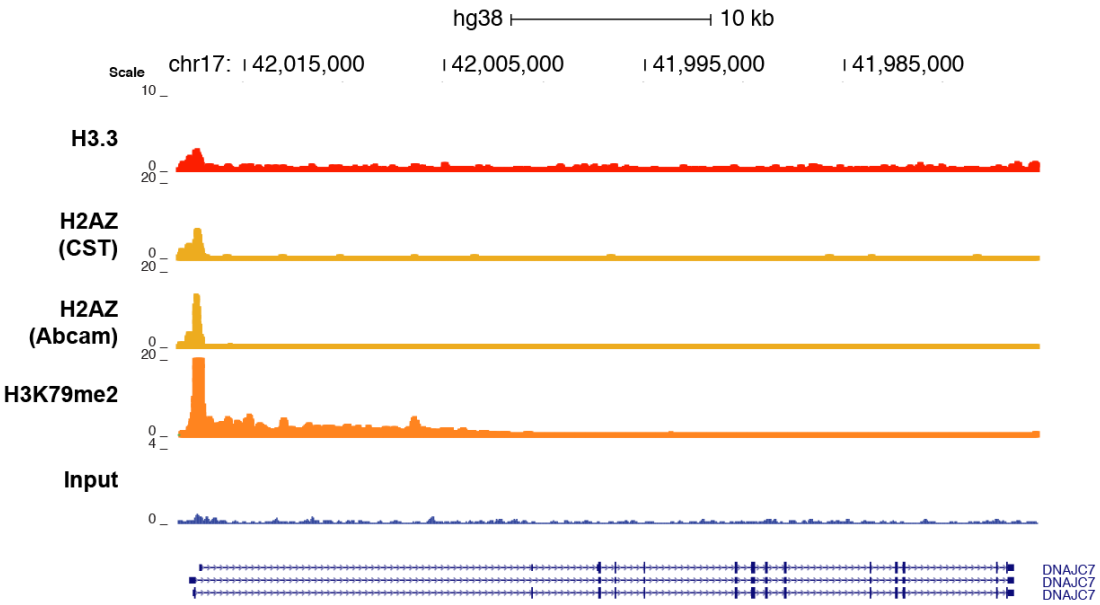
Supplemental Figure 16. Nuc-MS data of a synthetic nucleosome ubiquitinated at H2AK119 serves as a positive control to detect H2A ubiquitination in endogenous samples. (a) MS^1: full charge state distribution of a synthetic mono-ubiquitinated nucleosome (O, observed average mass and precision at 1σ; T, theoretical mass for a homotypic nucleosome carrying two mono-ubiquitins on H2AK119). **(b)** MS^2: spectral region reporting monoisotopic neutral masses for the ubiquitinated histone H2A and ejected proteoforms of the other core histones. **(c)** MS^3: graphical fragment map of ubiquitinated H2A, consistent with mono-ubiquitination at lysine 119. Fragments were manually validated using TDValidator. P-scores were calculated using ProSight Lite.

Supplemental Discussion on ubiquitinated nucleosomes. H3.3K27M is presumed to dysregulate Polycomb Repressive Complex (PRC) recruitment by blocking or tethering models,[21] which might be expected to impact local levels of H2AK119ub. As a positive control,

