

Systems Level Analysis of Histone H3 Post-translational Modifications (PTMs) Reveals Features of PTM Crosstalk in Chromatin Regulation*

Veit Schwämmle[†]**, Simone Sidoli^{‡§§}**, Chrystian Ruminowicz[‡], Xudong Wu^{‡‡}, Chung-Fan Lee[§], Kristian Helin[¶], and Ole N. Jensen[‡]

Histones are abundant chromatin constituents carrying numerous post-translational modifications (PTMs). Such PTMs mediate a variety of biological functions, including recruitment of enzymatic readers, writers and erasers that modulate DNA replication, transcription and repair. Individual histone molecules contain multiple coexisting PTMs, some of which exhibit crosstalk, *i.e.* coordinated or mutually exclusive activities. Here, we present an integrated experimental and computational systems level molecular characterization of histone PTMs and PTM crosstalk. Using wild type and engineered mouse embryonic stem cells (mESCs) knocked out in components of the Polycomb Repressive Complex 2 (PRC2, *Suz12*^{-/-}), PRC1 (*Ring1A/B*^{-/-}) and (*Dnmt1/3a/3b*^{-/-}) we performed comprehensive PTM analysis of histone H3 tails (50 aa) by utilizing quantitative middle-down proteome analysis by tandem mass spectrometry. We characterized combinatorial PTM features across the four mESC lines and then applied statistical data analysis to predict crosstalk between histone H3 PTMs. We detected an overrepresentation of positive crosstalk (codependent marks) between adjacent mono-methylated and acetylated marks, and negative crosstalk (mutually exclusive marks) among most of the seven characterized di- and tri-methylated lysine residues in the H3 tails. We report novel features of PTM interplay involving hitherto poorly characterized arginine methylation and lysine methylation sites, including H3R2me, H3R8me and H3K37me. Integration of the H3 data with RNAseq data by coabundance clustering analysis of histone PTMs and histone modifying enzymes re-

vealed correlations between PTM and enzyme levels. We conclude that middle-down proteomics is a powerful tool to determine conserved or dynamic interdependencies between histone marks, which paves the way for detailed investigations of the histone code. Histone H3 PTM data is publicly available in the CrossTalkDB repository at <http://crosstalkdb.bmb.sdu.dk>. *Molecular & Cellular Proteomics* 15: 10.1074/mcp.M115.054460, 2715–2729, 2016.

Chromatin is a dynamic fiber mainly composed of DNA and histone proteins. Although DNA stores the genomic information of the organism, histones and their post-translational modifications (PTMs) are responsible for maintaining the chromatin structure. Histone PTMs play a fundamental role in transcriptional regulation and epigenetics, as they contribute in recruiting enzymes involved in chromatin remodeling and gene activity, and can be inherited through the cell cycle thereby maintaining a memory of gene expression profiles (1). Histone PTMs have been extensively studied, and a large number of histone readers/writers/erasers have been characterized (2). It has been known for about 20 years now that combinations of histone modifications can affect histone-enzyme interaction, where nearby or distant PTMs interdependently recruit or release a given enzyme. This phenomenon is described as PTM crosstalk (3, 4). Originally, crosstalk is the disturbance caused by the electric or magnetic fields of one telecommunication signal affecting a signal in an adjacent circuit. Thus, it defines something that generates interference to a functional mechanism. In chromatin biology, crosstalk can be used also in positive terms, as one single PTM might require the presence of a second distinct PTM within the same protein in order to mediate a biological function. This phenomenon is typically described as positive crosstalk (5). For instance, the combinatorial histone mark H3S10phK14ac within the p21 (*CDKN1A*) gene activates transcription, which does not occur when only one of the two PTMs are present (6), implying that the two modifications need to be deposited in a coordinated manner. Also, the histone lysine demethylase

From the [†]Centre for Epigenetics and VILLUM Center for Bioanalytical Sciences, Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark; [§]Biotech Research and Innovation Centre and Centre for Epigenetics, University of Copenhagen, DK-2200, Copenhagen, Denmark; [¶]The Danish Stem Cell Centre (Danstem), University of Copenhagen, DK-2200, Copenhagen, Denmark

Received August 7, 2015, and in revised form, June 13, 2016

Published, MCP Papers in Press, June 14, 2016, DOI 10.1074/mcp.M115.054460

Author contributions: V.S., S.S., K.H., and O.N.J. designed research; V.S., S.S., X.W., and C.L. performed research; V.S., S.S., and C.R. analyzed data; V.S., S.S., and O.N.J. wrote the paper.

PHF8 has its highest binding efficiency to the nucleosome when the three marks H3K4me3K9acK14ac are present (7). On the other hand, negative crosstalk involves two PTMs that are mutually exclusive, either in the same amino acid residue or on two different residues. For instance, the protein HP1 recognizes H3K9me2/me3 and it spreads this mark along chromatin to compact it, but it releases the binding in case H3S10 is phosphorylated (8). Also, acetylation of H3K4 was found to inhibit the binding of the protein spChp1 to H3K9me2/me3 in *S. pombe* (9).

The Polycomb Repressive Complex (PRC) group proteins are involved in histone modifications and PRC2, involved in histone H3K27 methylation, was recently found to crosstalk with the other major histone H3 methyltransferases G9a/GLP, responsible for H3K9 methylation (10). Moreover, PRC1 recognizes H3K27me3 with its subunit CBX and catalyze ubiquitination of H2AK119 (11). The epigenetic machinery is also characterized by interconnection between numerous components; for instance, the silencing of the inactive X chromosome in female organisms is mediated by interacting DNA methylation, noncoding RNA and histone PTMs (12). Similarly, DNA methylation is required in embryonic stem cells (ESCs) for the deposition of H3K9me3 and it antagonizes deposition of H3K27me3 (13). Because of this, a number of engineered cell lines were developed with mESC genomes knocked out for crucial components of these complexes in order to investigate their role in embryonic cell development. Specifically, mESCs were knocked out in Suppressor of Zeste 12 (*Suz12*^{-/-}) to unravel its role in PRC2, revealing that this protein is essential for the complex activity and affects regulation of proliferation and embryogenesis (14). PRC1 was deprived of the protein Ring1A and its homolog Ring1B (*Ring1A/B*^{-/-}), showing that these proteins play a central role in genome-wide ubiquitylation of histone H2A, most likely functioning as H2A-specific E3 ligases (15). In particular, this study focuses in proving that the double knockout cells cannot maintain ubiquitination on histone H2A on the inactive X chromosome. Finally, mESCs were also knocked out in three fundamental DNA methyltransferases, Dnmt1, Dnmt3a, and Dnmt3b (*Dnmt* TKO) (16). These cells could maintain stem cells properties and proliferation in the complete absence of CpG methylation, including deposition of histone heterochromatic marks such as H3K9me3.

Quantitative proteomics analysis by mass spectrometry is a powerful tool to determine the existence and function of crosstalk patterns in proteins, as it can be used to determine the abundance of specific PTM combinations and thus potentially reveal the underlying rules for the deposition of these coexisting marks. Hence, deciphering crosstalk between histone PTMs might reveal the mechanisms underlying the biological processes responsible for the formation and maintenance of chromatin states. Accurate characterization of these mechanisms is achievable by combining large-scale studies of the wild-type cell system and a sufficiently large number of

perturbed states where global changes of the PTM landscape are induced. Such a study has not been carried out to date.

Mass spectrometry (MS) is extensively used to study histone PTMs (17), as it is currently the most suitable technology to characterize proteins containing unknown modifications (18). We recently optimized a middle-down proteomics workflow to quantify coexisting histone H3 marks from mouse stem cells (19) using high mass resolution MS equipped with electron transfer dissociation (ETD) MS/MS capability. ETD provides efficient fragmentation of histone N-terminal tails (20), which leads to precise mapping of PTMs within the protein sequence. We also developed bioinformatics tools to process MS/MS data from hypermodified peptides. We subsequently proposed a simple quantifier approach to estimate the crosstalk between two histone PTMs; by comparing the observed and theoretical frequency of two PTMs to coexist on the same histone tail the tool can identify conserved interplays between binary marks in different cell cultures (21). The interplay was calculated by predicting the frequency of two PTMs given their individual relative abundance as calculated from peptide ion intensities. Using our in-house developed data repository, CrossTalkDB (21), we can now analyze large data sets of extensively modified proteins and reveal PTM interdependencies and crosstalk.

Here, we report the application of an improved analytical and computational workflow for consistent systems level analysis of histone H3 PTMs for determination of the impact of epigenetic perturbations on coexisting PTMs and their interplay and crosstalk in histone H3 tails. We analyzed H3 tails isolated from mouse embryonic stem cells (mESCs, strain E14) and mESCs that were perturbed in three different components of the epigenetic reprogramming machinery, namely *Suz12*, *Ring1A* and *Ring1B*, and the three DNA methyltransferases *Dnmt1*, *Dnmt3a* and *Dnmt3b*. We demonstrate that our middle-down proteomics platform combined with advanced computational and statistical data analysis reveal histone PTM crosstalk at the systems level, including rarely studied marks such as arginine methylation. In addition, we performed a systems biology analysis by integrating RNAseq data with our histone-derived proteomics data to study whether observed histone PTM abundance correlated with the cognate enzyme abundances.

EXPERIMENTAL PROCEDURES

Cell Cultures—Cells were grown and prepared as previously described for mouse ESCs wild type, *Suz12*^{-/-} (14), *Ring1A/B*^{-/-} (15), and *Dnmt* triple knockout (16) (strain E14). Wild type and *Suz12*^{-/-} morphology was extensively described in (22). *Ring1A/B*^{-/-} were donated by Vidal and Koseki and thoroughly described in (23). *Dnmt* TKO cells were taken from ref. (24).

Histone Purification and Digestion—Histones were purified using a common protocol with minor modifications (25). Briefly, cell membrane was disrupted by 30 min incubation with hypotonic solution (10 mM TrisHCl pH 8.0, 1 mM KCl, 1.5 mM MgCl₂, and 1 mM DTT) and nuclei were precipitated by centrifugation at (10,000 × *g* for 10 min). Pellet was resuspended in 0.2 M H₂SO₄ and incubated for 30 mins at

room temperature. After centrifugation (10 min at $16,000 \times g$) supernatant was recovered and trichloroacetic acid (TCA) was added to the solution to a final concentration of 33%. After 30 min mix histones were spun down by centrifugation ($16,000 \times g$ for 10 min at 4 degrees). Pellet was washed twice with cold acetone. Samples were evaporated in speedvac, resuspended in $40 \mu\text{l}$ of NH_4HCO_3 100 mM and digested overnight with GluC with a ratio of 1:50 enzyme/sample. Digestion was interrupted by adding $1 \mu\text{l}$ of 100% trifluoroacetic acid (TFA).

