



Published in final edited form as:

*Nat Chem Biol.* 2021 April ; 17(4): 403–411. doi:10.1038/s41589-021-00738-1.

## Oncohistone mutations enhance chromatin remodeling and alter cell fates

John D. Bagert<sup>#1</sup>, Michelle M. Mitchener<sup>#1</sup>, Agata L. Patriotis<sup>#2</sup>, Barbara E. Dul<sup>#1</sup>, Felix Wojcik<sup>1,†</sup>, Benjamin A. Nacev<sup>2,3,4</sup>, Lijuan Feng<sup>2</sup>, C. David Allis<sup>2,\*</sup>, Tom W. Muir<sup>1,\*</sup>

<sup>1</sup>Department of Chemistry, Princeton University, Princeton, NJ, USA

<sup>2</sup>Laboratory of Chromatin Biology and Epigenetics, The Rockefeller University, New York, NY, USA

<sup>3</sup>Memorial Sloan Kettering Cancer Center, New York, NY, USA

<sup>4</sup>Weill Cornell Medical College, New York, NY, USA

# These authors contributed equally to this work.

### Abstract

Whole genome sequencing data mining efforts have revealed numerous histone mutations in a wide range of cancer types. These occur in all four core histones in both the tail and globular domains and remain largely uncharacterized. Here we used two high-throughput approaches, a DNA-barcoded mononucleosome library and a humanized yeast library, to profile the biochemical and cellular effects of these mutations. We identified cancer-associated mutations in the histone globular domains that enhance fundamental chromatin remodeling processes, histone exchange and nucleosome sliding, and are lethal in yeast. In mammalian cells, these mutations upregulate cancer-associated gene pathways and inhibit cellular differentiation by altering expression of lineage-specific transcription factors. This work represents a comprehensive functional analysis of the histone mutational landscape in human cancers and leads to a model in which histone mutations that perturb nucleosome remodeling may contribute to disease development and/or progression.

---

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:[http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

\*Correspondence and requests for materials should be addressed to C.D.A. (alliscd@mail.rockefeller.edu) and T.W.M. (muir@princeton.edu).

†Present address: Novo Nordisk A/S, Smørmosevej 17-19, 2880 Bagsværd, Denmark

#### Author contributions

M.M.M. prepared the DNA-barcoded MN library, purified ACF, performed all remodeling experiments, and prepared all samples for next-generation sequencing; J.D.B. prepared Nap1, performed all dimer exchange experiments, prepared heterotypic nucleosomes, and performed statistical analyses; B.E.D. prepared the humanized yeast library; A.L.P. performed all mammalian cell experiments; F.W. expressed ACF; M.M.M., J.D.B., B.E.D., A.L.P., C.D.A., and T.W.M. analyzed the data; M.M.M., J.D.B., B.E.D., A.L.P., F.W., C.D.A., and T.W.M. conceived the project with input from B.A.N. and L.F.; C.D.A. and T.W.M. supervised the study; M.M.M., J.D.B., and T.W.M. wrote the manuscript with input from all authors.

#### Competing interests

The authors declare no competing interests.

#### Supplementary Information

Supplementary information is available for this manuscript.

## Introduction

Establishment and maintenance of chromatin underlies transcription, DNA replication, and DNA repair, and disruption of these processes often occurs in human cancers<sup>1,2</sup>. More recently, mutations in histone proteins have been linked to rare cancers<sup>3,4</sup>, including H3.3K27M (pediatric gliomas), H3.3G34W/L (giant cell tumors of bone), and H3.3K36M (chondroblastomas and sarcomas)<sup>5–11</sup>. A key feature of these founding “oncohistone” mutations is their occurrence at or near sites of histone post-translational modification (PTM). Indeed, mutations at H3K27, H3G34, and H3K36 disrupt regulatory mechanisms by which key PTMs are installed and maintained<sup>6,11–14</sup>. Histone mutations also have been found in other cancers, including diffuse large B-cell lymphoma, head and neck squamous cell carcinoma, acute myeloid leukemia, and uterine and ovarian carcinosarcoma, although these remain less studied<sup>15–18</sup>. Recent efforts to catalogue and characterize the histone missense mutational landscape of human cancers<sup>19,20</sup> have revealed thousands of somatic mutations in genes encoding all four core histones across diverse tumor types (Fig. 1a). Many of these histone mutations occur with a frequency similar to that of somatic mutations in well-established, cancer-associated genes, such as *BRCA2*, *TET2*, *SMAD4*, and *NOTCH1*. Some of these histone mutations occur at or near sites of PTM (e.g. the aforementioned oncohistone mutations), however, the majority of high-frequency mutations occur in the folded, globular domains. Relatively few of these globular domain sites are known to be post-translationally modified, raising the question of whether these mutations function through mechanisms distinct from those of established oncohistones. Furthermore, although a select few of these histone mutations have been the subject of recent studies<sup>20–22</sup>, the vast majority remain entirely uncharacterized.

Here we report the systematic biochemical and cellular characterization of the newly identified oncohistone mutational landscape. We developed two complementary high-throughput screening tools for the rapid interrogation of these mutations: a DNA-barcoded mononucleosome (MN) library, which enabled single-pot biochemical evaluation of the effects of these mutations using next-generation sequencing (NGS), and a “humanized” yeast library, which allowed evaluation of the phenotypic effects of human histone mutations in a cellular context. Together, these libraries enabled identification of mutations that alter fundamental chromatin remodeling processes, namely histone exchange and nucleosome sliding, and are lethal in yeast. Notably, unlike the aforementioned classical “oncohistone” mutations, many of these mutations occur in the histone globular domains. Expression of these mutant histones in mammalian cells caused significant changes in transcription, dysregulating cell-signaling and cancer-associated gene pathways, and impaired cellular differentiation. Collectively, this study presents an experimental platform for rapid biochemical and cellular interrogation of histone mutations and demonstrates the utility of these tools by identifying mutations that affect fundamental chromatin processes and alter cell fates. Moreover, this work suggests that cancer-associated histone globular domain mutations may contribute to disease development via perturbation of nucleosome remodeling.

## Results

### Generation of libraries to profile histone mutations

To enable expeditious interrogation of cancer-associated histone mutations, we narrowed the set of histone mutations found in over 3000 cancer patient samples to an experimentally tractable set likely to have functional effects in vitro and in cells (Extended Data Fig. 1, see methods for discussion of selection criteria). Importantly, for every mutated site chosen for analysis, we included an Ala mutation in addition to any specific missense mutations that were overrepresented in the dataset. This approach offered both experimental redundancy as well as the ability to distinguish between loss and gain of function at a given amino acid residue. The resulting list contained ~160 distinct histones, including experimental controls, which were incorporated into subsequent high-throughput screening approaches (Fig. 1b, Supplementary Table 1). These mutations are distributed throughout all four core histones (including variants), providing broad biochemical and structural coverage of the nucleosome. Therefore, the experimental libraries described herein not only enable analysis of disease-associated histone mutations but also facilitate coarse-grained profiling of nucleosome structure-activity relationships. Thus, these libraries may additionally offer insight into the role of specific nucleosome regions or histone residues in normal chromatin physiology.

We then designed a DNA-barcoded mononucleosome (MN) library to enable high-throughput biochemical evaluation of the effects of these histone mutations using next-generation sequencing. We previously reported DNA-barcoded MN libraries to profile the effects of histone PTMs on various chromatin effectors<sup>23,24</sup>. Here we used an analogous approach to evaluate the impact of cancer-associated histone mutations on chromatin remodeling. Utilizing a streamlined MN reconstitution procedure (see methods), we assembled a MN library from purified recombinant mutant and wild-type human histones and a set of double-stranded DNAs, each containing the Widom 601 nucleosome positioning sequence flanked by a unique 6-bp molecular identifier (i.e. barcode) (Extended Data Fig. 2a–c). Importantly, the presence of these barcodes allowed all subsequent biochemical experiments to be performed on the pooled mixture, exploiting the sensitivity and throughput afforded by an NGS-based readout. Additionally, we included five wild-type MN controls to assess potential variation in nucleosome behavior due to barcode sequence-specific effects and/or differences in nucleosome assembly. We confirmed the faithful assembly of this library via deep sequencing and demonstrated that barcode-histone scrambling does not occur, even under harsh conditions (Extended Data Fig. 2d, e).

We also sought to develop an additional screening tool to assess the effects of histone mutations in more complex cellular environments. Previous studies have employed yeast (*S. cerevisiae*) as a model system for performing comprehensive mutagenesis of histones, wherein endogenous histone genes were replaced with yeast histone mutant sequences of interest<sup>25,26</sup>. To provide consistency between our biochemical and cellular libraries and avoid instances where the yeast corollary of human histone mutants was unclear, we adapted a screening platform that allows yeast histones to be fully replaced with human histones using a 5-fluoroorotic acid (5FOA) plasmid shuffle approach<sup>27</sup>. Yeast with humanized

chromatin (hereafter “humanized yeast”) were originally generated to probe the fundamental relationship between chromatin, histone variants, and cellular functions. We envisioned that this system could be repurposed to assess the effects of cancer-associated histone mutations on cell viability (Extended Data Fig. 3a). With this in mind, we generated a library of 130 yeast strains containing a set of mutations similar to that present in our barcoded library (Extended Data Fig. 3b, Supplementary Table 2). Importantly, phenotypic effects are not attributable to differential growth rates prior to 5FOA selection (Extended Data Fig. 3c), suggesting that observed growth defects instead reflect histone mutants that are compromised in an essential function. A notable advantage of this yeast screening approach is that, unlike in polygenic mammalian cells, the mutated histone gene is the only copy present in the organism, likely amplifying phenotypic effects and making them more readily detectable. Additionally, the structural effects of mutations that perturb nucleosome assembly, stability, and remodeling are likely preserved and detectable in this system (vide infra). Together, these libraries offered a versatile approach for rapidly assessing the biochemical and cellular effects of cancer-associated histone mutations.

### Identification of mutations affecting dimer exchange

Although some newly identified histone mutations occur at or near sites of post-translational modification on histone tails, as do classical oncohistone mutations, the majority appear in the folded globular domains (Fig. 1a). Relatively few of these globular domain sites are known to be post-translationally modified, raising the question whether these mutations function through mechanisms distinct from those of established oncohistone mutations. We hypothesized that histone core mutations might affect the structural stability of the nucleosome and rates of histone exchange. Therefore, we adapted an H2A/H2B dimer exchange assay<sup>28</sup> in which the histone chaperone nuclear assembly protein-1 (Nap1) was used to mediate the exchange of tagged wild-type H2A/H2B dimers into mutant MNs (Extended Data Fig. 4a, b). As Nap1 also serves as an H3/H4 tetramer chaperone, we verified that under our assay conditions, histone tetramer exchange does not occur (Extended Data Fig. 4c). We then applied this approach to our DNA-barcoded MN library, using biotinylated dimers and streptavidin affinity purification to enrich for less stable mutant MNs (Fig. 2a, Extended Data Fig. 4d, Supplementary Table 3). Note that, including biological replicates, this corresponded to over 600 biochemical experiments and utilized only 3 pmol (or 0.3%) of the library. Overall, histone mutations tended to increase dimer exchange compared to their wild-type counterparts (Fig. 2b), with mutations in the globular domains exhibiting the most pronounced effects (Fig. 2c). The locations of highly destabilizing mutations within the nucleosome structure revealed a striking convergence at the dimer-tetramer interface between histones H2B and H4, including mutations at H2BD68, F70, E71, E76, and H4D68 and K91 (Fig. 2d, e). Other sites of enhanced exchange included H2AR29 (a histone-DNA contact), H2AE56 (an acidic patch residue), and H3E50, E97, and H4G42 (sites buried within the H3/H4 tetramer). Interestingly, other acidic patch mutations at H2AE61, E91, E92 and H2BE113 did not increase dimer exchange, suggesting a distinct role for the H2AE56 residue in nucleosome stability (Extended Data Fig. 4e).

To validate results from the barcoded library, we prepared individual nucleosome substrates containing histone mutants that showed both increased dimer exchange and high mutational

frequency (Fig. 1a, Extended Data Fig. 4f, g). All of these MNs were assembled using identical, Widom 601 DNA, thus eliminating any potential sequence-specific effects. Mutations at H2AE56 and R29, H2BD68, E71, and E76, and H3E50 and E97 resulted in increased dimer exchange, as measured by in-gel fluorescence, which corresponded with the respective increases observed in the barcoded MN library experiments (Extended Data Fig. 4h, i). These nucleosomes also exhibited decreased thermal stability, with the exception of those with mutations at H3E50 (Fig. 2f, Extended Data Fig. 5a). Notably, histone mutations H2BD68A/N, H2BE76K, H3E50K, and H3E97A/K did not form stable octamer complexes with wild-type histones under standard assembly conditions, providing additional evidence of the perturbative effects of these mutations on nucleosome structure (Extended Data Fig. 5b). Mutants that enhanced dimer exchange and reduced thermal stability in MNs exhibited significant growth defects in humanized yeast (Fig. 2g, Extended Data Fig. 3b). Agreement between the biochemical and yeast datasets extended to the level of amino acid identity. For example, charge-swap mutations H2BE71K and H2BE76K did not support yeast growth, whereas charge-neutralizing mutations H2BE71Q and H2BE76Q were better tolerated.

### Asymmetric mutant nucleosomes are destabilized

Due to the polygenic nature of mammalian histones, mutations in these genes represent only a subset of the total histone pool<sup>11</sup>. Stochastic incorporation of histone mutants into mammalian chromatin predominantly gives rise to asymmetric nucleosomes, containing one mutant and one wild-type histone (as opposed to two mutant histones per nucleosome). To determine the destabilizing effects of histone mutations on asymmetric MNs, we developed a method for the synthesis of heterotypic nucleosomes using split inteins (Fig. 3a, see methods for further discussion). To generate heterotypic nucleosomes, mixtures of hexasomes and nucleosomes were produced by under-titrating H2A/H2B dimers relative to H3/H4 tetramers during nucleosome assembly. This approach generates uniformly oriented hexasomes due to the inherent asymmetry of histone-DNA binding strength in the Widom 601 sequence<sup>29</sup>. Hexasomes were then complemented by the addition of wild-type dimers fused to the N-terminal fragment of the engineered split intein, CfaN<sup>30</sup>, forming tagged heterotypic nucleosomes (Extended Data Fig. 6a). Tagged MNs were subsequently affinity purified and eluted via on-resin split intein complementation and thiolysis<sup>31</sup>, tracelessly yielding purified heterotypic nucleosomes. As a proof of concept, we generated asymmetric ubiquitinated H2B MNs, which can be discriminated from their homotypic counterparts by size on both native and denaturing gels (Extended Data Fig. 6b).

We then synthesized heterotypic MNs containing nucleosome-destabilizing mutants H2BE71K and H2BE76K (Fig. 3b, Extended Data Fig. 6c). Asymmetric mutant nucleosomes showed intermediate (between wild type and homotypic mutant) effects both in terms of Nap1-mediated dimer exchange and thermal stability (Fig. 3c, d, Extended Data Fig. 6d, e). Therefore, the destabilizing effects observed for homotypic nucleosomes (assessed in the barcoded MN and humanized yeast libraries) also occur in a heterotypic context.

## Identification of mutations affecting nucleosome sliding

ATP-dependent nucleosome remodeling complexes share a common mechanism in which ATP hydrolysis is utilized to translocate (slide) DNA around the nucleosome, facilitating chromatin assembly, access, and editing<sup>32</sup>. Histone core residues form many key electrostatic interactions with DNA, and these interfaces must be disrupted for efficient nucleosome sliding to occur. We therefore reasoned that histone globular domain mutations might affect nucleosome sliding. To determine the effects of histone mutations on nucleosome remodeling rates, we performed a restriction enzyme accessibility (REA) assay<sup>24</sup> on the DNA-barcoded MN library using ACF as a representative ATP-dependent remodeler (Fig. 4a, Extended Data Fig. 7a, Supplementary Table 4). All told, this experiment corresponded to over 6500 measurements when factoring in replicates and controls, again highlighting the advantages of the barcoded MN library platform, in particular the ability to multiplex experiments. Analysis of the remodeling data revealed that cancer-associated histone mutations can lead to both increased and decreased rates of ACF remodeling (Fig. 4b, Extended Data Fig. 7b). Notably, mutations at sites of histone-DNA contact, including H2AR29, K74, K75, and R77, H2BG53, and H4R45, primarily showed enhanced remodeling rates (Fig. 4c, d). The exception to this trend occurred at H3R83, however, where mutations showed decreased rates of sliding. Consistent with previous results<sup>24,33</sup>, mutations in the nucleosome acidic patch reduced or completely abolished ACF activity, with the exception of mutations at H2AE56 and H2BE113 which enhanced activity. Likewise, nucleosomes bearing the R17/19A mutations in the basic patch of the H4 tail were poor ACF substrates<sup>24,34,35</sup>.