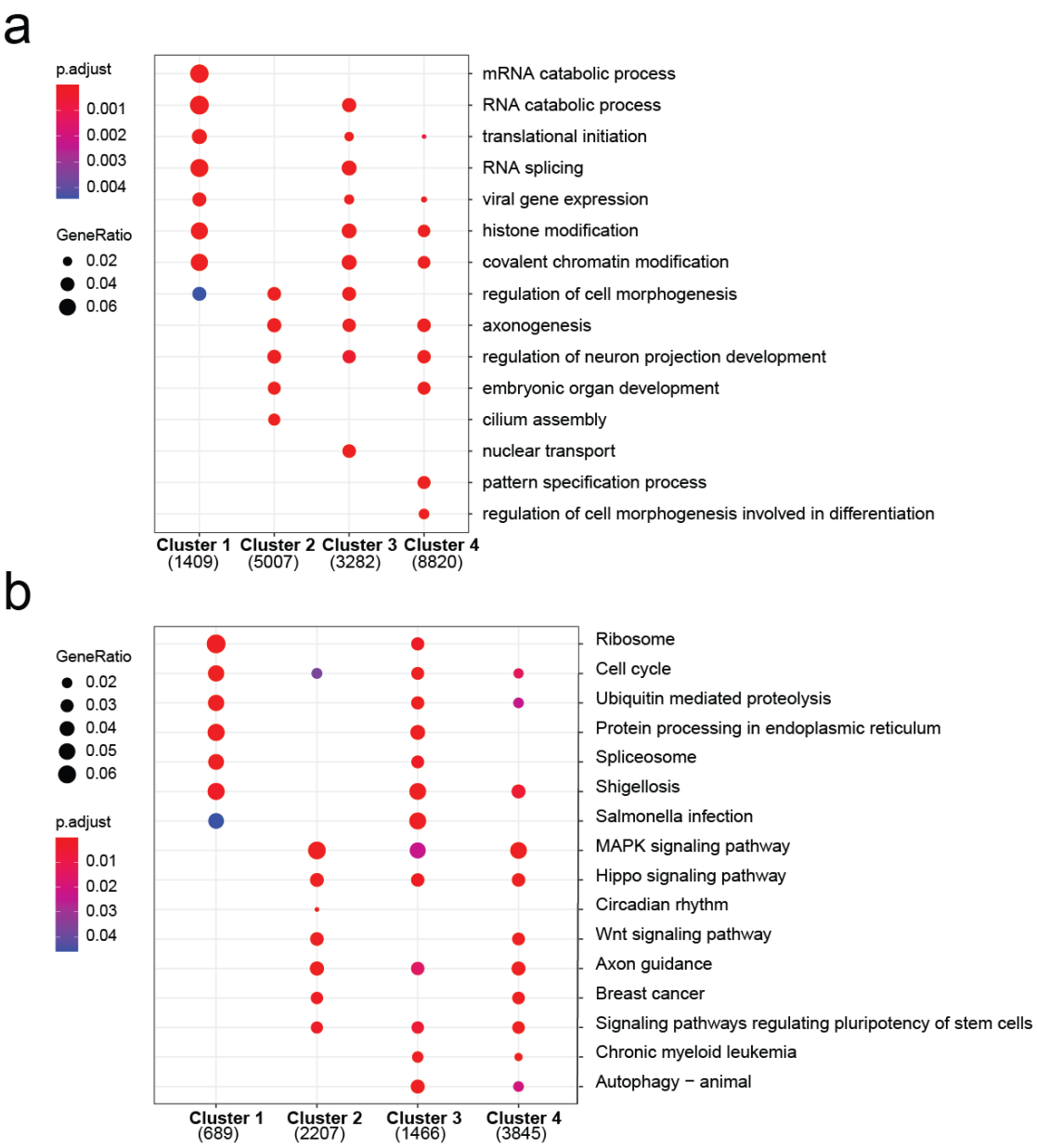
we analyzed mono-ubiquitinated nucleosomes created *in vitro* and found that the new method efficiently detects this large PTM at both the nucleosome and histone levels; however, we did not detect ubiquitinated histones in the endogenous samples. Therefore, we have confidence in the sensitivity of Nuc-MS and can assert that H2A or H2B ubiquitination does not occur on >1% of H3.3WT or K27M nucleosomes in any of the HEK293T cells analyzed by Nuc-MS.



Supplemental Figure 17. (a) Western blot analyses of input and IP samples confirming detection of H3K79me2 and H4K16ac in immuno-enriched samples. Consistent with previous reports on the H3.3K27M system, the H3K27me3 mark is abolished in H3.3K27M samples but is detected in H3.3WT. Moreover, there is a global increase of H3K27ac in the pre-IP sample of H3.3K27M-expressing cells as shown previously.[20] **(b)** DNA gel showing extraction of mononucleosomes from corresponding cell lines.



Supplemental Figure 18. Expansion of a representative gene showing peak overlap for H3.3, H2A.Z and H3K79me2. Moreover, the H3K79me2 track shows that this mark is propagated into the gene body, unlike H3.3 and H2A.Z signals, which are mostly confined to the transcription start site.



Supplemental Figure 19. Gene composition of each cluster presented in the *ChIP-seq* heatmaps in main text **Figs. 2** and **3** and **Supplemental Fig. 10**, described by the gene ontology terms **(a)** biological process or **(b)** molecular function. These dot plots were made using clusterProfiler package and they represent the results of an Over Representation Analysis (ORA; Boyle et al. 2004). ORA helps determine if known biological functions or processes are over-represented in an experimentally derived gene list.

Supplemental Table 1. List of monoisotopic (MI) masses ejected from HeLa mononucleosomes and assigned to >100 histone proteoforms. This list provides information on proteoforms and their defining post-translational modifications (PTMs) detected by Nuc-MS and supports some of the assertions in this work. Data and mass assignments for proteoforms were manually curated; a minority of proteoforms were characterized by tandem MS (MS^3).