Liquid Chromatography—Digested histones were analyzed as previously described (19) with few modifications. Briefly, histones were loaded and separated in a Dionex nanoLC Ultimate 3000 (Thermo Scientific). The two column system consisted on a 5 cm pre-column (100 μm ID) packed with C_{18} bulk material (ReproSil, Pur C18AQ 5 μm ; Dr. Maisch) and a 22 cm analytical column with pulled needle (75 μm ID) packed with Polycat A resin (PolyLC, Columbia, MD, 3 μm particles, 1500 Å). Loading buffer was 0.1% formic acid. Buffer A was 75% acetonitrile, 20 mM propionic acid (Fluka), adjusted to pH 6.0 using ammonium hydroxide (Sigma-Aldrich), and solvent B was 25% acetonitrile adjusted to pH 2.5 with formic acid. Histones were run at least in four replicates at a flowrate of 250 nL/min, with a gradient of 5 min 100% solvent A, followed by 55 to 85% solvent B in 150 min and 85–100% in 10 min for column washing.

Tandem Mass Spectrometry—For detection an LTQ-Orbitrap Velos with ETD source (Thermo Scientific) was coupled online with the nanoLC. Nano-electrospray (Proxeon) was used with a spray voltage of 2.2 kV. No sheath, sweep, and auxiliary gases were used, and capillary temperature was set to 270 °C. Acquisition method was set for not using dynamic exclusion. Acquisition was performed in the Orbitrap for both precursors and products, with a resolution of 60,000 (full-width at half-height) for MS and 30,000 for MS/MS. Precursor charge state 1+, 2+ and 3+ were excluded. Isolation width was set at 2 m/z . The six most intense ions were isolated for fragmentation using ETD with an activation Q value of 0.25, activation time of 90 ms, and supplementary activation. m/z window was set at 450–750 to include only charge states 8–10.

Identification of Peptide Species—Spectra were deconvoluted with Xtract using the following parameters: S/N threshold 0; resolution at 400 m/z 30,000 and monoisotopic mass only true. Database search was performed with Proteome Discoverer 1.4.0.288 (Thermo Scientific). Mascot (v2.5, Matrix Science) was chosen as search engine, with the following search parameters: MS mass tolerance 2.2 Da; MS/MS tolerance 0.01 Da; enzyme GluC with 0 missed cleavages allowed, database was manually curated by filtering for histone proteins from mouse in Uniprot (downloaded October 2012, 43 entries). Variable modifications were: mono- and dimethylation (KR), trimethylation (K) and acetylation (K). Raw files and annotated spectra are available at the ProteomeXchange database (Accession: PXD002560). CSV result files from Mascot were imported and processed in 'isoScale slim' by using a tolerance of 30 ppm. This tool includes filtering of ambiguously assigned PTMs (only PTMs with at least one ion before and after the assigned modification site were accepted) and quantification (principle of quantification described in Sidoli *et al.*, (26)). isoScale slim is available at <http://middle-down.github.io/Software>.

Quantification of Crosstalk/Interplay Values—isoScale slim files can be directly uploaded to the new version of CrossTalkDB. The server summarizes peptides with identical sequences and PTM codes. Abundances are calculated from the summed intensities by normalizing them to the total of a protein family within the data set.

Experimental Design and Statistical Rationale—The overall aim of this work is to reveal conserved cross-talk patterns of histone PTMs independently of PTM abundance. Therefore, we performed intact histone tails analysis across differently mutated mESCs, which were previously characterized as with altered epigenome and phenotype.

As control, we used the mESCs wild type in proliferation. Peptides with similar sequences (Levenshtein distance smaller than 3) are considered members of the same protein family. This normalization allows quantification of the total of e.g. histone H3 variants while preserving their relative abundance and revealing the statistical properties of individual histone variants.

Cofrequencies of binary marks are calculated by summing the abundances of peptides containing the respective marks. Interplay values result from cofrequencies according to $I = \log\left(\frac{f_{12}}{f_1 f_2}\right)$ where f_{12}, f_1, f_2 denote the relative abundance of the binary mark, and both single marks 1 and 2, respectively. This formula yields positive/negative values when the double mark is over-/under-represented in the data set (21). Conservation of interplay values was assessed by a t test. p values were corrected for multiple testing according to Benjamini-Hochberg.

Comparison of the results to randomized data sets provides information about conservation of the detected patterns. The *Ring1A/B*^{-/-} data set was taken and two stages of randomization were introduced. ESRand set: Shuffling of intensity values. ESRandPTM set: Shuffling of PTM codes where only lysine modifications were taken into account. The distribution for the number of PTMs per peptide was maintained. Unrealistic peptides with more than one PTM per residue were removed. Both artificial data sets are available at CrossTalkDB.

Data Integration—Transcriptomics data sets were downloaded from (27) (wild type, *Suz12*^{-/-} and *Ring1A/B*^{-/-}), and GEO data set GSE61457, which is described in (28). Expression profiles over the all four cell lines were merged with abundances of single and binary histone marks and clustered by the fuzzy c-means algorithm (29). Both parameters of the clustering methods were determined according to (30). Results were filtered for a minimum membership value of 0.5 to avoid poor assignments. For each of the four clusters, known enzymes from Histome (31) and ref (32), were selected. Moreover, gene names of each cluster were submitted to DAVID (33) using the R library clusterProfiler (34) for standard analysis (p value cutoff 0.05). GO terms were simplified using the *simplify* function of clusterProfiler.

RESULTS

We set out to detect and reveal novel features of combinatorial histone H3 PTM regulation and crosstalk by studying four different mESC cell lines, three of which were perturbed in the epigenetic machinery. Perturbations of chromatin-modifying enzymes often lead to large-scale changes of chromatin features and markers, such as distinct histone PTMs and DNA methylation. We focused our study on histone H3 which carry many distinct PTMs, including coexisting and combinatorial functional PTMs. Thus, histone H3 is an excellent model for systematic investigations of PTM crosstalk and interplay in proteins.

In order to investigate the global changes of the PTM landscape in mESCs, we selected four model systems, *i.e.* wild type mESC, *Suz12*^{-/-} (PRC2 perturbation), *Ring1A/B*^{-/-} (PRC1 perturbation), and Dnmt TKO (DNA methylation perturbation). Specifically, PRC2 and PRC1 are involved in gene repression by directly catalyzing H3K27me2/me3 and H2AK119ub deposition (29), whereas DNA methylation is also involved in gene repression and for the deposition of H3K9me3 (13). As previously reported, *Suz12*^{-/-} mESCs survive only up to 10.5 days post coitus, although they are much smaller than wild type ESCs (14); *Ring1A/B*^{-/-} has defects in

the maintenance of ESCs identity, as the silencing of genes that govern differentiation of ESCs is hampered (35); finally, the growth of undifferentiated Dnmt TKO cells was comparable to that of wild-type cells, but their growth was delayed upon differentiation by formation of embryoid bodies (16). Dnmt TKO cells also expressed Oct4, Rex1, Fgf4 and Nanog at the same or higher levels than wild-type, typical markers of undifferentiated cells. Wild type, *Suz12*^{-/-} and *Ring1A/B*^{-/-} mESC phenotypes are described elsewhere (15, 16). **Supplemental Fig. S1** shows the morphology of the Dnmt TKO cells grown in 2i culture condition. Western blots confirmed a drastic decrease in expression of Dnmt1. The low expression levels of Dnmt3a and 3b in 2i condition were already shown in ref (36).

Overview of the Analytical and Computational Workflow—The complete analytical and computational workflow for histone analysis is outlined in Fig. 1. Histone purification from cultured cells was performed as described (37). Purified histones were digested using endoproteinase GluC, generating an N-terminal histone H3 polypeptide of length 50 amino acid residues (H3 tail). Tandem mass spectrometry (MS/MS) sequence analysis of the H3 tail was performed as previously described by us (26). Briefly, the LC was set up with a two column system consisting on a C18-AQ trap column, to allow sample loading in aqueous buffer, and a weak cation exchange - hydrophilic interaction chromatography (WCX-HILIC) analytical column, which currently provides the most efficient separation for histone N-terminal tails (38). We achieved reproducible chromatography between technical replicates and across different histone samples obtained from four stem cell lines (**supplemental Fig. S2**). Mass spectrometry data acquisition was performed at high mass resolution in both MS and MS/MS mode. MS/MS fragmentation was performed using ETD. Spectra were identified using traditional protein sequence database searching. However, common database searching engines are not optimized for long and hypermodified peptides, and allowing several (more than six) dynamic PTMs during database searching might lead to false positives because of the large number of molecular candidates allowed. Moreover, canonical histone H3 tails contain eight lysine and seven arginine residues, all potential modification sites; a confident determination of PTM localization might be challenging. Therefore, we implemented bioinformatics tools to filter database search results and for peptide quantification (26).

We have further developed two in-house software scripts, “Histone Coder” and “isoScale” (26), and integrated them into a user-friendly and automated tool “isoScale slim.” The isoScale slim software also provides relative quantification of isobaric peptides cofragmented in MS/MS spectra which share the same sequence but have distinct localizations of PTMs (principle described in (39)). As compared with other published scoring systems (e.g. (40)) the output of isoScale slim does not include score probabilities, but it includes the list of site determining ions for each assigned PTM, which can be used to assess the confidence of the assigned PTM local-

ization. Moreover, the possibility to select either CID or ETD MS/MS fragmentation mode and to define mass tolerances makes this software suitable for any kind of low or high mass resolution MS/MS spectra in middle-down or top-down proteomics experiments. CrossTalkDB, a data repository for multiply modified proteins (21), available at <http://crosstalkdb.bmb.sdu.dk> now takes isoScale slim files as input. It also enables label-free quantification based on peptide intensity instead of spectral counting, thereby yielding more accurate statistics.