We validated select histone mutations that increased (H2AR29P/Q, K74N, K75N, and H2BE113K/Q) and decreased (H2AE121K/Q and H3R83A) rates of ACF remodeling via gel-based quantification of REA assay products using individual MN substrates assembled on sequence-identical DNA (Fig. 4e, Extended Data Fig. 7c–e). Interestingly, we did not observe significant changes in nucleosome sliding rates for several mutations that had pronounced effects on dimer exchange (Extended Data Fig. 7f), suggesting that not all mutations affect various chromatin remodeling processes alike. Mutations at residues that alter ACF remodeling rates also conferred yeast growth defects, again demonstrating consistency between biochemical and cellular screening platforms (Fig. 4f, Extended Data Fig. 3b). Yeast were particularly sensitive to mutations in the nucleosome acidic patch, where 12 of 16 mutations exhibited a no-growth phenotype (Extended Data Fig. 3b). Importantly, we and others have recently shown that asymmetric disruption of the nucleosome acidic patch through mutation can lead to dysregulation of chromatin remodeling enzymes<sup>33,36</sup>.

## Mutations alter transcription and impair differentiation

Given the biochemical effects of cancer-associated histone mutations on chromatin remodeling processes *in vitro*, we hypothesized that these mutations might induce transcriptional changes in cells. As these prominent cancer-associated histone mutations that affect nucleosome stability (H2BE71K/Q and H2BE76K/Q) and nucleosome sliding (H2BE113K/Q) are present in a wide range of cancers<sup>19,20</sup>, we elected to express them in C3H10T1/2 murine mesenchymal progenitor cells (MPCs), a robust model system for



presented herein. Barcoded nucleosome libraries have been used to profile a variety of chromatin processes, including histone acetylation, ADP-ribosylation, phosphorylation, ubiquitylation, effector binding, and ISWI remodeling<sup>23,24,43–45</sup>. Similarly, the humanized yeast library can, in principle, be used to determine mutation-associated phenotypes beyond growth inhibition, including chromosomal instability, cell cycle inhibitor sensitivity, and DNA damage sensitivity. Furthermore, as a result of the relatively large number of mutations incorporated into these libraries and the broad coverage they provide throughout the tails and globular domains of all four core histones, these libraries also serve as tools for identifying nucleosome structure-activity relationships for a broad set of chromatin effectors, such as histone-modifying enzymes, remodelers, and transcription factors.

This work also presents a new approach for the streamlined, traceless assembly of heterotypic nucleosomes using split inteins. Previously reported methods for generating asymmetric nucleosomes involve laborious, multi-step purification protocols and/or technically challenging protein chemistry methods<sup>29,46–48</sup>. Our method avoids these potential difficulties and allows for the synthesis of asymmetric nucleosomes with histone mutations or post-translational modifications (Fig. 3a). This approach can be coupled with various genetic and semisynthetic protein production techniques and, since it can be used to synthesize nucleosomes with asymmetric H2A and H2B histones, is complementary to previously established protocols designed for assembling nucleosomes with asymmetrically modified H3 and H4 histones<sup>47,48</sup>. The need for such methods is imperative as accumulating evidence suggests that breakdown of nucleosome symmetry through heterotypic incorporation of mutant histones (or PTMs) can affect biochemical processes that act on chromatin<sup>48–50</sup>. Indeed, we and others have recently shown that asymmetric disruption of the nucleosome acidic patch through mutation (including cancer-associated mutations) can lead to dysregulation of chromatin remodeling enzymes<sup>33,36</sup>. Here we have extended these analyses to mutations that alter nucleosome stability, demonstrating that structurally perturbative dimer-tetramer interface mutants promote dimer exchange and decrease nucleosome thermal stability even in heterotypic contexts.

To date, our understanding of the role of histone mutations in cancers has primarily focused on the N-terminal tail of histone H3 either at or near sites of post-translational modification, where mutations inhibit lysine methyltransferase activity and subsequently disrupt local histone PTM installation and maintenance<sup>3,4</sup>. However, more recent studies have begun to suggest that mutations within the globular, folded domains of histones might also disrupt chromatin state. For example, structural and biochemical studies have shown that H2BE76K and H3.1E97K mutations compromise the structural integrity of the nucleosome, promote colony formation, and, in the case of the former, alter gene expression profiles and chromatin accessibility<sup>20,21</sup>. Consistent with these previous studies, we found that H2BE76K destabilizes the nucleosomes, even in a heterotypic context, and that this effect extends to several other mutations that map to the dimer-tetramer interface (e.g. those at H2BD68, H2BF70, H2BE71, H4D68, and H4K91) and elsewhere in the nucleosome core (e.g. H2AR29, H2AE56, H4G42, and H3.1E50). Moreover, our work indicates that expression of dimer-tetramer interface and acidic patch mutants in MPCs impacts gene expression profiles, upregulating several cancer-associated pathways, and leads to a profound differentiation blockade. Notably, this phenotype has been previously observed for



the well-established H3K36M oncohistone mutant, which, when stably expressed in MPCs, confers tumorigenic potential on these cells as observed using a mouse xenograft tumor model<sup>6</sup>. Collectively, these studies support a model in which histone core mutations that alter fundamental chromatin remodeling processes can promote human cancers (Fig. 6). These include mutations that disrupt nucleosome-binding interfaces (e.g. the acidic patch) as well as interactions that govern nucleosome structure and stability (e.g. histone-DNA contacts and the H2A/H2B dimer-H3/H4 tetramer interface). The effects of chromatin structure and dynamics on transcriptional potential are well established, and we propose that histone mutations that disrupt the underlying processes that control chromatin organization can cause gene dysregulation and contribute to a diseased state in human cancers. Additionally, the histone mutational landscape converges at critical functional and structural interfaces, raising the possibility that disparate histone mutations may function via shared mechanisms. While it remains to be determined how these mutations engender pathway-specific transcriptional profiles and what effects they have on genome architecture and accessibility, the model and tools developed here will serve to direct and expedite scientific investigations into this newly uncovered oncohistone landscape.

## Methods

### General lab protocols

Common lab reagents were purchased from Sigma-Aldrich unless otherwise stated. All commercial reagents were used as provided by the manufacturer without further purification. Oligonucleotides used for cloning were purchased from Integrated DNA Technologies. Plasmids and gene constructs were produced using standard cloning and site-directed mutagenesis techniques and were verified by Sanger sequencing (Genewiz). Amino acid derivatives for Fmoc solid-phase peptide synthesis (SPPS) were purchased from Novabiochem and coupling reagents were purchased from Matrix Innovation. Peptide synthesis was performed either by hand or with a CEM Liberty peptide synthesizer using standard Fmoc SPPS protocols.

Preparative-scale reverse-phase high-performance liquid chromatography (RP-HPLC) was performed using a Waters 2545 Binary Gradient Module with 2489 UV/Visible Detector with a Vydac C18 preparative column (19 × 250 mm, 10 μm) at a flow rate of 20 mL/min. Semi-preparative RP-HPLC was performed on either an Agilent 1100 Series or an Agilent 1260 Infinity system with a Waters XBridge Peptide BEH C18 OBD column (10 × 250 mm, 5 μm) at a flow rate of 4 mL/min at 40 °C. Analytical RP-HPLC was performed on either an Agilent 1100 Series or an Agilent 1260 Infinity system with a Vydac C18 column (4.6 × 150 mm, 5 μm) at a flow rate of 1 mL/min. Mobile phases for all histone purifications and analyses were 0.1% (v/v) TFA in water (HPLC solvent A) and 90% acetonitrile with 0.1% (v/v) TFA in water (HPLC solvent B). Mass spectrometric analysis was performed on a MicrOTOF-Q ESI-Qq-TOF mass spectrometer (Bruker).

Protein and DNA gels were imaged with an ImageQuant LAS-4000 (GE Healthcare Life Sciences) and densitometry measurements performed using Image Studio Lite (LI-COR Biosciences). Molecular graphics and analyses were performed with UCSF Chimera version 1.13.1, developed by the Resource for Biocomputing, Visualization, and Informatics at the

University of California, San Francisco<sup>51</sup>. Data analyses and visualization were performed with Python 2.7.10 and the Matplotlib 2.2.3 library.

### Selection of histone mutants to include in libraries

We developed a series of filtering criteria to reduce the number of candidate histone mutations to an experimentally tractable set. First, in an effort to distinguish mutations that serve as potential cancer drivers from those that are merely passengers, we focused our analysis on mutations from samples with a tumor mutational burden (TMB)  $\geq 10$  mutations per megabase. This TMB threshold significantly enriched known oncohistone mutations at H3K27, G34, and K36 above mutational background<sup>19</sup>. Second, we identified histone mutations that show statistical overrepresentation either when considering the total number of amino acid mutations at a given residue or only when considering mutation of a residue to a specific amino acid (e.g. Lys to Met) (Extended Data Fig. 1a, b). Third, we included mutations for which we had a priori hypotheses about their functional repercussions, including H2A/H2B acidic patch mutations, switch-independent (Sin-) mutations, candidate amino acid neomorphs, other Lys to Met mutations, as well as previously identified oncohistone mutations. We also included mutations that exhibit tumor type specificity and mutations at several mutationally silent residues. Finally, we included an Ala mutation at every site in addition to any specific missense mutations that were overrepresented in the dataset. This strategy offered both experimental redundancy as well as the ability to distinguish between loss and gain of function at a given amino acid residue.

### Production of recombinant histones

Recombinant human histones were expressed in and purified from *E. coli* BL21(DE3) cells. Standard site-directed mutagenesis techniques were used to generate histone mutant plasmids. *E. coli* were transformed with histone expression plasmids and grown in LB medium at 37 °C to an OD<sub>600</sub> of 0.6. Protein expression was induced by the addition of 1 mM IPTG for 2–3 hr at 37 °C. Cells were harvested by centrifugation at 4,000 *g* for 30 min at 4 °C and washed once with cold PBS (10 mL/L culture). Cell pellets were resuspended in cold lysis buffer (50 mM Tris, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.6 at 4 °C) (10 mL/L culture) and lysed with a Fisherbrand Model 505 Sonic Dismembrator. Inclusion bodies were isolated by centrifugation at 31,000 *g* for 30 min at 4 °C and washed twice with cold lysis buffer containing 1% (w/v) Triton X-100 (10 mL/L culture; 15 min centrifugations at 31,000 *g* and 4 °C). Inclusion bodies were washed one final time with lysis buffer (10 mL/L culture) and centrifuged at 31,000 *g* for 15 min at 4 °C. Inclusion bodies were resuspended in 15 mL/L culture resuspension buffer (6 M GuHCl, 20 mM Tris, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.6 at 4 °C) and nutated for 3 hr at room temperature. Resuspensions were acidified with 0.1% (v/v) TFA, centrifuged at 31,000 *g* for 30 min, and filtered through a 0.45  $\mu$ m filter prior to purification. Wild-type histones were purified by RP-HPLC on a preparative scale, and mutant histones were purified on a semi-preparative scale. Purified histones were characterized by analytical RP-HPLC and mass spectrometry (Supplementary Fig. 1).

## H2Aub and H2Bub synthesis

**Production of  $\alpha$ -thioesters for expressed protein ligation**—Synthesis of H2AK119ub(G76A) and H2BK120ub(G76A) was carried out using the general expressed protein ligation strategy that has been described previously<sup>23</sup>. An  $\alpha$ -thioester corresponding to the N-terminal portion of H2A (residues 1–112) was generated by thiolysis of the corresponding histone-intein fusion construct containing H2A(1–112)-GyrA-7xHis. Note that the histone amino acid numbering convention is used which excludes the initiator methionine. After a standard purification with HisPur™ Ni-NTA Resin (Thermo Fisher Scientific), eluted protein was refolded by stepwise dialysis from 6 M to 0.5 M GuHCl. Thiolysis was initiated by adding sodium 2-mercaptoethanesulfonate (MESNA, Sigma-Aldrich) to a concentration of 0.2 M. His-tagged intein was then removed by a reverse Ni-NTA column, and the H2A(1–112)-MES thioester was purified by preparative-scale RP-HPLC. Protein purity was confirmed by analytical HPLC and mass spectrometry (Supplementary Fig. 2a), and protein was subsequently lyophilized and stored at –80 °C until use. Thioesters containing the N-terminal portion of H2B (residues 1–116) and ubiquitin (residues 1–75) were produced analogously (Supplementary Fig. 2a).

**H2A K119Ub(G76A)**—Ubiquitinated H2A was assembled as previously described<sup>23</sup> by semi-synthesis from three pieces: recombinant truncated H2A (residues 1–112) containing an  $\alpha$ -thioester, a synthetic C-terminal H2A peptide (residues 113–129) with alanine 113 replaced by a cysteine protected as a thiazolidine derivative and a cysteine residue conjugated to the  $\epsilon$ -amino group of lysine 119 through an isopeptide bond, and a recombinant truncated ubiquitin (residues 1–75) containing an  $\alpha$ -thioester. The H2A peptide was synthesized using automated Fmoc SPPS. The  $\epsilon$ -amino group of lysine 119 was protected with an orthogonal alloc group, which allowed for the selective coupling of cysteine to this side chain, forming an isopeptide linkage. The N-terminal cysteine residue was protected as a thiazolidine derivative to ensure selectivity in the first ligation. The ubiquitinated H2A peptide was generated via expressed protein ligation with Ubiquitin(1–75)-MES. The thiazolidine group was deprotected using methoxylamine under acidic conditions, and a second expressed protein ligation was used to generate full-length ubiquitinated H2A. Radical-initiated desulfurization of the two cysteine residues yielded the final product, which was purified by semi-preparative HPLC and characterized by analytical RP-HPLC and mass spectrometry (Supplementary Fig. 1).

**H2B K120Ub(G76A)**—Ubiquitinated H2B was assembled as previously described<sup>23</sup> by semi-synthesis from three pieces: a recombinant truncated H2B (residues 1–116) containing an  $\alpha$ -thioester, a synthetic C-terminal H2B peptide (residues 117–125) with alanine 117 replaced by cysteine protected as a thiazolidine derivative and a cysteine residue conjugated to the  $\epsilon$ -amino group of lysine 120 through an isopeptide bond, and a recombinant truncated ubiquitin (residues 1–75) containing an  $\alpha$ -thioester. Semi-synthesis was carried out following the protocols as described for H2A K119ub(G76A). Purified ubiquitinated H2B was characterized by analytical RP-HPLC and mass spectrometry (Supplementary Fig. 1).

### H2B-biotin, H2B-fluorescein, and H3-Cy3 synthesis

H2B T115C and H3 G33C were expressed and purified as described previously for mutant histones. Maleimide reactions were performed with EZ-Link Maleimide-Peg<sub>2</sub>-Biotin (Thermo Fisher Scientific), Fluorescein-5-Maleimide (Thermo Fisher Scientific), and Cy3-Maleimide (Lumiprobe) to produce the corresponding H2B-biotin, H2B-fluorescein, and H3-Cy3 proteins, respectively. Lyophilized protein was dissolved in 6 M GuHCl, 20 mM Tris, 1 mM TCEP, pH 7.2 at a concentration of 3 mg/mL. A 10-fold molar excess of maleimide reagent was then added to the protein and allowed to react for 5 hr at 4 °C. Reactions were quenched with DTT and subsequently purified by semi-preparative RP-HPLC. Purified protein was characterized by mass spectrometry and analytical RP-HPLC (Supplementary Fig. 1, 2b, 2d).

### Assembly of histone octamers, tetramers, and dimers

Histone octamers used for assembly of the nucleosome library were generated as described previously<sup>24</sup>. To assemble histone octamers for all other experiments, purified histones were dissolved in histone unfolding buffer as described above and combined in ratios of 1:1:0.91:0.91 of wild-type or mutant H2A:H2B:H3:H4, such that all mutant octamers were composed of a single mutant histone mixed with the remaining wild-type histones. For histone tetramer assembly, histones H3 and H4 (or mutants) were mixed in equimolar quantities, and for histone dimer assembly, histones H2A and H2B (or mutants) were mixed in equimolar quantities. The total protein (histone) concentration was adjusted to 1 mg/mL and histone mixtures were placed in Slide-A-Lyzer dialysis cassettes (3.5 kDa MWCO, Thermo Fisher Scientific) and dialyzed 3 × 4 L against octamer refolding buffer as described above. Samples were then concentrated to volumes <500 µL in Vivaspin 500 centrifugal concentrators (10 kDa MWCO, Sartorius) and injected onto a gel filtration column (Superdex 200 Increase 10/300 GL, GE Healthcare Life Sciences) pre-equilibrated with octamer refolding buffer. Fractions containing octamer (or tetramer or dimer), as assessed by the FPLC UV chromatogram and SDS-PAGE analysis, were pooled and concentrated (Vivaspin 500 concentrators). Glycerol was added to a final concentration of 50% (v/v) and concentrations determined by UV spectroscopy at 280 nm. Octamers, tetramers, and dimers were stored at -20 °C until use in nucleosome assemblies.

### DNA preparation

**Barcoded 601 (BC-601) DNA**—Unique nucleosome identifier barcodes were added via PCR. First, the linear DNA template was generated by amplifying off of a KS BlueWhite screening vector bearing the 147-bp 601 sequence with a PstI site<sup>24</sup>. Primers were designed to amplify the 601 sequence with a 9-bp 5' overhang and a 15-bp 3' overhang generated from the surrounding vector.

