

# Quantitative Interaction Proteomics and Genome-wide Profiling of Epigenetic Histone Marks and Their Readers

Michiel Vermeulen,<sup>1,6,7,\*</sup> H. Christian Eberl,<sup>1,6</sup> Filomena Matarese,<sup>2,6</sup> Hendrik Marks,<sup>2</sup> Sergei Denisov,<sup>2</sup> Falk Butter,<sup>1</sup> Kenneth K. Lee,<sup>3</sup> Jesper V. Olsen,<sup>1,5</sup> Anthony A. Hyman,<sup>4</sup> Henk G. Stunnenberg,<sup>2,\*</sup> and Matthias Mann<sup>1,\*</sup>

<sup>1</sup>Department of Proteomics and Signal Transduction, Max-Planck-Institute of Biochemistry, D-82152 Martinsried, Germany

<sup>2</sup>Department of Molecular Biology, Faculty of Science, Nijmegen Centre for Molecular Life Sciences (NCMLS), Radboud University Nijmegen, Geert Grooteplein 26 Zuid, 6525 GA Nijmegen, The Netherlands

<sup>3</sup>Stowers Institute for Medical Research, 1000 East 50th Street, Kansas City, MO 64110, USA

<sup>4</sup>Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstrasse 108, 01307 Dresden, Germany

<sup>5</sup>Novo Nordisk Foundation Center for Protein Research, Faculty of Health Sciences, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen, Denmark

<sup>6</sup>These authors contributed equally to this work

<sup>7</sup>Present address: Department of Physiological Chemistry and Cancer Genomics Centre, University Medical Center Utrecht, Utrecht, The Netherlands

\*Correspondence: m.vermeulen-3@umcutrecht.nl (M.V.), h.stunnenberg@ncmls.ru.nl (H.G.S.), mmann@biochem.mpg.de (M.M.)

DOI 10.1016/j.cell.2010.08.020

## SUMMARY

Trimethyl-lysine (me3) modifications on histones are the most stable epigenetic marks and they control chromatin-mediated regulation of gene expression. Here, we determine proteins that bind these marks by high-accuracy, quantitative mass spectrometry. These chromatin “readers” are assigned to complexes by interaction proteomics of full-length BAC-GFP-tagged proteins. ChIP-Seq profiling identifies their genomic binding sites, revealing functional properties. Among the main findings, the human SAGA complex binds to H3K4me3 via a double Tudor-domain in the C terminus of Sgf29, and the PWWP domain is identified as a putative H3K36me3 binding motif. The ORC complex, including LRWD1, binds to the three most prominent transcriptional repressive lysine methylation sites. Our data reveal a highly adapted interplay between chromatin marks and their associated protein complexes. Reading specific trimethyl-lysine sites by specialized complexes appears to be a widespread mechanism to mediate gene expression.