Analysis of Histone PTMs in Mouse Embryonic Stem Cells—Our analysis of the four mESC systems by middle-down proteomics achieved quantification of 1092 combinatorial modifications in H3 tails from canonical histone H3 and H3.3 (**supplemental Table S1**). This set of combinatorial marks was composed of 44 different PTMs (type, position). PTMs were quantified by simply summing the abundances of all peptides carrying each particular type of PTM (**supplemental Table S1**). K9, K27 and K36 methylations are overall the most abundant modifications on histone H3, known to cover large regions along the chromatin fiber. Moreover, acetylations (K9, K14, K18, K23, and K27), associated with euchromatin, along with K4me1, being a marker of enhancers, are among the top PTMs on H3 tails with abundance above 1% (Fig. 2A). As modified lysine residues are ubiquitous in mouse stem cells, a majority of H3 histone molecules are modified in at least one of their lysine residues. Arginine methylation at residues R8 and R26 were less abundant but still reached up to 10% relative abundance. Specific properties of these PTMs were further investigated by studying their crosstalk, as shown below. In summary, the systems level global analysis of histone H3 tails in four mESC lines demonstrated that our analytical setup has the sensitivity and specificity to reveal a majority of the known and most abundant PTMs in H3 tails.

Most abundant single and binary PTMs of each cell line are shown in Table I and Table II, respectively. Alteration of the epigenetic machinery leads to rearrangement of the abundance-ranking of single marks but preserves to a large degree the same distinct PTMs within the top ten of single marks. However, changes occur within the entire H3 PTM landscape affecting most of the PTMs and their interactions. The ranking of binary PTMs gets altered to a higher extend. Hence, global rearrangements of the PTM landscape are mostly visible at the combinatorial PTM level. Fig. 2B–2D shows the fold changes for the most abundant PTMs and quantifies the changes of single and binary PTMs upon perturbation. Comparison of the H3 PTMs in knock out stem cell lines versus wild type mESC confirmed known or expected changes but also provided insight into PTM alterations induced by yet undetected mechanisms. Comparative analysis of *Suz12*^{-/-} versus wild type cells was previously reported (26), and we here confirmed the drastic reduction of K27me2 and K27me3. This effect was accompanied by increasing K9me2, K27me1, and K36me1. Dnmt TKO and *Ring1A/B*^{-/-} cells showed de-

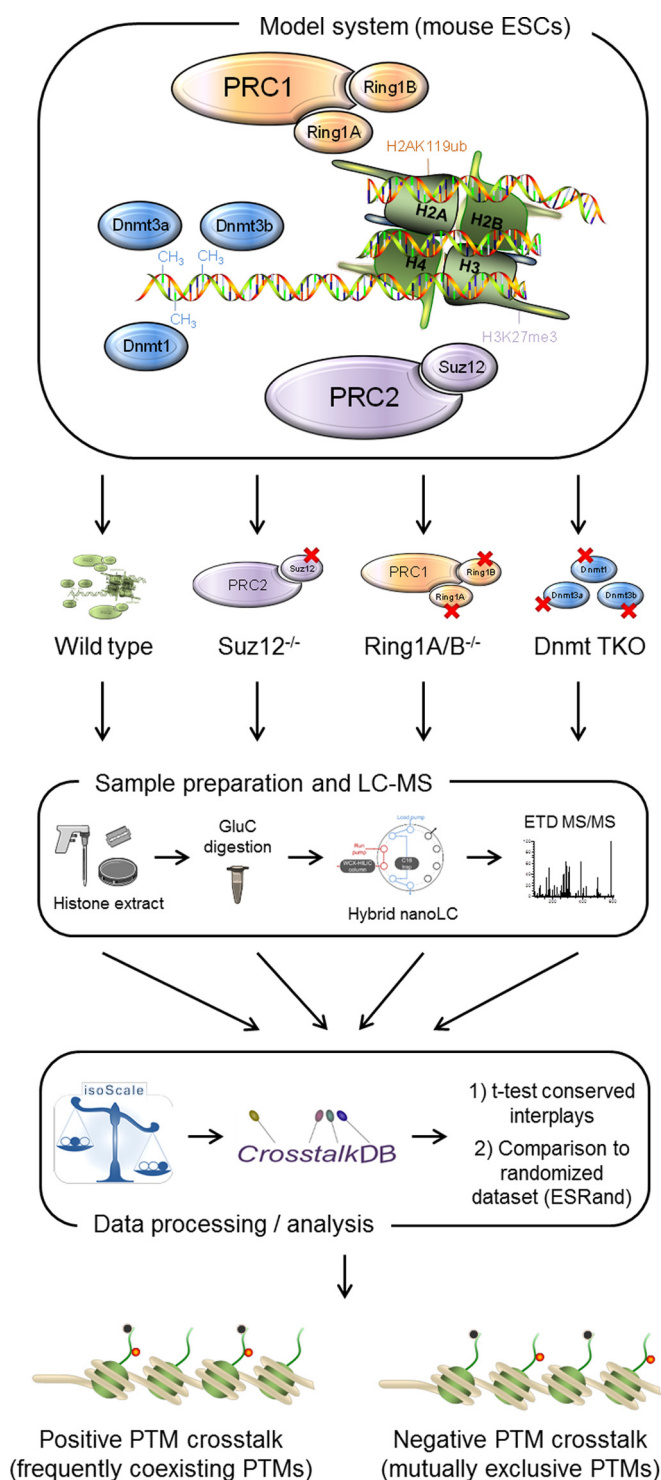


FIG. 1. Workflow for systemic analysis of mouse embryonic stem cells and their conserved crosstalk patterns. We grew mESCs wild type and knockout for the indicated genes to investigate the conserved PTM crosstalk patterns. Histones were extracted and digested by Endoproteinase GluC for analysis of intact N-terminal tails by the middle-down proteomics mass spectrometry platform. Spectra identification (Mascot, Matrix Science) was followed by data processing using isoScale slim and data analysis using CrossTalkDB. Positive crosstalk was defined for combinations of histone PTMs that

created K9me3 and K36me2 abundance, the former potentially resulting from positive crosstalk between the two PTMs as well as from negative crosstalk with K27me3. Murphy *et al.* (13) demonstrated that there is a distinct crosstalk between DNA methylation and K9/K27me3, as H3K9me3 levels are increased by DNA methylation for placement, while K27me3 antagonizes DNA methylation. Our data also reflects this, as the Dnmt TKO cell line presents a reduction in K9 methylation (Fig. 2D) and increase of K27me3. Acetylation features showed less dramatic changes; K14, K18, and K23 acetylations were less abundant in Dnmt TKO cells. In summary, our results highlight that inactivation of *Suz12*^{-/-}, *Ring1A/B*^{-/-} and Dnmt TKO leads to extensive and highly correlated changes in the relative abundance of lysine PTMs on H3.

We also identified common trends for a series of less characterized histone H3 PTMs, including arginine methylation and lysine methylation (K23me2 and K37me1). R8 mono- and di-methylation exhibit only slight changes across the different cell lines, with R8me2 displaying the smallest variation between conditions within the 26 most abundant PTMs (Fig. 2A). Changes of K23me2, R26me1 and R26me2 are highly correlated to K27me2 and K27me3 marks in *Suz12*^{-/-} cells (Fig. 2B). K37me1 stands out as it increases strongly in *Suz12*^{-/-} but decreases in *Ring1A/B*^{-/-} and Dnmt TKO systems, exhibiting the most severe PTM response among the detected histone H3 marks. K37me1 decreases in combination with K27me3 and K23ac whereas the abundance of the latter two marks is not altered significantly. This means that there is suppression of the combinatorial marks K23acK37me1, K27me3K37me1, and K23acK27me3K37me1. Generally, cofrequencies, *i.e.* the abundance of peptides containing two particular PTMs, cannot simply be extrapolated from changes of single PTMs. Supplemental Fig. S3 shows the PTMs with the highest cofrequencies for the different cell lines. Such cofrequencies were subsequently corrected and converted into interplay values in order to estimate the degree of PTM crosstalk (see below).

Characteristic Profiles of Histone PTMs and Histone Modifier Enzymes—With our quantitative approach, we obtain relative frequencies of single and binary histone H3 PTMs. Frequency profiles of single and binary marks along the four cell lines can be further investigated by grouping them according to their common patterns. PTMs with similar profiles are likely to be correlated because of positive crosstalk provoked for example by the same enzyme being responsible for PTM deposition. We determined which histone PTM frequency profiles exhibit similarities with expression profiles of gene transcripts. We furthermore focused on the enzymes reported

coexist with a frequency higher than random. Negative crosstalk was defined as mutually exclusive combinations. Conservation of crosstalk values was determined by calculating conservation of PTM interplay scores across the analyzed data sets (*t* test) and comparing these with a randomized data set (ESRand).