Histone Type	Proteoform Assignment / with PTMs	Deconvoluted MI Mass	Theoretical MI Mass	UniProt Accession
H2A	H2A.Z	13412.93	13413.51	P0C0S5
	H2A.X	15003.3	15004.37	P16104
	H2A.X+Ac	15047.18	15046.37	
	H2A.J	13878.41	13879.89	Q9BTM1
	H2A.2-C/Q	13846.52	13848.8	Q16777
	H2A.2-C/Q+Ac	13890.62	13890.8	
	H2A.2-C/Q+2xAc	13934.73	13932.8	
	H2A.2-C/Q+3xAc	13973.79	13974.8	
	H2A.2-A/B	13955.74	13955.85	Q6FI13/Q8IUE6
	H2A.1-C	13967.68	13965.9	Q93077
	H2A.1-C+Ac	14008.84	14007.91	
	H2A.1-C+2xAc	14049.85	14049.91	
	H2A.3	13980.7	13981.89	Q7L7L0
	H2A.3+Ac	14023.73	14023.89	
	H2A.1-B/E/O	13995.78	13995.91	P04908
	H2A.1-B/E/O+Ac	14037.77	14037.91	
	H2A.1A	14090.73	14093.87	Q96QV6
	H2A.V	13369.4	13369.49	Q71UI9
	H2A.Bbd type 1	12689.72	12689.61	P0C5Y9
H2B	H2Bpfr1	13691.91		
	H2Bpfr2	13707.97		
	H2Bpfr3	13719.97		
	H2Bpfr4: 1-K (-16.0)	13735.53	13734.52	O60814
	H2B.1-A	14025.74	14027.7	Q96A08
	H2B.1-C/E/F/G/I/O	13765.44	13766.52	P62807
	H2B.1-B	13812.41	13810.54	P33778
	H2B.2-F	13780.43	13780.53	Q5QNW6
	H2B.1-K	13750.38	13750.52	O60814

	H2B.1-D	13796.42	13796.53	P58876
	H2B.1-L	13812.41	13812.56	Q99880
	H2B.1-M	13850.46	13849.59	Q99879
	H2B.1-H	13750.38	13752.50	Q93079
	H2B.2-E1	13750.38	13751.95	A0A2R8Y619
	H2B.F-M (A-trunc)	16790.48	16788.60	P0C1H6
	H2B.F-M (A-trunc) + Ac	16834.50	16830.61	
H4	H4	11227.24	11229.34	P62805
	H4K20me1	11242.22	11243.35	
	H4K20me2	11255.26	11257.36	
	H4-NtAc	11271.15	11271.37	
	H4K20me1-NtAc	11285.34	11285.38	
	H4K20me1-NtAc+Ph	11363.23	11365.28	
	H4K20me2-NtAc	11299.27	11299.39	
	H4K20me2-NtAc+Ph	11381.27	11379.29	
	H4K20me2-NtAc+me1	11314.26	11313.4	
	H4K20me2-NtAc+me2	11325.35	11327.41	
	H4K20me2K16ac-NtAc	11341.26	11341.42	
	H4K20me2K16ac-NtAc	11341.26	11341.42	
	H4-NtAc+me6	11355.24	11355.43	
	H4-NtAc+me9	11396.27	11397.46	
	H4-NtAc+me10	11411.26	11411.47	
	H4-NtAc+me11	11425.3	11425.48	
	H4-NtAc+me12	11438.28	11439.49	
H3	H3.1	15262.41	15263.44	P68431
	H3.1me1	15276.38	15277.45	
	H3.1me2	15290.38	15291.46	
	H3.1me3	15303.45	15305.47	
	H3.1me4	15318.36	15319.48	
	H3.1me5	15332.38	15333.49	
	H3.1me6	15346.41	15347.5	
	H3.1me7	15361.4	15361.51	
	H3.1me8	15374.38	15375.52	
	H3.1me9	15389.46	15389.53	
	H3.1me10	15403.44	15403.54	
	H3.1me11	15416.39	15417.55	
	H3.1me12	15430.35	15431.56	

	AR-trunc H3.1	15037.31	13036.3	
	AR-trunc H3.1me1	15049.51	13050.31	
	AR-trunc H3.1me2	15063.54	13064.32	
	AR-trunc H3.1me3	15077.19	13078.33	
	AR-trunc H3.1me4	15092.38	13092.34	
	AR-trunc H3.1me5	15105.26	13106.35	
	AR-trunc H3.1me6	15120.22	13120.36	
	AR-trunc H3.1me7	15134.27	13134.37	
	AR-trunc H3.1me8	15148.25	13148.38	
	AR-trunc H3.1me9	15161.26	13162.39	
	ARTK-trunc H3.1	14807.08	14807.16	
	ARTK-trunc H3.1me1	14821.08	14821.17	
	ARTK-trunc H3.1me2	14834.06	14835.18	
	ARTK-trunc H3.1me3	14847.17	14849.19	
	ARTK-trunc H3.1me4	14863.17	14863.2	
	ARTK-trunc H3.1me5	14877.09	14877.21	
	ARTK-trunc H3.1me6	14891.07	14891.22	
	ARTK-trunc H3.1me7	14905.14	14905.23	
	ARTK-trunc H3.1me8	14919.06	14919.24	
	ARTK-trunc H3.1me9	14934.1	14933.25	
	ARTKQT-trunc H3.1	14578.08	14578.06	
	ARTKQT-trunc H3.1me1	14589.27	14592.06	
	ARTKQT-trunc H3.1me2	14603.25	14606.06	
	ARTKQT-trunc H3.1me3	14623.1	14620.06	
	ARTKQT-trunc H3.1me4	14632.21	14634.06	
	ARTKQT-trunc H3.1me5	14649.25	14648.06	
	ARTKQTAR-trunc H3.1	14350.86	14350.92	
	ARTKQTAR-trunc H3.1me1	14363.06	14364.93	
	ARTKQTAR-trunc H3.1me2	14376.86	14378.94	
	ARTKQTAR-trunc H3.1me3	14389.88	14392.95	
	ARTKQTAR-trunc H3.1me4	14405.85	14406.96	
	ARTKQTAR-trunc H3.1me5	14421.88	14420.97	
	ARTKQTAR-trunc H3.1me6	14432.86	14434.98	
	ARTKQTAR-trunc H3.1me7	14448.88	14448.99	
	H3.2	15246.39	15247.47	Q71DI3
	H3.2me1	15259.49	15261.48	
	H3.3	15186.25	15187.46	P84243
	H3.3me1	15201.21	15201.47	
	H3.3me2	15214.23	15215.48	

