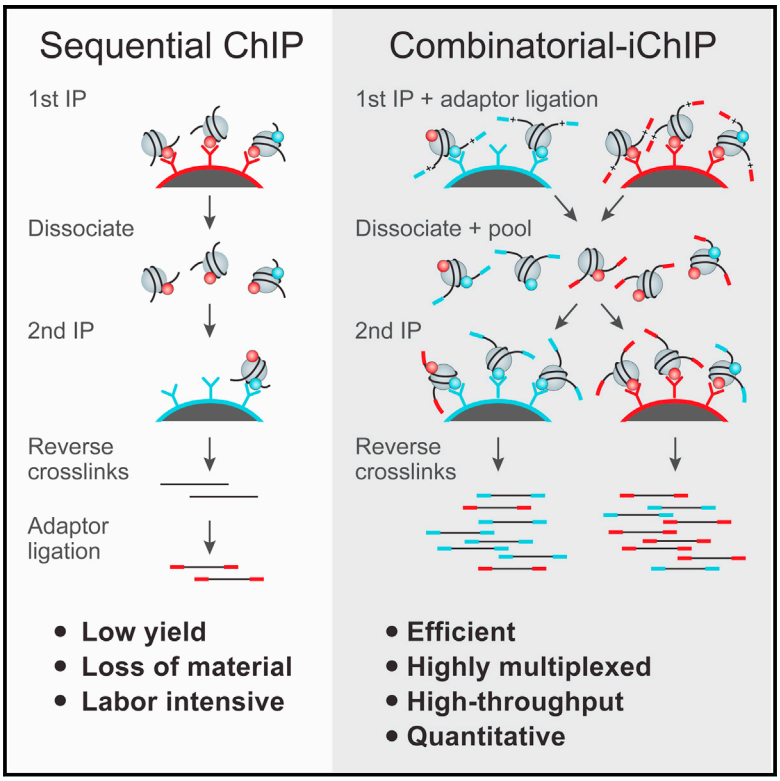
Technology

Molecular Cell

Elucidating Combinatorial Chromatin States at Single-Nucleosome Resolution

Graphical Abstract



Highlights

- High-throughput method for mapping the genome-wide co-occurrence of histone marks
- H3K36me3 and H3K79me3 co-occur more frequently than would be expected by chance
- Initiating and elongating Pol II-deposited marks co-occur on the same nucleosomes
- Disruption of RPD3 increases the co-occurrence of H3K36me3 and H3 acetylation

Authors

Ronen Sadeh, Roee Launer-Wachs, Hava Wandel, Ayelet Rahat, Nir Friedman

Correspondence

ronensadeh@gmail.com (R.S.), nir@cs.huji.ac.il (N.F.)

In Brief

Standard chromatin immunoprecipitation followed by sequencing reports on the average occupancy of individual histone marks in a population of cells. As such, it cannot ascertain whether two marks co-exist on the same nucleosome. Sadeh et al. describe methodology to determine the genome-wide co-occurrence of histone marks at single-nucleosome resolution.

Accession Numbers

GSE84240

Sadeh et al., 2016, Molecular Cell *63*, 1080–1088
September 15, 2016 © 2016 The Author(s). Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.molcel.2016.07.023

CellPress

Molecular Cell
Technology

Elucidating Combinatorial Chromatin States at Single-Nucleosome Resolution

Ronen Sadeh,[1,3,*] Roee Launer-Wachs,[1] Hava Wandel,[1] Ayelet Rahat,[1] and Nir Friedman[1,2,*]
[1]School of Computer Science and Engineering and Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel
[2]Broad Institute of MIT and Harvard, Cambridge, MA 02139, USA
[3]Lead Contact
*Correspondence: ronensadeh@gmail.com (R.S.), nir@cs.huji.ac.il (N.F.)
http://dx.doi.org/10.1016/j.molcel.2016.07.023

SUMMARY

Chromatin immunoprecipitation followed by sequencing (ChIP-seq) has been instrumental to our current view of chromatin structure and function. It allows genome-wide mapping of histone marks, which demarcate biologically relevant domains. However, ChIP-seq is an ensemble measurement reporting the average occupancy of individual marks in a cell population. Consequently, our understanding of the combinatorial nature of chromatin states relies almost exclusively on correlation between the genomic distributions of individual marks. Here, we report the development of combinatorial-iChIP to determine the genome-wide co-occurrence of histone marks at single-nucleosome resolution. By comparing to a null model, we show that certain combinations of overlapping marks (H3K36me3 and H3K79me3) co-occur more frequently than would be expected by chance, while others (H3K4me3 and H3K36me3) do not, reflecting differences in the underlying chromatin pathways. We further use combinatorial-iChIP to illuminate aspects of the Set2-RPD3S pathway. This approach promises to improve our understanding of the combinatorial complexity of chromatin.

INTRODUCTION

Nucleosomal histones, the fundamental packaging units of DNA, are extensively decorated by a large number of posttranslational modifications (PTMs) or "marks." These marks are highly conserved and play key roles in all genomic transactions (Rivera and Ren 2013). Enzymes that deposit, remove, or bind histone marks, as well as modified residues in histones, are frequently mutated in diseases such as cancer (Baylin and Jones 2011; Maze et al., 2014; Chi et al., 2010; Lewis et al., 2013; Lu et al., 2016). Chromatin immunoprecipitation followed by next-generation sequencing (ChIP-seq) has been widely used to determine the genome-wide location of nucleosomes bearing specific histone marks and has been instrumental to our understanding of chromatin architecture, structure, and function in many cell types (ENCODE Project Consortium 2012; Weiner et al., 2015). ChIP-seq studies in a variety of organisms identify combinations of spatially correlated histone marks; these combinatorial patterns demarcate biologically relevant domains such as actively transcribed or polycomb repressed genes, heterochromatin, paused and active regulatory elements, and enhancers. These patterns can be used to annotate the genome and predict unknown genomic functionalities (Guttman et al., 2009; Heintzman et al., 2007; Ernst and Kellis, 2012). Moreover, several chromatin-binding proteins that contain multiple recognition domains were shown to bind to specific combinations of histone marks (Ruthenburg et al., 2011), suggesting that mark co-occurrence can have functional implications.

A typical ChIP-seq experiment reports on the average position and occupancy of a single modification averaged over a large population of potentially heterogeneous cells. As a result, our current understanding of the combinatorial nature of chromatin states relies almost exclusively on genomic correlations between chromatin features that may or may not occur in the same cell. Biochemical studies have identified dozens of different histone marks, as well as multiple proteins that deposit, erase, and bind them. However, efforts to probe the complexity of chromatin in various cell types have identified a limited number of combinations of histone marks that specify defined genomic regions (ENCODE Project Consortium 2012; Weiner et al., 2015). Mass spectrometry has proven to be a powerful tool for identification of histone mark complexity; however, it lacks genomic localization and is limited to marks co-residing on a single, relatively short, peptide (Garcia et al., 2007; Young et al., 2009). Recently, single-molecule imaging allowed visualization of combinations of histone modifications; however, it has limited genomic context (Shema et al., 2016). Importantly, it is generally unknown if genomically correlated histone marks coexist or, alternatively, if they represent different chromatin states occurring in different subsets of a population of cells (Figure 1A). Sequential ChIP, where histone marks are sequentially immunoprecipitated, can report on the actual combinatorial nature of histone marks (Bernstein et al., 2006), yet such experiments are surprisingly rare and fraught with technical challenges. One factor that may impair the robustness and reliability of such experiments is the large amount of biological material required to provide sufficient input to the second ChIP, which can result in low signal-to-background ratio. Here, we report the development of a combinatorial indexed ChIP method (combinatorial-iChIP) to map the genome-wide co-occurrence of histone marks at single-nucleosome resolution (Figure 1B)

 © 2016 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

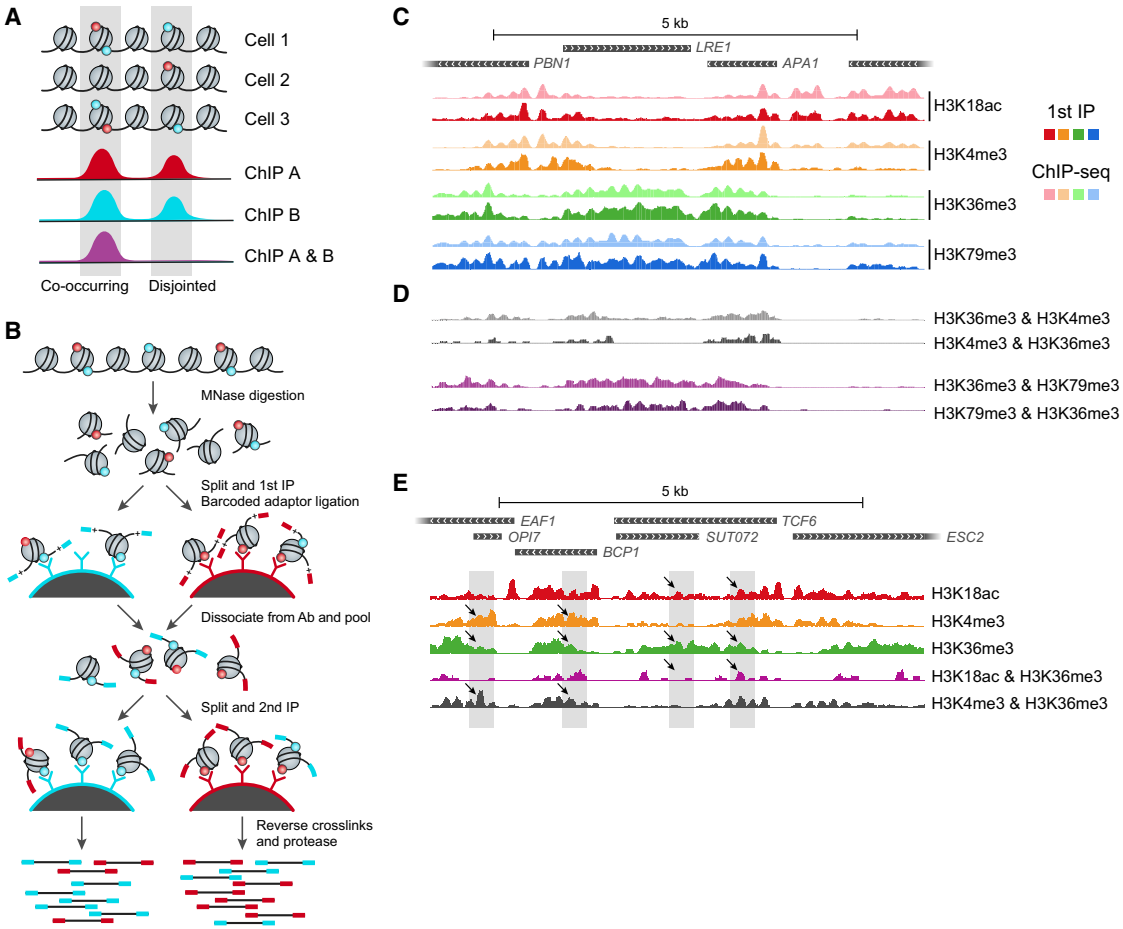


Figure 1. Combinatorial-iChIP Protocol for Assaying Combinations of Histone Modifications

(A) Overlapping signal of standard ChIP (red and cyan) can be due to co-occurrence of the two marks on the same nucleosomes (left peak, co-occurring), but can also be due to disjoint occurrence in the same location in different cells (right peak, disjointed). Combinatorial signal (purple) allows us to distinguish the two scenarios.