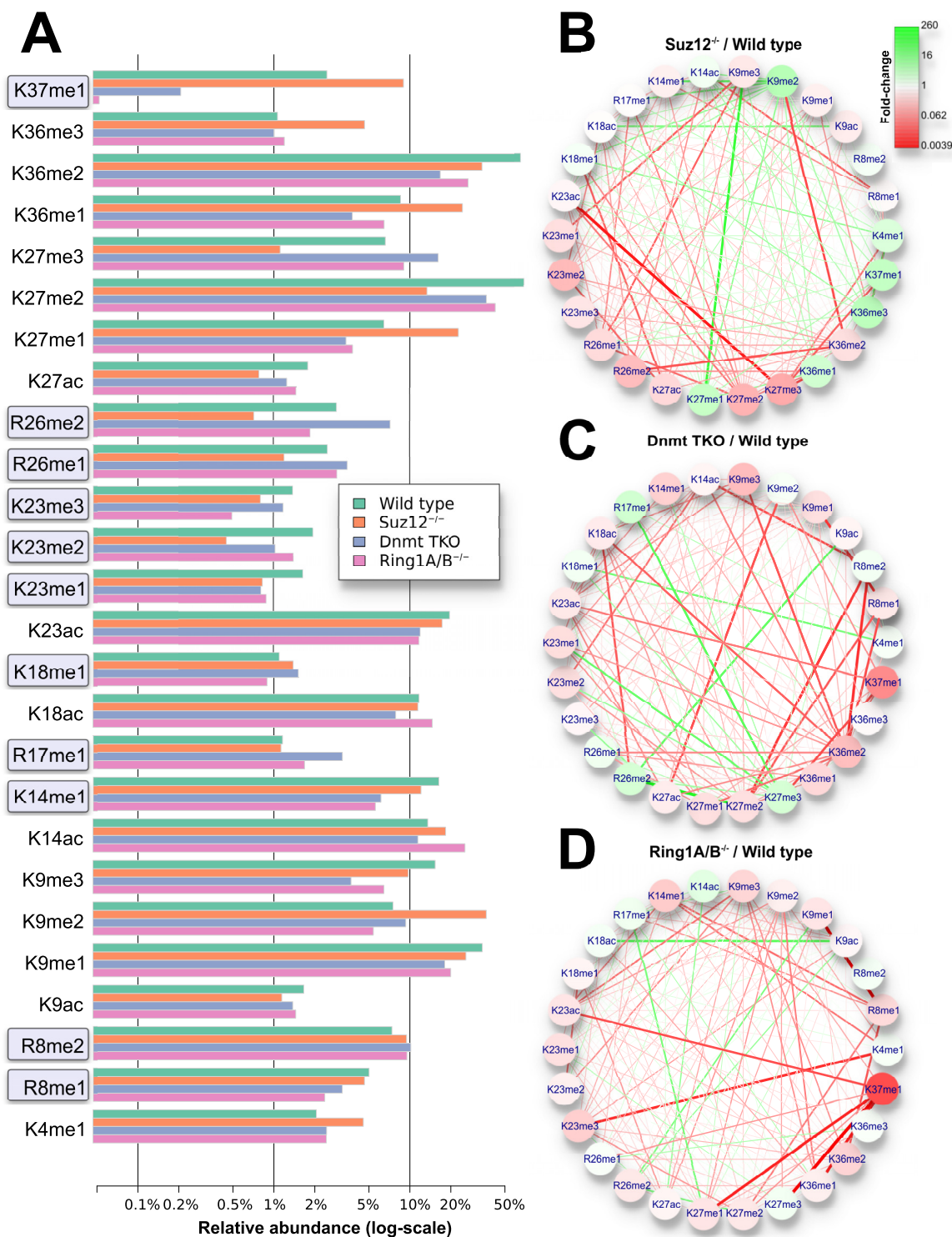


FIG. 2. **Relative abundance of single and binary PTMs in H3 tails from four mESC lines.** A, Comparison of frequencies of the most abundant individual marks in wild type cells (logarithmic scale). Commonly found marks include methylation and acetylation of K9, K27, and K36. In addition, less well characterized marks were detected at rather high abundance, including R8me1/2, K14me1, R17me1, K18me1, K23me1/2/3, R26me1/2, and K37me1 (gray boxes). These marks also exhibited abundance changes upon perturbations, except for R8me1/2. B–D, Abundance fold changes for individual marks (nodes) and pairs of marks (edges) for three perturbed ESC systems compared with wild type cells. Only PTM marks with wild type abundance larger than 1% are shown. Green and red colors denote increase or decrease *versus* wild type, respectively. Suz12 depletion mostly leads to decrease of K27me2 and K27me3 and increase of competing marks such as K9me2. Ring1A/B^{-/-} and Dnmt TKO have strong impact on K37me1.

cluster including up-regulated transcripts in Dnmt TKO and *Ring1A/1B*^{-/-} showed high enrichment for DNA related functions such as nucleic acid binding, chromatin binding and DNA-directed RNA polymerase activity.

We extracted common frequency profiles of single and binary marks together with the expression profiles of their histone modifier enzymes (Fig. 3). Associations between histone PTMs and enzymes are given by edges whereas enzymes without associated PTM in the same cluster are not shown. Concurrent enzyme expression with respect to related histone PTMs occurred for mono-methylated marks and K9 methylations with exception of K18ac and K27ac. These PTMs are highly correlated to the abundance of their modifying enzymes despite of the observed global rearrangements within the PTM landscape. This result suggests that mono-methylated histones and histones methylated at K9 function as a well-maintained core structure of the epigenetic landscape such as the fundamental organization of heterochromatic regions. Other PTMs present a highly intertwined nature where changes of a single mark lead to a global adjustment of histone PTM levels, therefore preventing detection of concurrent expression profiles between histone PTMs and their enzymes. Thus their PTM levels result from a combination of enzyme abundance and crosstalk with other PTMs.

Determination of PTM Crosstalk—We next quantified the interplay between histone PTM marks by calculating over- and under-represented combinatorial marks and assess their conservation at the systems level across the four conditions. The resulting interplay values provide information about the underlying crosstalk rules leading to global regulatory changes in the perturbed cell lines. For example, a low interplay value for the two marks H3K9me3 and H3K27me3 denotes negative crosstalk as both PTMs appear in combination with low frequency compared with independent PTMs without crosstalk. Consequences of the perturbations of the epigenetic machinery can be observed by comparing the entire set of cofrequencies (supplemental Fig. S5). In the following, we compare the correlation among measured and randomized data. Higher correlation between measured samples compared with random samples indicates preservation of the observed patterns. Pearson's correlation assesses the similarity between cofrequencies across the four model systems and varies between 0.5 and 0.8, including the ESRand data set with randomized intensities (supplemental Fig. S6). Hence, PTM cofrequencies get altered to a high degree on global scale in cell lines where the epigenetic machinery is perturbed to directly affect distinct individual PTMs. In contrast, the comparison of interplay values demonstrates a high degree of conservation of interplay values across all 4 cell lines (Fig. 4). As previously proposed, the interplay score calculated for binary PTMs can be used to estimate positive and negative crosstalk (21). Here, we assess changes of the potential overall PTM crosstalk patterns conserved across the four cell lines by comparing interplay values of the observed data sets and

two randomized data sets (Fig. 4). Correlations of interplay values between the four cell lines fell in the range 0.47 to 0.68, while after randomizing the quantitative values of the identified *Ring1A/1B*^{-/-} peptides (ESRand) this correlation decreased to 0.34, strongly indicating conservation of interplay values across the four cell lines. As the data set with randomized intensity values still contains exactly the same peptides and their PTMs, it is expected to be similar to the original set as abundant binary PTMs are naturally present in several modified peptides, thus leading to a high "peptide counting," and the opposite for rare marks. Thus, even though their abundance has been scrambled, they maintain their information about observed PTM patterns because of their presence in multiple peptides. On the other hand, randomization of PTM codes (ESRandPTM) resulted in no correlation at all with experimental data, confirming that arbitrary PTMs do not crosstalk. Specifically, by randomizing the PTM codes, the interplay values become distributed in a bimodal shape (Fig. 4 and supplemental Fig. S5).

Different PTM types, e.g. phosphorylation, acetylation and methylation, have different chemical properties. We therefore investigated whether these pattern become detectable on a global level. Fig. 5A depicts general trends for interplay between PTMs of the same type with respect to their localization on the amino acid chain. We find that crosstalk within arginine mono-methylations, lysine acetylations and mono-methylations is mostly positive at short distances (1–5 amino acids) but becomes negative for larger distances (>10 amino acids) within the H3 tail. Thus it is likely that the enzymes responsible for depositing these marks are not highly specific and can also modify neighboring residues. Many neighboring mono-methylated lysine residues were found to exhibit positive crosstalk that was conserved across the four cell lines (Fig. 5B). Higher methylation states (di/tri) do not show this particular pattern *i.e.* adjacent di/tri-methylated residues mostly interact in a negative fashion. Mutual exclusion of PTMs of the same type should be expected considering the chemistry of such modifications; in fact, an excess of acetylated lysines on the same histone protein would remove the positive charges (at neutral pH), impairing the interaction between DNA and nucleosomes. This hypothesis is supported by the observation of negative crosstalk between distant acetylation sites. Moreover, most trimethyl marks tend to not co-occur on the same histone tail, as this would lead to an excess of positive charges. In addition, lysine trimethylations on histone H3 have different biological functions, explaining why they mostly do not co-occur on the same protein (K4me3, active promoters; K9me3, constitutive heterochromatin; K27me3, facultative heterochromatin; K36me3, active gene bodies). The shape of the interplay trends becomes different for combinations of different PTMs (supplemental Fig. S8), where they often assume a u-like shape, with positive crosstalk for distant residues as compared with negative interplay at short distances. For example, combinations of adjacent lysine methylation