	H3.3me3	15229.23	15229.49	
	H3.3me4	15243.22	15243.5	
	H3.3me5	15259.29	15257.51	
Unknown		13316.91		
		12527.74		
		12771.8		
		12789.89		
		12804.94		
		12812.78		
		12828		
		12842.95		
		12150.7		
		12164.35		
		12178.57		
		12194.52		
		17110.03		
		18778.63		
		19348.85		
		25047.4		
		25113.59		
		27757.71		

Supplemental Table 2. Calculated *p*-values for abundance differences between histone proteoforms derived from HEK bulk chromatin and H3.3-FLAG enriched nucleosomes. Statistical significance was evaluated using two-sided, two-sample t-tests. Red *p*-values denote statistically significant differences for a given proteoform's intensity between the samples with Bonferroni-corrected α-values.

Histone	Proteoform	Bonferroni-corrected α-value	*p*-value
H4	1	0.0056	0.093
	2	0.0056	**0.0035**
	3	0.0056	**7.5E-06**
	4	0.0056	**2.3E-06**
	5	0.0056	**2.2E-04**
	6	0.0056	0.0139
	7	0.0056	**6.8E-05**
	8	0.0056	**1.0E-05**
	9	0.0056	**0.0036**
H2A	Z	0.0071	**2.7E-07**

Supplemental Table 3. Calculated *p*-values for abundance differences between histone proteoforms derived from H3.3WT and H3.3K27M nucleosomes. Statistical significance was evaluated using two-sided, two-sample t-tests. Red *p*-values denote statistically significant differences for a given proteoform's intensity between WT and K27M samples with Bonferroni-corrected α-values.

Note on H3.3 methyl distribution in Fig. 3: H3.3WT contains an average of 4.4 ±0.02 methyl equivalents compared to H3.3K27M, which has an average of 3.9 ±0.1 (n = 3; average ±1 standard deviation).

Histone	Proteoform	Bonferroni-corrected α-value	*p*-value
H4	1	0.0056	0.033
	2	0.0056	0.032
	3	0.0056	**0.0031**
	4	0.0056	0.016
	5	0.0056	**0.0012**
	6	0.0056	0.026
	7	0.0056	0.016
	8	0.0056	**1.0E-04**
	9	0.0056	**0.0011**
H3.3	1 methyl equivalent	0.0071	**0.0044**
	2 methyl equivalents	0.0071	**5.8E-04**
	3 methyl equivalents	0.0071	**0.0043**
	4 methyl equivalents	0.0071	0.065
	5 methyl equivalents	0.0071	**0.0028**
	6 methyl equivalents	0.0071	**1.9E-05**
	7 methyl equivalents	0.0071	0.84
H2A	Z	0.0071	**1.7E-4**
	2C+Ac	0.0071	0.42
	3	0.0071	0.53
	1-B/E	0.0071	0.066
	1C+Ac	0.0071	**0.0014**
	3+Ac	0.0071	0.66
	1-B/E+Ac	0.0071	0.095