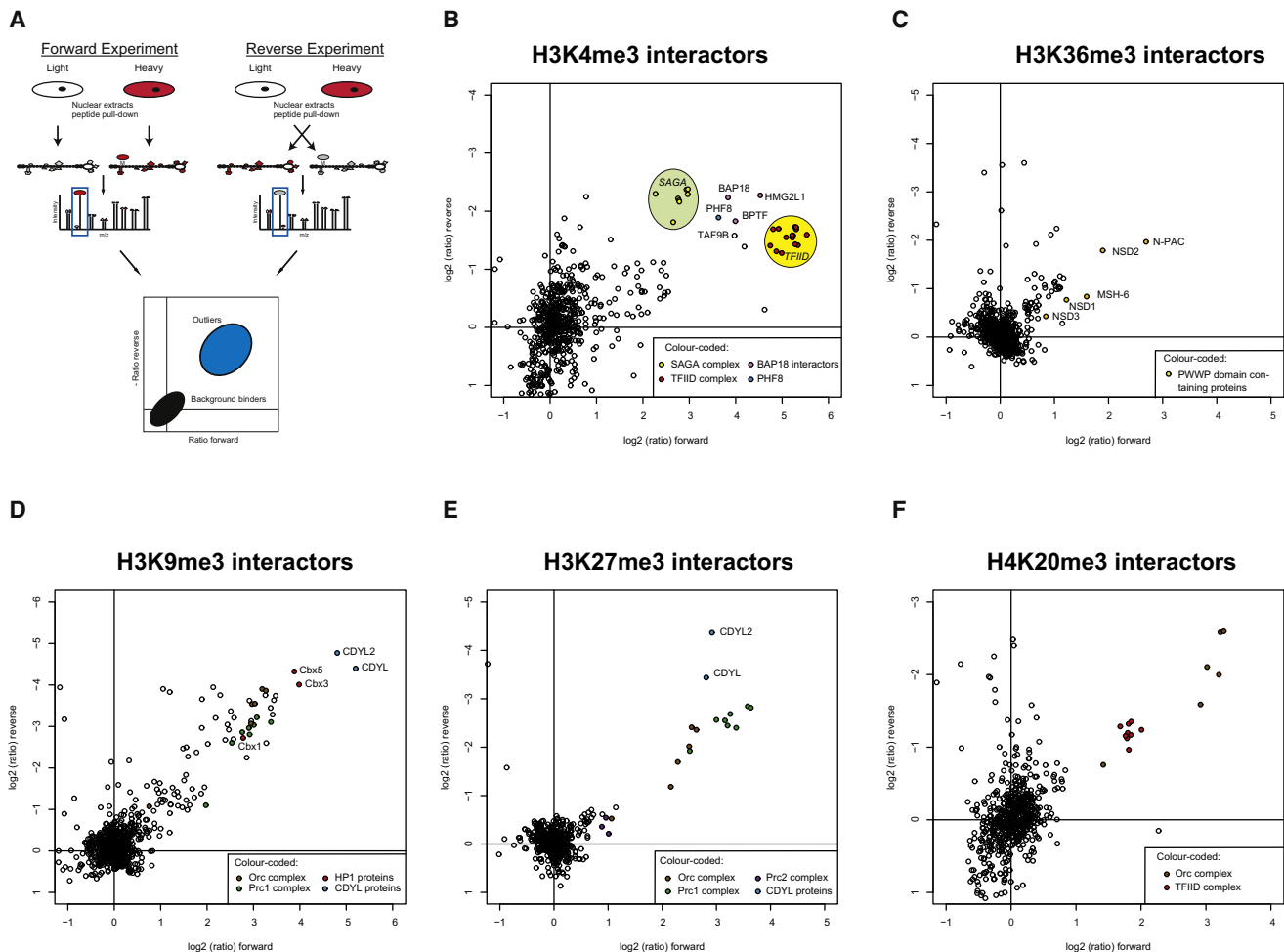
## INTRODUCTION

In the eukaryotic nucleus, DNA is wrapped around an octamer of histone proteins, which constitute the nucleosomes. Rather than merely serving as a means to store genetic material, nucleosomes play an active role in regulating processes such as transcription, DNA repair, and apoptosis. The N-terminal tails of the four core histones that protrude from the core structure of the nucleosome are subject to a variety of posttranslational

modifications such as acetylation, methylation, and phosphorylation. One role of these modifications is the recruitment of regulatory proteins that in turn exert their function on chromatin (Jenuwein and Allis, 2001; Kouzarides, 2007).

The major lysine methylation sites on the N terminus of histone H3 and histone H4 with a clearly defined biological function are H3K4me3, H3K9me3, H3K27me3, H3K36me3, and H4K20me3, which are associated with different functional states of chromatin. H3K4me3 is almost exclusively found on promoter regions of actively transcribed genes while H3K36me3 is linked to transcription elongation. H3K9me3, H3K27me3, and H4K20me3 are generally found on silent heterochromatic regions of the genome. Part of the functional distinction between these methylation sites relates to the proteins interacting with them. A number of these “chromatin readers” for various histone methyl lysine sites have already been identified and characterized (Kouzarides, 2007; Shilatifard, 2006; Taverna et al., 2007), but this list is unlikely to be exhaustive. To obtain a comprehensive map of the histone methyl lysine interactome, unbiased screening methods are required.

Mass spectrometry (MS)-based proteomics is increasingly used in functional biological studies and has proved to be a powerful tool to characterize histone modifications (Garcia et al., 2007; Vermeulen and Selbach, 2009). For protein-protein interactions a quantitative format is desirable, as this enables to distinguish specific and background binders (Vermeulen et al., 2008). In particular, the technology of stable isotope labeling by amino acids in cell culture (SILAC) (Ong et al., 2002) can be used to expose peptide baits bearing a posttranslational modification to “heavy” SILAC-labeled cell extracts, whereas the unmodified peptide is exposed to “light” labeled cell extract. Binders specific to the modified form of the peptide appear in mass spectra with a significant ratio between heavy and light form of the protein. Using this approach, we discovered that TFIID binds to H3K4me3, thereby providing a link between



**Figure 1. A Histone Peptide Pulldown Approach Using SILAC Technology**

(A) Schematic representation of the experimental approach (M indicates trimethyl lysine).

(B) The H3K4me3 interactome. Proteins are plotted by their SILAC-ratios in the forward (x axis) and reverse (y axis) SILAC experiment. Specific interactors should lie close to the diagonal in the upper right quadrant. The two major transcriptional coactivator complexes that were found to interact with this mark (TFIIID and SAGA) are circled. TAF9b, which is localized between TFIIID and SAGA in the figure, is a shared subunit between these two complexes.

(C) The H3K36me3 interactome. Proteins carrying a PWWP domain are colored yellow.

(D–F) The interactome of H3K9me3, H3K27me3 and H4K20me3, respectively. Note that the ORC complex, including LRWD1, binds to these three marks.

See also Figure S1 and Table S1.

this modification and activation of transcription (Vermeulen et al., 2007).

Here, we refine this technology and perform an unbiased interaction screen for the known activating and repressive trimethyl histone marks on H3 and H4. We apply the BAC-GFP transgeneOmics technology (Poser et al., 2008) to characterize chromatin readers and their complexes. Chromatin immunoprecipitation followed by massive parallel sequencing (ChIP-Seq) with the same BAC-GFP lines identifies the *in vivo* target genes, which are found to overlap with the histone marks they interact with. This integrative approach provides not only an interactome of the studied histone marks, including many previously uncharacterized factors, but also mechanistic insights into epigenetic regulation of gene expression.