(B) Outline of the combinatorial-iChIP protocol: MNase-digested chromatin is immobilized on magnetic beads coated with antibodies of interest (first IP). Immobilized nucleosomes are ligated to barcoded adaptors to specify the 1st antibody. Following antibody inactivation and nucleosome release, samples are pooled, re-divided, and subjected to a second ChIP. Nucleosomes are reverse crosslinked and NGS sequences are added by PCR to barcoded DNA to generate NGS-compatible libraries. At this stage a second barcode denoting the second ChIP pool is added to the fragments.

(C) The signal from the first ChIP step of the MNase combinatorial-iChIP protocol (first ChIP, solid colors) is in close agreement to standard MNase-ChIP (faded colors) (Weiner et al., 2015). Shown are coverage tracks for a representative genomic region.

(D) Reciprocal combinatorial-iChIP signals are in good agreement. Same locus as (C).

(E) Combinatorial-iChIP uncovers co-occurrences and disjoint occurrences that are not available from individual ChIP. Shown is a representative genomic region. The gray boxes highlight locations that are similar in terms of individual ChIP but different in combinatorial-iChIP for these individual marks (black arrows). See also Figures S1 and S2.

and provide experimental and analytical tools for efficient, reliable, and reproducible detection of combinations of histone marks.

DESIGN

To investigate combinations of histone marks on individual nucleosomes, we need to address several concerns:

(1) To obtain single-nucleosome resolution, we need to be able to determine the chromatin fragments that undergo IP.

(2) Once we consider successive IP stages, we need to deal with the issue of low amounts of material.

(3) We need to deal with multiple samples, each probed by a large number of IP pair combinations.

(4) We must ensure that the method is quantitative and reproducible.

(5) We would like the experimental design to be as compatible as possible with high-throughput practices.

Following our experience with MNase-ChIP (Liu et al., 2005; Weiner et al., 2015), we chose to deal with the first challenge by using MNase digestion to break chromatin into nucleosomal

particles. The DNA fragment size allows us to distinguish mononucleosomal particles from multi-nucleosomal ones. In addition, unlike sonication, MNase digestion can be readily adapted to high-throughput practices.

Performing library construction on small amounts material following IP has been the primary challenge in applying ChIP-seq to small populations of cells. Several recent works in the literature avoid this problem by performing early adaptor ligation on chromatin before the IP step (van Galen et al., 2016; Lara-Astiaso et al., 2014; Rhee and Pugh 2011; Singh et al., 2014). This implies that, following IP, the material can be amplified and sequenced without additional processing. The use of early adaptor ligation enables barcoding of samples. Once samples are barcoded, they can be pooled together and demultiplexed after sequencing (Lara-Astiaso et al., 2014). Specifically, the iChIP method (Lara-Astiaso et al., 2014) combines these ideas to multiplex ChIP from multiple samples with limited numbers of cells.

Sample pooling can multiplex many samples, which allows us to use small amounts of input material per sample and to process a combinatorial number of samples. Second, by performing IP for multiple samples in a single tube, we reduce technical variability between samples.

RESULTS

Combinatorial-iChIP Can Detect Co-occurrence of Histone Marks

To test the feasibility of chromatin barcoding for detecting co-occurrence of two histone marks on the same nucleosome, we selected well-established antibodies against promoter and gene 5′ region (H3K4me3, and H3K18ac) and gene-body (H3K36me3, and H3K79me3) histone marks. We used exponentially growing *S. cerevisiae* (budding yeast) as a model in studying the interactions of these marks (Rando and Winston, 2012).

In combinatorial-iChIP the first IP and barcoding step provide the first layer of specificity (Figure 1B). To test the specificity provided by this step, we sequenced material obtained from our first IP and barcoding steps. The resulting coverage maps are in good agreement with previously published MNase-ChIP-seq datasets (Weiner et al., 2015) at both local (Figure 1C) and genomic (Figure S1) scales.

Following the first IP, we pooled the barcoded chromatin and used it for the second IP step with the same battery of antibodies (Figure 1B), thus reading out pairwise co-occurrence of histone modifications. The use of MNase-digested chromatin ensures mononucleosome resolution for this assay. Combinatorial-iChIP produced clear signal that was distinguishable from the parental ChIP experiments. Independent combinatorial-iChIP experiments showed highly similar enrichment patterns (Figure S1B). Importantly, combinatorial-iChIP signal was highly similar between reciprocal experiments in which the order of the two antibodies was reversed (Figures 1D and S1B) and exhibited different patterns compared to either of the relevant individual ChIPs (Figures 1E and S2), implying that combinatorial-iChIP captures genomic location where the tested histone marks co-reside on a single nucleosome.

To further determine whether combinatorial- iChIP signal is specific to its target modification, we repeated these assays in cells that express histones mutated at the antibody target residue (e.g., histone 3 lysine 18 replaced by an arginine, H3K18R). This experiment allows one to measure specificity of both IP steps (Figure 2A). After the first IP step, we sequenced the pooled material and subjected all samples to the same amplification and sampling depth. Comparing the number of unique reads, we observed a dramatic reduction in yield when the target residue is mutated compared to WT sample (Figure 2B). The effect was somewhat smaller in H3K36R, potentially due to cross-reactivity of the specific batch of antibody used and the lack of competition with high-affinity target of the antibody in IP reaction. The dramatic loss of material in the mutant backgrounds confirms that the barcoding step during the first IP is highly specific to its target residue. To measure yield in the second IP stage, we compared the number of unique reads from a specific first IP (e.g., H3K4R-H3K36me3) before and after the second IP. This yield was dramatically and significantly reduced when the target residue of the second IP was mutated (Figure 2C and Table S1). We stress that, unlike canonical ChIP, our second IP was carried in the same tube for all samples, thus eliminating any noise due to sample handling. Together, these observations suggest that the signal obtained by combinatorial-iChIP is specific and is not the result of nonspecific background interactions during the first and second IPs.

Combinatorial-iChIP Is a Quantitative Assay

We next turned to examine the quantitative nature of the combinatorial-iChIP signal. At any genomic locus, we define the co-abundance of nucleosomes with a combination of two marks (e.g., H3K36me3 and H3K79me3) as the fraction of cells in the population with a nucleosome with both marks at this locus. This abundance cannot be larger than the abundance of individual marks at the same locus (Figure 3A). Specifically, the co-abundance of two marks cannot exceed the abundance of each of the individual marks, leading to two constraints (Figures 3B and 3C) on the co-abundance. On the other hand, if both the individual marks are highly abundant, then the co-abundance must be above a linear constraint (Figure 3D). These constraints hold for absolute co-abundances. However, the actual read counts in each combinatorial-iChIP library depend not only on the co-abundance but also on other factors, such as antibody yield (fraction of targets retained in the IP steps) and the sequencing depth.

We reasoned that if the signal is quantitative, we would expect to observe these constraints in the data up to an (unknown) amplification and measurement noise. In other words, the signal in each library should have a linear relation with the true abundances. This hypothesis leads to a testable prediction: there is a multiplicative scaling coefficient that would make the combinatorial-iChIP signal obey the underlying constraints. To test this prediction, we searched for a scaling coefficient that minimizes the violations of these constraints in each (first-second) IP pair (Figures 3B–3D, Experimental Procedures). Indeed, for each pair we find a scaling coefficient (one parameter) that agrees with the constraints (less than 2.5% constraint violations). Additionally, the complementary set (>97.5% of nucleosomes) spans the range of allowed interactions (Figures 3B–3D and S3).