Forward: 5'-CACCGCGTGACAGG-3'

Reverse: 5'-CCTAGGTCTCTGATGCTGG-3'

171-bp DNA fragment generated (bold, 601 sequence, bold underlined, PstI site):

5'-

**CACCGCGTGACAGGATGTATATATCTGACACGTGCCTGGAGACTAGGGA**

**GTAATCCCCTTGGCGGTAAAACGCGGGGGACAGCGCGTACGTGCGT  
TTAAGCGGTGCTAGAGCTGTCTACGACCAATTGAGCGGCTGCAGCACC  
GGGATTCTCCAGCATCAGAGACCTAGG-3'**

The resultant PCR product was subjected to DpnI (New England Biolabs) digest and purified using a QIAquick PCR Purification Kit (Qiagen). This linear DNA fragment (above) was then used as a template for all PCR reactions for the addition of unique barcodes. Primers were designed to amplify this fragment and add on a unique hexanucleotide nucleosome identifier barcode, yielding DNA with a 45-bp 5' overhang (inclusive of the 6-bp barcode) and a 15-bp 3' overhang. Hexanucleotide MN identifier barcode sequences can be found in Supplementary Table 1.

Forward (underlined, unique barcode):

5'-  
TACGGCGACCACCGAGATCTACACCAGCGTNNNNNNCACCGCGTGACAGG-  
3'

Reverse (common to all reactions): 5'-ATCCAAAAACTGTGCCGCAGTCGG-3'

207-bp barcoded DNA sequence generated (bold, 601 sequence; bold underlined, PstI site; underlined, unique hexanucleotide nucleosome identifier barcode):

5'-  
TACGGCGACCACCGAGATCTACACCAGCGTNNNNNNCACCGCGTGACAGG  
**ATGTATATATCTGACACGTGCCTGGAGACTAGGGAGTAATCCCCTTGGC**  
**GGTAAAACGCGGGGGACAGCGCGTACGTGCGTTAAGCGGTGCTAG**  
**AGCTGTCTACGACCAATTGAGCGGCTGCAGCACC****GGATTCTCCAGCA**  
TCAGAGACCTAGG-3'

In total, 157 unique DNA sequences (of the above general format) were produced. PCR products were purified using a QIAquick PCR Purification Kit (Qiagen). DNA was quantified by UV spectroscopy at 260 nm and stored at -20 °C.

**Biotinylated MMTV 'buffer' DNA**—Biotinylated MMTV 'buffer' DNA was prepared as previously described<sup>24</sup>.

**207-bp 601 DNA**—The 207-bp, 601-containing DNA used in gel-based (non-library) ACF remodeling assays on individual nucleosomes was generated via PCR using the 171-bp DNA template described above and the following primers:

Forward:

5'-  
TACGGCGACCACCGAGATCTACACCAGCGTGCAGTCACCGCGTGACAGG-  
3'

Reverse: 5'-CCTAGGTCTCTGATGCTGG-3'

207-bp DNA sequence generated (bold, 601 sequence; bold underlined, PstI site):

5'-  
TACGGCGACCACCGAGATCTACACCAGCGTGTCTCAGTCACCGCGTGACAGG  
ATGTATATATCTGACACGTGCCTGGAGACTAGGGAGTAATCCCCTTGGC  
GGTAAACGCGGGGGACAGCGCGTACGTGCGTTTAAAGCGGTGCTAG  
AGCTGTCTACGACCAATTGAGCGGCTGCAGCACCGGGATTCTCCAGCA  
TCAGAGACCTAGG-3'

**147-bp 601 DNA**—DNA containing the 147 bp 601 positioning sequence was produced and purified as described previously with some modifications<sup>52</sup>. *E. coli* DH5 $\alpha$  cells were grown with a pUC19 vector containing 32  $\times$  601 sites separated by EcoRV restriction sites<sup>53</sup>, pelleted, and subjected to alkaline lysis. The supernatant was clarified by centrifugation, filtered, and isopropanol was added to 0.52 volumes to precipitate plasmid DNA. Precipitate was collected by centrifugation at 5000 *g* for 20 min at 4 °C, resuspended in TE 10/50 buffer (10 mM Tris, 50 mM EDTA, pH 8.0), and treated with RNase A (Qiagen) at 30  $\mu$ g/mL overnight at 37 °C. DNA was isolated via three sequential phenol:chloroform:isoamyl-alcohol (25:24:1) extractions, and then precipitated by addition of 0.2 volumes of 4 M NaCl and 0.4 volumes of 40% (w/v) PEG 6000 and incubation at 4 °C for 30 min. DNA was pelleted by centrifugation at 3000 *g* for 20 min at 4 °C and dissolved in TE 10/0.1 buffer (10 mM Tris, 0.1 mM EDTA, pH 8.0). PEG was removed by two chloroform:isoamyl alcohol (24:1) extractions, and DNA was precipitated by addition of 0.1 volumes of 3 M sodium acetate (pH 5.2) and 2.5 volumes of cold absolute ethanol. DNA was pelleted by centrifugation and dissolved in TE 10/0.1. Plasmid DNA was then digested overnight with EcoRV-HF (New England Biolabs) at 37 °C. Digested 601 fragments were purified by PEG precipitation at a PEG 6000 concentration of 9% (v/v, a percent empirically determined by titration to separate digested 601 fragments from plasmid vector). PEG was removed by two chloroform:isoamyl alcohol (24:1) extractions followed by cold ethanol precipitation, as performed previously, to give purified 601 DNA.

### Nucleosome reconstitution for the MN library

Nucleosomes were reconstituted as previously described<sup>24</sup> with the following modification. Following peristaltic pump addition of nucleosome assembly end buffer, samples were removed from dialysis devices, placed into microcentrifuge tubes, and incubated in a 37 °C water bath for 30 min to promote uniform positioning of octamers on the DNA. All 155 unique nucleosomes were prepared in this fashion with the exception of the nucleosome containing H2B-biotin, which was prepared as described in the following section. Nucleosome quality was assessed by native polyacrylamide gel electrophoresis (5% acrylamide gel, 0.5x TBE, 190 V, 35 min) followed by SYBR<sup>TM</sup> Gold Nucleic Acid Gel Stain (Invitrogen) staining. Nucleosome bands appeared to migrate around 600–700 bp, depending on the nucleosome variant (Extended Data Fig. 2b). To prepare the final nucleosome library, normalized amounts (9 pmol) of each nucleosome were combined and concentrated using a Vivaspin 500 centrifugal concentrator (10 kDa MWCO, Sartorius). For storage, PMSF and glycerol were added to final concentrations of 0.5 mM and 20% (v/v), respectively. The final MN library was analyzed by native gel electrophoresis as described above and determined to have a concentration of approximately 1.2  $\mu$ M by UV spectroscopy (260 nm). Library aliquots were flash frozen in liquid nitrogen and stored at –80 °C.

### MN formation for follow-up assays

Purified octamers were combined with 601 DNA (either without overhangs in dimer exchange and thermal stability experiments or with 45/15-bp overhangs for remodeling experiments) in empirically determined ratios in octamer refolding buffer. In cases where a histone mutation prevented octamer formation, H2A/H2B dimers and H3/H4 tetramers were used in stoichiometric ratios (2:1). Mixtures were placed in Slide-A-Lyzer MINI dialysis devices (3.5 kDa MWCO, Thermo Fisher Scientific) and dialyzed against 200 mL nucleosome assembly start buffer for 1 hr at 4 °C. 330 mL nucleosome assembly end buffer was then added via peristaltic pump at a rate of 1 mL/min, after which samples were removed to tubes and heated at 37 °C for 30 min to promote repositioning, as described above. Samples were then placed on ice for 10 min and spun down for 10 min at 17,000 *g* and 4 °C. Supernatants were placed back into dialysis devices and dialyzed 2 × 200 mL against nucleosome assembly end buffer (first dialysis for 4 hr, second dialysis overnight) at 4 °C. Samples were then removed to microcentrifuge tubes, centrifuged at 17,000 *g* for 10 min at 4 °C to remove any precipitate, and final supernatants transferred to fresh tubes. Nucleosomes were quantified by UV spectroscopy (260 nm) and stored at 4 °C. Nucleosome quality was assessed by native gel electrophoresis as described above.

### Multiplex barcode and Illumina adaptor addition

Approximately 10 pg DNA from each purified experimental sample from either dimer exchange or remodeling assays on the nucleosome library (described in the corresponding sections below) were amplified using the following primers:

Forward: 5'-AATGATACGGCGACCACCGAG-3'

Reverse (XXXXXX, unique hexanucleotide multiplexing barcode):

5'-

CAAGCAGAAGACGGCATAACGAGATXXXXXXXXCCTAGGTCTCTGATGCTGGA  
GA-3'

Hexanucleotide multiplexing barcodes used in this study can be found in Supplementary Table 7. Full-length (un-cut) DNA from each experiment was amplified using a Phusion High-Fidelity PCR Kit (New England Biolabs) (0.2 mM dNTPs, 10 pg DNA, 0.5 μM each primer, 0.50 μL polymerase, 50 μL total reaction volume) with the above primers and the following amplification conditions: step 1: 50 °C for 2 min; step 2: 95 °C for 2 min; step 3: 95 °C for 15 sec; step 4: 60 °C for 1 min. Steps 3–4 were repeated for a total of 12 cycles. To ensure reactions remained in the exponential phase, thus preserving the relative abundances of DNA fragments, qPCR analysis was performed using 10 pg of select library experimental samples and PowerUp™ SYBR™ Green Master Mix reagent (Applied Biosystems) with detection on a ViiA 7 Real-Time PCR System (Applied Biosystems). Identical conditions were used as described above and a mean *C<sub>t</sub>* value of 13.8 was determined (compare to 12 cycles used for amplification of DNA prior to Illumina sequencing). The following is the general amplicon sequence that was generated to be compatible with Illumina sequencing (bold, 601 sequence; bold underlined, PstI site; NNNNNN, unique hexanucleotide nucleosome identifier barcode; XXXXXX, unique hexanucleotide multiplexing barcode):

5'-  
 AATGATACGGCGACCACCGAGATCTACACCAGCGTNNNNNNCACCGCGTG  
 ACAGGATGTATATATCTGACACGTGCCTGGAGACTAGGGAGTAATCCCC  
 TTGGCGGTTAAACGCGGGGACAGCGCGTACGTGCGTTTAAGCGGT  
 GCTAGAGCTGTCTACGACCAATTGAGCGGCTGCAGCACCGGGATTCTC  
 CAGCATCAGAGACCTAGGXXXXXXTCAGATCTGGTACCCAGCTTTTGT-3'

The full forward adaptor sequence is 5' of the nucleosome identifier barcode (NNNNNN), and the full reverse adaptor sequence is 3' of the multiplexing barcode (XXXXXX). For each sample, correct amplicon size (242 bp) was verified by gel electrophoresis and staining with SYBR™ Gold Nucleic Acid Gel Stain (Invitrogen). Approximately equal amounts of DNA samples with multiplexing barcodes were pooled and purified using a QIAquick PCR Purification Kit (Qiagen). After purification, samples were subjected to high-throughput DNA sequencing as described in the following section.

### Illumina sequencing

Single-read sequencing of barcoded DNA libraries (starting from the forward adaptor and covering the unique nucleosome identifier barcode) was performed on an Illumina MiSeq instrument with a read length of 185 bp. Due to the substantial sequence homogeneity present in all library samples, DNA sequencing libraries were diluted with a PhiX control library to enable sequencing. Sequencing reads from library experiments were imported into Galaxy (Princeton University installation) and separate FASTQ files generated based on multiplexing barcodes using the Galaxy Barcode Splitter tool. FASTQ files were exported from Galaxy, and the number of MN barcode reads were counted in each experiment (i.e. each FASTQ file) using a custom R script<sup>24</sup>. The resulting output provided read counts for each individual MN in each multiplexed experiment.

### Assessment of MN library assembly using next-generation sequencing

MN fidelity in the final, assembled MN library was determined using a modified version of the restriction enzyme accessibility assay. This assay, described in the ACF remodeling assays section below, enables determination of rates of MN sliding based on liberation of a PstI restriction site embedded in the nucleosome 601 positioning sequence. The restriction enzyme digestion employed in this assay can also be used to quantify the fraction of disassembled nucleosomes for each individual MN in the library. Aliquots of the DNA-barcoded MN library were treated with or without 2 U/μL PstI-HF (New England Biolabs) at 37 °C for 64 min. Barcoded DNA was amplified and prepared for and subjected to next-generation sequencing as described above. MN barcode read counts were first normalized among samples to the 'uncuttable' (PstI site absent) non-nucleosomal DNA barcode read counts. The 'cuttable' (PstI site present) non-nucleosomal DNA barcode served as a positive control for PstI digestion. These data confirmed that MNs in the library were well assembled (98 ± 5%) (Extended Data Fig. 2d).

### Assessment of MN library stability

Nucleosomes were assembled as described above for the nucleosome library. H2BE76K-containing nucleosomes were assembled on 207-bp DNA. H2B-fluorescein-containing



nucleosomes were assembled on 192-bp DNA. This difference in DNA length (207-bp vs. 192-bp) enabled differentiation of these nucleosome species by native gel polyacrylamide electrophoresis (Extended Data Fig. 2e). To assess whether barcodes would exchange between nucleosomes or if nucleosomes would fall apart under standard assay (37 °C) or stringent (50 °C) conditions, nucleosomes were incubated either alone, in 1:1, or in 1:10 ratios at concentrations similar to what might be used in a library assay for 1 hr at the indicated temperatures. Following incubation, samples were analyzed by native gel polyacrylamide electrophoresis along with MNs containing H2B-fluorescein generated on 207-bp DNA as a positive control for exchange. Neither barcode exchange nor denaturation of nucleosomes was observed (Extended Data Fig. 2e).

### Humanized yeast library

The yDT51 yeast strain<sup>27</sup> was used for the development of the humanized yeast library. yDT51, which contain a URA3 plasmid bearing yeast histones (pDT105), were transformed with an ‘8-swap’ human histones TRP plasmid (pDT130) containing either wild-type or mutant histones. Standard site-directed mutagenesis techniques were used to generate histone mutant plasmids, and transformations were accomplished using a lithium acetate and DMSO approach as described previously<sup>54</sup>. Transformations were plated on SC-URA-TRP and individual colonies isolated after four days. Liquid SC-URA-TRP cultures were then grown overnight. Cultures were diluted and allowed to double twice to mid/late log phase, and then 10-fold serial dilutions were spotted on SC-URA-TRP + 5-FOA plates. Plates were imaged 6–7 days after spotting using an ImageQuant LAS-4000 (GE Healthcare Life Sciences). Images shown (Fig. 2g, 4f, and Extended Data Fig. 3b) are representative of two separate transformations (biological replicates), each performed in technical duplicate. All histone mutants included in the humanized yeast library, along with their phenotypes, are listed in Supplementary Table 2.

### Nap1 dimer exchange assays

The mammalian Nap1 protein (*Mus musculus*, NP\_056596.1) was expressed and purified as described previously<sup>55</sup>. Nap1-mediated H2A/H2B dimer exchange assays were performed with fluorescein-labeled dimers for in-gel fluorescence experiments and with biotinylated dimers for MN library experiments. Assays were carried out at a MN concentration of 50–100 nM in 20 mM Tris, 100 mM NaCl, 1 mM DTT, pH 7.5. Nap1, H2A/H2B dimers, and MN substrates ratios were fixed at 50:5:1 for all assays. Prior to addition to the MN substrates, Nap1 and H2A/H2B dimers were first co-incubated at 4 °C for 1 hr to allow for complex formation. Dimer exchange assays were initiated by addition of complexed Nap1 and H2A/H2B dimers to MN substrates and samples incubated at 37 °C for 1 hr.

For in-gel fluorescence analysis of dimer exchange, an experimental sample volume corresponding to 0.5 pmol of MN was immediately separated on 2% UltraPure™ Agarose (Invitrogen) native gels. Fluorescence from fluorescein-labeled H2A/H2B dimers was measured using a Typhoon FLA 9500 biomolecular imager (GE Healthcare Life Sciences). Native gels were then stained with ethidium bromide (Thermo Fisher Scientific) for 1 hr to detect MN substrates and densitometry measurements determined as described under general lab protocols.