## RESULTS

### A Large-Scale Methyl Lysine Interactome

To characterize the interactome of trimethyl-lysine chromatin marks, we developed an interaction screen based on a recently described technology (Vermeulen et al., 2007). In brief, nuclear extracts derived from HeLaS3 cells grown in “light” or “heavy” medium were incubated with immobilized biotinylated histone peptides (Figure 1A). After incubation, beads from both pull-downs were pooled, run on a one-dimensional PAGE gel, and subjected to in-gel trypsin digestion. The resulting peptide mixtures were measured by high-resolution on-line electrospray MS on a hybrid linear ion trap, Orbitrap (see Experimental Procedures). Computational analysis was done with the MaxQuant

algorithms (Cox and Mann, 2008), which enabled sub parts-per-million mass assignment and accurate quantitation even for very low abundance SILAC pairs. Eluates from methylated and non-methylated peptides each contained hundreds of proteins and are visually indistinguishable on 1D gels (Figure S1A available online). Nevertheless, the SILAC-ratios reliably retrieved specific binders even when they were hundred-fold less abundant than background binders (Figure S1B). We determined the interactome of the two activating marks H3K4me3 and H3K36me3 and three repressive marks, H3K9me3, H3K27me3, and H4K20me3 (Table S1; Figures 1B–1F). Each measurement identified between 600 and 1200 proteins at a confidence level of 99%. Of these, between 10 and 60 had highly significant ratios indicating specific binding to the respective marks.

In our previous study, we identified interactions of members of the TFIID complex with H3K4me3. Here, we performed the interaction screen in the “forward” and “reverse” format to obtain higher discrimination between specific baits and background. The forward experiment consists of incubating the modified peptide with heavy labeled cell lysate and the nonmodified peptide with light labeled cell lysate, whereas in the reverse format the labels are switched. These two experiments also constitute a biological replicate. With a minimum of two quantification events, every significant interactor is supported by at least four quantitative measurements. Plotting interaction data for H3K4me3 in a two-dimensional space and inverting the SILAC-ratios of the reverse experiment places the true interactors into the top right quadrant (Figure 1A). Nonlabeled contaminants, such as keratin and proteins derived from the medium will not change the ratio in the reverse experiment and are located in other quadrants. Furthermore, a number of other proteins, such as polypyrimidine tract-binding protein 2, were automatically filtered out because they show significant ratios only in one of the labeling experiments, and are color coded accordingly in Table S1. In some cases, interactions may be biophysically correct but they may not occur in vivo because of compartmentalization in the cell (for example, mitochondrial hsp60 binding to H3K9me3). We noticed that the entire TFIID protein complex clustered together in the two-dimensional plot, indicating very similar SILAC ratios in the forward and reverse experiments (Figure 1B). This prompted us to inspect the interaction plots for other protein complexes binding to specific chromatin marks.

### Sgf29 Links the Human SAGA Complex to H3K4me3

The measured H3K4me3 interactome contained eight subunits of the human SAGA complex, which tightly clustered together in the two-dimensional plot (green circle in Figure 1B). Inspection of the sequences of all known SAGA subunits revealed a double Tudor domain in the C terminus of Sgf29 (Figure 2A). Double Tudor domains are known to have affinity for H3K4me3 (Huang et al., 2006). We therefore speculated that Sgf29 could be the subunit within the SAGA complex that directly binds to H3K4me3. To address this question, we used RNAi to knock down Sgf29 in HeLa cells (Figure 2B). The nuclear extracts derived from these cells as well as nuclear extracts derived from cells transfected with control oligonucleotides were used for peptide pull-downs. Western blotting shows that the SAGA subunit GCN5 only binds to H3K4me3 and not to H3K4me0