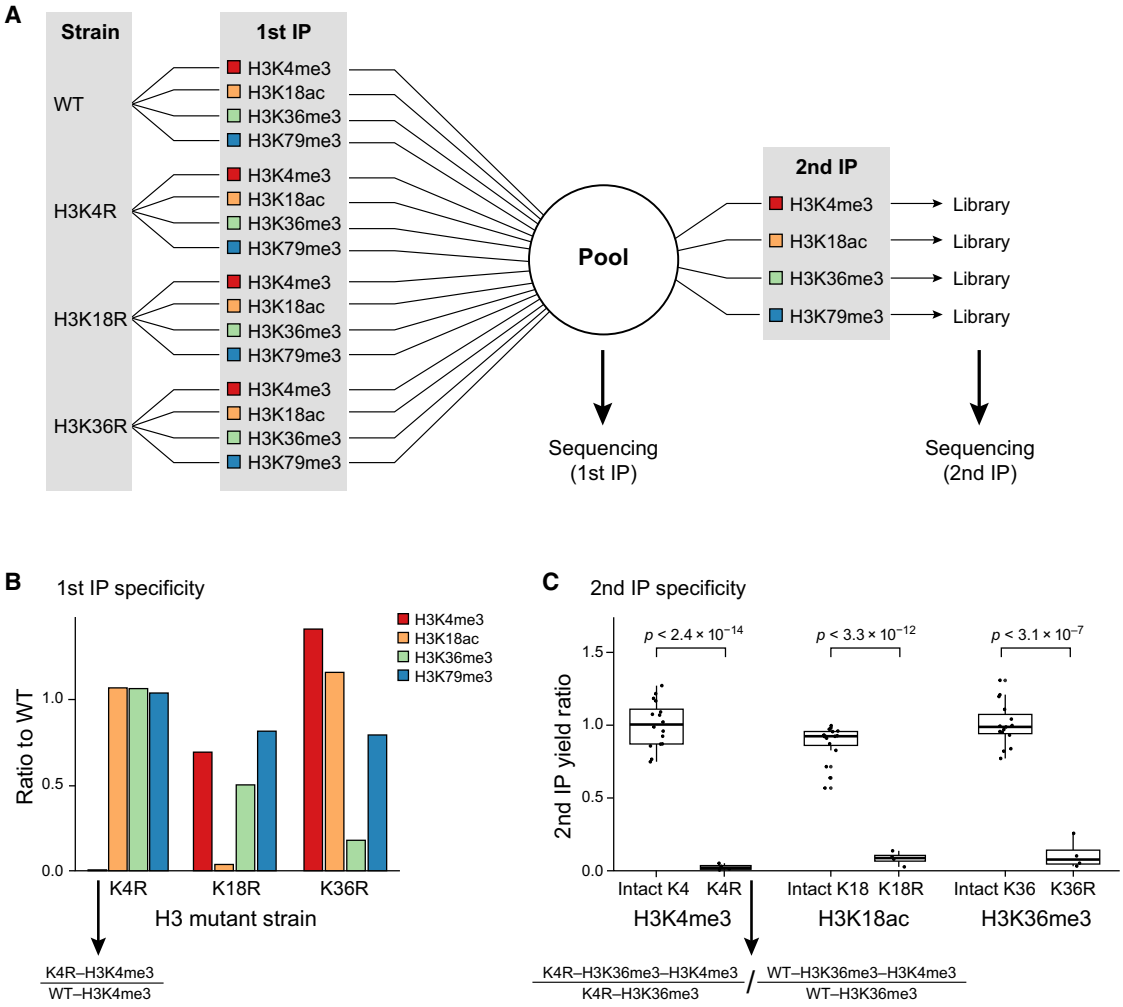


Figure 2. Combinatorial-iChIP Signal Is Specific

(A) Schematic diagram of the flow of material in our histone KO experiment. Samples from four strains (first column from the left) were each subjected to first IP step with four different antibodies (listed on the second column) and each barcoded with a different index. We pooled the resulting material. A fraction was directly sequenced (first IP), and the rest was divided into the second IP step with four antibodies.

(B) Number of unique sequenced reads in different first IP steps shown relative to the same IP in the WT sample. Observed numbers are 0.5%, 3.7%, and 17.8% for H3K4me3, H3K18ac, and H3K36me3, respectively.

(C) Number of unique sequenced reads in different second IP libraries. For each index, we calculate the yield as the ratio between the output (number reads after the second IP) and input (number of reads in the pooled input). We report yield normalized to the yield of the same sample from the WT strain. Results are shown as boxplots (25%, 50%, and 75% percentile) with individual indexes (strain/first IP combinations) as dots. p values are computed using t test. Median yield is 2.5%, 9.2%, and 8.2% for H3K4me3, H3K18ac, and H3K36me3, respectively. See also Table S1.

Finding simple scaling rules for our samples indicates that the combinatorial-iChIP signal is approximately linear in the actual co-abundances. The scaled signal, for any given location in the genome, provides a quantitative estimate of the abundance of single or combinatorial states at this locus. This allows us to distinguish nucleosomes whose combinatorial-iChIP signal is higher (or lower) than the expected levels given the single ChIP signal and an independent modification model (Figure 3E).

Co-occurrence of Gene-Body-Associated Marks

We observe that, for most nucleosomes, the co-occurrence of H3K36me3 and H3K79me3 is higher than we would expect from an independent model (Figure 3E). Both marks are known to accumulate in gene bodies in a pattern anticorrelated with nucleosome turnover rates (Venkatesh et al., 2012; Weiner et al., 2015; Dion et al., 2007). H3K36me3 is deposited by Set2, which is recruited by elongating RNA Pol II (Li et al., 2003), and its presence is currently viewed as protecting gene-body nucleosomes from eviction and thus reducing nucleosome turnover rates (Venkatesh et al., 2012). H3K79me3 is deposited by Dot1 in a manner that is dependent on H2B ubiquitylation (Ng et al., 2002). Moreover, there are no known histone demethylases that erase H3K79 methylation (unlike other lysine methylations). Thus, it is currently assumed that removal of

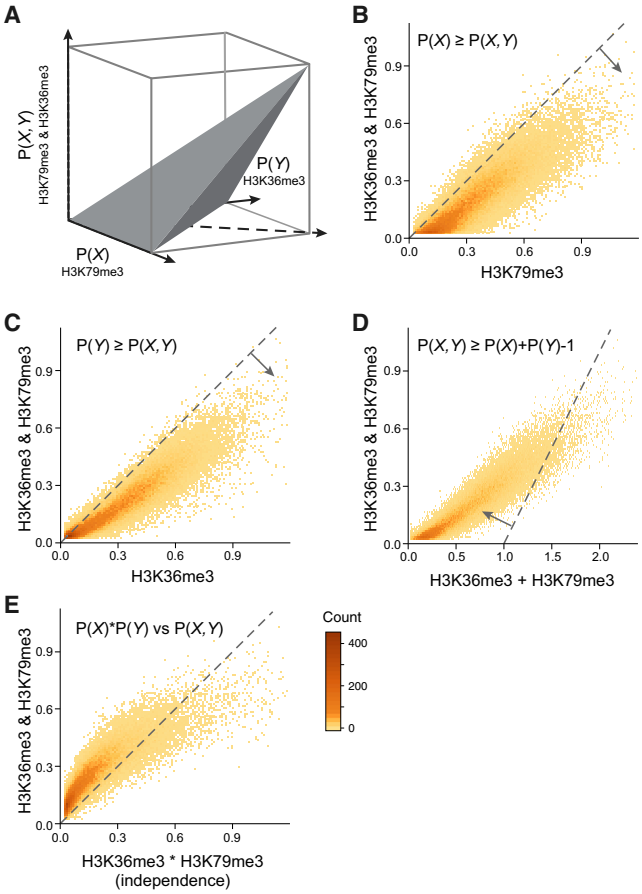


Figure 3. Combinatorial-iChIP Signal Is a Quantitative Measure

(A) Schematic description of the constraints relating the abundances of individual marks at a location to the co-abundance of the dual marks at the same location. If combinatorial-iChIP signal is quantitative, then it should obey these constraints (up to a multiplicative constant).

(B–D) Comparison of the individual ChIP of H3K36me3 and H3K79me3 to their combinatorial signal. Each panel interrogates one constraint (dashed lines). Occupancy levels are as assigned by the model's transformation of the data (Experimental Procedures).

(E) Comparison of observed combinatorial signal against the value expected based on independence between the two individual marks. See also Figure S3.

H3K79me3 is only through nucleosome eviction. The combinatorial-iChIP signal shows large co-occurrence of H3K36me3 and H3K79me3, even in nucleosomes with moderate levels of each individual mark. This observation is in agreement with the evidence that H3K36me3 slows nucleosome turnover, which will result in accumulation of H3K79me3 on nucleosomes marked with H3K36me3.

Co-occurrence of Transcription-Associated Marks

We next used this approach to gain insight into the relationship between H3K4me3 and H3K36me3 (Figure 4A). These marks are deposited by enzymes recruited by initiating (H3K4me3) and elongating (H3K36me3) forms of RNA Pol II that are differentially phosphorylated at the C-terminal domain (CTD) (Ng et al., 2002, 2003). Additionally, H3K36me3 has been implicated in suppressing transcription initiation within gene bodies (Carrozza et al., 2005; Venkatesh et al., 2012). These data suggest that H3K36me3 nucleosomes should be depleted of the initiation mark H3K4me3. This prediction is supported by the sparse overlap of the individual ChIP signals (Figure 4B). Examining the H3K4me3 and H3K36me3 combinatorial signal (Figures 4C and 4D), we see that a majority of the nucleosomes with noticeable signal for one of the marks do not display a combinatorial signal (12,735/23,832 and 22,165/30,721 of H3K4me3 and H3K36me3 nucleosomes, respectively). Focusing on the nucleosomes where both marks are present at the population level (Figure 4B, red outline), we observe a combinatorial signal that is proportional to the expected values from a multiplicative model that assumes independence between the marks (Figure 4E). This is in contrast to the behavior of H3K36me3 and H3K79me3 (Figure 3E). Examining the location of nucleosomes with combinatorial signal for these marks, we see that most of the combinatorial signal is in nucleosomes +3 to +5, which are in the overlap zone between the individual marks (Figures 4A and 4F). Moreover, this signal scales with expression level (Figure 4G). These results support a model where gene-body nucleosomes have a low chance to be modified at both lysines during passage of any individual RNA Pol II molecule through the nucleosomes. However, the border between the two modifications is fuzzy, either due to variation in timing of Pol II CTD modification or due to a large diameter of action of CTD-bound enzymes. Hence, the accumulation of both H3K4me3 and H3K36me3 on the same nucleosome is more likely to occur at highly expressed genes that experience repeated cycles of Pol II passages (Figure 4G).

Dissecting the Set2-RPD3S Pathway

One of the best-studied cases of chromatin regulation and crosstalk in yeast is the repression of "cryptic" transcription by the histone deacetylase Rpd3 small (RPD3S) complex (Figure S4A). RPD3S is recruited to active gene bodies by RNA Pol II and likely gets activated by binding of its Eaf3 subunit to H3K36me3 (Drouin et al., 2010; Govind et al., 2010). This, in turn, leads to hypoacetylation at gene-body nucleosomes of active genes. Thus, interference with H3K36 methylation or RPD3S activity results in hyperacetylation of these nucleosomes, which in turn increases transcription initiation from "cryptic" promoters found in gene bodies (Carrozza et al., 2005; Venkatesh et al., 2012). These previous findings predict that Eaf3 knockout strains will have increased co-occurrence of H3K36me3 and H3K18ac at gene-body nucleosomes.

We carried out a combinatorial-iChIP experiment in set2 and eaf3 deletion strains (Figure S4B). We reproduce previous reports (Carrozza et al., 2005; Venkatesh et al., 2012) showing an increase in gene-body H3K18 acetylation in cells lacking Eaf3 or Set2, with no apparent change in H3K36me3 in Eaf3 knockout cells (Figures 5A and 5B). We next tested the co-occurrence of H3K18ac and H3K36me3. As predicted, combinatorial signal of these marks increases specifically at gene bodies in Eaf3 knockout cells (Figure 5C), thus directly demonstrating this co-occurrence on individual nucleosomes.