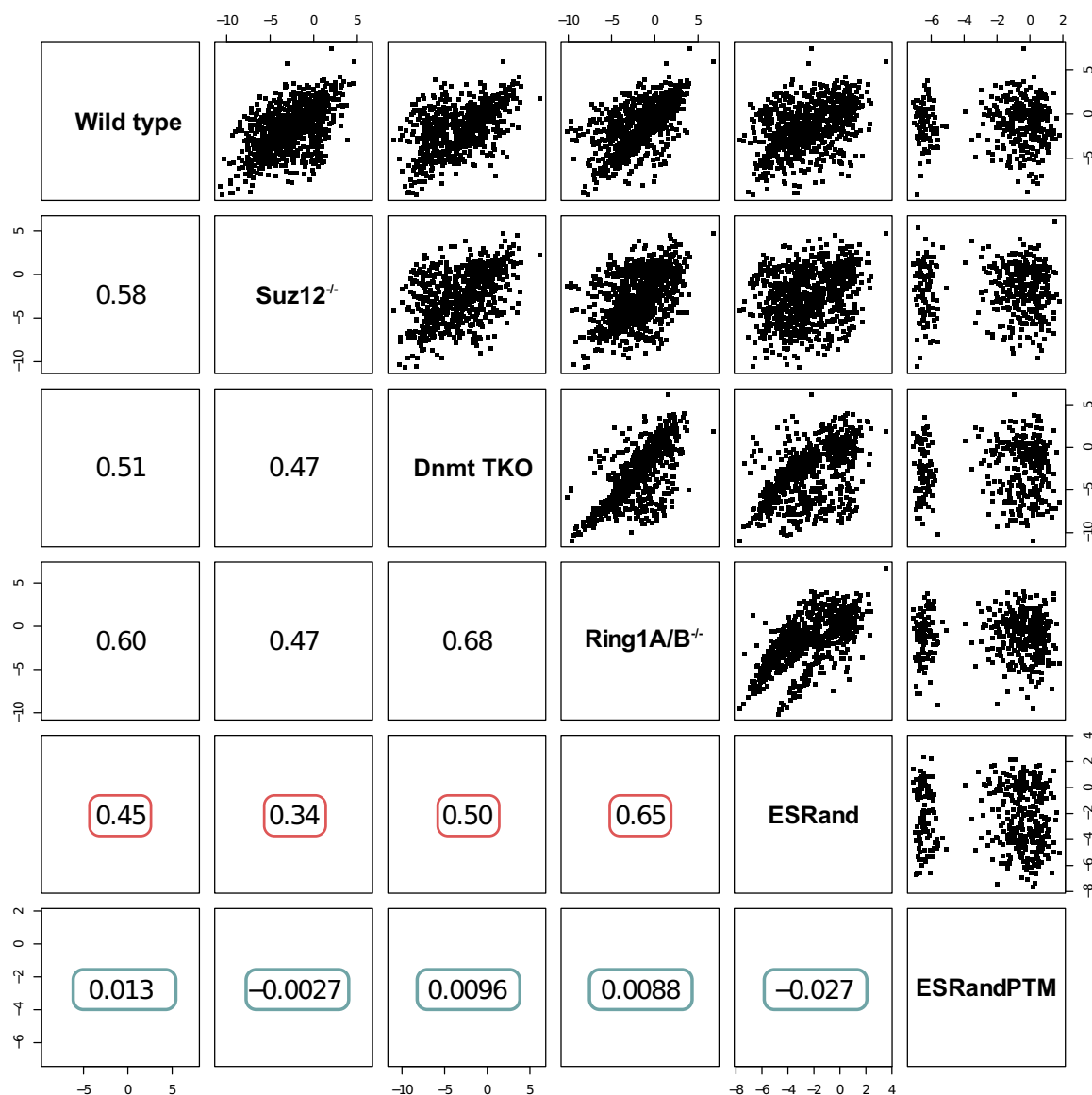


FIG. 4. Interplay values for the four ESC data sets versus two randomized data sets. Comparison of obtained interplay values allows assessing conservation of PTM crosstalk across perturbed cell states. Interplay values remain similar to a high degree within different mESC cell lines. The scatter plots compare data from the four cell lines and two randomized data sets. Lower panels show Pearson's correlation. In ESRand, intensities of the peptides were shuffled. The interplay values show lower correlation (red boxes) with other cell lines than *Ring1A/B*^{-/-}. Randomization of PTM codes (ESRandPTM) leads to no correlation (turquoise boxes) with all other data sets.

states (Kme1 with Kme2 and Kme2 with Kme3) tend to be more positively correlated at larger distances. The observed interplay trends could be confirmed for the tails of the histone variant H3.3 (supplemental Fig. S8).

We investigated crosstalk patterns of well-characterized histone PTMs to validate and extend previous findings. Fig. 6 shows cofrequencies and interplay values for selected groups of H3 marks. Crosstalk values for methylation events at K9, K27, and K36 exhibit overall negative values for trimethyl marks and a trend for positive crosstalk between adjacent lysine methylation states, as previously described (21). K14, K18, and K23 were the most abundant acetylated sites on

histone H3. These acetylation sites and K4 methylations show conserved positive or not conserved interplay values, which were expected as these are all markers of the euchromatic state and activated genes. Other acetylation sites, including K9ac and K27ac, have more specific trends for their interplay values by exhibiting conserved negative crosstalk for e.g. K9acK27ac and K14acK27ac. We further demonstrate crosstalk relationships for the different methylation states. High methylation states (di/tri) are associated with negative interplay values. Remarkably, for the very abundant histone marks such as K9me3 and K27me3 we did not detect peptides that carried both of these marks, *i.e.* they exhibit low interplay

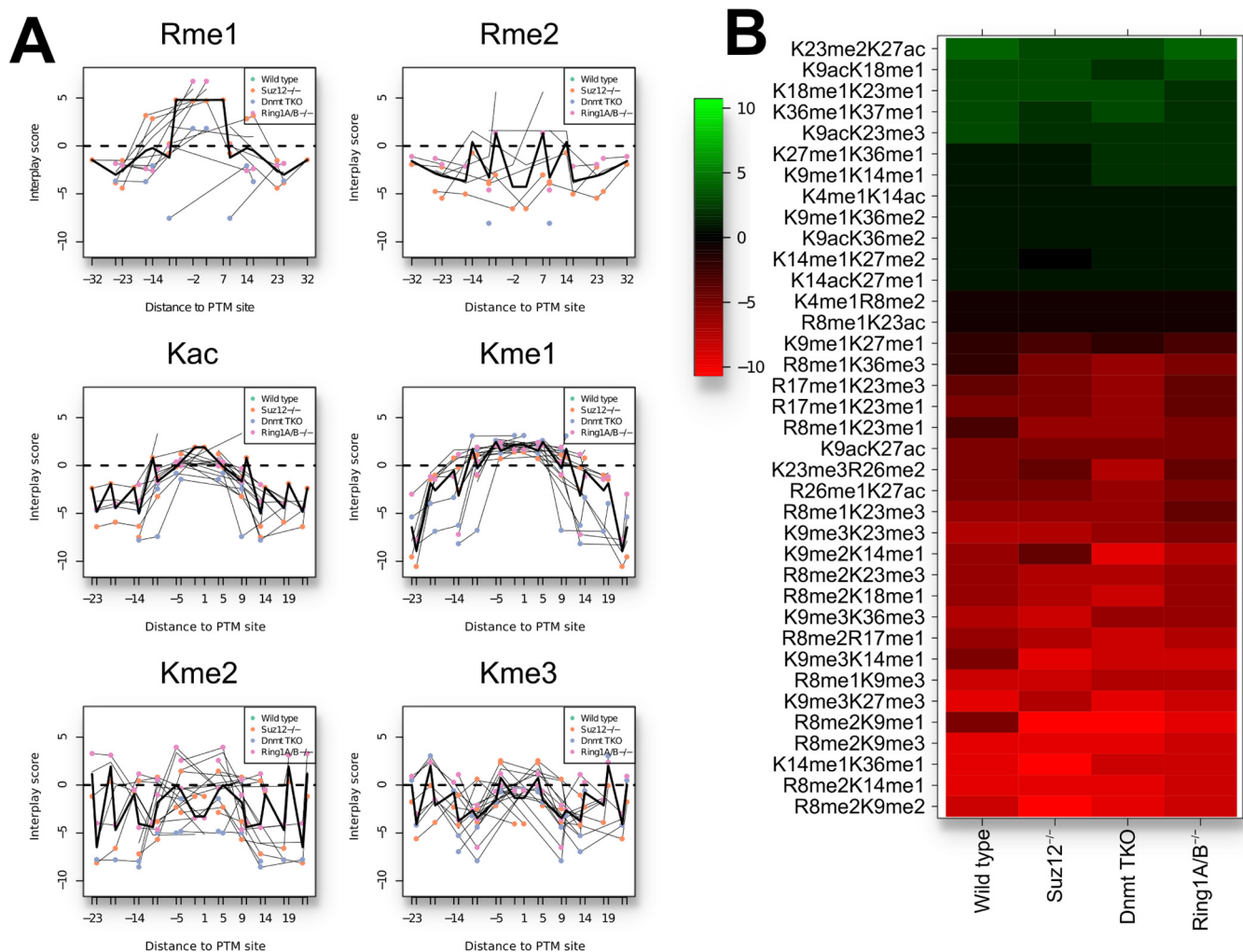


FIG. 5. Overview of histone H3 tail PTM interplay. General patterns of PTM crosstalk on H3 are revealed by investigation of the trend lines with respect to the mutual distance on the protein sequence, as well as by determining the conservation of crosstalk across the different cell lines. **A**, Averaged crosstalk within PTM types arginine methylation and lysine methylation and acetylation. For each PTM, interplay scores with other PTMs versus the distance from their residues are shown (thin lines). The thick line denotes mean values taken over all interplay scores at a distance. Thus the panels show general trends within PTM types, such as positive crosstalk between nearby mono-methylations and acetylations. **B**, Conserved interplay values between cell lines according to a standard *t* test (*p* value corrected for multiple testing according to Benjamini-Hochberg, *p* < 0.1). Only PTMs with wild type abundance above 2% were considered.

scores. We therefore conclude that each of these individual marks is associated with distinct chromatin states. In contrast, we confirm positive crosstalk between near-vicinity mono-methylated residues. Overall, our study reveals a complex network of relationships between mono-, di- and trimethylations, including the actual methylation state and the proximity of adjacent modified residues, which might provide further insights about the mechanism of methylation turnover and order of deposition of these marks.

Arginine Methylation—As shown in Fig. 2A, several arginine methylation features are among the 20 most abundant PTMs on histone H3 tails. However, many more details remain to be elucidated in order to determine the biological functions of arginine methylation. We therefore investigated how these arginine methylation marks relate to other histone H3 marks. From

Fig. 7A and supplemental Fig. S9 follows that three PTMs (R2me1, R8me1, and R8me2) stand out by having a high number of conserved negative interplay values and no conserved positive crosstalk. PTMs on the second residue of the H3 tail (R2me1, R2me2) were found to be unique marks, *i.e.* they did not coexist with any other PTMs in the H3 tail. This is remarkable, and it suggests that there are subsets of histone H3 tails with a well-defined but simple PTM state that largely excludes any other PTMs. R8 methylations are rather abundant PTMs in wild type mESC (R8me1: 5%, R8me2: 7.3%). By using CrossTalkDB we calculated the statistical properties of the data subset containing peptides with R8 methylation and displayed them as word clouds (Fig. 7B and 7C). These PTMs possess well-defined patterns of coexistence with other PTMs. Although R8me2 defines a single mark that is only found alone (as for R2