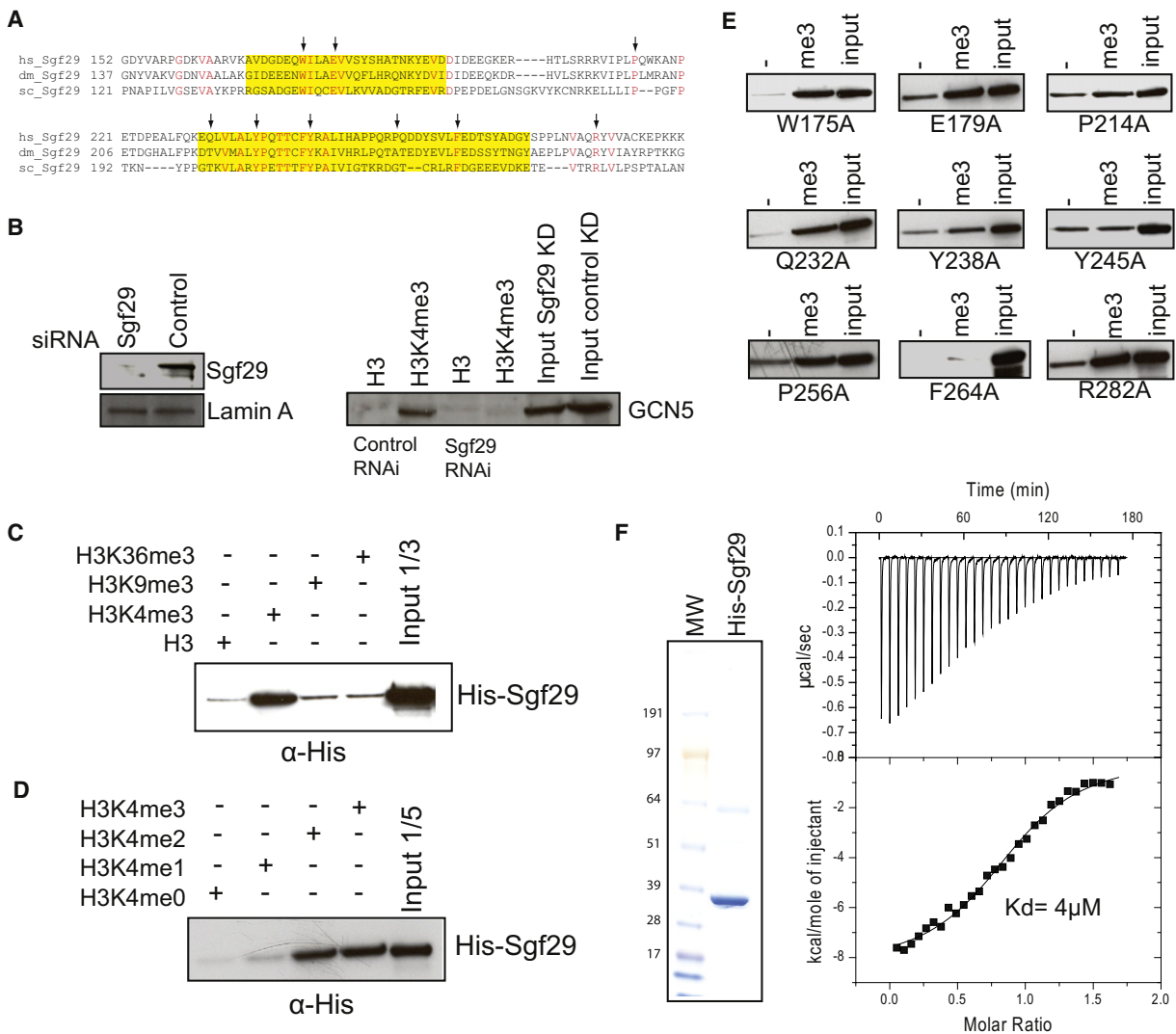
(Figure 2B). This binding is abolished upon knockdown of Sgf29, while GCN5 levels in these cells are similar to those in the cells treated with mock siRNA. These experiments also imply that, at least in mammalian cells, Sgf29 is responsible for the observed interaction between H3K4me3 and SAGA, and not CHD1, as has been suggested in yeast (Pray-Grant et al., 2005).

To biophysically characterize this interaction, we expressed Sgf29 as a recombinant protein in *E. coli* and used the induced bacterial lysates for histone peptide pull-downs. As shown in Figure 2C, Sgf29 binds to histone H3 peptides, with a clear preference for H3K4me3. This binding is specific as no interaction with other histone lysine methylation sites such as H3K9me3 or H3K36me3 was observed. Sgf29 binds to both H3K4me2 and H3K4me3 with a slight preference for H3K4me3 (Figure 2D). Based on sequence alignments between yeast, *Drosophila* and human Sgf29 we selected conserved and nonconserved residues for mutational analyses (Figure 2A). Results of nine pull-down experiments revealed that conserved residues in the second Tudor domain of Sgf29 are particularly important for H3K4me3 binding. As expected, mutating nonconserved residues did not affect the binding (Figure 2E). We used isothermal calorimetry experiments to measure the affinity of the interaction between Sgf29 and H3K4me3 (Figure 2F). The binding constant of 4  $\mu\text{M}$  is comparable to that of other trimethyl-lysine marks to their readers and in particular to the interaction constant of the Tudor domain of JMJD2A, which is 10  $\mu\text{M}$  (Huang et al., 2006). No affinity between Sgf29 and the unmethylated histone H3 peptide could be observed. Together, these results demonstrate that the human SAGA complex binds to H3K4me3 and that the double Tudor domain in its subunit Sgf29 is both necessary and sufficient to mediate this interaction.

### Functional Insights into Chromatin Readers Using BAC transgeneOmics

Our screening of the H3K4me3 and H3K36me3 interactome, two lysine methylations associated with actively transcribed genes, revealed a large number of chromatin readers of unknown function. To gain insight into the molecular mechanism of their interaction with the lysine methylation sites, we tagged a selection of these proteins with GFP using the recently developed BAC transgeneOmics technology (Poser et al., 2008). In this strategy, a GFP-tagged fusion of the protein of interest is stably integrated preserving the endogenous genomic context—in HeLa cells by recombineering (Zhang et al., 1998). Fusion proteins are therefore expressed at near endogenous levels, as demonstrated previously (Poser et al., 2008). Furthermore, we tested expression levels of several of the GFP-tagged BAC lines and found very similar expression levels to the endogenous proteins (Figures S2E–S2H).

Quantitative SILAC-based GFP pull-downs employing wild-type parental cells as control were optimized such that protein complexes can be identified and visualized in a single two hour MS analysis without the need to separate proteins on an SDS PAGE gel (Hubner et al., 2010). As a proof of principle we applied this workflow to the K4me3 binding protein Sgf29, which is known to assemble into either the SAGA or the ATAC complex (Nagy et al., 2010). Both SAGA and ATAC complex subunits copurified with GFP-Sgf29 demonstrating the applicability of single



**Figure 2. Sgf29 Links the SAGA Complex to H3K4me3**

(A) Alignment of the C-terminal part of human, Drosophila, and yeast Sgf29. Tudor domains are indicated in yellow.