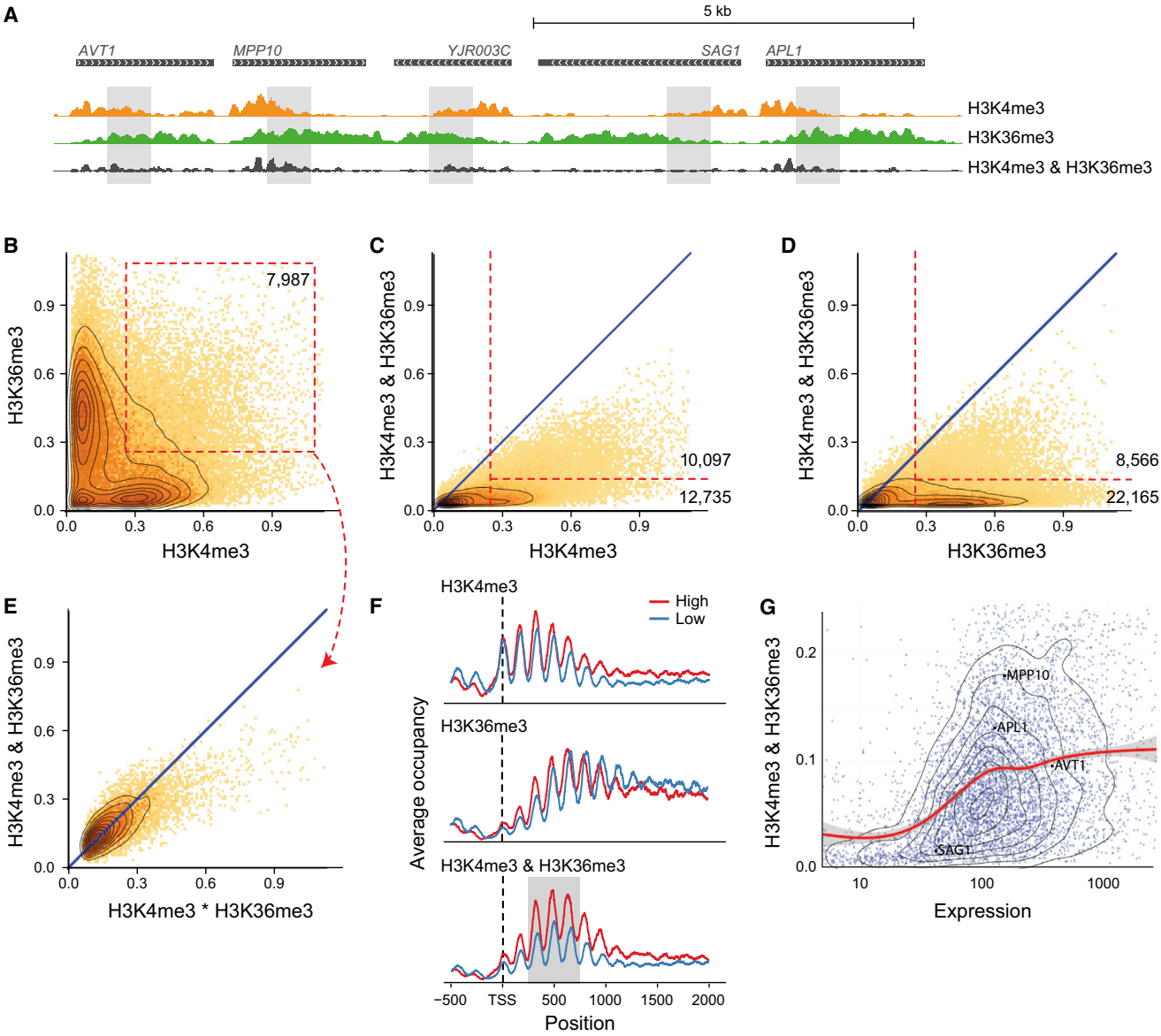


Figure 4. Co-occurrence of H3K4me3 and H3K36me3 Scales with Expression

(A) Polymerase-deposited marks H3K4me3 and H3K36me3 appear at gene start and body, respectively. Shown is a representative genomic region. Overlapping regions of the individual signals show combinatorial-iChIP signal (gray area).

(B) Scatter of normalized H3K4me3 and H3K36me3 individual ChIP levels on all nucleosomes. Most nucleosomes have strong signal for one or the other marks. A subpopulation of nucleosomes show co-enrichment for both marks (red box).

(C and D) Comparison of individual ChIP to combinatorial-iChIP (as in Figures 3B and 3C). Red lines denote population of nucleosomes with high levels of the individual ChIP signal and either low or high levels of combinatorial-iChIP signal.

(E) Comparison of expected co-occurrence signal by chance versus combinatorial-iChIP signal (as in Figure 3E) for the subpopulation marked in (B).

(F) Meta genes of ChIP signal of H3K4me3, H3K36me3, and their combinatorial-iChIP. "High" and "low" denote averages on genes in the 80%–100% and the 20%–40% quantiles of expression, respectively. Combinatorial-iChIP signal is highest in nucleosomes +3 to +5 (gray background).

(G) Comparison of expression levels of genes to the average combinatorial-iChIP signal on nucleosomes +3 to +5 (area marked in gray in F). Red line marks the smoothed mean (gray area, confidence interval in the mean). The location of genes shown in (A) are marked on the scatterplot.

DISCUSSION

The experimental and analytic framework presented here represents an important step toward determining the genome-wide co-occurrence of histone marks at a single-nucleosome resolution. We demonstrate the power of combinatorial-iChIP in resolving population correlations of histone marks into a functional understanding of chromatin states. Using combinatorial-iChIP we were able to demonstrate that a certain combination of histone marks (H3K36me3 and H3K79me3) tends to co-occur

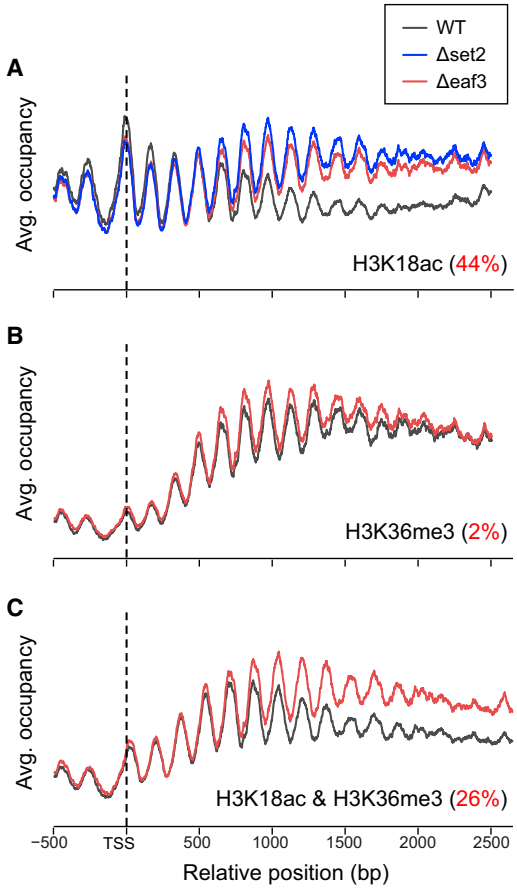


Figure 5. Gain of Co-occurrence of Marks as a Result of Genetic Perturbation

(A–C) Metagene profiles over long genes (ORFs of 2,000 bp or longer). Average occupancy (arbitrary units) versus location relative to TSS in different strains. Median increase in gene body signal in Eaf3 knockout (relative to the increase in 5′ signal for each gene) is reported in red. See also Figure S4.

while another (H3K4me3 and H3K36me3) appears to be deposited independently during transcription. These differences likely reflect underlying mechanisms that drive the co-accumulation of these histone marks. We further used combinatorial-iChIP to shed light on longstanding questions in chromatin biology. By applying combinatorial-iChIP to cells compromised on the Set2-Rpd3 pathway, we found that a mark for H3 acetylation (H3K18ac) coexists with H3K36me3, directly supporting the current theory about the working of this pathway.

Combinatorial-iChIP provides a powerful extension of the widely used ChIP-seq assays. It is not limited to histone marks and can be readily adapted for detecting the co-occurrence of transcription factors as well as other chromatin-associated molecules. Combinatorial-iChIP is amenable to simultaneous processing of many samples. It requires much less biological material compared to canonical ChIP and re-ChIP. This is mainly achieved by early barcoding of chromatin in the first IP step and pooling samples prior to the second IP step. Another advantage of early barcoding and sample pooling is that the second IP step is done in a single tube, simultaneously for all samples, which reduces handling noise and contributes to the quantitative nature of combinatorial-iChIP (Figure 3). By distinguishing co-occurring from merely correlating histone marks, combinatorial-iChIP can help to pinpoint marks that co-specify distinct chromatin states (Ruthenburg et al., 2011).

Limitations

As with ChIP-seq, combinatorial-iChIP relies on antibodies, with the known caveats of antibody specificity and sensitivity. In addition, due to differences in abundance of modifications and in antibody yield efficiencies, the read counts can vary dramatically between different antibodies. As result, much of the sequenced reads can be dominated by few samples. To avoid this, calibration of antibody concentrations and the amount of chromatin is required prior to experiments (Experimental Procedures). We noticed that the barcoding efficiency can differ between different antibodies used during the first IP step. This could be a result of steric interference and/or the binding constants of the antibody. This problem can be overcome either by barcoding more material when using "problematic" antibodies or by barcoding nucleosomes in solution (van Galen et al., 2016) prior to the first IP.

In summary, MNase combinatorial-iChIP is a robust, straightforward methodology that can be easily adjusted to robotic frameworks to probe the combinatorial nature of chromatin. We believe that combinatorial-iChIP can greatly improve our understanding of the compositional and functional complexity of chromatin.

EXPERIMENTAL PROCEDURES

See Supplemental Information for detailed experimental steps.

Yeast Strains and Growth

Yeast strains were obtained from the yeast KO collection (BY4741 with KanMX cassette replacing the deleted gene) and the histone substitution and deletion library (Dai et al., 2008). As "WT" strains we used Bar1 knockout from the yeast KO collection and the H3 WT from the histone substitution and deletion library. Yeast cells were grown in YPD media at 30°C with constant shaking to OD 0.6–0.8.

Cell Growth, Fixation, and MNase Digestion

Cells were fixed with 1% formaldehyde and treated with zymolyase to prepare spheroplasts. Spheroplasts were treated with MNase to digest chromatin into nucleosomes. In all cases MNase pattern showed ~80% mononucleosomal fragments.

Chromatin Immobilization

MNase-digested chromatin was allowed to bind to antibodies for 2 hr. Protein G dynabeads were added for an additional hour, and the beads were extensively washed.

Chromatin Barcoding and Release

Chromatin barcoding and release was performed as previously reported by Lara-Astiaso et al. (2014), with minor modifications.