References for Supplemental Information

1. Lowary, P. T.; Widom, J., New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol* **1998,** *276* (1), 19-42.
2. Ettig, R.; Kepper, N.; Stehr, R.; Wedemann, G.; Rippe, K., Dissecting DNA-histone interactions in the nucleosome by molecular dynamics simulations of DNA unwrapping. *Biophys J* **2011,** *101* (8), 1999-2008.
3. Fornelli, L.; Srzentić, K.; Huguet, R.; Mullen, C.; Sharma, S.; Zabrouskov, V.; Fellers, R. T.; Durbin, K. R.; Compton, P. D.; Kelleher, N. L., Accurate Sequence Analysis of a Monoclonal Antibody by Top-Down and Middle-Down Orbitrap Mass Spectrometry Applying Multiple Ion Activation Techniques. *Analytical Chemistry* **2018,** *90* (14), 8421-8429.
4. DeHart, C. J.; Fellers, R. T.; Fornelli, L.; Kelleher, N. L.; Thomas, P. M., Bioinformatics Analysis of Top-Down Mass Spectrometry Data with ProSight Lite. In *Protein Bioinformatics: From Protein Modifications and Networks to Proteomics*, Wu, C. H.; Arighi, C. N.; Ross, K. E., Eds. Springer New York: New York, NY, 2017; pp 381-394.
5. Schachner, L. F.; Ives, A. N.; McGee, J. P.; Melani, R. D.; Kafader, J. O.; Compton, P. D.; Patrie, S. M.; Kelleher, N. L., Standard Proteoforms and Their Complexes for Native Mass Spectrometry. *J Am Soc Mass Spectr* **2019,** *30* (7), 1190-1198.
6. Belov, M. E.; Damoc, E.; Denisov, E.; Compton, P. D.; Horning, S.; Makarov, A. A.; Kelleher, N. L., From protein complexes to subunit backbone fragments: a multi-stage approach to native mass spectrometry. *Anal Chem* **2013,** *85* (23), 11163-73.
7. Pesavento, J. J.; Mizzen, C. A.; Kelleher, N. L., Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: human histone H4. *Anal Chem* **2006,** *78* (13), 4271-80.
8. Kornacki, J. R.; Stuparu, A. D.; Mrksich, M., Acetyltransferase p300/CBP associated Factor (PCAF) regulates crosstalk-dependent acetylation of histone H3 by distal site recognition. *ACS Chem Biol* **2015,** *10* (1), 157-164.
9. Kuo, Y.-M.; Andrews, A. J., Quantitating the Specificity and Selectivity of Gcn5-Mediated Acetylation of Histone H3. *PLOS ONE* **2013,** *8* (2), e54896.
10. Schiltz, R. L.; Mizzen, C. A.; Vassilev, A.; Cook, R. G.; Allis, C. D.; Nakatani, Y., Overlapping but Distinct Patterns of Histone Acetylation by the Human Coactivators p300 and PCAF within Nucleosomal Substrates. *Journal of Biological Chemistry* **1999,** *274* (3), 1189-1192.
11. Piunti, A.; Shilatifard, A., Epigenetic balance of gene expression by Polycomb and COMPASS families. *Science* **2016,** *352* (6290), aad9780.
12. Swalm, B. M.; Knutson, S. K.; Warholic, N. M.; Jin, L.; Kuntz, K. W.; Keilhack, H.; Smith, J. J.; Pollock, R. M.; Moyer, M. P.; Scott, M. P.; Copeland, R. A.; Wigle, T. J., Reaction Coupling between Wild-Type and Disease-Associated Mutant EZH2. *ACS Chem Biol* **2014,** *9* (11), 2459-2464.
13. Laugesen, A.; Højfeldt, J. W.; Helin, K., Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation. *Molecular Cell* **2019,** *74* (1), 8-18.
14. Kafader, J. O.; Melani, R. D.; Durbin, K. R.; Ikwuagwu, B.; Early, B. P.; Fellers, R. T.; Beu, S. C.; Zabrouskov, V.; Makarov, A. A.; Maze, J. T.; Shinholt, D. L.; Yip, P. F.; Tullman-Ercek, D.; Senko, M. W.; Compton, P. D.; Kelleher, N. L., Multiplexed mass spectrometry of individual ions improves measurement of proteoforms and their complexes. *Nature Methods* **2020,** *17* (4), 391-394.
15. Jooß, K.; Schachner, L. F.; Watson, R.; Gillespie, Z. B.; Howard, S. A.; Cheek, M. A.; Meiners, M. J.; Licht, J. D.; Keogh, M.-C.; Kelleher, N. L., Separation and Characterization of Endogenous Nucleosomes by Native Capillary Zone Electrophoresis – Top-Down Mass Spectrometry (nCZE-TDMS). *bioRxiv* **2020**, 2020.11.25.398925.