(B) siRNA experiments followed by peptide pull-downs show that Sgf29 links the SAGA complex to H3K4me3.

(C and D) Bacterial lysates expressing recombinant his-tagged Sgf29 were incubated with the indicated peptides. Following incubation and washes, the amount of bound Sgf29 protein was determined by western blotting using an anti-His antibody.

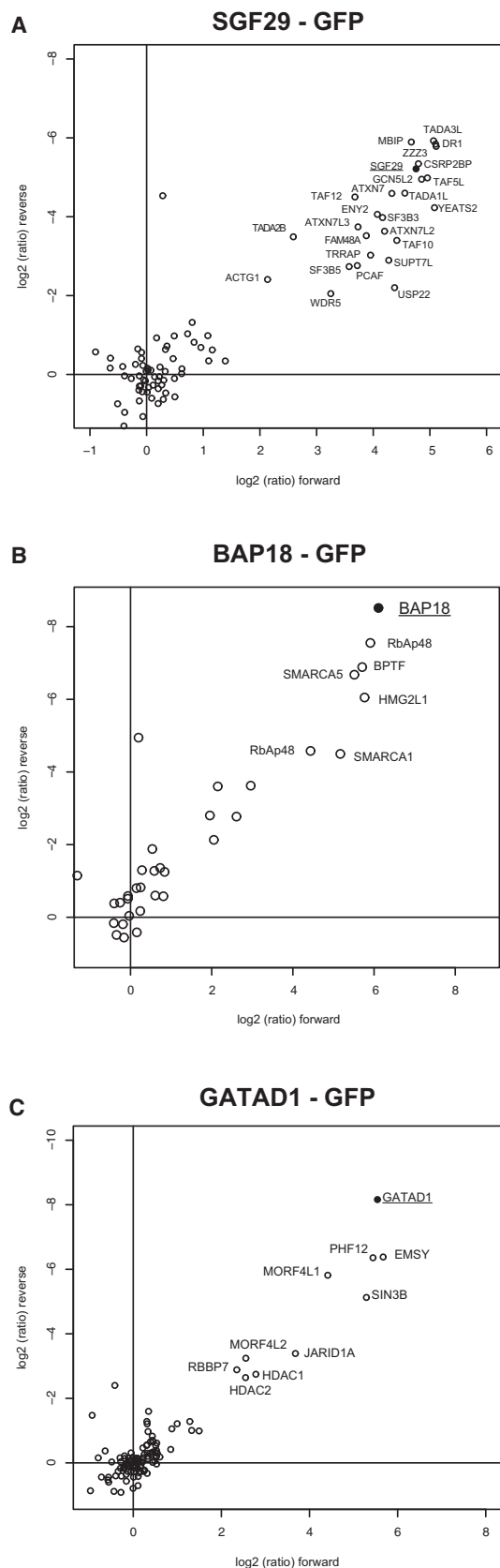
(E) Bacterial lysates expressing the indicated Sgf29 mutants were used for histone peptide pull-downs to determine their binding affinity for H3K4me3. The first lane represents peptides without the me3 modifications.

(F) Isothermal calorimetry experiment revealing the affinity of the full-length Sgf29 protein for H3K4me3.

step GFP affinity purification to identify protein-protein interactions for chromatin readers (Figure 3A; Table S2). We then applied this approach to the as-yet uncharacterized protein C17orf49, which we had found to interact with H3K4me3 (Figure 1B). C17orf49 is an 18 kDa protein that carries a SANT domain, which commonly occurs in chromatin associated proteins. Pull-down of the GFP fusion protein from stably transfected HeLa cells specifically copurified subunits of the human NuRF/BPTF complex (Figure 3B; Table S2). Strikingly, HMG2L1, another highly significant interactor of H3K4me3 (Figure 1B) is one of the most prominent interactors of

C17orf49. Thus, this experiment established C17orf49 and HMG2L1 as subunits of the human NuRF/BPTF complex. Their association with H3K4me3 is explained by their interaction with the H3K4me3 reader BPTF. We name the uncharacterized open reading frame C17orf49 as “BPTF associated protein of 18 kDa” (BAP18).

GATA zinc finger domain containing 1 (GATAD1) is another protein of unknown function that was identified as a H3K4me3 interactor. Using the GFP pull-down approach, we identified subunits of the Sin3b/HDAC complex, the H3K4me3-specific lysine demethylase Jarid1A/RBBP2, and the breast cancer