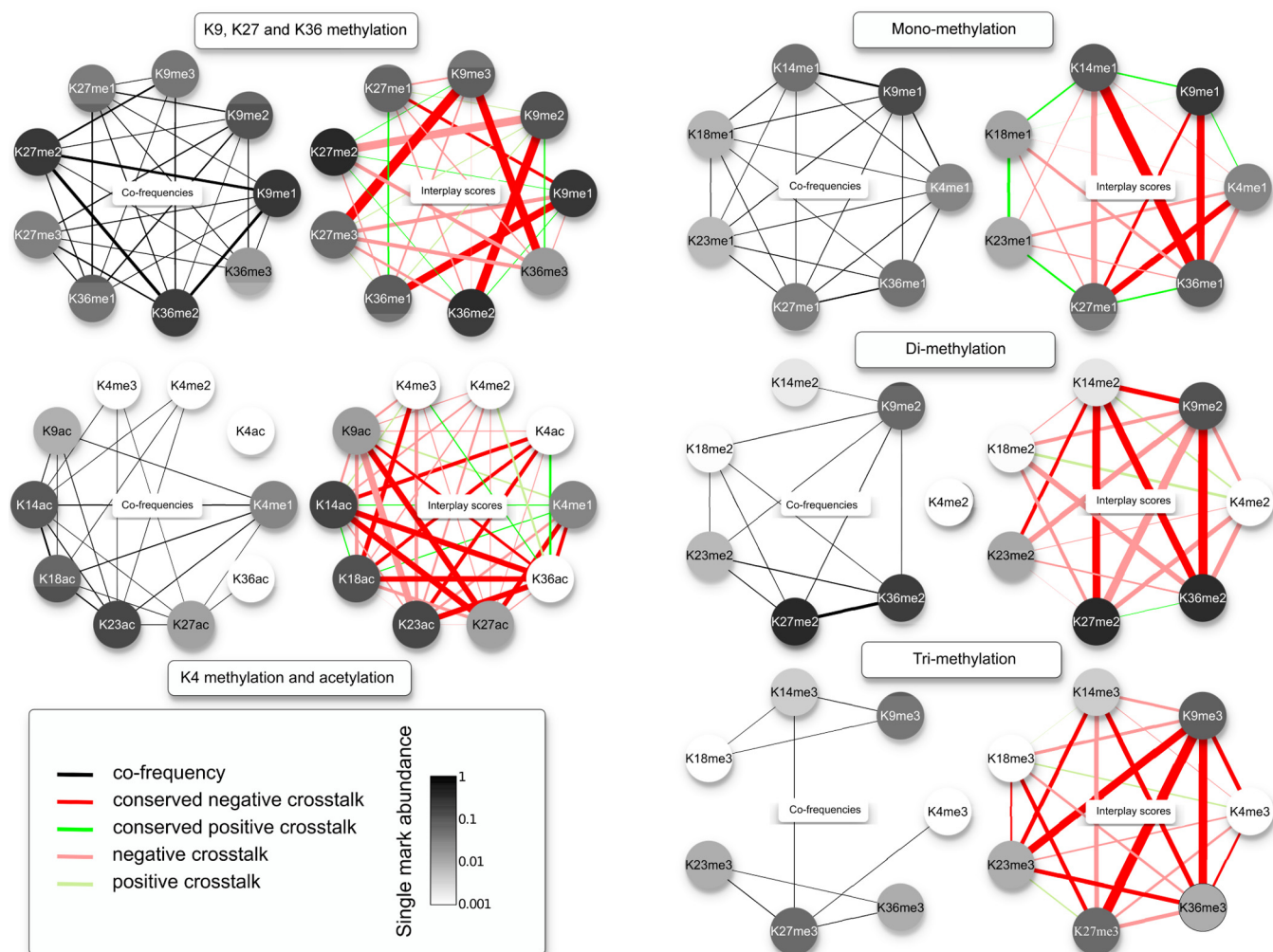


FIG. 6. Fine-structure of conserved crosstalk between selected PTMs. Graph representations of PTM subsets show averaged frequencies of single and double marks and quantitative crosstalk rules. PTM abundance is represented on grayscale. Thickness of black lines denotes cofrequencies. Red and green lines represent conserved interplay values (t test, $p < 0.1$) and light red and green lines correspond to interplay values not found to be conserved. For all values, the average over the four cell lines was taken.

methylation), R8me1 shows a different behavior comprising essentially 2 polypeptide forms: R8me1 and R8me1K9me1K27me2K36me2. Previous studies demonstrate that R8me2 prevents K9ac deposition (42) and leads to gene repression. However, not much is known about R8me2 interplay with other PTMs and how it promotes gene repression. Our data indicate that this mark is mostly acting alone on the histone H3 tail. Moreover, word clouds generated from all H3 tail proteoforms within a cell line demonstrate that the H3R8me2 tail proteoform is the most abundant H3 tail species in all three perturbed ESC lines, but not in the wild type ESCs (Fig. 7D).

We extended our interplay analysis to other less-studied marks including histone H3 methylation sites R17, K18, R26, and K37 (Fig. 7E and supplemental Fig. S10–S18). Most histones containing H3R26me2 only carried one additional PTM mark, most frequently K9me2, but also K27me2 or K36me2 (supplemental Fig. S17). The R26me2 mark is deposited by CARM1 (PRMT4), which is also re-

sponsible for the deposition of H3R17me2 (43). Similarly, no high abundance levels of R17me2 and R26me2 were found (supplemental Fig. S11). Our data suggest that the H3R17me2 and H3R26me2 marks are not deposited within the same histone H3 molecule. Not much is known about other marks such as K37me1, also previously observed by others (44). Our data show that K37me1 is strongly regulated in *Ring1A/B*^{-/-} and Dnmt TKO cells (Fig. 2C and 2D), and that about 50% of all H3 histones that contain this mark are simultaneously modified by K9me2, K27me1, and K36me1 in wild type cells (Fig. 7E). Remarkably, this peptide with a quadruple mark re-emerged in the other cell lines as the most abundant peptide with altered methylation states.

In summary, our analysis of coexpression and crosstalk of single histone H3 marks highlights distinct PTM patterns and the fine-tuning of their relative abundance and interplay as a response to perturbations of the epigenetic machinery.

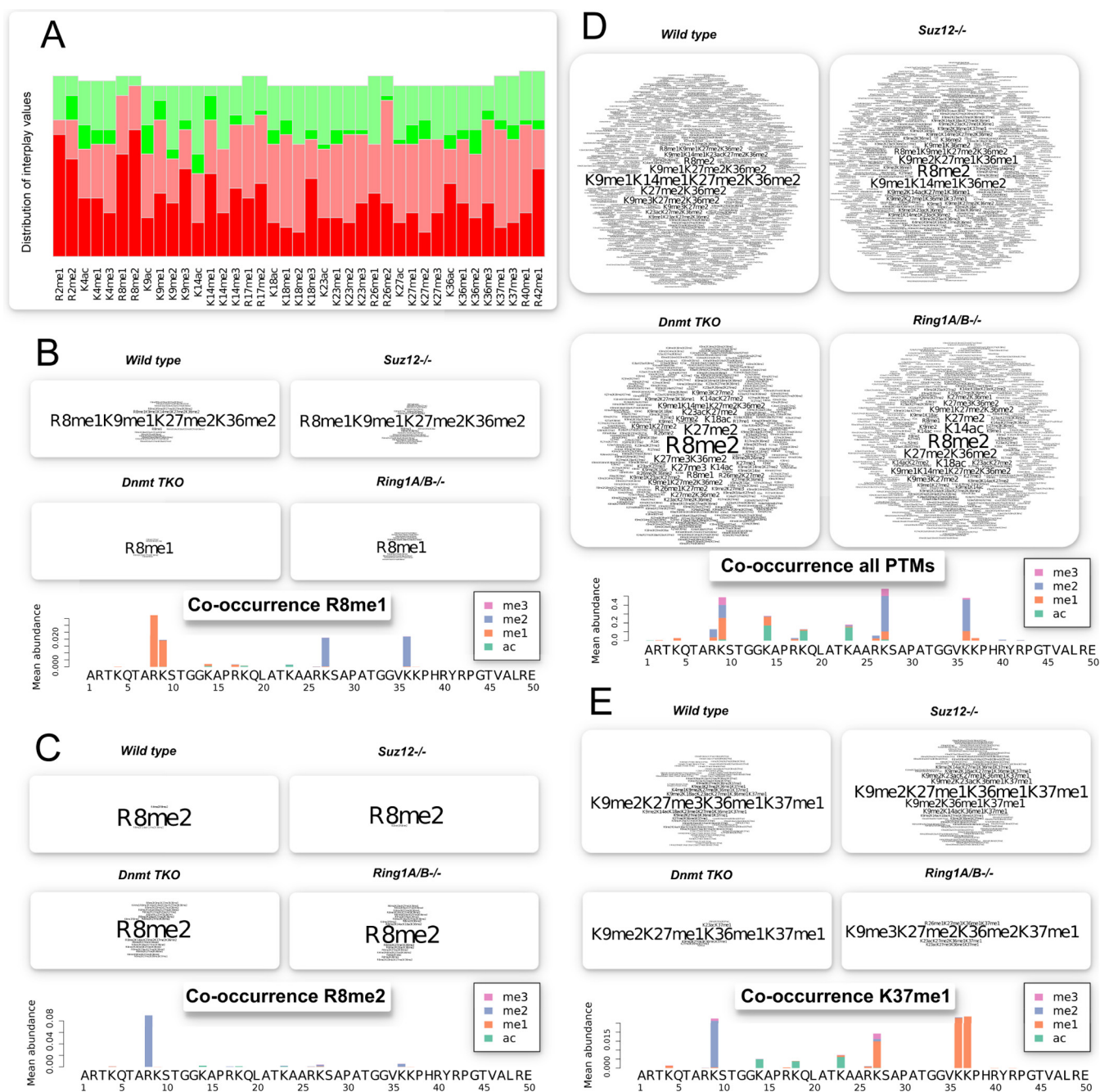


FIG. 7. H3 PTMs with distinctive crosstalk have highly specific PTM codes. Details on interplay values, peptide isoforms and cooccurring PTMs of selected marks provide information about their role in setting the global epigenetic landscape. *A*, Distribution of interplay values of the PTM for wild type cells. Red and green colors denote negative and positive interplay values. Darker colors correspond to interplay values that are conserved in the 4 cell lines. R2me1 and R8 methylations show highly negative crosstalk. *B–E*, Relative abundance of peptides of the four cell lines containing R8me1, R8me2, all PTMs and K37me1, respectively. Upper panels depict a word cloud with all peptides containing these marks. Word size corresponds to abundance. Lower panels present the distribution of PTMs coexpressed with the particular PTM along H3 tails. The height of the bars corresponds to PTM abundance averaged over the four cell lines.