Second ChIP and Next-Generation Sequencing

The barcoded chromatin was pooled and divided into fresh tubes as the number of antibodies used for the second IP step. Chromatin immobilization and washing was done essentially as for the first ChIP. Chromatin elution and library preparation were performed as reported by Lara-Astiaso et al. (2014). DNA libraries were paired-end sequenced by Illumina NextSeq 500.

Read Mapping

Pair-end reads were mapped to the yeast genome (sacCer3) using bowtie2 with maximal fragment size of 2,000 bp. We removed duplicate fragments, as they are potential PCR artifacts. We defined mononucleosome fragments as these shorter than 220 bp.

Nucleosome Coverage

We used the nucleosome location atlas defined by Weiner et al. (2015). We measured nucleosome coverage by counting the number of fragments overlapping a window of size 150 bp around the center of the nucleosome.

Model Normalization

Consider two modifications, X and Y. Let X_l and Y_l denote the event that a nucleosome at location l has either mark. Thus, the abundance in the population of each mark, or the combination, is $P(X_l)$, $P(Y_l)$, and $P(X_l,Y_l)$. From the laws of probabilities, we have three constraints:

(1) $P(X_l) \geq P(X_l,Y_l)$
(2) $P(Y_l) \geq P(X_l,Y_l)$
(3) $P(X_l,Y_l) \geq P(X_l) + P(Y_l) - 1$

These three constraints define the boundaries of the allowed region shown in Figure 2A. We assume that the number of reads, N_l^X, N_l^Y, and $N_l^{X,Y}$, in our libraries is related to the abundance of each of the combination. The simplest assumption is that

$$N_l^X = \alpha^X P(X_l) + noise \text{ for all locations } l$$

Similarly, such equations can be used for other marks, each with its own multiplicative scaling coefficient.

To test whether we can assign such scaling coefficients, we did the following steps:

(1) We assigned ChIP and combinatorial-iChIP coverage (raw counts) per nucleosome (as described above).
(2) The coverage counts were divided by nucleosome occupancy levels, as measured by Weiner et al. (2015).
(3) Each single ChIP nucleosome coverage vector (value per nucleosome position) was transformed to the range [0,1] by dividing by its 99.5% quantile value. This provides the single ChIP scaling coefficient. The scaled value defines our estimate of occupancy $P(X_l)$ for mark X.
(4) For each combinatorial-iChIP, we performed a line search for the scaling coefficient that minimizes the sum of the deviations from the allowed region constraints (as shown in Figures 2B–2D). Formally, we define the loss of a coefficient α as

$$Loss(\alpha) = \sum_l S\left(N_l^{X,Y} - \alpha\, min(P(X_l), P(Y_l))\right) + S\left(\alpha(P(X_l) + P(Y_l) - 1) - N_l^{X,Y}\right),$$

where $N_l^{X,Y}$ is the normalized value of co-occurrence counts for X and Y from Step 2, and $S(z) = max(z, 0)$. The value of α that minimizes this loss is chosen for normalizing the counts for the combinatorial-iChIP of X and Y. In all cases, the fraction of nucleosome violations did not exceed 2.5%.

We repeated this procedure with different values of quantiles in Step 2. While the actual values were somewhat different, the relative conclusions, including differences from expected value (Figure 2E), were fairly robust to this choice.

Density Plots and Smoothing

All plots were generated using the ggplot2 library of R (version 3.2.3). Density scatters were generated using the geom_bin2d(), contours by geom_density2d(), and smoothed averages by geom_smooth().

ACCESSION NUMBERS

The data in this paper are deposited at Gene Expression Omnibus (GEO) under accession number GEO: GSE84240.

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, one table, and a Detailed Combinatorial-iChIP Protocol and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2016.07.023.

AUTHOR CONTRIBUTIONS

R.S. and N.F. conceived and designed this study. R.S., R.L.-W., and H.W. conducted the experiments with support from A.R. N.F. performed data analysis. R.S. and N.F. wrote the manuscript with contributions from all authors.

ACKNOWLEDGMENTS

We thank D. Lara-Astiaso and I. Amit for providing guidance and support setting up the iChIP protocol. We thank A. Appleboim, C. de Boer, J. Engreitz, S. Grossman, N. Habib, S. Myers, O.J. Rando, K. Shekhar, and I. Tirosh for critical comments on the manuscript and L. Friedman and L. Gaffney for help with the graphics. This work was supported by ERC Grant 340712 and by the ISF I-Core on Chromatin and RNA in Gene Regulation.

Received: June 24, 2016
Revised: July 21, 2016
Accepted: July 25, 2016
Published: August 2, 2016

REFERENCES

Baylin, S.B., and Jones, P.A. (2011). A decade of exploring the cancer epigenome - biological and translational implications. Nat. Rev. Cancer *11*, 726–734.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315–326.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.-J., Anderson, S., Yates, J., Washburn, M.P., and Workman, J.L. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581–592.

Chi, P., Allis, C.D., and Wang, G.G. (2010). Covalent histone modifications–miswritten, misinterpreted and mis-erased in human cancers. Nat. Rev. Cancer *10*, 457–469.

Dai, J., Hyland, E.M., Yuan, D.S., Huang, H., Bader, J.S., and Boeke, J.D. (2008). Probing nucleosome function: a highly versatile library of synthetic histone H3 and H4 mutants. Cell *134*, 1066–1078.

Dion, M.F., Kaplan, T., Kim, M., Buratowski, S., Friedman, N., and Rando, O.J. (2007). Dynamics of replication-independent histone turnover in budding yeast. Science *315*, 1405–1408.

Drouin, S., Laramée, L., Jacques, P.-É., Forest, A., Bergeron, M., and Robert, F. (2010). DSIF and RNA polymerase II CTD phosphorylation coordinate the recruitment of Rpd3S to actively transcribed genes. PLoS Genet. *6*, e1001173.

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. Nat. Methods *9*, 215–216.

Garcia, B.A., Pesavento, J.J., Mizzen, C.A., and Kelleher, N.L. (2007). Pervasive combinatorial modification of histone H3 in human cells. Nat. Methods *4*, 487–489.

Govind, C.K., Qiu, H., Ginsburg, D.S., Ruan, C., Hofmeyer, K., Hu, C., Swaminathan, V., Workman, J.L., Li, B., and Hinnebusch, A.G. (2010). Phosphorylated Pol II CTD recruits multiple HDACs, including Rpd3C(S), for methylation-dependent deacetylation of ORF nucleosomes. Mol. Cell *39*, 234–246.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., et al. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223–227.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. *39*, 311–318.

Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretsky, I., Jaitin, D.A., David, E., Keren-Shaul, H., Mildner, A., Winter, D., Jung, S., et al. (2014). Immunogenetics. Chromatin state dynamics during blood formation. Science *345*, 943–949.

Lewis, P.W., Müller, M.M., Koletsky, M.S., Cordero, F., Lin, S., Banaszynski, L.A., Garcia, B.A., Muir, T.W., Becher, O.J., and Allis, C.D. (2013). Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. Science *340*, 857–861.

Li, B., Howe, L., Anderson, S., Yates, J.R., 3rd, and Workman, J.L. (2003). The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. J. Biol. Chem. *278*, 8897–8903.

Liu, C.L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S.L., Friedman, N., and Rando, O.J. (2005). Single-nucleosome mapping of histone modifications in S. cerevisiae. PLoS Biol. *3*, e328.

Lu, C., Jain, S.U., Hoelper, D., Bechet, D., Molden, R.C., Ran, L., Murphy, D., Venneti, S., Hameed, M., Pawel, B.R., et al. (2016). Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. Science *352*, 844–849.

Maze, I., Noh, K.-M., Soshnev, A.A., and Allis, C.D. (2014). Every amino acid matters: essential contributions of histone variants to mammalian development and disease. Nat. Rev. Genet. *15*, 259–271.

Ng, H.H., Xu, R.-M., Zhang, Y., and Struhl, K. (2002). Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79. J. Biol. Chem. *277*, 34655–34657.

Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol. Cell *11*, 709–719.

Rando, O.J., and Winston, F. (2012). Chromatin and transcription in yeast. Genetics *190*, 351–387.

Rhee, H.S., and Pugh, B.F. (2011). Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. Cell *147*, 1408–1419.

Rivera, C.M., and Ren, B. (2013). Mapping human epigenomes. Cell *155*, 39–55.

Ruthenburg, A.J., Li, H., Milne, T.A., Dewell, S., McGinty, R.K., Yuen, M., Ueberheide, B., Dou, Y., Muir, T.W., Patel, D.J., and Allis, C.D. (2011). Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. Cell *145*, 692–706.

Shema, E., Jones, D., Shoresh, N., Donohue, L., Ram, O., and Bernstein, B.E. (2016). Single-molecule decoding of combinatorially modified nucleosomes. Science *352*, 717–721.

Singh, S.S., Singh, N., Bonocora, R.P., Fitzgerald, D.M., Wade, J.T., and Grainger, D.C. (2014). Widespread suppression of intragenic transcription initiation by H-NS. Genes Dev. *28*, 214–219.

van Galen, P., Viny, A.D., Ram, O., Ryan, R.J.H., Cotton, M.J., Donohue, L., Sievers, C., Drier, Y., Liau, B.B., Gillespie, S.M., et al. (2016). A Multiplexed System for Quantitative Comparisons of Chromatin Landscapes. Mol. Cell *61*, 170–180.

Venkatesh, S., Smolle, M., Li, H., Gogol, M.M., Saint, M., Kumar, S., Natarajan, K., and Workman, J.L. (2012). Set2 methylation of histone H3 lysinec36 suppresses histone exchange on transcribed genes. Nature *489*, 452–455.

Weiner, A., Hsieh, T.-H.S., Appleboim, A., Chen, H.V., Rahat, A., Amit, I., Rando, O.J., and Friedman, N. (2015). High-resolution chromatin dynamics during a yeast stress response. Mol. Cell *58*, 371–386.

Young, N.L., DiMaggio, P.A., Plazas-Mayorca, M.D., Baliban, R.C., Floudas, C.A., and Garcia, B.A. (2009). High throughput characterization of combinatorial histone codes. Mol. Cell. Proteomics *8*, 2266–2284.