Nap1 dimer exchange assays on the MN library were performed in quadruplicate, with samples collected both before and after streptavidin affinity purification for Illumina sequencing (8 samples total). Barcoded DNA reads pre- and post-enrichment were used to control and normalize barcode counts intra-sample, allowing for robust calculation of barcode enrichment (indicative of H2A/H2B dimer exchange). Specifically, 2% of each Nap1 dimer exchange assay sample was quenched into 500  $\mu$ L Qiagen PB buffer and barcoded DNA was later purified using the QIAquick PCR Purification Kit (Qiagen). The remainder of the samples were incubated with magnetic streptavidin resin (Dynabeads™ MyOne™ Streptavidin T1, Invitrogen) and allowed to bind for 1 hr at 4 °C. The resin was then washed three times with 10x volume (relative to resin slurry volume) MN wash buffer (20 mM Tris, 500 mM NaCl, 0.1% (v/v) IPEGAL CA-630, 0.1% BSA, pH 7.5). These wash conditions effectively removed non-specific MN-resin interactions without disrupting biotin-captured MNs. Affinity purified nucleosomes were eluted with 5x volume elution buffer (100 mM Tris, 10 mM EDTA, 1% SDS, 10 mM BME, 1/1000 (v/v) Proteinase K (New England Biolabs), pH 7.5) at 50 °C for 1 hr at 750 rpm in a thermomixer (Eppendorf). DNA was then purified using the QIAquick PCR Purification Kit (Qiagen) and prepared as described for Illumina sequencing. Barcode reads were first normalized among samples by the uncuttable, non-nucleosomal DNA standard. Relative enrichment values were calculated for each of the four experimental replicates by dividing MN barcode reads in post-enrichment samples by their corresponding values before streptavidin enrichment. Enrichment values were then divided by the average value of the five wild-type MN controls to give enrichment measurements of mutant MNs as a fold-change relative to wild-type. Nap1 dimer exchange values were reported as the average of these values (with standard deviations) from the four experimental replicates. Raw read counts and subsequent calculations can be found in Supplementary Table 3.

Assays to determine whether H3/H4 tetramers swap between nucleosomes during Nap1 dimer exchange assays were performed using similar conditions with the following modifications. Mixtures of wild-type MNs and MNs containing H2BE76K and a fluorescently-labeled H3 (H3-Cy3) were subjected to the Nap1 dimer exchange assay at either a 1:1 or 1:10 ratio. Total MN concentration was fixed at 100 nM in all assays. Samples were separated on a 5% TBE native gel and Cy3 and ethidium bromide fluorescence were measured by a Typhoon FLA 9500 biomolecular imager (GE Healthcare Life Sciences).

### Thermal stability assays

Nucleosome stability was measured using the SYPRO Orange (Sigma-Aldrich) thermal stability assay<sup>56</sup>. Nucleosomes were assembled with 147 bp 601 DNA and stored in 10 mM TEK (10 mM Tris, 0.1 mM EDTA, 10 mM KCl, pH 7.4) at 4 °C for less than a week prior to use. SYPRO Orange was supplied as a 5000x concentrated solution from the manufacturer and was used at a final concentration of 5x in the assay. Assays were performed in a 10  $\mu$ L reaction volume in 10 mM TEK with a MN concentration of 0.5–1  $\mu$ M and SYPRO Orange concentration of 5x. Samples were analyzed in triplicate in 384-well plates in a ViiA 7 Real-Time PCR System (Applied Biosystems).  $T_{\text{half}}$  values were calculated by fitting fluorescence measurements at temperatures between 45 and 83 °C, inclusive, to a standard sigmoid function and calculating the inflection point.  $T_{\text{half}}$  values statistics were reported as

the average and standard deviation from three experimental replicates, each fit independently. Fits were performed in Python 2.7.10 using the `optimize.curve_fit` function in the SciPy library (0.15.1).

### ACF remodeling assays

The human ACF complex was expressed and purified from Sf9 cells as previously described<sup>24</sup>. All remodeling assays were conducted using a restriction enzyme accessibility (REA)-based strategy and performed similarly to those previously described<sup>24</sup>. Each remodeling reaction utilized 10 nM nucleosome substrate (MN library or individual MNs) and 0.4 nM ACF in a total reaction volume of 60  $\mu$ L, concentrations that resulted in approximately 60% of the wild-type nucleosomes remodeled in 64 min. Remodeling assays were carried out in REA buffer (12 mM HEPES, 4 mM Tris, 60 mM KCl, 10 mM MgCl<sub>2</sub>, pH 7.6, 10% (v/v) glycerol, and 0.02% (v/v) IGEPAL CA-630) in the presence of 2 U/ $\mu$ L PstI-HF restriction enzyme (New England Biolabs). All reaction components except ATP were pre-incubated together at 30 °C for 10 min to allow for cutting of any free DNA. Then ATP, or equivalent buffer volume for no ATP control samples, was added to a final concentration of 2 mM to initiate the remodeling reaction. Remodeling assays were carried out for a total of 64 min at 30 °C, with 6  $\mu$ L aliquots removed at 0, 2, 4, 8, 16, 32, and 64 min and each quenched with 9  $\mu$ L quench buffer (20 mM Tris, 70 mM EDTA, pH 8.0, 10% (v/v) glycerol, 2% SDS).

For gel-based analysis of ACF remodeling, reactions were performed in triplicate with seven time points taken per reaction ( $t = 0, 2, 4, 8, 16, 32, 64$  min). Additionally, one sample was prepared without ATP and incubated for 64 min at 30 °C to serve as a negative control. Samples (7  $\mu$ L) were analyzed directly on a 5% polyacrylamide gel (0.5x TBE, 195 V, 35 min) and staining was performed with SYBR<sup>TM</sup> Gold Nucleic Acid Gel Stain (Invitrogen). Gel imaging and densitometry calculations were performed as described under general lab protocols.

ACF remodeling experiments on the MN library were performed in triplicate, with and without ATP, over the same seven time points. Samples were then deproteinized with 30 U/mL proteinase K (New England Biolabs) for 1 hr at 37 °C. DNA was subsequently purified using a QIAquick PCR Purification Kit (Qiagen) and prepared as described for Illumina sequencing. Barcode reads were first normalized among samples by the uncuttable, non-nucleosomal DNA standard. In order to compare remodeling rates of different MNs, barcode reads for each MN were re-scaled by their corresponding values at  $t = 0$  min (as in Extended Data Fig. 7b). Remodeling rates were determined from a collective fit of the time-course data to a single one-phase exponential decay function with a y-intercept fixed at a value of 1, yielding  $k_{\text{obs}}$  rate values and their corresponding errors. Fits were performed with Python 2.7.10 using the `optimize.curve_fit` function in the SciPy library (0.15.1). Raw read counts and subsequent calculations can be found in Supplementary Table 4.

### Production of H2B-CfaN-TwinStrepTag-HisTag

The fusion construct **H2B-CfaN-TwinStrepTag-HisTag** was produced using standard cloning techniques in the same pET expression vectors described previously. The final construct had the following amino acid sequence:

**MPEPAKSAPAPKKGSKKAVTKAQKKDGGKRRKRSRKESYSVYVYKVLKQ**  
**VHPDTGISSKAMGIMNSFVNDIFERIAGEASRLAHYNKRSTITSREIQTAVR**  
**LLLPGELAKHAVSEGTKAVTKYTSKCLSYDTEILTVEYGFPLPIGKIVEERIE**  
**CTVYTVDKNGFVYTQPIAQWHNRGEQEVFEYCLEDGSIIRATKDHKFMTTDG**  
**QMLPIDEIFERGLDLKQVDGLPWSHPQFEKGGGSGGGSGGSAWSHPQFEKGS**  
**HHHHHH**

Protein was expressed in *E. coli* BL21(DE3) cells and purified with standard Ni-NTA affinity chromatography protocols under denaturing conditions. Protein was then further purified by semi-preparative RP-HPLC, characterized by analytical HPLC and mass spectrometry (Supplementary Fig. 2c), and subsequently lyophilized and stored at  $-80^{\circ}\text{C}$  until use. H2A/H2B-CfaN dimers were assembled and purified by FPLC as described above with the addition of TCEP (5 mM) to all buffers to prevent disulfide formation.

### Generation of heterotypic MNs

601 DNA, H3/H4 tetramers, and H2A/H2B dimers were mixed in ratios of 1:1:1.2, which produced approximately equal amounts of hexasomes and nucleosomes. During this step, non-wild-type dimers were used (e.g. mutant or ubiquitinated dimers), since wild-type dimers were used throughout the rest of the protocol. Hexasome-nucleosome mixtures were then diluted to a concentration of  $0.6\ \mu\text{M}$  in complementation buffer (20 mM Tris, 50 mM KCl, 5 mM TCEP, pH 7.6) and H2A/H2B-CfaN-Streptag dimers were added at a DNA:dimer ratio of 1:1 and allowed to incubate for 60 min at room temperature. Note that the optimal quantity of intein-tagged dimers was determined empirically, as over-addition of these dimers caused nucleosome precipitation. Nucleosome complementation was complete after 60 min, as assessed by native gel electrophoresis on a 5% TBE gel. To aid in resin purification of intein-tagged asymmetric nucleosomes, 0.02% (v/v) IGEPAL CA-630 (Sigma-Aldrich) was added. For thermal stability experiments using SYPRO Orange dye, however, use of IGEPAL CA-630 resulted in high background signals. Therefore, for these experiments 0.01% (w/v) octylglucoside (Alfa Aesar) was used as a replacement for IGEPAL CA-630. Intein-tagged nucleosomes were then bound to Strep-Tactin Superflow Plus resin (Qiagen) pre-washed with complementation buffer, using approximately 200  $\mu\text{L}$  resin slurry for 300 pmol 601 DNA. Resin was washed with 4 volumes of nucleosome wash buffer (20 mM Tris, 0.5 M NaCl, pH 7.6, and 0.02% (v/v) IGEPAL CA-630 or 0.01% (w/v) octylglucoside) and 2 volumes of nucleosome splicing buffer (10 mM Tris, 250 mM KCl, 0.1 mM EDTA, 1 mM TCEP, 10 mM 2-mercaptoethanol, pH 7.6, and 0.02% (v/v) IGEPAL CA-630 or 0.01% (v/v) octylglucoside). Nucleosomes were eluted via on-resin thiolysis with the catalytically dead C-intein CfaC<sup>30</sup> (Supplementary Fig. 2c). Bolded residues denote mutations in CfaC that prevent intein splicing:

**VKIISRKSLGTQNVYDIGVEKDHNFLKNGLVASAAFN**

Dead CfaC was diluted to 20  $\mu$ M in nucleosome splicing buffer and two volumes were added to the resin and allowed to incubate for 8 hr with periodic mixing. Eluted heterotypic nucleosomes were separated from the resin using empty Micro Bio-Spin columns (Bio-Rad) and subsequently concentrated and buffer exchanged into 10 mM TEK buffer using 30 kDa MWCO ultrafiltration centrifugal concentrators (Sartorius). Heterotypic nucleosomes were visualized by native gel electrophoresis on a 5% TBE gel and concentrations were determined by UV spectroscopy at 260 nm.

### Generation of mutant H2B-expressing cell lines

HEK293T (CRL-11268, ATCC) and C3H10T1/2 (CCL-226, ATCC) were cultured in Dulbecco's Modified Eagle Medium (DMEM, Corning) with 10% fetal bovine serum (FBS, Atlanta Biologicals) at 37 °C, 20% oxygen, and 5% CO<sub>2</sub>. Cells were dissociated with TrypLE Express (Thermo Fisher Scientific) during passaging and regularly tested and found free of mycoplasma.

The murine HIST2H2B cDNA sequence (shown below) was used to generate all H2B constructs:

```

ATGCCTGAACTGGCCAAATCTGCCCGGCTCCCAAGAAGGGTTCCAAGAA
GGCTGTACCAAGGCGCAAAAGAAAGATGGCAAGAAGCGCAAGCGCAGC
CGCAAGGAGAGCTACTCCATCTATGTGTACAAAGTGCTGAAGCAGGTGCAC
CCGGACACCGGCATCTCGTCCAAGGCCATGGGCATCATGAACTCGTTTGTG
AATGACATCTTTGAGCGCATAGCAAACGAGGCTTCTCGCCTGGCGCATTAC
AACAAGCGGTCTACAATCACATCGCGGGAGATCCAGACGTCGGTGCGCCTT
CTACTGCCCGGAGAACTGGCCAAGCACGCCGTGTCGGAGGGCACCAAGGC
CGTCACCAAGTACACCAGCGCCAAGGCGGCCGCTGGAGGAGACTACAAGG
ACGACGATGACAAGTCGGCCGCTGGAGGATACCCCTACGACGTGCCCCGACT
ACGCCTAG

```

HA-FLAG-tagged wild-type and mutant H2B constructs were cloned into pCDH-EF1-MCS-Puro lentiviral vectors (Systems Biosciences) as previously described<sup>11</sup>. To produce lentivirus, HEK293T cells (ATCC) were transfected with the lentiviral vectors and helper plasmids (psPAX2 and pMD2.G). After 72 hr, the supernatant (containing lentivirus) was collected, filtered, and concentrated. C3H10T1/2 cells (ATCC) were transduced with concentrated lentivirus ( $2 \times 10^7$  IFU) and grown under puromycin selection (2  $\mu$ g/mL) 48 hr after transduction.

### Acid histone extraction and immunoblotting

Cells were lysed with hypotonic lysis buffer (10 mM Hepes, 10 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.5 mM DTT, protease inhibitors) for 1 hr at 4 °C. The nuclei pellet was resuspended in 0.2 N H<sub>2</sub>SO<sub>4</sub> overnight at 4 °C with agitation. After centrifugation (16,000 g, 10 min at 4 °C), supernatants were collected and histone proteins were precipitated in 33% TCA. The histone pellet was washed with acetone and resuspended in deionized water. Sample concentrations were quantified using a protein quantification assay (BioRad, cat #5000006) using a mixture of histones H2A, H2B, H3 and H4 as a standard. Before gel analysis, samples were resuspended in SDS-Laemmli buffer and boiled for 10 min. Samples were run on Novex™

16% Tris-Glycine gels (Thermo Fisher Scientific) and subsequently transferred to nitrocellulose membranes. Membranes were then blocked (PBS with 5% milk and 0.1% Tween-20) and incubated with the following primary antibodies overnight in PBS with 2% milk and 0.1% Tween-20: rabbit anti-histone H2B antibody (Cell Signaling Technology, 12364) and mouse anti-HA tag antibody (BioLegend, 901501) at 1:1000 dilutions. HRP-coupled secondary anti-mouse (GE Healthcare Life Sciences, NA931) and anti-rabbit (Bio-Rad, 1706515) antibodies were used at 1:5000 dilutions and detected by enhanced chemiluminescence (Immobilon Western, Millipore) using an Amersham Imager 600. To quantify relative expression levels of mutant histones, a standard curve of untagged murine H2B (mH2B) and tagged mH2B (mH2B-FLAG-HA) were loaded adjacent to samples from cellular histone fractions, allowing the estimation of relative histone expression by comparison of anti-HA and anti-H2B signals (Extended Data Fig. 8a, Supplementary Fig. 2e).

### Immunofluorescence

C3H10T1/2 cells were seeded onto sterile coverslips. After 24 hr, cells were washed with PBS and fixed with 4% paraformaldehyde for 20 min. Cells were permeabilized with 0.1% Triton X-100 in PBS for 5 min, washed twice with PBS, and then blocked with 5% BSA in PBS for 20 min. Coverslips were rinsed with PBS and incubated overnight with mouse Alexa Fluor 488 anti-HA antibody (BioLegend, 901509) at a dilution of 1:100. Coverslips were then washed five times with PBS and samples stained with DAPI (4',6-diamidino-2-phenylindole). Fluorescent images were obtained with a Zeiss fluorescence microscope.

### RT-qPCR and RNA-seq analyses

RNA was isolated using an RNeasy Mini Kit (Qiagen). For quantitative real-time PCR (RT-qPCR), extracted RNA was treated with DNase and cDNA was synthesized using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). qPCR was performed using Power SYBR Green PCR Master Mix (Applied Biosystems) and data normalized to a GAPDH reference gene. All qPCR primers are listed in Supplementary Table 8. For RNA sequencing (RNA-seq), libraries were prepared according to the NEBNext Ultra II RNA Library Prep Kit (New England Biolabs) protocol and sequenced on an Illumina NextSeq instrument.

Quantification of RNA transcripts was performed with the k-mer counting software Salmon<sup>57</sup>, and a reference genome produced using Illumina's iGenomes UCSC mm10 package. Expression levels were estimated by counting reads with featureCounts (v1.4.4) using default settings. Read summarization was performed at the gene symbol level using UCSC mm10 known gene exons obtained from the UCSC table browser. The gene symbol-summarized read counts were normalized using the variance stabilized transformation technique (VST), and expression levels clustered using Euclidean distance and hierarchical clustering with the Ward variance minimization algorithm to produce heatmaps. Differentially expressed genes were identified using the DESeq Package<sup>58,59</sup>. Differentially expressed genes were determined by a  $\log_2$  fold-change cutoff of  $\pm 1$  with a 5% FDR. Gene ontology (GO) analysis was performed with DAVID and top GO and KEGG pathway terms were plotted against their Benjamini FDR values on a logarithmic scale.

### Cellular differentiation assays

Adipocyte differentiation was performed by stimulating confluent C3H10T1/2 cells with a cocktail of 0.5 mM isobutylmethylxanthine, 1  $\mu$ M dexamethasone, 5  $\mu$ g/ml insulin, and 5  $\mu$ M troglitazone (all from Sigma-Aldrich) in DMEM with 10% FBS. After two days, cells were maintained in DMEM with insulin for an additional five days. On day seven, cells were washed in PBS and then fixed with 3% formaldehyde for 20 min at room temperature. Cells were then stained with Oil Red O solution (Sigma-Aldrich) and the extent of differentiation quantified by measuring absorbance at 500 nm.