**Figure 3. GFP Pulldowns for H3K4me3 Readers**

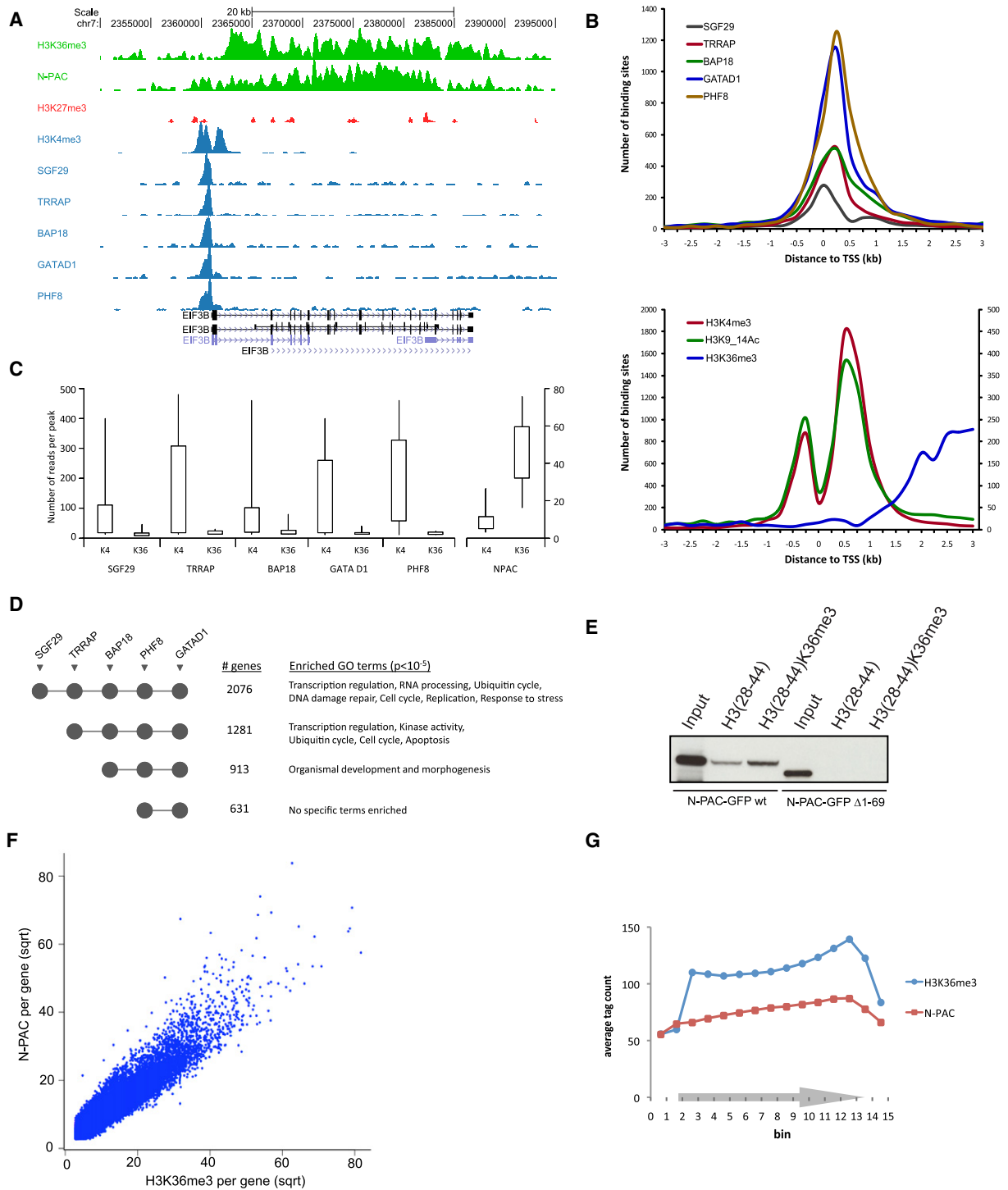
HeLa Kyoto cells expressing GFP-Sgf29 (A), GFP-C17orf49/BAP18 (B), and GFP-GATAD1 (C) were SILAC-labeled and subjected to single-step affinity purifications in a “forward” and “reverse” pulldown using GFP nanotrap beads. In each panel the ratio of the identified proteins in the forward and reverse pulldown is plotted. Proteins interacting with the baits are indicated. See also Figure S2 and Table S2.

associated protein EMSY (Hughes-Davies et al., 2003) as interactors for GATAD1 (Figure 3C; Table S2). Because all of the subunits in this complex were identified as H3K4me3 readers with similar ratios, we hypothesized that they form an as-yet uncharacterized chromatin reading complex (Figures S2A–S2D). Jarid1a was recently reported to bind tightly to H3K4me3 with a  $K_d$  of 0.75  $\mu$ M (Wang et al., 2009a) and therefore forms the direct link between the complex and the chromatin mark. Further evidence for our hypothesis comes from a subsequently published *Drosophila* Lid complex (Lee et al., 2009; Moshkin et al., 2009). Lid is the *Drosophila* homolog of the mammalian Jarid1 family of proteins, consisting of Jarid1a, Jarid1b and Jarid1c. The complex furthermore contains homologs of the Sin3 proteins, as well as an EMSY and GATAD1 homolog. In mammals, interactions between the Sin3/HDAC complex and Jarid1a have also been reported (van Oevelen et al., 2008). However, EMSY has not been tied to any of these proteins yet. EMSY is known to be a repressor of transcription (Hughes-Davies et al., 2003) but the mechanisms underlying this repressive activity are poorly understood. The identification of the above-described complex provides important clues as to how EMSY represses transcription. We hypothesize that gene repression involves histone deacetylation coupled with H3K4me3 demethylation.