16. Svensson, J. P.; Shukla, M.; Menendez-Benito, V.; Norman-Axelsson, U.; Audergon, P.; Sinha, I.; Tanny, J. C.; Allshire, R. C.; Ekwall, K., A nucleosome turnover map reveals that the stability of histone H4 Lys20 methylation depends on histone recycling in transcribed chromatin. *Genome Res* **2015,** *25* (6), 872-883.
17. Huang, C.; Zhang, Z.; Xu, M.; Li, Y.; Li, Z.; Ma, Y.; Cai, T.; Zhu, B., H3.3-H4 tetramer splitting events feature cell-type specific enhancers. *PLoS Genet* **2013,** *9* (6), e1003558-e1003558.
18. Wirbelauer, C.; Bell, O.; Schübeler, D., Variant histone H3.3 is deposited at sites of nucleosomal displacement throughout transcribed genes while active histone modifications show a promoter-proximal bias. *Genes & development* **2005,** *19* (15), 1761-1766.
19. Herz, H.-M.; Morgan, M.; Gao, X.; Jackson, J.; Rickels, R.; Swanson, S. K.; Florens, L.; Washburn, M. P.; Eissenberg, J. C.; Shilatifard, A., Histone H3 lysine-to-methionine mutants as a paradigm to study chromatin signaling. *Science* **2014,** *345* (6200), 1065-1070.
20. Piunti, A.; Hashizume, R.; Morgan, M. A.; Bartom, E. T.; Horbinski, C. M.; Marshall, S. A.; Rendleman, E. J.; Ma, Q.; Takahashi, Y.-h.; Woodfin, A. R.; Misharin, A. V.; Abshiru, N. A.; Lulla, R. R.; Saratsis, A. M.; Kelleher, N. L.; James, C. D.; Shilatifard, A., Therapeutic targeting of polycomb and BET bromodomain proteins in diffuse intrinsic pontine gliomas. *Nature Medicine* **2017,** *23* (4), 493-500.
21. Lowe, B. R.; Maxham, L. A.; Hamey, J. J.; Wilkins, M. R.; Partridge, J. F., Histone H3 Mutations: An Updated View of Their Role in Chromatin Deregulation and Cancer. *Cancers (Basel)* **2019,** *11* (5), 660.

Supplemental References for ChIP-seq methods

-Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

-Boyle, E. et al. (2004). GO::TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics, 20(18):3710–3715.

- Edgar R, Domrachev M, Lash AE. (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 30(1):207-10.

-Heinz S, Benner C, Spann N, Bertolino E et al. (2010) Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Mol Cell 38(4):576-589.

-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.

-Lee et al., 2006. “Chromatin immunoprecipitation and microarray-based analysis of protein location.” Nature Protocols.

-Ramírez, Fidel et al. (2016). deepTools2: A next Generation Web Server for Deep-Sequencing Data Analysis. Nucleic Acids Research 44: W160–W165.

-Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, Ali S, Chin S, Palmieri C, Caldas C, Carroll JS (2012). “Differential oestrogen receptor binding is associated with clinical outcome in breast cancer.” Nature, 481, -4.

-Vo et al., 2017. Cell Reports. “Inactivation of Ezh2 Upregulates Gfi1 and Drives Aggressive Myc-Driven Group 3 Medulloblastoma”

-Yu G, Wang L, Han Y, He Q (2012). “clusterProfiler: an R package for comparing biological themes among gene clusters.” OMICS: A Journal of Integrative Biology, 16(5), 284-287.

-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.

Supplemental Protocol for Nuc-MS Sample Preparation, Data Acquisition and Analysis

Schachner, L. et al. "Decoding the Protein Composition of Whole Nucleosomes with Nuc-MS." Nature Methods (2020)

This protocol is available on the Nature Protocol Exchange repository (DOI 10.21203/rs.3.pex-1288/v1).

Sample Preparation

Mononucleosome Extraction from Cells and Purification

1. With 2.5 mL cell pellet, resuspend cell pellet with 2.5X PCV of Buffer A (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 340 mM sucrose (12% w/v), 10% glycerol (v/v), 0.5 mM DTT, and 1X *Roche* cOmplete EDTA-free protease inhibitor cocktail).
2. Lyse cells by adding equal volume of Buffer A supplemented with 2% Triton X-100
    a. Note: Final concentration of Triton X-100 in cell suspension is 1%.
3. Incubate on ice for 10 mins with occasional mixing.
4. Pellet nuclei by centrifuging at 1,300 x*g* for 5 minutes at 4C.
5. Resuspend nuclei pellet with 6X PCV of Buffer A.
6. Transfer nuclei suspension onto 35 mL sucrose cushion (10 mM HEPES pH 7.9, 30% (w/v) sucrose, 1.5 mM $MgCl_2$).
7. Pellet nuclei by centrifuging at 1,300 x*g* for 12 minutes at 4C to separate nuclei from cell debris.
8. Resuspend nuclei pellet with 2X PCV of Buffer A.
9. Incubate at room temperature for 5 mins.
10. Supplement nuclei suspension with 1 mM $CaCl_2$.
11. Supplement nuclei suspension with 2 µL of *NEB* micrococcal nuclease per 1 mL of cell suspension.
12. Incubate at 37C for 15 minutes with occasional mixing.
13. To quench micrococcal nuclease digestion, incubate on ice and supplement nuclei suspension with 2 mM EGTA and 1 mM EDTA.
14. Supplement nuclei suspension with final concentration of 100 mM KCl and 0.05% Triton X-100.
15. Incubate 15 minutes on ice with occasional mixing.
16. Centrifuge at 20,000 x*g* for 20 mins at 4C and collect resulting supernatant containing nucleosomes.
17. Mononucleosomes are concentrated and buffer exchanged into Buffer A supplemented with 650 mM NaCl using 30 kDa MWCO spin filter (*Millipore-Sigma*).