Chondrocyte differentiation was performed by stimulating confluent C3H10T1/2 cells with a cocktail of 10  $\mu$ g/mL insulin (Sigma-Aldrich), 30 nM sodium selenite (Sigma-Aldrich), 10  $\mu$ g/mL human transferrin (Sigma-Aldrich), 10 nM dexamethasone (Sigma-Aldrich), and 100 ng/mL rhBMP-2 (R&D Systems) in DMEM with 1% FBS. Medium was replaced with fresh cocktail every 2–3 days until day nine when cells were washed in PBS and then fixed with 4% formaldehyde for 5 min at room temperature. Cells were stained with 1% Alcian Blue 8GX (pH 1.0, Sigma-Aldrich) for 1 hr at room temperature. For quantification, cells were lysed in 1% SDS and absorbance measured at 605 nm.

### Statistics and Reproducibility

Single preparations of enzymes (Nap1 and ACF) were used for all relevant experiments. Histone octamer, tetramers, and dimers, as well as individual MNs and the MN library were assembled once. MN library experimental results were validated on independently prepared MNs (formed from their independently purified constituent octamers/tetramers/dimers). Cellular data reflect representative results of at least two independent experiments. All data are reported as the mean of experimental replicates with error bars designating either standard deviations or standard error of the mean (denoted in figure legends). Measurements were taken from biologically independent samples in all experiments. Statistical tests, the number of experimental replicates ( $n$ ), and significance thresholds are described in the figure legends. Poisson distributions were fit in Python (2.7.10) using the `optimize.curve_fit` function in the SciPy library (0.15.1). Two-sample Kolmogorov-Smirnov tests were performed in Python using the `stats.ks_2samp` function in the SciPy library. Two-sample  $t$ -tests were performed in Excel. An F-test of equality of variances was used to justify cases where equal-variance  $t$ -tests were performed. Correction for multiple hypotheses testing was performed using the Benjamini-Hochberg procedure for false discovery rates<sup>60</sup>.

### Code Availability

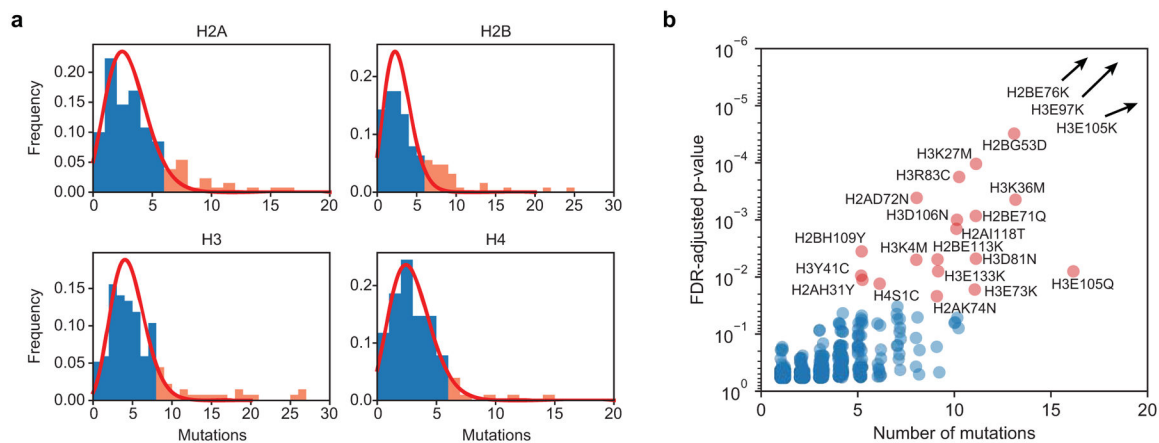
A custom R script was used to count individual DNA barcodes from sequencing data in MN library experiments and was previously reported<sup>24</sup>. Custom Python (2.7.10) scripts were used to fit and determine  $T_{\text{half}}$  values in MN thermal stability experiments and to calculate  $k_{\text{obs}}$  values in ACF remodeling experiments are available at <https://github.com/jdbagert/oncohistonelibrary>.

### Data Availability

RNA-seq raw data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE146959.

All other data are included within this paper, Extended Data, and Supplementary Information files.

## Extended Data



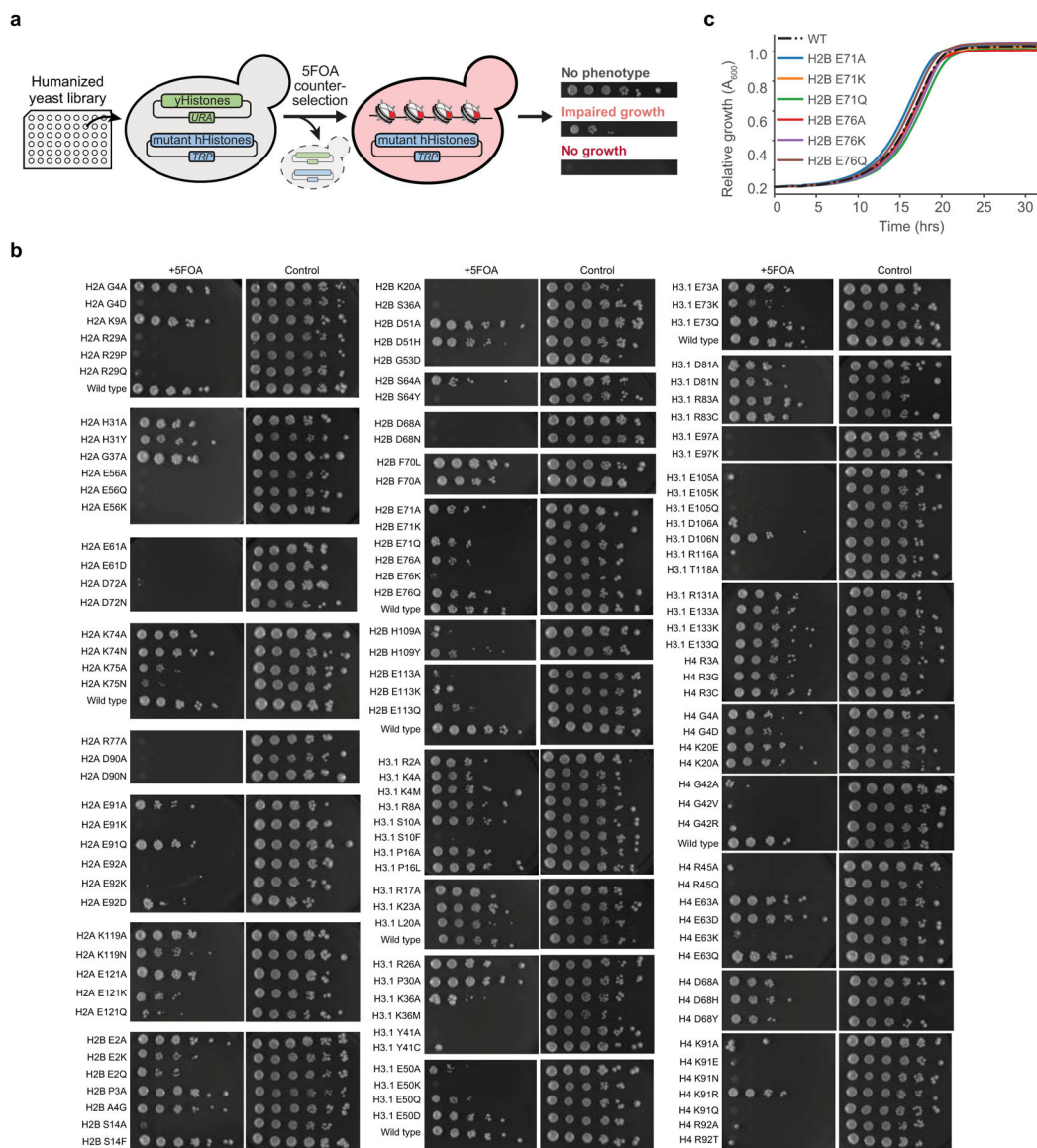
### Extended Data Figure 1 | Selection criteria for histone mutations to include in high-throughput libraries.

**a**, Distributions of histone mutations (tumor mutational burden  $\geq 10$ ) across the four canonical histones were fit to a Poisson distribution (red line) to identify sites of high mutational frequency for inclusion in the nucleosome library. Mutation values for each histone in the 99<sup>th</sup> percentile of the Poisson distribution are marked in red. **b**, The set of amino acid mutations observed at each histone site was assessed for overrepresentation of specific histone mutations by calculating the probability of selecting the observed number of specific mutations randomly, adjusting for codon bias. Probabilities were then adjusted for false discovery rates by the Benjamini-Hochberg method. FDR-adjusted  $p$ -value  $< 0.03$  are shown in red.



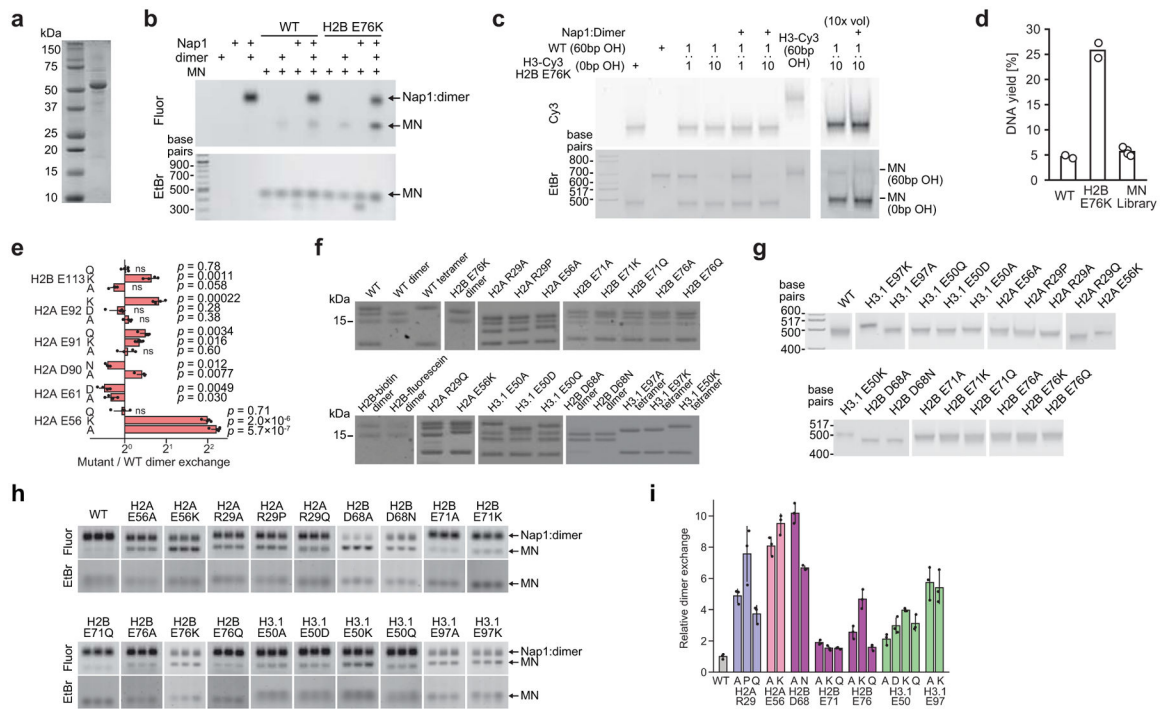


assembled in the library. Data are presented as the mean  $\pm$  SD ( $n = 3$  biologically independent samples). **e**, Unstable nucleosomes do not exhibit spontaneous exchange of DNA with fluorescently labeled nucleosomes, even under conditions of biased stoichiometry and high temperature. Samples containing individual or mixtures of unlabeled H2BE76K MNs and fluorescently tagged (H2B-Fluor) MNs were incubated at the indicated temperatures for 1 hr and then assessed by native gel electrophoresis to determine whether exchange of fluorescent histones occurs. Numbers indicate ratios of MNs in a given sample. H2BE76K MNs were assembled on DNA with a 15 bp 3' overhang (15 bp 3'OH), whereas H2B-Fluor MNs were assembled on DNA with no 3' overhang (0 bp 3'OH) so the two species could be distinguished based on differential gel migration. Additionally, H2B-Fluor MNs were assembled on DNA with a 15 bp 3' overhang as a positive control for DNA exchange.



**Extended Data Figure 3 | Structurally perturbative histone mutations cause growth defects in humanized yeast.**

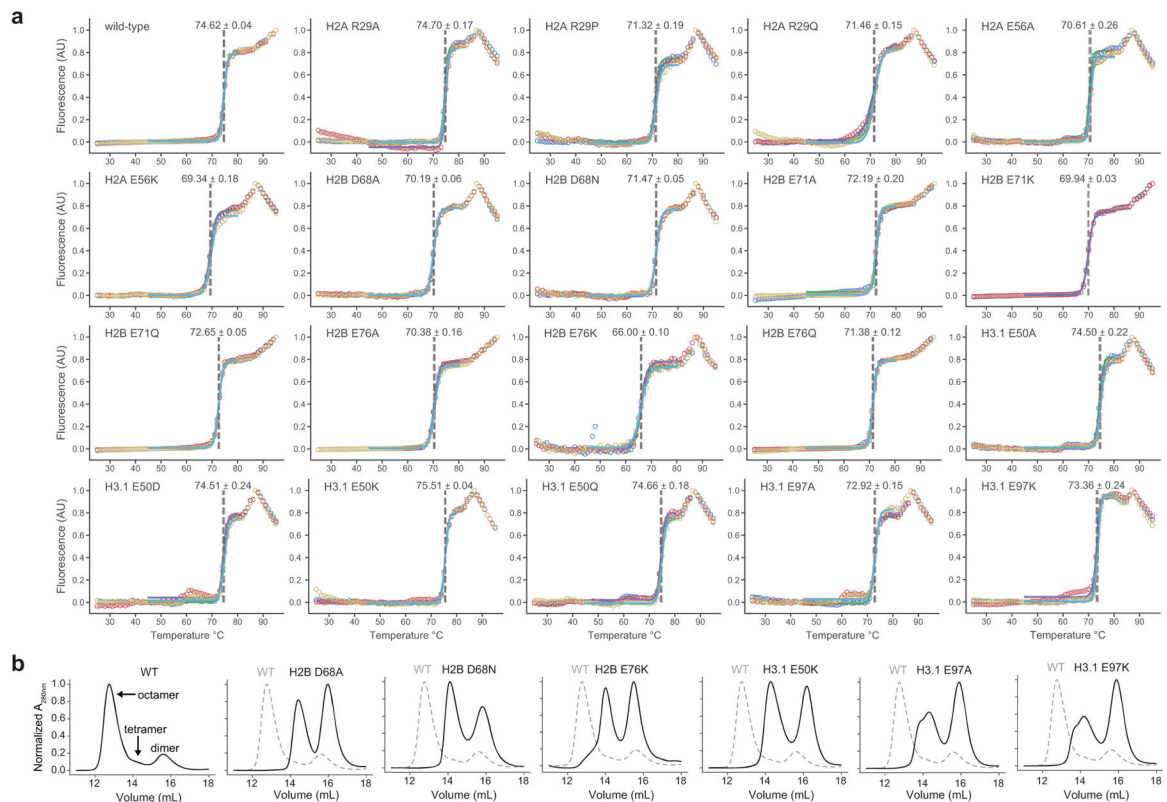
**a**, A humanized yeast library facilitates high-throughput analysis of the effects of histone mutations on yeast viability. A plasmid shuffle assay is used to select for yeast that can subsist solely on mutant human histones. **b**, Serial dilutions of humanized yeast with mutant histones after 5FOA selection. Control cells are grown without 5FOA to show strain growth prior to selection for human histones. **c**, Yeast carrying wild-type and select histone mutants show similar growth rates prior to 5FOA selection. This indicates that growth differences observed on 5FOA plates are not attributable to differential growth rates prior to selection for cells able to subsist solely on mutant human histones.