Molecular Cell, Volume *63*

Supplemental Information

Elucidating Combinatorial Chromatin States at Single-Nucleosome Resolution

Ronen Sadeh, Roee Launer-Wachs, Hava Wandel, Ayelet Rahat, and Nir Friedman

Contents

Supplemental Figures and Table

Figure S1: combinatorial-iChIP recapitulates standard ChIP-seq results and is reproducible. Related to Figure 1. 1

Figure S2: combinatorial-iChIP genic patterns. Related to Figure 1. 3

Figure S3: Quantitative combinatorial-iChIP patterns. Related to Figure 3 4

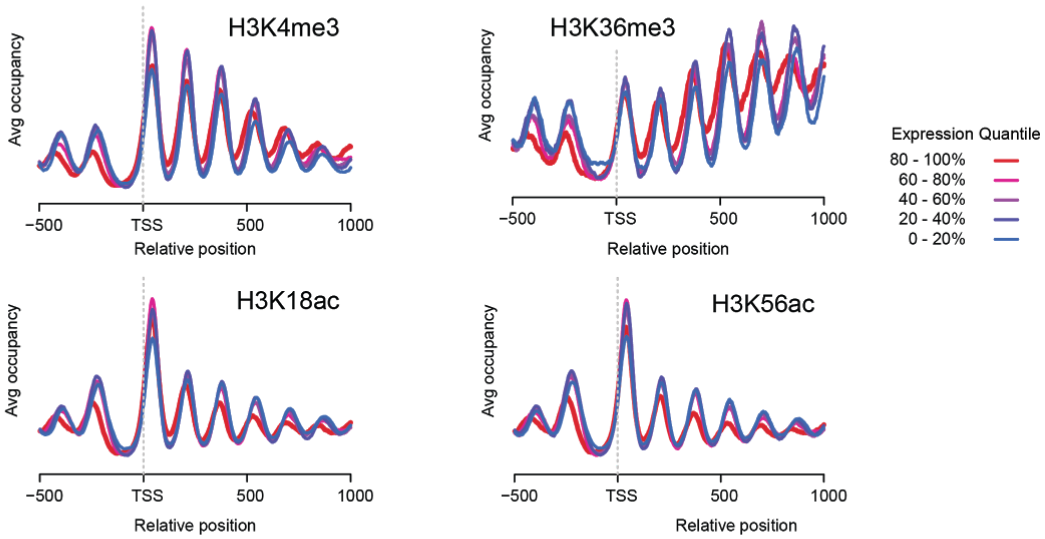
Figure S4: The Set2-RPD3 pathway. Related to Figure 5. 5

Table S1: Unique read counts for each sample. Related to Figure 2. 6

Detailed Combinatorial-iChIP Protocol 7

Supplemental Figures and Table

A



B

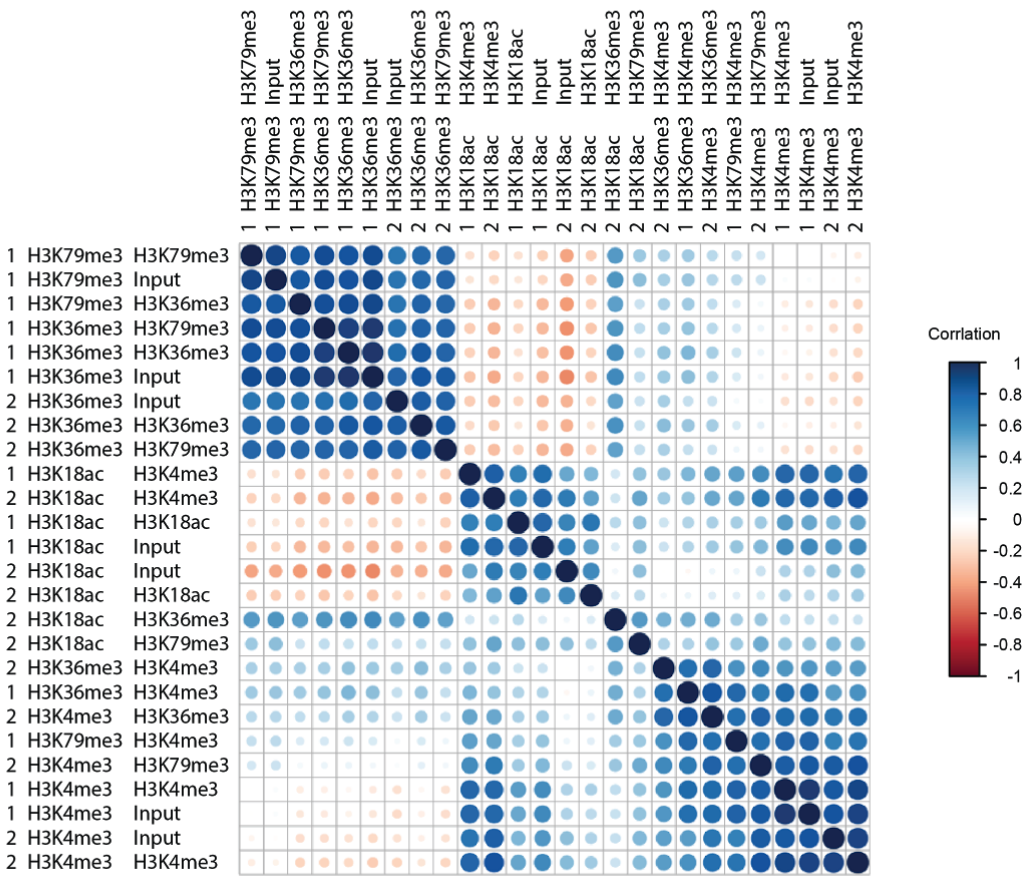


Figure S1: combinatorial-iChIP recapitulates standard ChIP-seq results and is reproducible. Related to Figure 1.

A Metagene profiles of the individual combinatorial-iChIP samples ("input"). Each group of genes (sorted by expression quantiles) are averaged in TSS-aligned manner. **B** Correlation plot of nucleosome coverage of different combinatorial-iChIP tracks from WT strain. Each name consists of batch number (1 - histone mutants, 2 - SET2/RPD3 KOs), first IP, and second IP ("input" first IP without second step). Correlation is denoted by color and by magnitude of the circle. Correlation was computed between coverage counts over nucleosomes (see Methods).

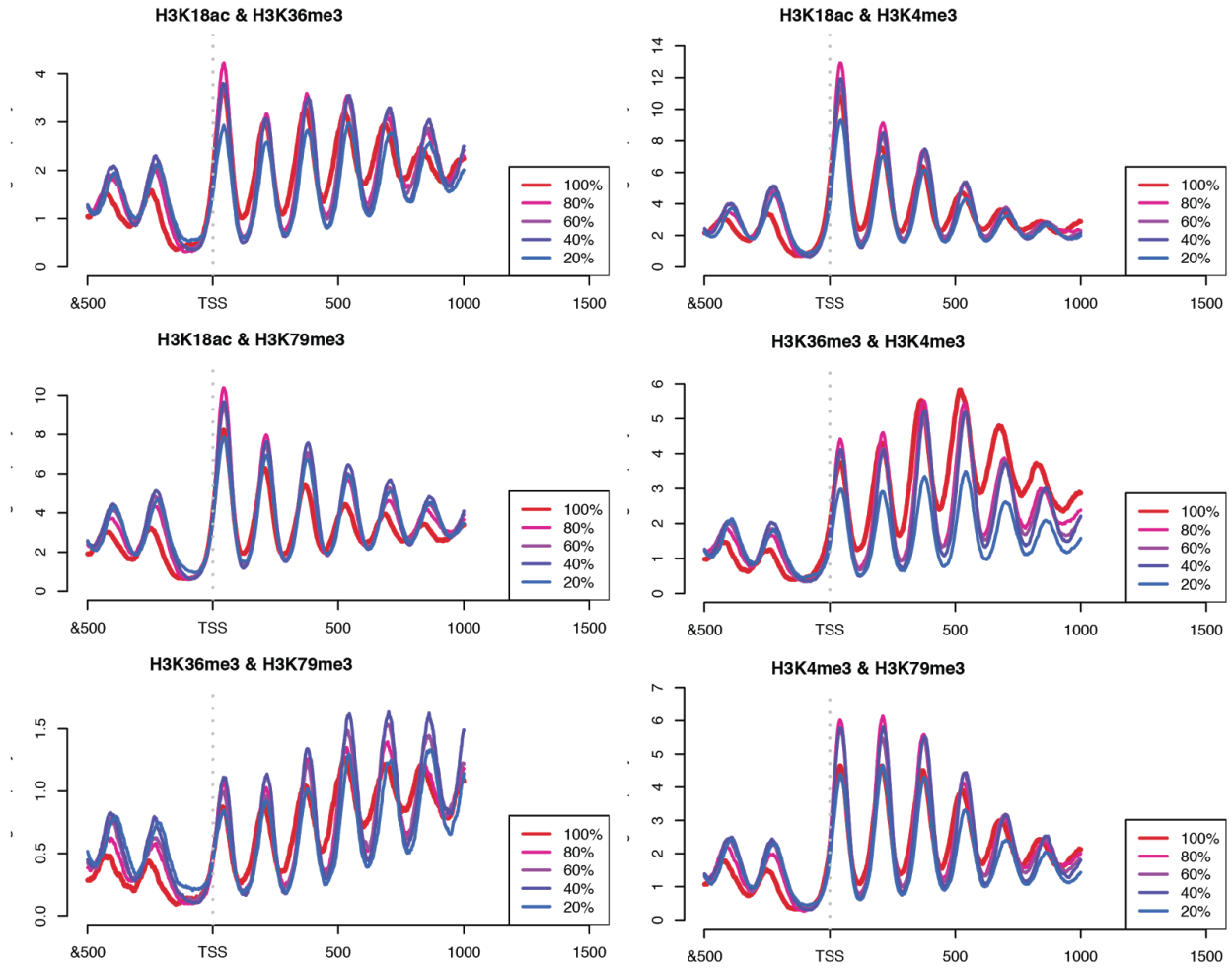


Figure S2: combinatorial-iChIP genic patterns. Related to Figure 1.

Metagene profiles of combinatorial-iChIP. Each group of genes (sorted by expression quantiles) are averaged in TSS-aligned manner.