#### Localizing the Chromatin Readers on the Genome

To further investigate the function of our proteins of interest in vivo, we performed ChIP-Seq profiling using an anti-GFP antibody on the BAC-GFP lines. Figure 4A shows a representative snapshot of the ChIP-Seq data. Profiling of GFP-tagged proteins interacting with H3K4me3 and H3K36me3 was performed on biological replicates and showed that the approach is highly reproducible (Pearson correlation >0.85; Figures S3F and S3G). In agreement with our peptide pulldown data, the identified H3K4me3 readers Sgf29, TRRAP, PHF8, GATAD1, and BAP18, are associated mainly with promoters (Figures S3A and S3B) and coincide with H3K4me3 marking (Figures 4B and 4C; Figure S3C). We also identified a small number of binding sites of H3K4me3 readers outside of annotated promoters (Figure S3A). As these are not associated with H3K4me3 (Figure S3B), the interactor proteins are apparently recruited to these loci by H3K4me3 independent mechanisms. Nevertheless, for each of these five proteins we observed a good genome-wide correlation with H3K4me3 (Pearson correlation BAP18: 0.71, GATAD1: 0.71, PHF8: 0.66, TRRAP: 0.66, SGF29: 0.55).

For Sgf29, TRRAP, and BAP18, it was expected that they would localize to promoters, as they are part of conserved complexes associated with active transcription—SAGA/ATAC, SAGA/NuA4, and BPTF/NuRF, respectively (Nagy et al., 2010; Wysocka et al., 2006). PHD finger protein 8 (PHF8) belongs to



**Figure 4. ChIP Sequencing of H3K4me3 and H3K36me3 Readers**

(A) ChIP-Seq profiles of three histone modifications and the interactors across the Eif3B gene on human chromosome 7. (B) Distance distribution of the binding sites for the H3K4me3 interactors and the three histone modifications relative to the closest transcription start site (TSS). x axis is in 1000 bp; on the y axis the number of binding sites is indicated. Values for H3K36me3 are plotted on a separate scale (right side). (C) Number of reads for H3K4me3 and H3K36me3 (indicated with K4 and K36, respectively) within the binding sites for the H3K4me3 interacting proteins. The ends of the whiskers represent the 9th and 91st percentile, respectively. Values for SGF29, TRRAP, BAP18, PHF8, and GATAD1 are on the scale on the left side of the plot, while values for N-PAC are on a separate scale on the right. (D) Promoters clustered by the binding sites for the H3K4me3 interacting proteins (Figure S3). Co-occurrence of binding sites is indicated with gray circles under

the JmJc domain-containing family of proteins that can remove methyl groups from arginine or lysine residues (Cloos et al., 2008). PHF8 can remove the repressive mark H3K9me2 (Horton et al., 2010), associating it with activation of transcription, which is in agreement with our ChIP-Seq analyses.

We found GATAD1 to interact with Jarid1a/EMS1/Sin3 (Figure 3C). Jarid1a is a JmJc domain-containing protein that demethylates H3K4me3 (Cloos et al., 2008). In addition, the GATAD1 purification enriched for components of the Sin3/HDAC transcriptional corepressor complex, including two histone deacetylases, HDAC1 and HDAC2. Despite the repressive enzymatic activities associated with GATAD1, our ChIP-Seq analysis reveals that this complex binds to promoters marked with H3K4me3. These data may be explained by invoking a mechanism of cyclical recruitment of “writers” and “erasers” to sites of active transcription (Wang et al., 2009c).

Interestingly, our ChIP-Seq analyses showed that many target genes can be occupied by each of the five H3K4me3 readers. Analysis of all identified target genes resulted in four discrete clusters (Figures S3D and S3E; Table S3). PHF8 and GATAD1 were the only factors found to be common to all clusters and therefore are likely to have a general role in transcription. The two largest clusters combined genes whose promoters were bound by Sgf29 and/or TRRAP, indicating that transcriptional regulation of these genes involves SAGA/NuA4-related complexes. Gene ontology (GO) annotation of the genes in these clusters revealed a number of highly enriched ( $p < 10^{-5}$ ) functional terms that agree very well with the biological functions of these complexes (Figure 4D). For example, SAGA/ATAC and NuA4 complexes are crucial regulators of transcription, DNA repair, DNA replication, and the cell cycle (Squatrito et al., 2006). Distinct GCN5/PCAF-containing complexes function as coactivators and are involved in transcription factor and global histone acetylation (Nagy and Tora, 2007). SAGA was shown to regulate various stress-response genes (Huisinga and Pugh, 2004; Nagy et al., 2010), while TRRAP-containing complex NuA4 regulates apoptosis (Ikura et al., 2000; Tyteca et al., 2006). Thus, each functional category of the GO analysis corresponds to an established function of the SAGA and NuA4 complex, which independently validates the connection between the activating histone mark and its reader found in our experiments.

N-PAC, MSH-6, and NSD1 as well as NSD2 were identified as H3K36me3 interactors (Figure 1C; Table S2). Interestingly, these four proteins share a PWWP domain which is part of the Tudor domain “Royal Family” and includes the Tudor, chromo and MBT domains that can interact with methylated lysine residues. The PWWP domain of Set9 was recently identified as a reader for H4K20me1 (Wang et al., 2009b). Our peptide pulldown data

suggest that this domain is also capable of recognizing H3K36me3, which is associated with elongation of transcription and peaks in coding regions of genes (Shilatfard, 2006). Very recently the PWWP domain of Brpf1 was shown to bind specifically to H3K36me3 (Vezzoli et al., 2010). Indeed, deletion analyses revealed that the PWWP domain of N-PAC is necessary for H3K36me3 binding (Figure 4E). This PWWP domain mediated K36me3 binding is most likely direct, since purification of N-PAC-GFP from a BAC line did not reveal protein-protein interactions (data not shown). To investigate the genomic binding pattern of N-PAC, we generated the corresponding BAC-GFP line and performed ChIP-Seq analysis. Consistent with our peptide pulldown data, N-PAC binds to coding regions of active genes correlating with the presence of H3K36me3 (Figures 4C and 4F). N-PAC and H3K36me3 increase toward the 3' end (Figures 4A and 4G). Together our data establish the PWWP domain as a putative binder of H3K36me3. In addition to a PWWP domain, N-PAC also contains an AT-hook that is often found in proteins that are associated with elongation of transcription and an enzymatic domain of unknown function. Our ChIP-Seq analysis revealed that both H3K36me3 and N-PAC are present almost exclusively over gene bodies (data not shown), and that the vast majority of H3K36me3 marked regions are also bound by N-PAC, indicating a broad or universal function of this protein in transcriptional elongation.