**Extended Data Figure 4 | A histone dimer exchange assay reveals mutations that destabilize nucleosomes.**

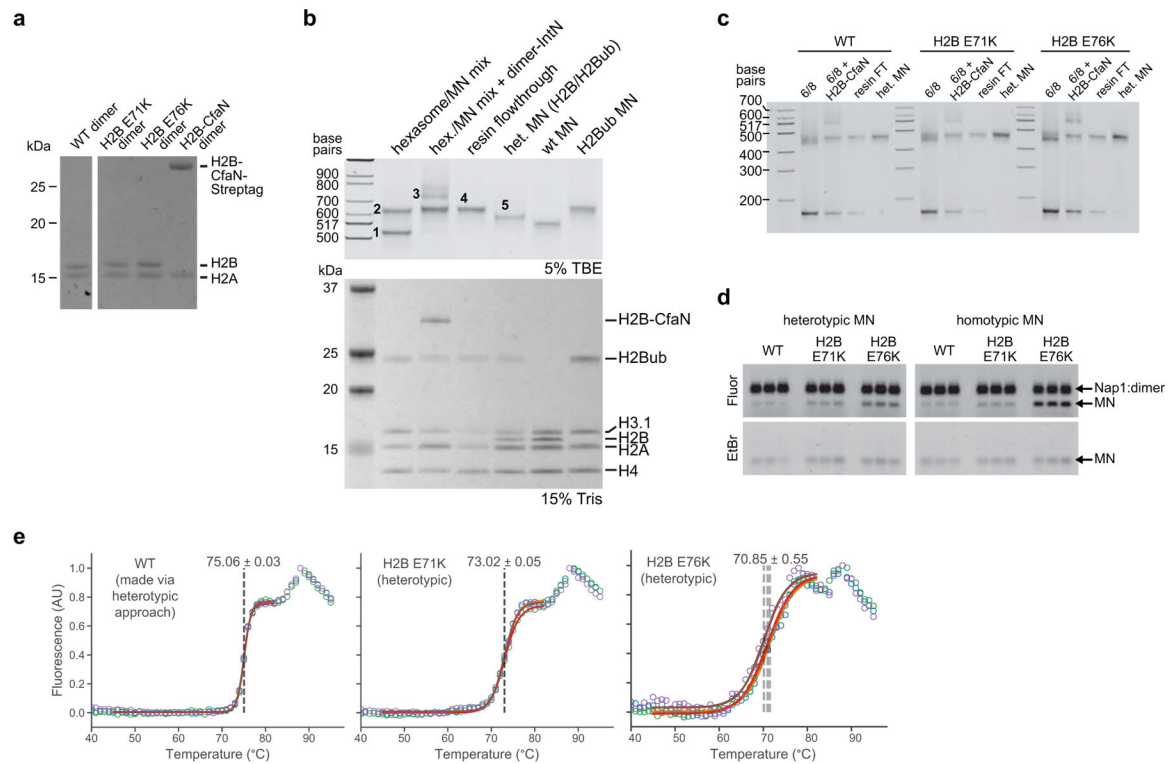
**a**, Denaturing gel analysis of purified mouse Nap1. **b**, Proof-of-concept demonstration of the dimer exchange assay on wild-type and H2BE76K mutant MNs using fluorescently labeled dimers and in-gel fluorescence readout. The lower band in the ethidium bromide stain shows partially disassembled nucleosomes, which are observed only when Nap1 is added to nucleosomes without additional H2A/H2B dimers. Fluor, fluorescein. EtBr, ethidium bromide. **c**, Nucleosomes do not exhibit exchange of H3/H4 tetramers during the Nap1 H2A/H2B dimer exchange assay. Samples containing wild-type MNs or H2BE76K and fluorescently tagged H3 (H3-Cy3) MNs were incubated under our Nap1 H2A/H2B dimer exchange conditions (see methods) and then assessed by native gel electrophoresis to determine whether exchange of fluorescent histones occurs. Numbers indicate the relative ratios of MNs in each sample. Wild-type MNs were assembled on DNA with a 60 bp overhang (60 bp OH), whereas H2BE76K/H3-Cy3 MNs were assembled on DNA with no overhang (0 bp OH), allowing discrimination of the two nucleosome species by gel

migration. Additionally, H2BE76K/H3-Cy3 MNs were assembled on DNA with a 60 bp overhang as a positive control for DNA exchange. **d**, Qubit quantification of MN DNA after affinity purification of nucleosomes that were subjected to the Nap1-mediated dimer exchange assay with biotinylated H2A/H2B dimers. DNA yield displays the fraction of DNA from affinity purified samples relative to input samples, indicating the extent of H2A/H2B dimer exchange. Individual WT and H2BE76K MNs were used as a proof-of-principle demonstration that destabilized MNs result in increased dimer exchange and, by extension, increased MN DNA after affinity purification. DNA yields from barcoded MN library experiments are similar to wild-type MNs, which is consistent with the resulting library data that shows that the majority of mutations have no or only moderate effects on dimer exchange (Fig. 2b). Data are presented as the mean ( $n = 2$  for individual MNs,  $n = 4$  for MN library). **e**, H2AE56A/K greatly enhance dimer exchange compared to other acidic patch mutations in the histone mutant library. Data are presented as the mean  $\pm$  SD ( $n = 4$  biologically independent samples). Values from mutant MNs exhibit statistically significant differences relative to wild-type ( $p < 0.05$ ) in a two-sample, unequal variances *t*-test, unless otherwise noted. ns, not significant. **f**, Denaturing gel analysis of purified octamers, dimers, and tetramers used to generate individual MNs for follow-up stability assays. **g**, Native gels of assembled mutant MNs used in follow-up dimer exchange and thermal stability assays. **h**, Raw data showing enhanced fluorescence for destabilizing mutants in the Nap1-mediated dimer exchange assay ( $n = 3$  biologically independent samples). **i**, Validation of select dimer exchange-enhancing histone mutant MNs via quantification of the in-gel fluorescence data shown in panel h. Data are presented as the mean  $\pm$  SD ( $n = 3$  biologically independent samples).



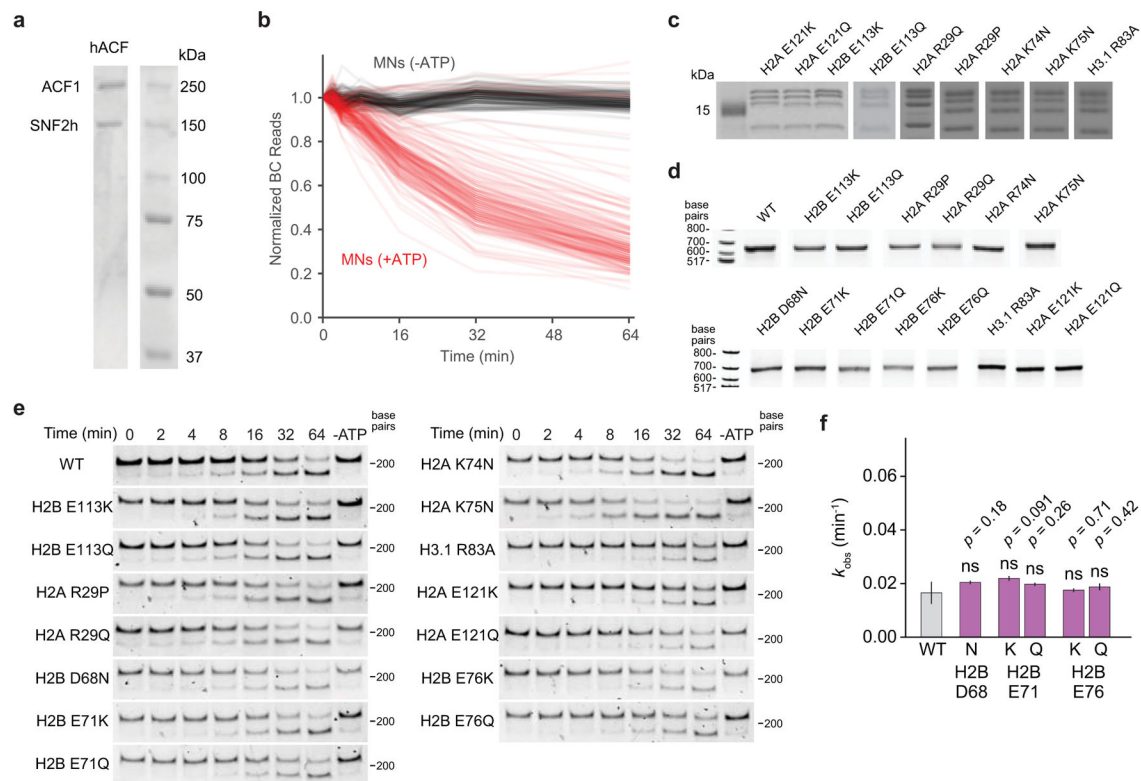
**Extended Data Figure 5 | Destabilizing histone mutations perturb nucleosome thermal stability and octamer assembly.**

**a**, Raw data for Fig. 2f showing the thermal stability of MNs with histone mutations. Data are shown in triplicate with a sigmoidal fit for the dimer melt (see methods). The  $T_{\text{half}}$  is shown for each fit as a dotted vertical line and reported as the mean  $\pm$  SD ( $n = 3$ ). **b**, Fast protein liquid chromatography traces of histone mutants that did not form octamers under standard assembly conditions.



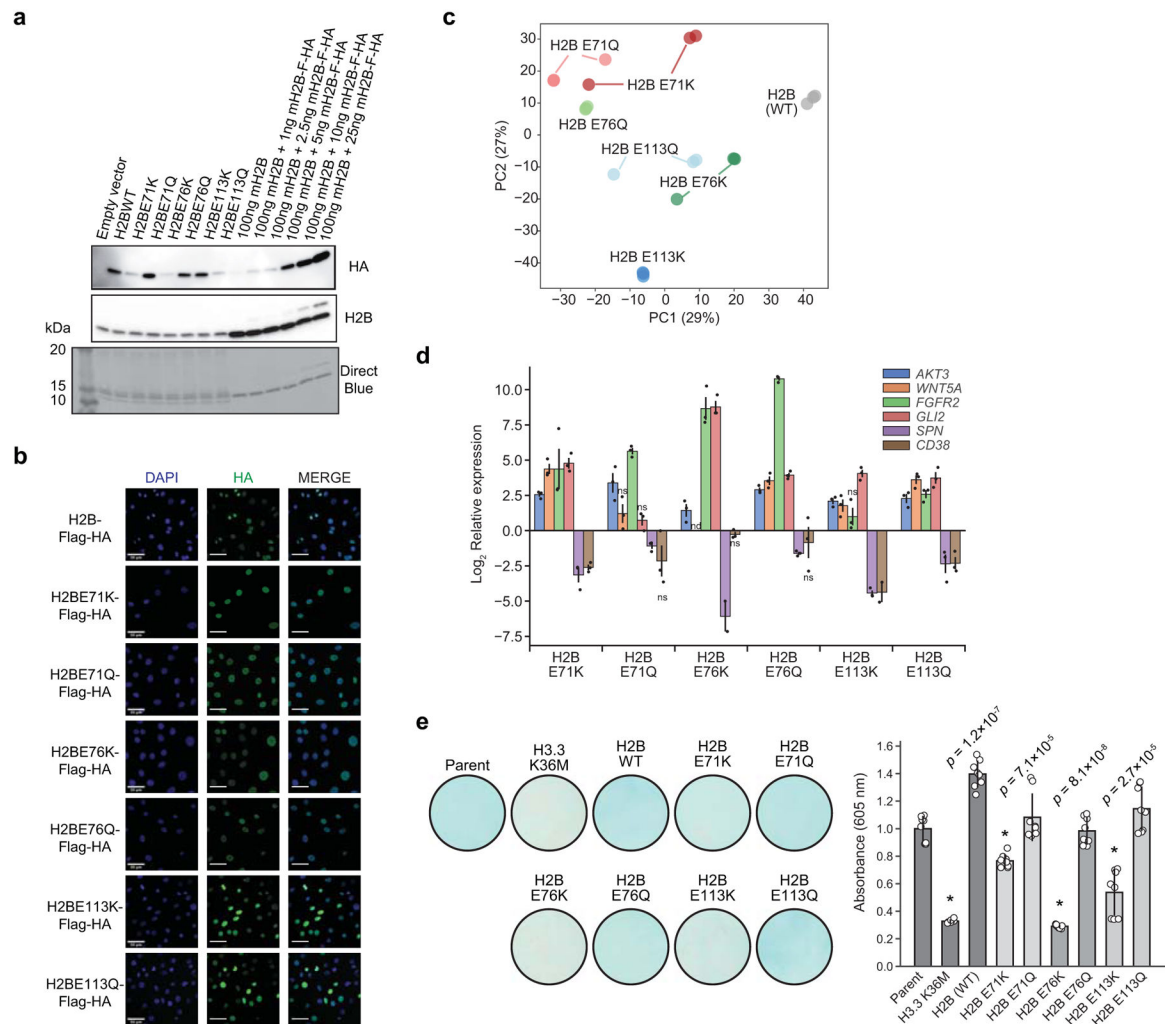
### Extended Data Figure 6 | Asymmetric mutant H2B nucleosomes are destabilized.

**a**, Denaturing gel analysis of purified wild-type, mutant, and H2A/H2B-CfaN-streptag dimers. **b**, Proof-of-concept assembly and purification of heterotypic nucleosomes with ubiquitinated H2B (H2Bub), which allows asymmetric nucleosomes to be distinguished by size. Numbers indicate molecular species shown in Fig. 3a. Hex., hexasomes. Het., heterotypic. **c**, Assembly and purification of heterotypic (het.) nucleosomes containing H2BE71K and H2BE76K. Assembly of wild-type (effectively homotypic) nucleosomes served as a control for the assembly process. 6/8 refers to hexasomes/octasomes. **d**, Raw data showing enhanced fluorescence for both homotypic and heterotypic destabilizing mutant MNs in the Nap1-mediated dimer exchange assay ( $n = 3$  biologically independent samples). **e**, Raw data showing the thermal stability of homotypic and heterotypic mutant MNs. Data are shown in triplicate with a sigmoidal fit for the dimer melt. The  $T_{\text{half}}$  is shown for each fit as a dotted vertical line and reported as the mean  $\pm$  SD ( $n = 3$ ).



**Extended Data Figure 7 | Mutations at the histone-DNA interface and the acidic patch alter nucleosome sliding.**

**a**, Denaturing gel showing the two subunits of the purified hACF complex. **b**, Time-course of ACF remodeling reactions from all MNs in the library both with ATP (red) and without ATP (black) demonstrate that the remodeling rates measured are ATP-dependent. Data are presented as the mean of individual MN values at each time point ( $n = 3$ ). **c**, Denaturing gel analysis of histone mutant octamers used to generate individual MNs for follow-up sliding assays. A subset of octamers, also utilized for follow-up dimer exchange and thermal stability assays, were already depicted in Extended Data Fig. 4f. **d**, Native gels of assembled MNs used in sliding assays. **e**, Representative gel images showing a time-course of DNA cleavage during ACF sliding in REA assays ( $n = 3$  biologically independent samples). **f**, Observed rate constants for dimer-tetramer interface mutant MNs do not exhibit statistically significant differences relative to wild-type ( $p > 0.05$ ) in a two-sample, unequal variances  $t$ -test. Remodeling rates were determined from a collective fit of the time-course data to a single one-phase exponential decay function, yielding  $k_{obs}$  values and their corresponding errors. Data are presented as the fit  $k_{obs}$  parameter  $\pm$  SD ( $n = 3$  biologically independent samples). ns, not significant.



### Extended Data Figure 8 | Globular domain mutations alter gene expression and impair cell differentiation.

**a**, Representative western blot analysis of acid extracted histones from C3H10T1/2 cells stably expressing HA-tagged histone mutants shows mutant incorporation into chromatin ( $n = 3$  separately generated cell lines). A standard curve of mixtures of mammalian H2B (mH2B) and tagged H2B (mH2B-F-HA) was used to estimate the relative expression of mutant histones, which ranged from 1–17% of the total H2B protein. **b**, Representative immunofluorescence images of tagged histone mutants expressed in C3H10T1/2 cells ( $n = 3$  separately generated cell lines). Scale bar, 50  $\mu\text{m}$ . **c**, Principal component analysis (PCA) of gene expression data from histone mutant cell lines. PCA distinguishes wild-type from mutant cell lines, as well as individual mutations from each other. **d**, Expression of select cancer-associated genes representing the dysregulated molecular pathways in histone mutant-expressing cell lines as assessed by RT-qPCR (relative to wild type). Data presented are the mean  $\pm$  SEM ( $n = 3$  biologically independent samples) and values from mutant cell lines exhibit significant differences relative to wild-type ( $p < 0.05$ ) in a two-sample  $t$ -test, unless otherwise noted. ns, not significant. nd, not detected. **e**, Mutant and wild-type H2B cell lines were induced to undergo chondrocyte differentiation for 7 days. Representative





## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

We thank Thomas Carroll from the Bioinformatics Resource Center at The Rockefeller University and Jacek Majewski and Eric Bareke from McGill University for help and guidance with the RNA-seq analyses. We thank Wei Wang and the Princeton Genomics Core Facility for designing the Illumina adaptor sequences and performing next-generation sequencing on DNA-barcoded MN library experimental samples. We thank Antony J. Burton for providing the dead CfaC peptide and Hengyuan Liu for providing Cy3-tagged histone H3.