18. Concentrated nucleosomes are purified using HiPrep™ 26/60 Sephacryl S-300 HR column (Size Exclusion Chromatography) equilibrated with Buffer A supplemented with 650 mM NaCl using AKTA Prime Plus FPLC (*GE Lifescience*).
19. Fractions containing mononucleosomes are validated by purifying 10 ug of material using Qiagen DNA clean up kit and resolving it on 2% agarose gel in 0.5X TBE.
    a. Note: Fractions showing ~150 DNA bp are pooled together for Nuc-MS analysis.

Sample Desalting for Nuc-MS

1. Purified mononucleosomes and synthetic Nucs (such as EpiCypher *designer Nucs*) must be concentrated and desalted into 150 mM ammonium acetate solution. Best results are achieved using the 0.5 mL 30 kDa MWCO spin filter (*Millipore-Sigma*). A concentration >2 µM in a volume of 50 µL yields adequate signal intensity for MS[1-3].
2. Filter is first conditioned by spinning 500 µL of 150 mM ammonium acetate for 3 mins at 13,000 xg. Upon completion of spin, empty the filter of any remaining liquid.
3. Add up to 500 µL of Nuc sample into the filter – topping off with ammonium acetate – and spin for 5 mins at 13,000 xg or until sample is concentrated at or below 100 µL.
4. For sample desalting, add ammonium acetate up to the 500 µL mark and spin sample for 5 mins at 13,000 xg. Repeat desalting process 10-15 times for best results.
5. For final spin, centrifuge for 10 minutes to concentrate sample >2 µM in <50 µL.

Nuc-MS

Data Acquisition

1. *Instrumentation*: Given the high mass of Nuc particles, the Q Exactive HF mass spectrometer with Extended Mass Range (QE-EMR) and Q Exactive HF Ultra-High Mass Range (QE-UHMR) instrumentation are best for Nuc-MS analysis. Other instruments with up to 8000 *m/z* transmission and detection range may also be suited for Nuc-MS.
2. XCalibur QualBrowser 4.0.27.10 (*Thermo Fisher Scientific*) is used for MS data acquisition.
3. Native electrospray ionization (nESI) can be achieved with commercial Nanospray and Nanospray Flex Ion Sources with a static NSI probe (*Thermo Fisher Scientific*) as well as a capillary-based ion source as described previously.[1]
4. Once stable electrospray is achieved, the MS^1 (intact mass of Nuc particle) can be collected using the following parameters in positive ion mode:

Parameter	**Q Exactive HF EMR**	**Q Exactive UHMR**
Injection time	Fixed at 50 ms	Fixed at 200 ms
Number of microscans	10	10
Resolution (MS^1)	15,000 at 400 *m/z*	17,500 at 200 m/z

Ultra-High Vacuum (UHV) Pressure	4 V: 1.65x10^{-9} mBar	2-4 V: 7.7 x10^{-11} to 1.55x10^{-10} mBar
Source Induced Dissociation (SID)	25 V	5 V
In-Source Trapping (IST)	Not available	25 to 50 V
HCD trapping energy	10 V	0 V
Capillary temperature	330 ºC	300 ºC

5. Ejection of histones (MS^2) is achieved with HCD activation of an isolated nucleosome charge state for a synthetic Nuc or isolation of the entire charge state distribution for endogenous mononucleosomes. The isolation window using the quadrupole mass filter should be optimized based on heterogeneity of the sample and signal intensity output. The following parameters are recommended once the desired species/population is isolated:

Parameter	**Q Exactive HF EMR**	**Q Exactive UHMR**
Injection time	Fixed at 500 ms	Fixed at 500 ms
Number of microscans	10-20	10-20
Resolution (MS^2)	120,000 at 400 *m/z*	100,000 at 200 m/z
Ultra-High Vacuum (UHV) Pressure	2 V: 7.11x10^{-10} mBar	2 V: 7.7 x10^{-11} mBar
Source Induced Dissociation (SID)	5 V	0 V
In-Source Trapping (IST)	Not available	25 to 50 V
HCD trapping energy	120 V	5-30 V (Normalized Collision Energy, NCE)

6. For histone ejection at the source in preparation for pseudo-MS^3, the following parameters are recommended:

Parameter	**Q Exactive HF EMR**	**Q Exactive UHMR**

Injection time	Fixed at 100-300 ms	Fixed at 100-300 ms
Number of microscans	2	2
Resolution (MS^2)	120,000 at 400 *m/z*	100,000 at 200 m/z
Ultra-High Vacuum (UHV) Pressure	2 V: $7.11x10^{-10}$ mBar	1-2 V: $4.67x10^{-11}$ - $7.7\ x10^{-11}$ mBar
Source Induced Dissociation (SID)	50-150 V*	0 V
In-Source Trapping (IST)	Not available	150-200 V
HCD trapping energy	0 V	0 V

*Note that histone ejection at the source using SID on the QE-EMR is not as efficient as IST on the QE-UHMR. A 50-150 V SID range is provided given that some DNA fragmentation occurs which must be balanced with histone ejection.