DISCUSSION

The presented systems level study characterized histone PTM levels in cell lines perturbed in the epigenetic machinery by knocking out proteins controlling the deposition of particular histone PTMs. After integrating our data with RNA-seq

experiments we observed that the rearrangements of the PTM levels and changes of the expression levels of their enzymes only coincide for mono-methylated marks and K9 methylations. This suggests that fine tuning or modulation of the levels of hypermethylated marks, *i.e.* marks with the most

pronounced biological effects, is regulated not simply by enzyme abundance but also by other processes. Local intracellular availability of the substrate S-adenosyl methionine (SAM) might influence the methylation levels of proteins by histone methyltransferase enzymes. However, a variety of up- and downregulated methylation states were detected in our data set and a simple conclusion on the effects of SAM levels cannot be made. Nevertheless, we inspected transcriptomics data set for mRNA expression levels corresponding to enzymes involved in SAM production. Interestingly, the enzyme *Mat2a* exhibited a higher expression level in *Dnmt* TKO and *Ring1A/B*^{-/-} cells as compared with the wild type (supplemental Fig. S19). However, the transcriptomics data analysis included clusters with downregulated methylation levels (Fig. 3, cluster 3) and up-regulated methylation levels (Fig. 3, cluster 4) in *Dnmt* TKO and *Ring1A/B*^{-/-} cells, so no definitive conclusions on the role of SAM and these enzymes can be drawn.

Our middle-down proteomics approach distinguishes positive and negative crosstalk for all the measured histone H3 PTMs, given the number of quantified peptide isoforms is sufficiently large for statistical assessment. Hence, the workflow provides a method to quantify crosstalk between PTMs on a large scale. This approach is intrinsically more direct as compared with other strategies such as measuring peptide coexistence (45), gene mutation (46) or investigation of structural proximity and evolutionary conservation of sequences (41). These latter approaches either cannot distinguish different methylation states or assess the cofrequency of the PTMs within the same peptide. Because the main aim of this study is to reveal the conserved cross-talk patterns across different cell lines we used the different conditions to reveal strong features of conserved PTM cross-talk because we intrinsically have more fluctuation of single PTM abundances in individual cell states than in multiple perturbed states (quadruplicate analysis) of the same cell lines.

Fine-structure of PTM Crosstalk—Global patterns of PTM crosstalk determined in the four mESC lines show strong negative crosstalk, *i.e.* strong mutual exclusive behavior of most me2/me3 methylation states and frequent positive crosstalk between adjacent or near-proximity lysine acetylation and mono-methylation sites. We observed crosstalk relationships between K9, K27 and K36 methylation events that were previously demonstrated in the investigated cell lines (21, 47, 48). Mono-methylation events were found to be highly cooperative and may form the core structure and starting points for the rearrangement of chromatin regions into regions with more specific function given by higher methylation states. Especially the observation of positive crosstalk between neighboring mono-methylated lysine residues suggests that these marks might be set by unspecific methyltransferases.

Di- and tri-methylated states of K9 and K27 are known to define facultative and constitutive heterochromatin where most genes are silenced. Our study suggests an even more complex scenario involving additional marks and PTM states.

Although K9me2K27me2 and K9me3K27me3 represent highly forbidden combinations, the two binary marks K9me2K27me3 and K9me3K27me2 appear with high frequency and the constituting PTMs exhibit positive crosstalk. Such observations are of relevance in order to understand enzyme-enzyme interactions for histone modifications, as recently discovered. For instance, PRC2 involved in K27 methylation, cooperates with HP1 and G9a enzymes responsible for the reading of and catalysis of K9 methylation, respectively (10, 49). Given the highly mutually exclusive nature of most H3 lysine tri-methylation marks, we argue for a classification of chromatin regions by tri-methylation states of the different lysine residues, including hitherto neglected PTMs such as K14me3 and K23me3.

H3 K14, K18, and K23 acetylation events exhibited positive interplay or nonconserved negative interplay between each other and toward other activating marks (K4me), whereas K9ac and K27ac were involved in more complex PTM relationships by exhibiting negative crosstalk with subgroups of activating marks. H3K9ac is indeed a multifunctional PTM; for instance, it is up-regulated in ESCs as compared with cells initiating differentiation, but H3K9ac also increases during neural differentiation to activate neurodevelopmental genes (50). Much needs to be revealed regarding H3K9ac crosstalk with nearby and more distant marks in H3. H3K27ac is a widely studied histone mark, known to be enriched in the promoter regions of active genes (51) and identifies active enhancers by colocalizing with K4me1 as compared with poised enhancers, where only K4me1 is present (52).

A Focus on Arginine Methylation—We detected various types of H3 tail molecules carrying arginine methylation marks. Arg-methylation can be of comparable abundance to lysine modifications, constituting up to 10% of H3 histones (R8me2) and more than 1% for R8, R17 and R26 methylation events. However, our data on arginine methylation is less extensive as data on lysine methylation events. This is probably because of their lower abundance, but also because the standardized MS-based strategies, although highly efficient, are biased toward investigation of lysine modifications (53). Moreover, arginine dimethylation adds an additional layer of complexity, as it exists in symmetric and asymmetric configurations which are difficult to distinguish using standard analytical workflows. We show that arginine methylation interacts with other PTMs, indicating that Arg methylations play important roles in histone function.

R2me1, R2me2 and R8me2 represent PTMs that do not co-occur with other PTMs in the H3 tail. Histones carrying these marks are likely to fulfill specific tasks. For instance, R2 methylations were found to antagonize K4 tri-methylation (54). R8me2 was reported to antagonize deposition of K9 methylation (55), K4me3 (56), K9ac and K14ac (57). PTMs on the outer H3 tail appear to have well-defined crosstalk patterns and we propose that these PTMs have a strong impact on biological function by mediating access of enzymes to write PTMs further

up the histone sequence. We hypothesize that R2 and R8 methylation states are indicative of distinct chromatin states.

We observed several combinatorial marks that seemed specific and may convey distinct biological functions. R8me1 occurred either as a single PTM or in combination with K9me1K27me2K36me2. Observation of negative crosstalk between R8me1/me2 and K9 di- and tri-methylations (55, 58) support our findings of high co-occurrence frequency of R8me1/me2 with K9 mono-methylation as compared with K9me2/me3. R26me2 was frequently found in combination with one di-methylated lysine in the H3 tail. In this case, the PTM could be either a side-product of the enzymes setting the di-methyl marks on lysines 9, 27, and 36 or marker of a specific state involving all four PTM marks.

In conclusion, our optimized middle-down proteomics workflow combined with computational and statistical tools enables the detailed characterization of histone PTM marks, their interplay and crosstalk. Our study of histone H3 in four mouse stem cell lines confirms known crosstalk features involving acetylated and methylated lysine residues in H3 and supports the hypothesis that different histone methylation states have distinct functions. Investigation of potential novel functional PTM marks, such as arginine methylation and K37 methylation, revealed specific patterns of coexisting PTMs. To our knowledge, this is the first time that single, exclusive R2me1/me2 and R8me1 marks are reported in histone H3 tails, *i.e.* in the absence of other detectable PTMs. With additional data sets on hand it will be possible to obtain insights into crosstalk patterns on the levels of cell lines, within species or even among similar organisms. This raises the possibility to use PTM cofrequencies and interplay values as “biomarkers” to classify specific cell states, dissect perturbations or disease mechanisms. Our middle-down proteomics strategy and associated computational tools are suitable for characterization of PTM crosstalk in protein domains and in small intact proteins, and can be extended to top-down protein analysis by mass spectrometry. As mass spectrometers can now distinguish many proteoforms originating from a single gene we expect that characterization of PTM marks, interplay and crosstalk, will become a standard approach in functional proteomics and biomarker research.

Acknowledgments—We thank Miguel Vidal and Haruhiko Koseki for the *Ring1A/B*^{-/-} mouse ESCs, and Masaki Okano for the Dnmt1 KO mouse ES cells.

* VS was funded by The Danish Council for Independent Research and the EU ELIXIR consortium (Danish ELIXIR node). SS and ONJ acknowledge generous financial support from the Danish Council for Research, Natural Sciences (FNU) and support from the VILLUM Foundation to the Center for Bioanalytical Sciences. The Center for Epigenetics at University of Copenhagen (K.H.) and the University of Southern Denmark (O.N.J.) is financed by a generous grant from the Danish National Research Foundation (DNRF82). KH would also like to thank the European Research Council (294666_DNAME1), and the Novo Nordisk Foundation for financial support.

§ This article contains supplemental material.

|| To whom correspondence should be addressed: University of Southern Denmark, Campusvej 55, Odense 5230 Denmark. Tel.: +45 6550 2483; Fax: +45 6550 2467; E-mail: veits@bmb.sdu.dk.

** These authors contributed equally to this work.

§§ Current address: Epigenetics Program, Department of Biochemistry and Biophysics, Perelman Medicine, University of Pennsylvania, Philadelphia, PA 19104.

‡‡ Current address: Department of Cell Biology, Tianjin Medical University, Tianjin 300070, China.

REFERENCES

- Kouzarides, T. (2007) Chromatin modifications and their function. *Cell* **128**, 693–705
- Bannister, A. J., and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–395
- Fischle, W., Wang, Y., and Allis, C. D. (2003) Histone and chromatin cross-talk. *Current Opinion Cell Biol.* **15**, 172–183
- Lee, J. S., Smith, E., and Shilatifard, A. (2010) The language of histone crosstalk. *Cell* **142**, 682–685
- Hunter, T. (2007) The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol. Cell* **28**, 730–738
- Simboeck, E., Sawicka, A., Zupkovitz, G., Senese, S., Winter, S., Dequiedt, F., Ogris, E., Di Croce, L., Chiocca, S., and Seiser, C. (2010) A phosphorylation switch regulates the transcriptional activation of cell cycle regulator p21 by histone deacetylase inhibitors. *J. Biol. Chem.* **285**, 41062–41073
- Vermeulen, M., Eberl, H. C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K. K., Olsen, J. V., Hyman, A. A., Stunnenberg, H. G., and Mann, M. (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* **142**, 967–980
- Hirota, T., Lipp, J. J., Toh, B. H., and Peters, J. M. (2005) Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438**, 1176–1180
- Xhemalce, B., and Kouzarides, T. (2010) A chromodomain switch mediated by histone H3 Lys 4 acetylation regulates heterochromatin assembly. *Genes Development* **24**, 647–652
- Mozzetta, C., Pontis, J., Fritsch, L., Robin, P., Portoso, M., Proux, C., Margueron, R., and Ait-Si-Ali, S. (2014) The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. *Mol. Cell* **53**, 277–289
- Margueron, R., and Reinberg, D. (2011) The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343–349
- Heard, E. (2005) Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome. *Current Opinion Genetics Development* **15**, 482–489
- Murphy, P. J., Cipriani, B. R., Wallin, C. B., Ju, C. Y., Szeto, K., Hagarman, J. A., Benitez, J. J., Craighead, H. G., and Soloway, P. D. (2013) Single-molecule analysis of combinatorial epigenomic states in normal and tumor cells. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 7772–7777
- Pasini, D., Bracken, A. P., Jensen, M. R., Lazzarini Denchi, E., and Helin, K. (2004) Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J.* **23**, 4061–4071
- de Napoles, M., Mermoud, J. E., Wakao, R., Tang, Y. A., Endoh, M., Appanah, R., Nesterova, T. B., Silva, J., Otte, A. P., Vidal, M., Koseki, H., and Brockdorff, N. (2004) Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Developmental Cell* **7**, 663–676
- Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H. R., Nakayama, J., and Okano, M. (2006) Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* **11**, 805–814
- Sidoli, S., Cheng, L., and Jensen, O. N. (2012) Proteomics in chromatin biology and epigenetics: Elucidation of post-translational modifications of histone proteins by mass spectrometry. *J. Proteomics* **75**, 3419–3433
- Walther, T. C., and Mann, M. (2010) Mass spectrometry-based proteomics in cell biology. *J. Cell Biol.* **190**, 491–500
- Jung, H. R., Sidoli, S., Haldbø, S., Sprenger, R. R., Schwammle, V., Pasini, D., Helin, K., and Jensen, O. N. (2013) Precision mapping of coexisting