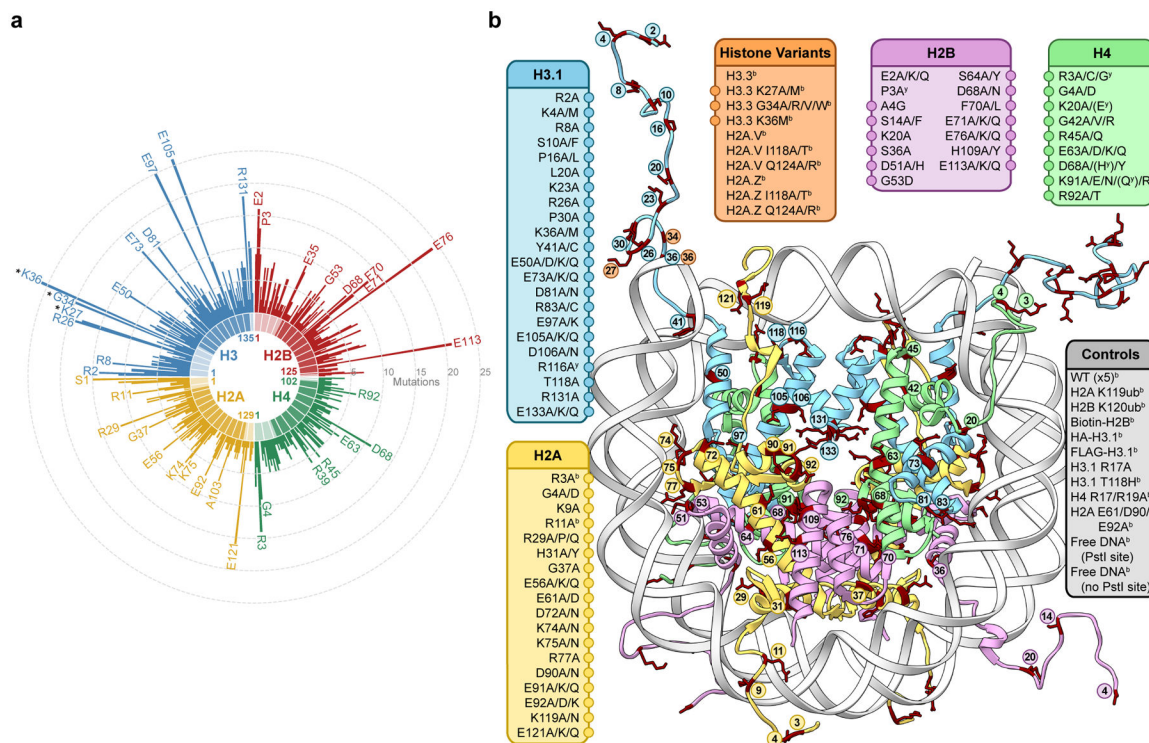
Funding support includes P01CA196539 (T.W.M, C.D.A.) and R37-GM086868 (T.W.M.), and we thank members the Allis and Muir laboratories especially those members of our collective P01 team. J.D.B. was funded by a postdoctoral fellowship from the U.S. National Institutes of Health (GM123659). M.M.M. was funded by a postdoctoral fellowship from the U.S. National Institutes of Health (GM131632). A.L.P. was funded by an Anderson Graduate Fellowship in Cancer Research from the Anderson Cancer Center. F.W. was supported by a postdoctoral fellowship from the German Research Foundation (WO 2039/1-1). B.A.N. was funded by the National Cancer Institute (1K08CA245212) and we acknowledge the Memorial Sloan Kettering Cancer Center Support Grant/Core Grant (P30 CA008748). L.F. was supported by the C. H. Li Memorial Scholar Fund Award.

## References

- Allis CD & Jenuwein T The molecular hallmarks of epigenetic control. *Nat. Rev. Genet* 17, 487–500 (2016). [PubMed: 27346641]
- Dawson MA & Kouzarides T Cancer epigenetics: From mechanism to therapy. *Cell* 150, 12–27 (2012). [PubMed: 22770212]
- Mohammad F & Helin K Oncohistones: drivers of pediatric cancers. *Genes Dev.* 31, 2313–2324 (2018).
- Weinberg DN, Allis CD & Lu C Oncogenic mechanisms of histone H3 mutations. *Cold Spring Harb. Perspect. Med* 7, a026443 (2017). [PubMed: 27864305]
- Behjati S et al. Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. *Nat. Genet* 45, 1479–1482 (2013). [PubMed: 24162739]
- Lu C et al. Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science* 352, 844–849 (2016). [PubMed: 27174990]
- Presneau N et al. Diagnostic value of H3F3A mutations in giant cell tumour of bone compared to osteoclast-rich mimics. *J. Pathol. Clin. Res* 1, 113–123 (2015). [PubMed: 27499898]
- Schwartzentruber J et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* 482, 226–231 (2012). [PubMed: 22286061]
- Sturm D et al. Hotspot Mutations in H3F3A and IDH1 Define Distinct Epigenetic and Biological Subgroups of Glioblastoma. *Cancer Cell* 22, 425–437 (2012). [PubMed: 23079654]
- Wu G et al. Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat. Genet* 44, 251–253 (2012). [PubMed: 22286216]
- Lewis PW et al. Inhibition of PRC2 Activity by a Gain-of-Function H3 Mutation Found in Pediatric Glioblastoma. *Science* 340, 857–861 (2013). [PubMed: 23539183]
- Brown ZZ et al. Strategy for ‘detoxification’ of a cancer-derived histone mutant based on mapping its interaction with the methyltransferase PRC2. *J. Am. Chem. Soc* 136, 13498–13501 (2014). [PubMed: 25180930]
- Chan KM et al. The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. *Genes Dev.* 27, 985–990 (2013). [PubMed: 23603901]
- Fang D et al. The histone H3.3K36M mutation reprograms the epigenome of chondroblastomas. *Science* 352, 1344–1348 (2016). [PubMed: 27229140]
- Lohr JG et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc. Natl. Acad. Sci. U.S.A* 109, 3879–3884 (2012). [PubMed: 22343534]

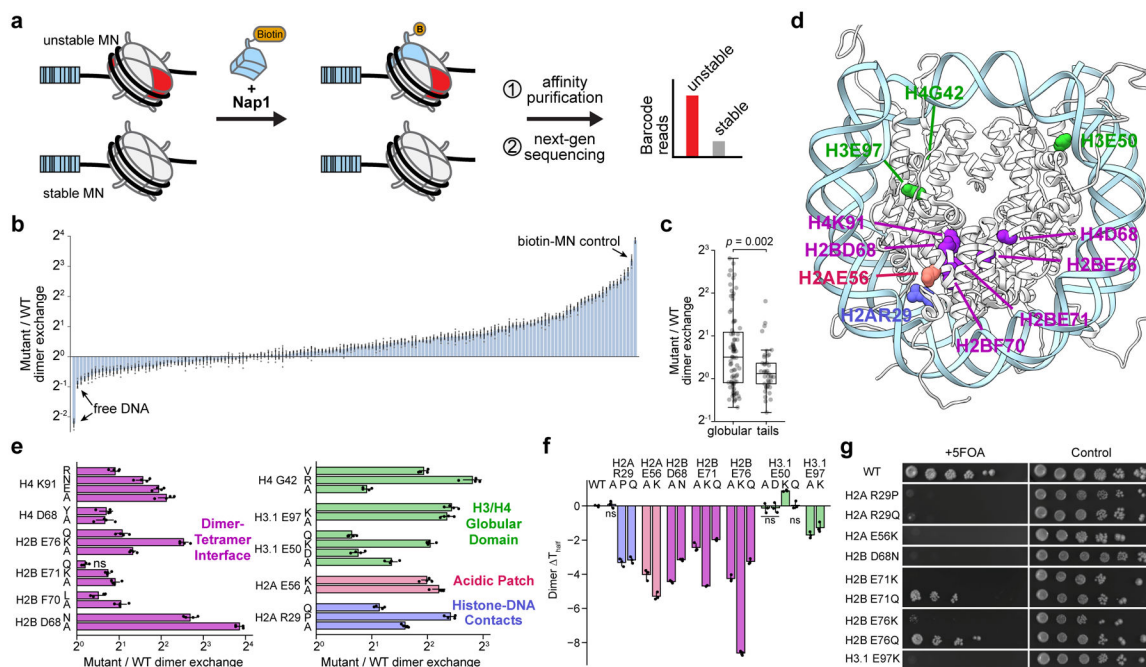
16. Papillon-Cavanagh S et al. Impaired H3K36 methylation defines a subset of head and neck squamous cell carcinomas. *Nat. Genet* 49, 180–185 (2017). [PubMed: 28067913]
17. Zhao S et al. Mutational landscape of uterine and ovarian carcinosarcomas implicates histone genes in epithelial–mesenchymal transition. *Proc. Natl. Acad. Sci* 113, 12238–12243 (2016). [PubMed: 27791010]
18. Boileau M et al. Mutant H3 histones drive human pre-leukemic hematopoietic stem cell expansion and promote leukemic aggressiveness. *Nat. Commun* 10, 2891 (2019). [PubMed: 31253791]
19. Nacev BA et al. The expanding landscape of ‘oncohistone’ mutations in human cancers. *Nature* 567, 473–478 (2019). [PubMed: 30894748]
20. Bennett RL et al. A Mutation in Histone H2B Represents a New Class of Oncogenic Driver. *Cancer Discov.* 9, 1438–51 (2019). [PubMed: 31337617]
21. Arimura Y et al. Cancer-associated mutations of histones H2B, H3.1 and H2A.Z.1 affect the structure and stability of the nucleosome. *Nucleic Acids Res.* 46, 10007–18 (2018). [PubMed: 30053102]
22. Wan YCE et al. Cancer-associated histone mutation H2BG53D disrupts DNA – histone octamer interaction and promotes oncogenic phenotypes. *Signal Transduct. Target. Ther* 5, 27 (2020). [PubMed: 32296031]
23. Nguyen UTT et al. Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries. *Nat. Methods* 11, 834–840 (2014). [PubMed: 24997861]
24. Dann GP et al. ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference. *Nature* 548, 607–611 (2017). [PubMed: 28767641]
25. Nakanishi S et al. A comprehensive library of histone mutants identifies nucleosomal residues required for H3K4 methylation. *Nat. Struct. Mol. Biol* 15, 881–888 (2008). [PubMed: 18622391]
26. Dai J et al. Probing Nucleosome Function: A Highly Versatile Library of Synthetic Histone H3 and H4 Mutants. *Cell* 134, 1066–1078 (2008). [PubMed: 18805098]
27. Truong DM & Boeke JD Resetting the Yeast Epigenome with Human Nucleosomes. *Cell* 171, 1508–1519 (2017). [PubMed: 29198523]
28. Park YJ, Chodaparambil JV, Bao Y, McBryant SJ & Luger K Nucleosome assembly protein 1 exchanges histone H2A-H2B dimers and assists nucleosome sliding. *J. Biol. Chem* 280, 1817–1825 (2005). [PubMed: 15516689]
29. Levodosky RF, Sabantsev A, Deindl S & Bowman GD The Chd1 chromatin remodeler shifts hexasomes unidirectionally. *eLife* 5, e21356 (2016). [PubMed: 28032848]
30. Stevens AJ et al. Design of a Split Intein with Exceptional Protein Splicing Activity. *J. Am. Chem. Soc* 138, 2162–2165 (2016). [PubMed: 26854538]
31. Vila-Perelló M et al. Streamlined expressed protein ligation using split inteins. *J. Am. Chem. Soc* 135, 286–292 (2013). [PubMed: 23265282]
32. Clapier CR, Iwasa J, Cairns BR & Peterson CL Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol* 18, 407–422 (2017). [PubMed: 28512350]
33. Dao HT, Dul BE, Dann GP, Liszczak GP & Muir TW A basic motif anchoring ISWI to nucleosome acidic patch regulates nucleosome spacing. *Nat. Chem. Biol* 16, 134–142 (2020). [PubMed: 31819269]
34. Clapier CR & Cairns BR Regulation of ISWI involves inhibitory modules antagonized by nucleosomal epitopes. *Nature* 492, 280–284 (2012). [PubMed: 23143334]
35. Shogren-Knaak M et al. Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions. *Science* 311, 844–847 (2006). [PubMed: 16469925]
36. Levodosky RF & Bowman GD Asymmetry between the two acidic patches dictates the direction of nucleosome sliding by the ISWI chromatin remodeler. *eLife* 8, e45472 (2019). [PubMed: 31094676]
37. Zhan T, Rindtorff N & Boutros M Wnt signaling in cancer. *Oncogene* 36, 1461–1473 (2017). [PubMed: 27617575]
38. Massagué J TGF $\beta$  in Cancer. *Cell* 134, 215–230 (2008). [PubMed: 18662538]

39. Wong KK, Engelman JA & Cantley LC Targeting the PI3K signaling pathway in cancer. *Curr. Opin. Genet. Dev* 20, 87–90 (2010). [PubMed: 20006486]
40. Reya T, Morrison SJ, Clarke MF & Weissman IL Stem cells, cancer, and cancer stem cells. *Nature* 414, 105–111 (2001). [PubMed: 11689955]
41. Zhuang L et al. Depletion of Nsd2-mediated histone H3K36 methylation impairs adipose tissue development and function. *Nat. Commun* 9, 1796 (2018). [PubMed: 29728617]
42. Farmer SR Transcriptional control of adipocyte formation. *Cell Metab.* 4, 263–273 (2006). [PubMed: 17011499]
43. Albig C et al. JASPer controls interphase histone H3S10 phosphorylation by chromosomal kinase JIL-1 in *Drosophila*. *Nat. Commun* 10, 5343 (2019). [PubMed: 31767855]
44. Liszczak G, Diehl KL, Dann GP & Muir TW Acetylation blocks DNA damage-induced chromatin ADP-ribosylation. *Nat. Chem. Biol* 14, 837–840 (2018). [PubMed: 30013063]
45. Wojcik F et al. Functional crosstalk between histone H2B ubiquitylation and H2A modifications and variants. *Nat. Commun* 9, 1394 (2018). [PubMed: 29643390]
46. Horikoshi N, Arimura Y, Taguchi H & Kurumizaka H Crystal structures of heterotypic nucleosomes containing histones H2A.Z and H2A. *Open Biol.* 6, 160127 (2016). [PubMed: 27358293]
47. Lechner CC, Agashe ND & Fierz B Traceless Synthesis of Asymmetrically Modified Bivalent Nucleosomes. *Angew. Chemie Int. Ed* 55, 2903–2906 (2016).
48. Voigt P et al. Asymmetrically modified nucleosomes. *Cell* 151, 181–193 (2012). [PubMed: 23021224]
49. Bernstein BE et al. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* 125, 315–326 (2006). [PubMed: 16630819]
50. Rhee HS, Bataille AR, Zhang L & Pugh BF Subnucleosomal structures and nucleosome asymmetry across a genome. *Cell* 159, 1377–1388 (2014). [PubMed: 25480300]
51. Pettersen EF et al. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem* 25, 1605–1612 (2004). [PubMed: 15264254]
52. Dyer PN et al. Reconstitution of Nucleosome Core Particles from Recombinant Histones and DNA. *Methods Enzymol.* 375, 23–44 (2003).
53. Whitcomb SJ et al. Histone monoubiquitylation position determines specificity and direction of enzymatic cross-talk with histone methyltransferases Dot1L and PRC2. *J. Biol. Chem* 287, 23718–23725 (2012). [PubMed: 22619169]
54. Soni R, Carmichael JP & Murray JAH Parameters affecting lithium acetate-mediated transformation of *Saccharomyces cerevisiae* and development of a rapid and simplified procedure. *Curr. Genet* 24, 455–459 (1993). [PubMed: 8299163]
55. Fierz B, Kilic S, Hieb AR, Luger K & Muir TW Stability of nucleosomes containing homogenously ubiquitylated H2A and H2B prepared using semisynthesis. *J. Am. Chem. Soc* 134, 19548–19551 (2012). [PubMed: 23163596]
56. Taguchi H, Horikoshi N, Arimura Y & Kurumizaka H A method for evaluating nucleosome stability with a protein-binding fluorescent dye. *Methods* 70, 119–126 (2014). [PubMed: 25220913]
57. Patro R, Duggal G, Love MI, Irizarry RA & Kingsford C Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419 (2017). [PubMed: 28263959]
58. Anders S, Reyes A & Huber W Detecting differential usage of exons from RNA-seq data. *Genome Res.* 22, 2008–2017 (2012). [PubMed: 22722343]
59. Love MI, Huber W & Anders S Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550 (2014). [PubMed: 25516281]
60. Benjamini Y, Hochberg Y & Publishing B Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B-Statistical Methodol* 57, 289–300 (1995).