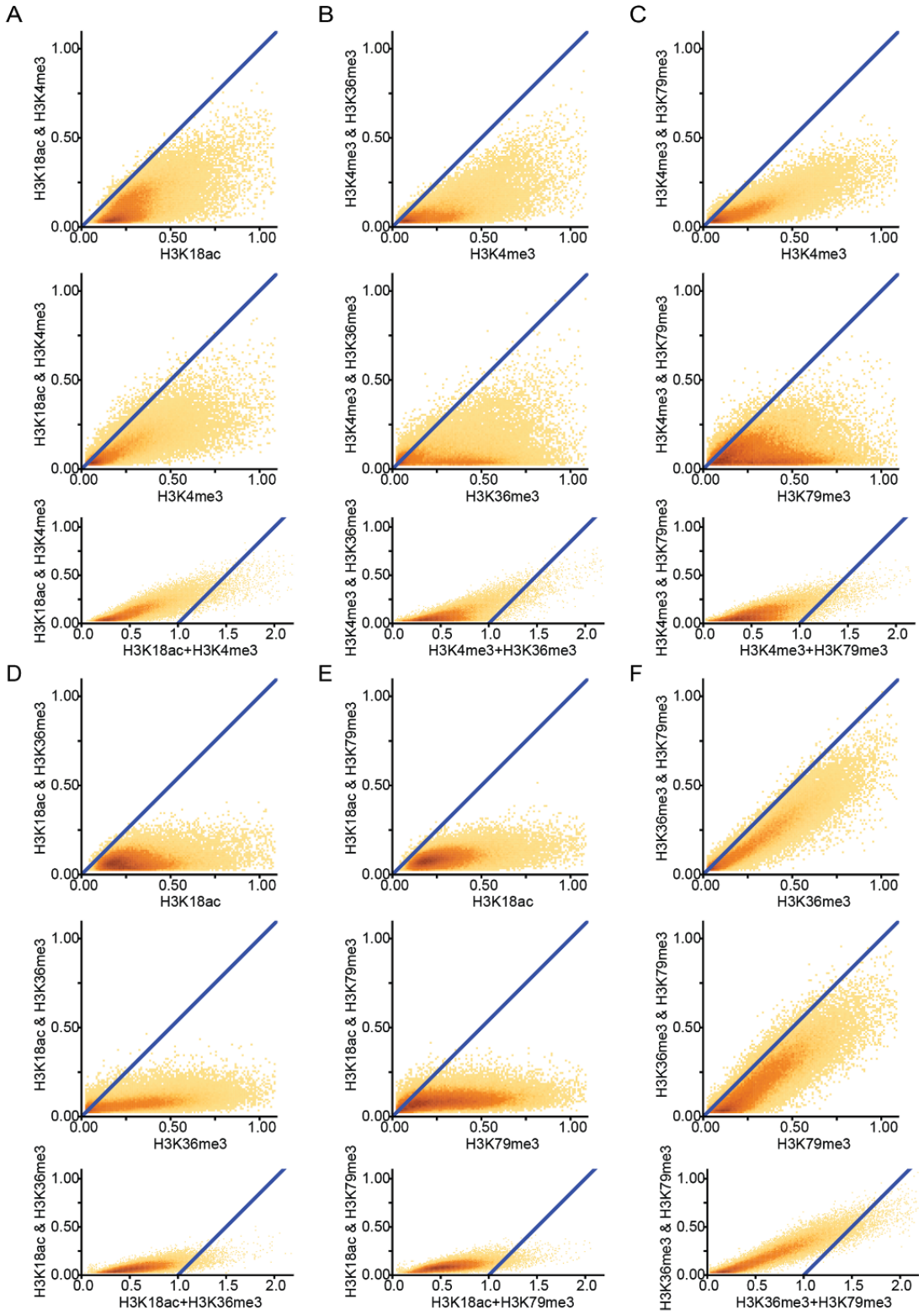


Figure S3: Quantitative combinatorial-iChIP patterns. Related to Figure 3

A-F Panels corresponding to Figure 2B-D for each pair of marks

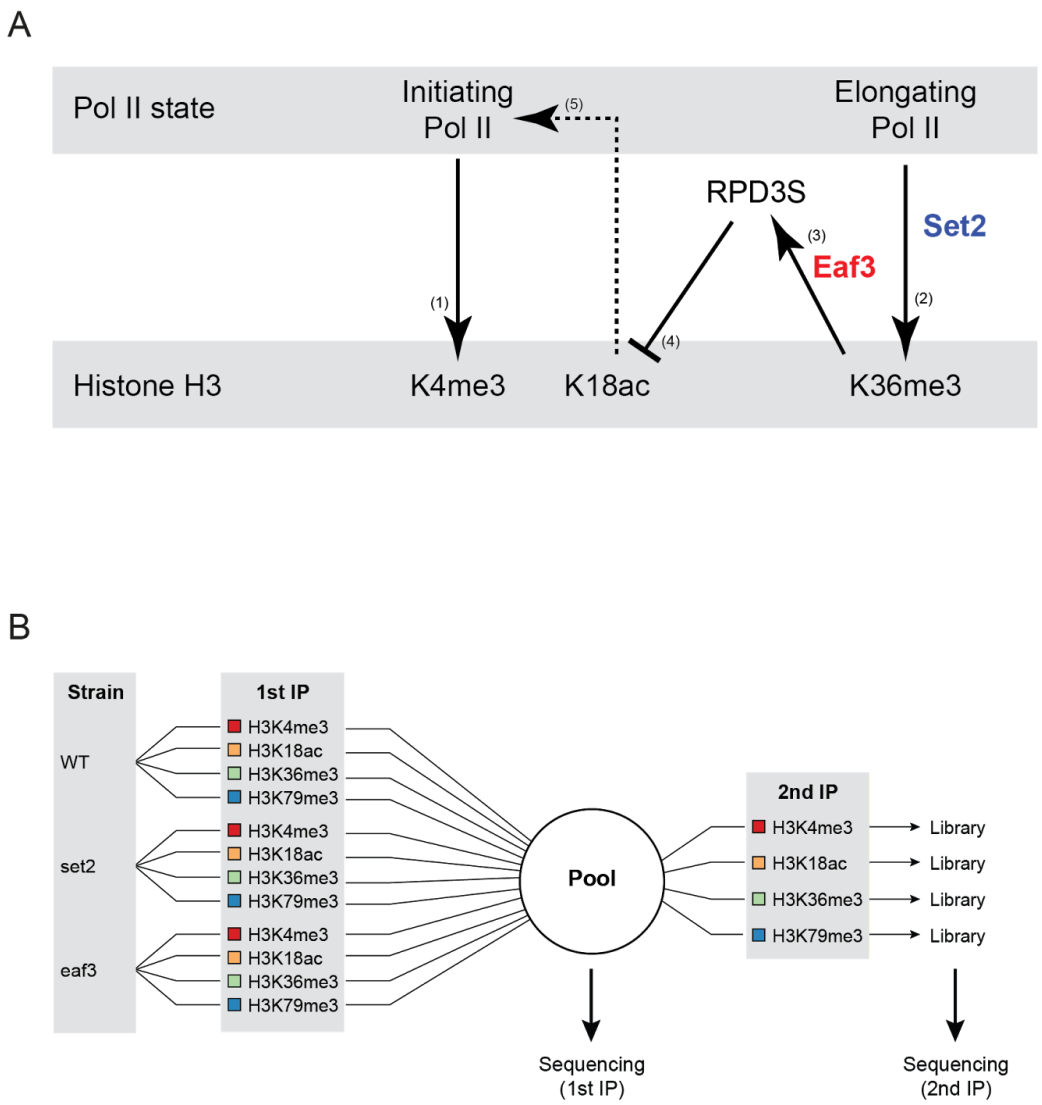


Figure S4: The Set2-RPD3 pathway. Related to Figure 5.

A Summary of the positive and negative relations between Pol II states, RPD3S recruitment, Histone H3 modifications, and turnover (Rando and Winston 2012). Solid lines represent established connections: (1) H3K4me3 is deposited by Set1 recruited to initiating Pol II; (2) H3K36me3 is deposited by Set2 recruited to elongating Pol II (Li et al. 2003); (3) RPD3S is recruited to active gene bodies by RNA Pol II and likely gets activated by binding of its Eaf3 subunit to H3K36me3 (Drouin et al. 2010; Govind et al. 2010); (4) RPD3S deacetylates H3 tail lysines; (5) Deacetylated H3 nucleosomes repress transcriptional initiation. **B** Outline of combinatorial-iChIP experiment to probe this pathway.

Table S1: Unique read counts for each sample. Related to Figure 2.

For each sample, listed the number of reads and the number of unique reads in two size ranges 50-220bp and 220-1000bp.

Detailed Combinatorial-iChIP Protocol

Yeast growth

- Yeast cells were grown in YPD media at 30ºC with constant shaking to OD 0.6-0.8.

Cells fixation

- Add formaldehyde (1% final concentration) directly to cells and rotate 15 minutes at RT.
- Add glycine (0.125M final concentration from 2.5M stock) to the fixed cells and rotate 5 minutes at RT.
- Pellet cells by centrifugation (4000g, 5 minutes, 4ºC).
- Resuspend the cells pellet in 1ml cold ddH2O supplemented with a EDTA-free protease inhibitors cocktail (Roche).
- Pellet cells by centrifugation (4000g, 5 minutes, 4ºC).

Note: It is possible at this point to flash freeze the cell pellet and store at -80ºC.

Spheroplasting

- Resuspend the cell pellet in 100µl buffer Z.
- Add zymolyase 20T (Seikagaku) at 0.3 - 1 units per 1 ml of original culture (2x10^7 cells/ml). Incubate cells at 30ºC for 20 minutes.
- To test spheroplasting efficiency, remove 1-5µl of the cells into 1% SDS solution and check the cells under the microscope. Estimate spheroplasting efficacy (counted cells/expected # of cells*100). This number should be lower than 5%. If it is higher than 5% continue incubating cells at 30ºC and test again.

MNase digestion

Note: It is important to calibrate your MNase digest by running a titration experiment to achieve the desired nucleosome pattern (see digest evaluation below). This calibration should be done for each batch of enzyme and each strain. We usually aim for ~ 80% mono nucleosomes with no or minimally apparent over digested nucleosomes.

- Pellet spheroplasts (6500 g, 10 minutes), remove the supernatant, and resuspended in NP buffer at final concentration of $0.2x10^7$ cells/µl.
- Prewarm samples to 37ºC for 5 minutes.
- Add 12.5 units/ml of MNase (Worthington) diluted in NP buffer for 20 minutes at 37ºC.
- Remove tubes to ice and add one volume of ice cold MNase stop buffer.

- Keep tubes on ice for 10 minutes, vortexed 3 x 10 seconds (this step increase chromatin yield but can be skipped when using 96 well plates to avoid sample spilling).
- Centrifuged samples (16,000g, 10 minutes, 4ºC - for 96 wells plates we centrifuge 30 minutes at 5000g).
- Remove the supernatant containing the chromatin to fresh tubes.