7. For fragmentation of source-ejected histones (MS^3 / pseudo-MS^3), the desired histone proteoform charge state is manually isolated using the quadrupole mass filter with an isolation window of 5-10 *m/z*. The following parameters are recommended once stable signal is achieved for the isolated species:

Parameter	**Q Exactive HF EMR**	**Q Exactive UHMR**
Injection time	Fixed at 1000-2000 ms	Fixed at 1000-2000 ms
Number of microscans	10	10
Resolution (MS^2)	120,000 at 400 *m/z*	100,000 at 200 m/z
Ultra-High Vacuum (UHV) Pressure	2 V: $7.11x10^{-10}$ mBar	2 V: $7.7\ x10^{-11}$ mBar
Source Induced Dissociation (SID)	50 V	0 V
In-Source Trapping (IST)	Not available	150-200 V
HCD trapping energy	120 V	30-40 V (Normalized Collision Energy, NCE)

Data Analysis

8. MS^1 and MS^2 data at medium-resolution (not isotopic) can be deconvoluted and visualized using MagTran 1.03[2] (mass range: 15-300 kDa; max no. of species: 10-15; S/N threshold: 1; mass accuracy: 0.05 Da; charge determined by: charge envelop only) and UniDec 3.2.0[3] (Charge Range: 20 – 50, Mass Range: 15 – 300 kDa, Sample mass every 0.5 Da).
9. Isotopically resolved MS^2 data can be analyzed using Xtract (Signal-to-Noise threshold ranging from 1-30, *Thermo Fisher Scientific*) and mMass 5.5.0 (www.mmass.org).
10. MS^3 data can be processed using Xtract (Signal-to-Noise threshold ranging from 1-30, *Thermo Fisher Scientific*), which provides the monoisotopic masses for the detected fragment ions. mMass 5.5.0 (www.mmass.org), ProSight Lite 1.4[4] (precursor mass type: average; fragmentation method: HCD; fragmentation tolerance: 10-15 ppm), and TDValidator 1.0[5] (max ppm tolerance: 25 ppm; cluster tolerance: 0.35; charge range: 1-10; minimum score: 0.5; S/N cutoff: 3; Mercury7 Limit: 0.0001; minimum size: 2) are used to assign recorded fragment ions to the primary sequence of the subunits.
11. For careful fragmentation analysis, ProSight Lite and TDValidator are used to analyze spectra in medium throughput to assign and validate *b* and *y* fragment ions to the histone sequences, and for generating a p-score. mMass generates *in silico* a list of theoretical fragment ions for a target proteoform and is thus used to interrogate individual fragment ions within a spectrum not identified by TDValidator or ProSight Lite.
12. Unexplained mass shifts (Δm) observed at the MS^1, MS^2, and MS^3 levels were manually interrogated using the UNIMOD database (http://www.unimod.org/modifications_list.php) as a reference for candidate modifications.

References for Nuc-MS Supplemental Protocol

1. Schachner, L. F.; Ives, A. N.; McGee, J. P.; Melani, R. D.; Kafader, J. O.; Compton, P. D.; Patrie, S. M.; Kelleher, N. L., Standard Proteoforms and Their Complexes for Native Mass Spectrometry. J Am Soc Mass Spectr 2019, 30 (7), 1190-1198.

2. Zhang, Z.; Marshall, A. G., A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. J. Am. Soc. Mass Spectrom. 1998, 9 (3), 225-33.

3. Marty, M. T.; Baldwin, A. J.; Marklund, E. G.; Hochberg, G. K. A.; Benesch, J. L. P.; Robinson, C. V., Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. Analytical Chemistry 2015, 87 (8), 4370-4376.

4. Fellers, R. T.; Greer, J. B.; Early, B. P.; Yu, X.; LeDuc, R. D.; Kelleher, N. L.; Thomas, P. M., ProSight Lite: graphical software to analyze top-down mass spectrometry data. Proteomics 2015, 15 (7), 1235-8.

5. Fornelli, L.; Srzentić , K.; Huguet, R.; Mullen, C.; Sharma, S.; Zabrouskov, V.; Fellers, R. T.; Durbin, K. R.; Compton, P. D.; Kelleher, N. L., Accurate sequence analysis of a monoclonal antibody by top-down and middle-down orbitrap mass spectrometry applying multiple Ion activation techniques. Anal. Chem. 2018, 90 (14), 8421-8429.