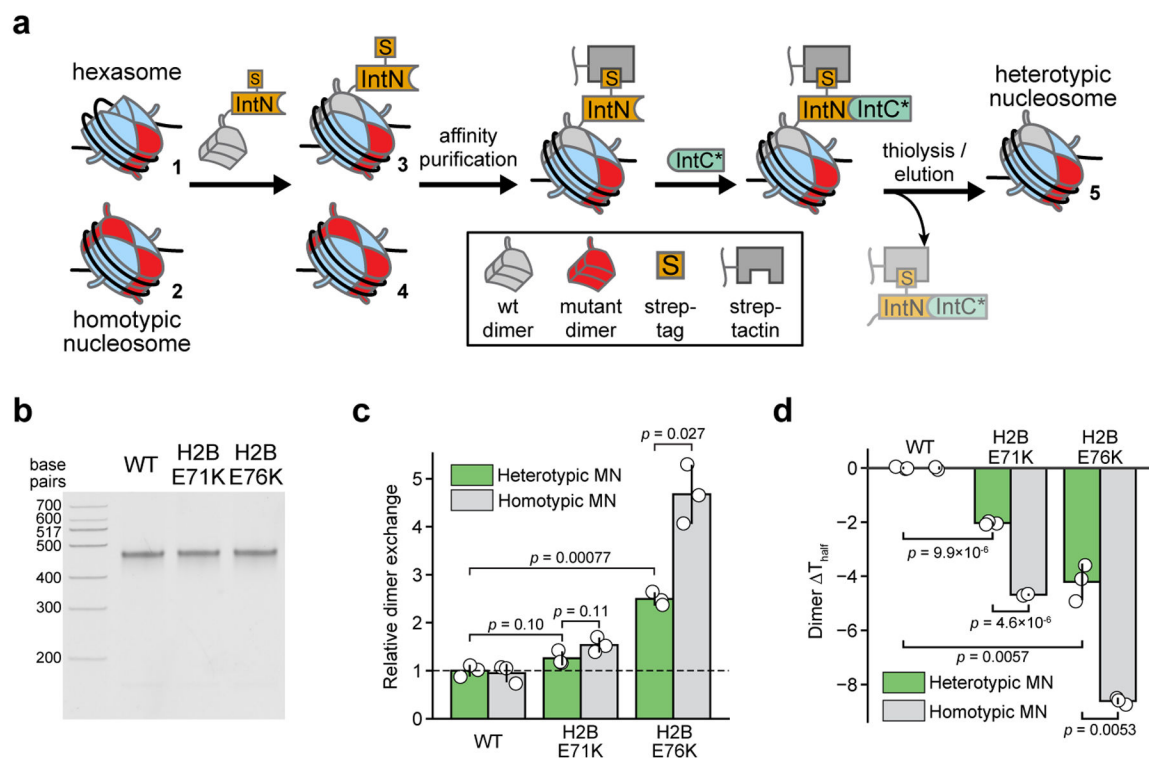
**Figure 1 |. A library of cancer-associated histone mutations.**

**a**, Circular barplot showing mutation counts in the four core histones in human cancers (data from Nacev et al., 2019)<sup>19</sup>. Select highly mutated sites are labeled, and established oncohistone mutations are denoted with an asterisk. Darker and lighter regions in each histone represent the globular and tail domains, respectively. **b**, Diagram depicting all histone mutants, variants, and controls included in the barcoded mononucleosome and humanized yeast libraries. Mutations with adjacent circles are displayed on the nucleosome structure with coordinated colors and labeled by residue. In addition to any specific missense mutations that were overrepresented in the dataset, alanine mutations were included at every site, offering both experimental redundancy as well as the ability to distinguish between loss and gain of function at a given amino acid residue. Indicated mutants are found only in the barcoded mononucleosome (<sup>b</sup>) or humanized yeast libraries (<sup>y</sup>), respectively (PDB 1kx5).



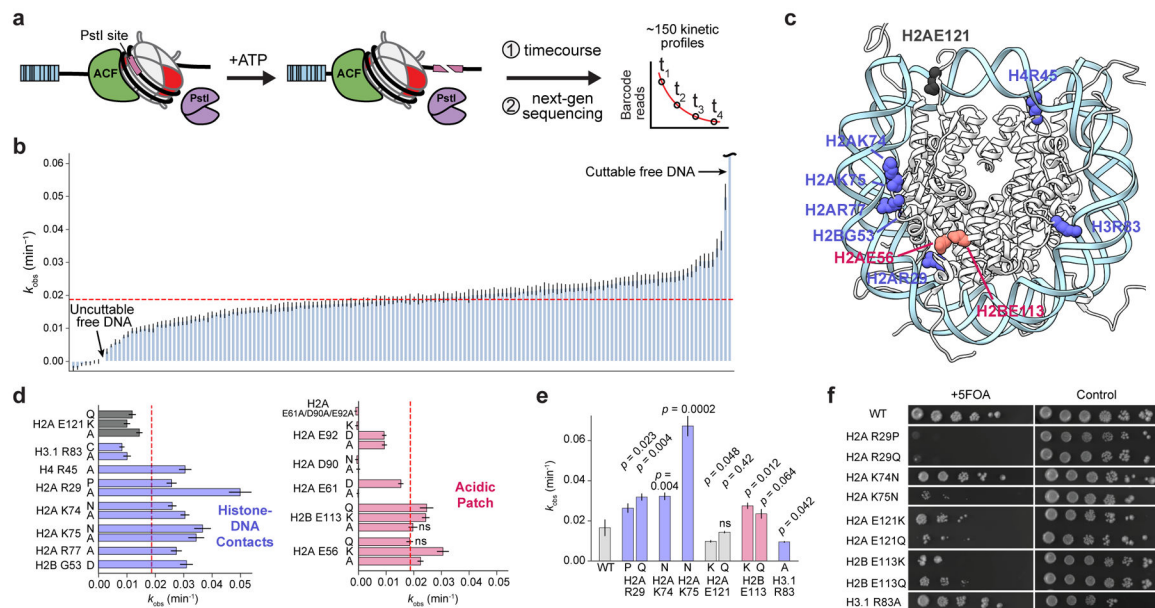
**Figure 2 |. An H2A/H2B dimer exchange assay identifies histone mutations that destabilize nucleosomes.**

**a**, Nap1-mediated dimer exchange assay schematic. **b**, Waterfall plot of results from the H2A/H2B dimer exchange assay on the barcoded MN library. Each bar represents a unique MN library member. Controls are denoted with arrows. Mutant and wild-type nucleosome barcode counts were normalized to their pre-enrichment counts and dimer exchange reported as the ratio of these values. Data are presented as the mean  $\pm$  SD ( $n = 4$  biologically independent samples). **c**, Box-and-whisker plot of average MN dimer exchange values from panel b, grouped by mutation location in either the histone globular ( $n = 84$  distinct MNs) or tail ( $n = 41$  distinct MNs) domains. Box, upper and lower quartiles; center line, median; whiskers, 1.5x interquartile range.  $p$ -value calculated via a two-sample Kolmogorov-Smirnov test. **d**, Locations of dimer exchange-enhancing histone mutant residues shown on the nucleosome structure (PDB 1kx5). **e**, Dimer exchange data for select nucleosomes from panel b, grouped by nucleosome location and indicated by color (see panel d): dimer-tetramer interface (purple), H3/H4 globular domain (green), acidic patch (red), and histone-DNA contacts (blue). Data are presented as the mean  $\pm$  SD ( $n = 4$  biologically independent samples). **f**, Thermal stability measurements of mutant relative to wild-type nucleosomes. Numbers shown are changes in  $T_{half}$  values, the temperature at which half of the H2A/H2B dimers are melted off the nucleosome (see methods). Data are presented as the mean  $\pm$  SD ( $n = 3$  biologically independent samples). Values from mutant MNs in panels e and f exhibit statistically significant differences relative to wild-type ( $p < 0.05$ ) in a two-sample, unequal variances  $t$ -test, unless otherwise noted. ns, not significant. **g**, Serial dilutions of humanized yeast with mutant (non-alanine) histones that decreased nucleosome thermal stability.



**Figure 3 | Asymmetric mutant H2B nucleosomes are destabilized.**

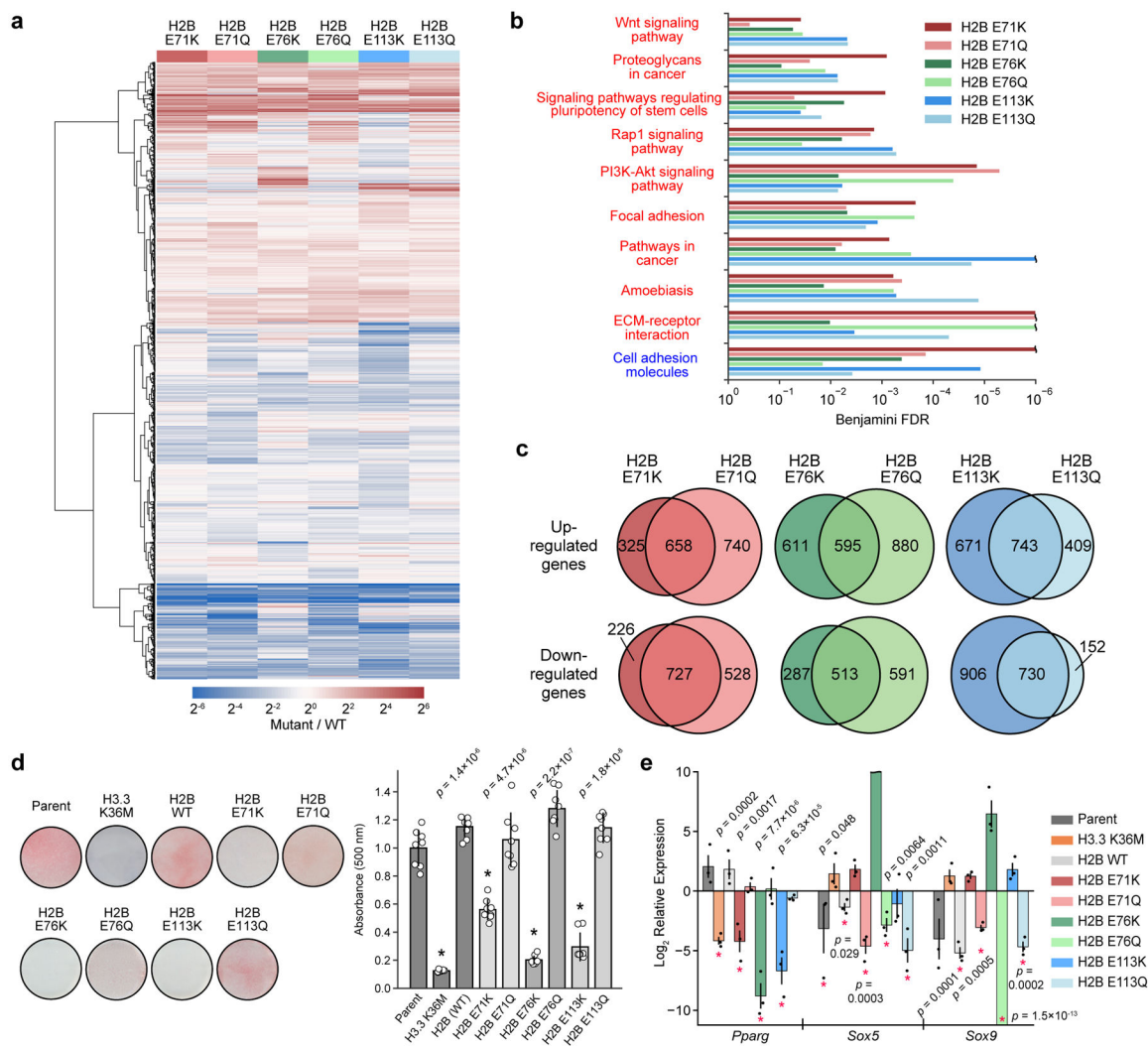
**a**, Schematic of the streamlined assembly and purification of heterotypic nucleosomes using split inteins (see methods). Mixtures of hexasomes and nucleosomes are treated with H2A/H2B dimers tagged with a split N-intein fragment (IntN), producing asymmetric MNs which are subsequently isolated via affinity purification. Captured MNs are then eluted by on-resin thiolysis using a catalytically dead C-intein fragment (IntC\*), yielding tracelessly purified heterotypic MNs. **b**, Native gel analysis of heterotypic nucleosomes containing cancer-associated histone mutations. The wild-type nucleosome is homotypic but assembled via the method in panel a to serve as a control for the process. **c**, Nap1-mediated dimer exchange of heterotypic mutant nucleosomes. Data are presented as the mean  $\pm$  SD ( $n = 3$  biologically independent samples). **d**, Thermal stability measurements for heterotypic mutant nucleosomes. Data are presented as the mean  $\pm$  SD ( $n = 3$  biologically independent samples). Values from heterotypic mutant MNs in panels c and d exhibit statistically significant differences relative to wild-type MNs or homotypic mutant MNs ( $p < 0.05$ ) in a two-sample, unequal variances  $t$ -test, unless otherwise noted.



**Figure 4 | A chromatin remodeling assay uncovers histone mutations that affect nucleosome sliding.**

**a**, Schematic of the restriction enzyme accessibility assay used to assess rates of nucleosome sliding by ACF. PstI cutting of DNA on remodeled nucleosomes manifests as fewer barcode reads since truncated DNA is not sequenced. Plots of barcode reads vs. time were modeled as single-phase exponential decays from which observed rate constants were extracted. **b**, Waterfall plot of observed nucleosome remodeling rates ( $k_{obs}$ ) for the barcoded MN library. Each bar represents a unique MN library member. Remodeling rates were determined from a collective fit of the time-course data to a single one-phase exponential decay function, yielding  $k_{obs}$  values and their corresponding errors. The dotted red line indicates the average remodeling rate for wild-type nucleosomes. Data are presented as the fit  $k_{obs}$  parameter  $\pm$  SD ( $n = 3$  biologically independent samples). **c**, Locations of mutant histone residues that alter nucleosome sliding depicted on the nucleosome structure (PDB 1kx5). **d**, Remodeling rates of select mutants from panel b, grouped by color based on nucleosome location (see panel c): histone-DNA contacts (blue) and acidic patch (red). Data are presented as the mean  $\pm$  SD ( $n = 3$  biologically independent samples). **e**, Observed rate constants for validated mutant nucleosomes. Data are presented as the fit  $k_{obs}$  parameter  $\pm$  SD ( $n = 3$  biologically independent samples). Values from mutant MNs in panels d and e exhibit statistically significant differences relative to wild-type ( $p < 0.05$ ) in a two-sample, unequal variances  $t$ -test, unless otherwise noted. ns, not significant. **f**, Serial dilutions of humanized yeast with mutant histones that altered nucleosome sliding rates.





**Figure 5 | Globular domain mutations alter gene expression and impair cell differentiation.**  
**a**, Heatmap depicting average differential gene expression in C3H10T1/2 cells stably expressing either wild-type or mutant H2B ( $n = 3$  separately generated cell lines each at distinct passage numbers). **b**, Gene ontology (GO) analysis showing significantly upregulated (red text) and downregulated (blue text) KEGG pathways for cells expressing mutant histones ( $n = 3$  separately generated cell lines). Pathways significantly affected (Benjamini FDR < 0.05) in at least 4 of the 6 mutants are shown. Expanded GO analyses for individual mutants are shown in Extended Data Fig. 9. **c**, Venn diagrams comparing significantly upregulated and downregulated genes (FDR < 0.05) for cells expressing Lys and Gln mutations at H2BE71, E76, and E113. **d**, Mutant and wild-type H2B cell lines were induced to undergo adipocyte differentiation for 7 days. Representative images are shown after Oil Red O staining. The barplot shows the mean  $\pm$  SD of quantified Oil Red O staining of cell extracts ( $n = 8$  biologically independent samples). An asterisk marks mutant cell lines that caused a significant reduction in staining compared to both the parent line and to cells expressing wild-type H2B ( $p < 0.05$ , two-sided unequal variances  $t$ -test). **e**, Expression of master transcriptional regulators at day 4 of adipocyte differentiation as assessed by RT-

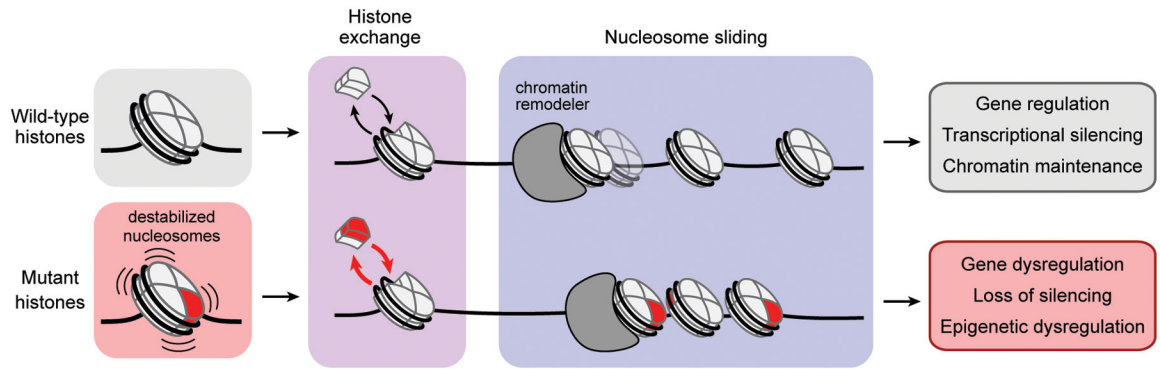
qPCR. Mutant H2B and H3 values are displayed relative to the wild-type H2B and parent expression levels, respectively, at day 0. Asterisks mark values with significant difference from reference genes at day 0 ( $p < 0.05$ , two-sided unequal variances  $t$ -test). Data are presented as the mean  $\pm$  SEM ( $n = 3$  biologically independent samples).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 6 |** Proposed model of how histone mutations may contribute to cancer development or progression by perturbing fundamental chromatin remodeling processes.