MNase digest evaluation

- Remove 8µl of MNase digested chromatin to a fresh tube.
- Add 2µl of 0.5ug/µl RNase A and incubate for for 30 minutes at 37ºC.
- Add 40µl of proteinase K solution (10 mM Tris pH 8.0, 5 mM EDTA, 300 mM NaCl, 0.6% SDS) containing 50 units of proteinase K.
- Incubate for 2 hours at 37ºC, and for 12-16 hours at 65ºC.
- Isolate DNA by 2X SPRI beads cleanup, resuspend DNA in 20µl of 10mM Tris pH-8.0, and measure DNA concentration by Qubit.
- Expect total of ~ 50-100 ng of DNA
- Visualize DNA by TapeStation (Agilent) or agarose gel. For nucleosomes ChIP we aim for ~ 75-80% mono nucleosomes.

Chromatin immobilization

Note: The amount of DNA used for ChIP can vary and depends on the abundance of the target and antibody yield and specificity. We find that using chromatin from ~$1x10^7$ cells (~0.5ml of yeast culture at log phase) gives good results for most chromatin marks.

- Remove MNase digested chromatin samples (from ~$1x10^7$ cells) to a fresh 96 well plate and adjust the volume to 80 µl with ice cold RIPA buffer and antibody (for specific details see antibodies section below).
- Incubate the samples with gentle tumbling for 2 hours at 4ºC.
- While samples are incubated wash protein G dynabeads three times in RIPA buffer (20µl beads per sample). Resuspend beads to the original volume with RIPA buffer.
- Centrifuge the samples shortly, add 20µl of protein G to each sample, and incubate the samples with gentle tumbling for an additional hour at 4ºC.
- Magnetize the samples and wash: 6 X RIPA buffer, 3 X RIPA 500 RIPA buffer containing 500 mM NaCl), 3 X LiCl wash buffer. It is possible to use vacuum for these washes.
- **Important**: From this point on DO NOT use vacuum to aspirate the supernatant.
- Wash the beads 3 X 10 mM Tris pH 7.5 supplemented with protease inhibitors.

Chromatin barcoding and release

These steps are a modified version of the iChIP protocol (Lara-Astiaso et al. 2014).

End repair:

- Resuspend the immobilized chromatin in 20 µl of 10 mM Tris pH 7.5.
- Add 40 µl of end repair and mix well by pipetting.
- Incubated for 22 minutes at 12°C followed by 22 minutes at 25°C.
- Magnetize beads and wash once in 150 µl 10 mM Tris pH 8.0 and resuspended the beads in 40 µl of 10 mM Tris pH 8.0.

A base addition:

- Add 20 µl of A-Base mix [10 mM Tris pH 8, 10 mM MgCl2, 50 mM NaCl, 1 mM DTT, 0.58 mM dATP, 0.75 units Klenow fragment (NEB)] to the beads and mix well by pipetting.
- Incubate samples at 37°C for 30 minutes.
- Magnetize beads and wash once in 150 µl 10 mM Tris pH 8.0 and resuspended the beads in 18 µl of 10 mM Tris pH 8.0.

Adapters ligation:

- Add 5µl of indexed adapters (Blecher-Gonen et al. 2013) to each sample mix well by pipetting. Add 34µl of ligation mix [29 µl of 2X quick ligase buffer (NEB), 5µl quick ligase (NEB)], mix well by pipetting, and incubate at 25°C for 45 minutes.
- Magnetize beads and wash once in 150 µl 10 mM Tris pH 8.0

Chromatin release:
This step releases bound chromatin and inactivates the antibodies used in the first ChIP.

Note: From this point it is important to keep samples at temperature higher than 15°C to prevent SDS precipitation.

- Resuspend beads in 12.5 µl fresh 0.1 M DTT and incubate at RT for 5 minutes.
- Add 12.5 ul of freshly prepared 2X Chromatin Release Buffer, mix well by pipetting, and incubate at t 37°C for 45 minutes.
- Pool all the samples that are going to be ChIPed together during the second ChIP into a 1.5ml tube.
- Magnetize the pooled samples and remove the supernatant into a fresh 1.5 ml tube.
- Centrifuged max speed, 5 minutes, 15-20°C, and remove the supernatant into a 15 ml falcon tube.
- Measure the volume of the pooled sample and dilute by adding 9 volumes of dilution buffer.
- Load the diluted samples on an amicon filter (Millipore UFC905024) containing 12 ml of Amicon buffer. We usually load ~2 ml of diluted sample per Amicon filter.

Note: In case that many samples are pooled from the first ChIP it is possible to leave the concentrated sample in the Amicon filter after sample concentration and reload it with additional 12 ml of Amicon buffer and 2 ml of pool sample. This will allow for a more concentrated chromatin for the second chIP.

- Centrifuge at 2000g, 15-20°C until ~ 0.25ml of concentrated sample is left in the filter (5-10 minutes).
- Pool the concentrated samples together and add 1 volume of equilibration buffer and vortex the samples for 10 seconds.

Note: At this point samples can be flash frozen and stored at -80°C or proceed to the next step.

Second ChIP:

Note: A critical point in this protocol is to use sufficient amount of barcoded chromatin from the first ChIP in the second ChIP step to obtain enough barcoded DNA for efficient library amplification. The amount of barcoded chromatin that should be used is dependent on factors such as antibody yield, modification abundance, and adapter ligation efficiency and should be determined empirically for each experiment. However, we find that pooling ~ 5 samples from the first ChIP gives good results in most cases. This means, for example, that 20 pooled samples from the first ChIP should be sufficient for ~ 4 different antibodies used during the second ChIP.

- Divide the pooled barcoded chromatin into fresh 96 well plate or PCR tubes according to the number of antibodies to be used for the second ChIP step.
- Adjust the volume to 80µl with RIPA buffer and the antibody.
- Chromatin immobilization and washes were done as for the first ChIP.
- Resuspend the beads in 23µl of chromatin elution buffer supplemented with 2µl of 0.5µg/µl RNase A and incubate for for 30 minutes at 37ºC.
- Add 24µl of chromatin elution buffer supplemented with 1 µl of proteinase K (50 units/µl, epicenter).
- Incubate for 2 hours at 37ºC, and for 12-16 hours at 65ºC.
- Isolate DNA by 0.8 X SPRI beads cleanup, resuspend DNA in 25µl of 10mM Tris pH-8.0.

Library amplification:

- Remove 23µl of the eluted chromatin into a fresh PCR tube. Add add 2µl of barcoded amplification primers mix (Different barcode for each antibody used in second ChIP, see primer sequence below) and 25µl of 2 X Kapa hifi hotstart ready mix.
- Run PCR for 12-16 cycles.
- The number of PCR cycles depends on the second ChIP yield. It is recommended to use the lowest number of cycles that yield sufficient library for next generation sequencing to reduce PCR duplicates.

- Isolate DNA by 0.8 X SPRI beads cleanup, resuspend in 20µl of 10mM Tris pH-8.0, and measure DNA concentration by Qubit.
- The total amount of amplified DNA can vary between 40-1000 ng of DNA.
- Visualize DNA by TapeStation. A good library should display discrete bands at MW higher in ~140bp than the original MNase digested DNA used for ChIP with no or minimal amount of adapter dimers at 140bp.

Forward amplification primer:
5' - AATGATACGGCGACCACCGAGATCTACAC [8bp barcode] ACACTCTTTCCCTACACGAC
Reverse amplification primer:
5' - CAAGCAGAAGACGGCATACGAGAT

Note: We find that different antibodies have varying chromatin barcoding efficiency that does not always correlate with IP efficiency. This is likely due to the mode of binding of the antibody to the nucleosome and/or the antibody-antigen binding affinity. As a result, we recommend testing the barcoding efficiency for each antibody. If barcoding efficiency is low, it is possible to perform several repeats of the 1st ChIP and pool the samples prior to the 2nd ChIP.

Buffers:

Buffer Z: 1 M sorbitol, 50 mM Tris 7.4, 10 mM β-mercaptoethanol (freshly added).

NP buffer: 10 mM Tris pH 7.4, 1 M sorbitol, 50 mM NaCl, 5 mM MgCl2, 1 mM CaCl2, and 0.075% NP-40, freshly supplemented with 1 mM β-mercaptoethanol, 500 µM spermidine, and EDTA-free protease inhibitor cocktail.

MNase stop buffer: 220 mM NaCl, 0.2% SDS, 0.2% sodium deoxycholate, 10 mM EDTA, 2%,Triton X-100, EDTA-free protease inhibitor cocktail.

RIPA buffer: 10 mM Tris pH 8.0, 140 mM NaCl, 1 mM EDTA, 0.1% SDS, 0.1% sodium deoxycholate, 1% Triton X-100, EDTA-free protease inhibitor cocktail.

LiCl wash buffer: 10 mM Tris pH 8.0, 0.25 M LiCl, 0.5% NP-40, 0.5% Sodium Deoxycholate, 1 mM EDTA, EDTA-free protease inhibitor cocktail

End Repair mix: 50 mM Tris pH 7.5, 10 mM MgCl2, 10 mM DTT, 10 mM ATP, 10 mM each dATP , dCTP, dGTP, dTTP, 0.375 units T4 polynucleotide kinase (NEB), 0.01 units T4 polymerase (NEB).

A-Base mix: 10 mM Tris pH 8, 10 mM MgCl2, 50 mM NaCl, 1 mM DTT, 0.58 mM dATP, 0.75 units Klenow fragment (NEB).

Chromatin Release Buffer: 500 mM NaCl, 2% Deoxycholate, 2% SDS, 2 mM EDTA, 2X EDTA-free protease inhibitor cocktail.

Dilution buffer: 100 mM Nacl, 10 mM Tris pH 8, 1 mM EDTA, EDTA-free protease inhibitor cocktail

Amicon buffer: 0.1% SDS, 0.1% Sodium Deoxycholate, 10 mM Tris pH 8,, 1 mM EDTA, 140 mM NaCl

Equilibration buffer: 2% Triton X-100, 0.1% SDS, 0.1% Sodium Deoxycholate, 10 mM Tris pH 8, 1 mM EDTA, 140 mM NaCl, 2X EDTA-free protease inhibitor cocktail.

Chromatin elution buffer: 10 mM Tris pH 8.0, 5 mM EDTA, 300 mM NaCl, 0.6% SDS

Antibodies

The following antibodies were used in this study:

*For each antibody we used qPCR to determine the amount of antibody that results in the best yield /to background ratio.

Antigen	Catlog #	µg antibody / ChIP
H3	ab1791 (Abcam)	3
H3K4me3	07-473 (Millipore)	1
H3K18ac	07-354 (Millipore)	3
H3K36me3	ab9050 (Abcam)	3
H3K79me3	ab2621 (Abcam)	1