- modifications in histone H3 tails from embryonic stem cells by ETD-MS/MS. *Anal. Chem.* **85**, 8232–8239
20. Mikesch, L. M., Ueberheide, B., Chi, A., Coon, J. J., Syka, J. E., Shabanowitz, J., and Hunt, D. F. (2006) The utility of ETD mass spectrometry in proteomic analysis. *Biochim. Biophys. Acta* **1764**, 1811–1822
 21. Schwammle, V., Aspalter, C. M., Sidoli, S., and Jensen, O. N. (2014) Large scale analysis of co-existing post-translational modifications in histone tails reveals global fine structure of cross-talk. *Mol. Cell. Proteomics* **13**, 1855–1865
 22. Pasini, D., Bracken, A. P., Hansen, J. B., Capillo, M., and Helin, K. (2007) The polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Mol. Cell. Biol.* **27**, 3769–3779
 23. Endoh, M., Endo, T. A., Endoh, T., Fujimura, Y., Ohara, O., Toyoda, T., Otte, A. P., Okano, M., Brockdorff, N., Vidal, M., and Koseki, H. (2008) Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. *Development* **135**, 1513–1524
 24. Sakaue, M., Ohta, H., Kumaki, Y., Oda, M., Sakaide, Y., Matsuoka, C., Yamagiwa, A., Niwa, H., Wakayama, T., and Okano, M. (2010) DNA methylation is dispensable for the growth and survival of the extraembryonic lineages. *Curr. Biol.* **20**, 1452–1457
 25. Vaidyanathan, G., and Zalutsky, M. R. (2007) Synthesis of N-succinimidyl 4-guanidinomethyl-3-[¹²⁵I]iodobenzoate: a radio-iodination agent for labeling internalizing proteins and peptides. *Nat. Protoc.* **2**, 282–286
 26. Sidoli, S., Schwammle, V., Ruminowicz, C., Hansen, T. A., Wu, X., Helin, K., and Jensen, O. N. (2014) Middle-down hybrid chromatography/tandem mass spectrometry workflow for characterization of combinatorial post-translational modifications in histones. *Proteomics* **14**, 2200–2211
 27. Riising, E. M., Comet, I., Leblanc, B., Wu, X., Johansen, J. V., and Helin, K. (2014) Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. *Mol. Cell* **55**, 347–360
 28. Li, Z., Dai, H., Martos, S. N., Xu, B., Gao, Y., Li, T., Zhu, G., Schones, D. E., and Wang, Z. (2015) Distinct roles of DNMT1-dependent and DNMT1-independent methylation patterns in the genome of mouse embryonic stem cells. *Genome Biol.* **16**, 115
 29. Bezdek, J. C. (1981) *Pattern recognition with fuzzy objective function algorithms*, Plenum Press, New York
 30. Schwammle, V., and Jensen, O. N. (2010) A simple and fast method to determine the parameters for fuzzy c-means cluster analysis. *Bioinformatics* **26**, 2841–2848
 31. Khare, S. P., Habib, F., Sharma, R., Gadwal, N., Gupta, S., and Galande, S. (2012) Histome—a relational knowledgebase of human histone proteins and histone modifying enzymes. *Nucleic Acids Res.* **40**, D337–D342
 32. Zhao, Y., and Garcia, B. A. (2015) Comprehensive Catalog of Currently Documented Histone Modifications. *Cold Spring Harb. Perspect. Biol.* **7**, a025064
 33. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57
 34. Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287
 35. Laphanasupkul, P., Feng, J., Mantesso, A., Takada-Horisawa, Y., Vidal, M., Koseki, H., Wang, L., An, Z., Miletich, I., and Sharpe, P. T. (2012) Ring1a/b polycomb proteins regulate the mesenchymal stem cell niche in continuously growing incisors. *Dev. Biol.* **367**, 140–153
 36. Leitch, H. G., McEwen, K. R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J. G., Smith, A., Surani, M. A., and Hajkova, P. (2013) Naive pluripotency is associated with global DNA hypomethylation. *Nat. Structural Mol. Biol.* **20**, 311–316
 37. Shechter, D., Dormann, H. L., Allis, C. D., and Hake, S. B. (2007) Extraction, purification and analysis of histones. *Nat. Protoc.* **2**, 1445–1457
 38. 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
 39. Pesavento, J. J., Mizzen, C. A., and Kelleher, N. L. (2006) Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: human histone H4. *Anal. Chem.* **78**, 4271–4280
 40. LeDuc, R. D., Fellers, R. T., Early, B. P., Greer, J. B., Thomas, P. M., and Kelleher, N. L. (2014) The C-score: a Bayesian framework to sharply improve proteoform scoring in high-throughput top down proteomics. *J. Proteome Res.* **13**, 3231–3240
 41. Minguez, P., Letunic, I., Parca, L., Garcia-Alonso, L., Dopazo, J., Huerta-Cepas, J., and Bork, P. (2015) PTMcode v2: a resource for functional associations of post-translational modifications within and between proteins. *Nucleic Acids Res.* **43**, D494–D502
 42. Rathert, P., Dhayalan, A., Murakami, M., Zhang, X., Tamas, R., Jurkowska, R., Komatsu, Y., Shinkai, Y., Cheng, X., and Jeltsch, A. (2008) Protein lysine methyltransferase G9a acts on non-histone targets. *Nat. Chem. Biol.* **4**, 344–346
 43. Bedford, M. T., and Richard, S. (2005) Arginine methylation an emerging regulator of protein function. *Mol. Cell* **18**, 263–272
 44. Huang, H., Sabari, B. R., Garcia, B. A., Allis, C. D., and Zhao, Y. (2014) SnapShot: histone modifications. *Cell* **159**, 458–458 e451
 45. Edwards, A. V., Edwards, G. J., Schwammle, V., Saxtorph, H., and Larsen, M. R. (2014) Spatial and temporal effects in protein post-translational modification distributions in the developing mouse brain. *J. Proteome Res.* **13**, 260–267
 46. Guan, X., Rastogi, N., Parthun, M. R., and Freitas, M. A. (2013) Discovery of histone modification crosstalk networks by stable isotope labeling of amino acids in cell culture mass spectrometry (SILAC MS). *Mol. Cell. Proteomics* **12**, 2048–2059
 47. Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. (2011) H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem.* **286**, 7983–7989
 48. Chandra, T., Kirschner, K., Thuret, J. Y., Pope, B. D., Ryba, T., Newman, S., Ahmed, K., Samarajiwa, S. A., Salama, R., Carroll, T., Stark, R., Janky, R., Narita, M., Xue, L., Chicas, A., Nunez, S., Janknecht, R., Hayashi-Takanaka, Y., Wilson, M. D., Marshall, A., Odum, D. T., Babu, M. M., Bazett-Jones, D. P., Tavare, S., Edwards, P. A., Lowe, S. W., Kimura, H., and Gilbert, D. M. (2012) Independence of repressive histone marks and chromatin compaction during senescent heterochromatic layer formation. *Mol. Cell* **47**, 203–214
 49. Boros, J., Arnoult, N., Stroobant, V., Collet, J. F., and Decottignies, A. (2014) Polycomb repressive complex 2 and H3K27me3 cooperate with H3K9 methylation to maintain heterochromatin protein 1alpha at chromatin. *Mol. Cell. Biol.* **34**, 3662–3674
 50. Qiao, Y., Wang, R., Yang, X., Tang, K., and Jing, N. (2015) Dual roles of histone H3 lysine 9 acetylation in human embryonic stem cell pluripotency and neural differentiation. *J. Biol. Chem.* **290**, 9949
 51. Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Peng, W., Zhang, M. Q., and Zhao, K. (2008) Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Gen.* **40**, 897–903
 52. Creighton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A., Boyer, L. A., Young, R. A., and Jaenisch, R. (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 21931–21936
 53. Garcia, B. A., Mollah, S., Ueberheide, B. M., Busby, S. A., Muratore, T. L., Shabanowitz, J., and Hunt, D. F. (2007) Chemical derivatization of histones for facilitated analysis by mass spectrometry. *Nat. Protoc.* **2**, 933–938
 54. Hyllus, D., Stein, C., Schnabel, K., Schiltz, E., Imhof, A., Dou, Y., Hsieh, J., and Bauer, U. M. (2007) PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation. *Genes Develop.* **21**, 3369–3380
 55. Daujat, S., Bauer, U. M., Shah, V., Turner, B., Berger, S., and Kouzarides, T. (2002) Crosstalk between CARM1 methylation and CBP acetylation on histone H3. *Curr. Biol.* **12**, 2090–2097
 56. Flanagan, J. F., Mi, L. Z., Chruszcz, M., Cymborowski, M., Clines, K. L., Kim, Y., Minor, W., Rastinejad, F., and Khorasanizadeh, S. (2005) Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* **438**, 1181–1185
 57. Rothbart, S. B., Krajewski, K., Nady, N., Tempel, W., Xue, S., Badeaux, A. I., Barsyte-Lovejoy, D., Martinez, J. Y., Bedford, M. T., Fuchs, S. M., Arrowsmith, C. H., and Strahl, B. D. (2012) Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat. Structural Mol. Biol.* **19**, 1155–1160
 58. An, W., Kim, J., and Roeder, R. G. (2004) Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53. *Cell* **117**, 735–748