### The Interactome of the Repressive Histone Methyl Marks

We next investigated the chromatin readers of H3K9me3, H3K27me3 and H4K20me3, histone methyl marks associated with gene repression (Figures 1D–1F). H3K9me3 yielded the richest set of interactors, including all three HP1 isoforms (CBX1, CBX3, and CBX5). The chromodomain-containing HP1 proteins are classical readers of H3K9me3 (Jenuwein and Allis, 2001) and our analysis confirms that they are restricted to this repressive modification. Two chromodomain proteins, CDYL and CDYL2, were identified as binders for both H3K9me3 and H3K27me3 but not H4K20me3. These proteins are members of a family of three chromodomain proteins, the third one being chromodomain Y protein, whose gene is located on the Y chromosome and whose expression is testis specific. Recently, direct binding of CDYL and CDYL2 to H3K9me3 and H3K27me3 has been reported (Fischle et al., 2008; Franz et al., 2009). As expected, Polycomb group proteins represent the major readers for H3K27me3, but many of these proteins were also identified as specific interactors for H3K9me3. Given the high degree of sequence identity surrounding H3K9 and H3K27 (TARKST and AARKSA for K9 and K27, respectively), it is not surprising to find Polycomb group proteins as interactors

the corresponding interactor names. Four major groups of promoters were identified, for which the number of genes within each group and highly enriched GO terms ( $p$  value  $< 10^{-5}$ ) are listed.

(E) Full-length N-PAC-GFP and  $\Delta$ 1-69 N-PAC-GFP were transfected into HeLa Kyoto cells. Extracts from these cells were subsequently used for K36/K36me3 peptide pulldowns. Unlike the wild-type protein,  $\Delta$ 1-69 N-PAC-GFP, that lacks most of the PWWP domain, does not bind to H3K36me3.

(F) Dotplot showing the correlation between H3K36me3 and N-PAC ( $R^2 = 0.86$ ). Every dot represents the number of N-PAC or H3K36me3 ChIP-Seq tags per gene.

(G) All genes containing H3K36me3 (>5 kb) were each divided in 15 bins followed by counting and averaging of the H3K36me3 and N-PAC ChIP-Seq tags within each bin.

See also Figure S3 and Tables S3 and S4.

for H3K9me3. Literature evidence also supports the interaction of Polycomb group proteins with H3K9me3, although their affinity for H3K27me3 is higher (Fischle et al., 2003b; Ringrose et al., 2004). Finally, we identified the origin recognition complex (ORC) as an interacting complex for all three repressive sites.

We purified complexes associated with the HP1 family members to ascertain if the H3K9me3 readers physically interact with them using BAC-GFP constructs (Figures 5A–5C). Among the specifically interacting proteins, known HP1 interactors were identified, such as chromatin assembly factors CHAF1A/CHAF1B and ADNP (Lechner et al., 2005; Mandel et al., 2007). Two uncharacterized proteins, POGZ and Znf828, consistently interacted with high ratios with all HP1 family members. We confirmed the binding of POGZ to H3K9me3 by western blotting (Figure S1C). POGZ and Znf828 have an interesting domain structure and multiple zinc fingers, suggesting that these proteins may specifically bind DNA sequences. POGZ or POGO transposable element with a ZNF domain is a 1410 amino acid protein containing two domains that are also present in the centromeric protein B (CenPB). Next, we generated BAC-GFP constructs for these proteins. Pulldowns with POGZ and Znf828 reciprocally confirmed interaction with HP1 and, interestingly, with each other (Figures 5D and 5E). Additionally, POGZ interacted specifically with mitotic spindle checkpoint protein, Mad2l2. To substantiate this possible connection to a prominent cell cycle protein, we performed a GFP pulldown with a cell line of this protein, which clearly demonstrated reciprocal binding (Figure 5F). Thus, a combination of repressive mark interactors and full-length protein interactomes allows us to deconstruct the majority of protein interactions involved in the biology of the repressive marks.

We noticed that LRWD1 clusters together in the two-dimensional interaction plots with the ORC complex in the pulldowns of each of the repressive marks (Figures 1D–1F). LRWD1 has not been characterized but obtains its name from a leucine-rich repeat and a stretch of WD40 domains. To test if this protein is a subunit of the ORC complex, we generated the BAC-GFP cell line of Orc2L. Pulldown with this ORC subunit indeed demonstrated specific interaction with LRWD1 (Figure 5G). Furthermore, ChIP-Seq of the BAC LRWD1-GFP line revealed a strong enrichment on satellite repeats, correlating with high levels of H3K9me3 which is known to be enriched over satellites (Figure 5H) (Martens et al., 2005).

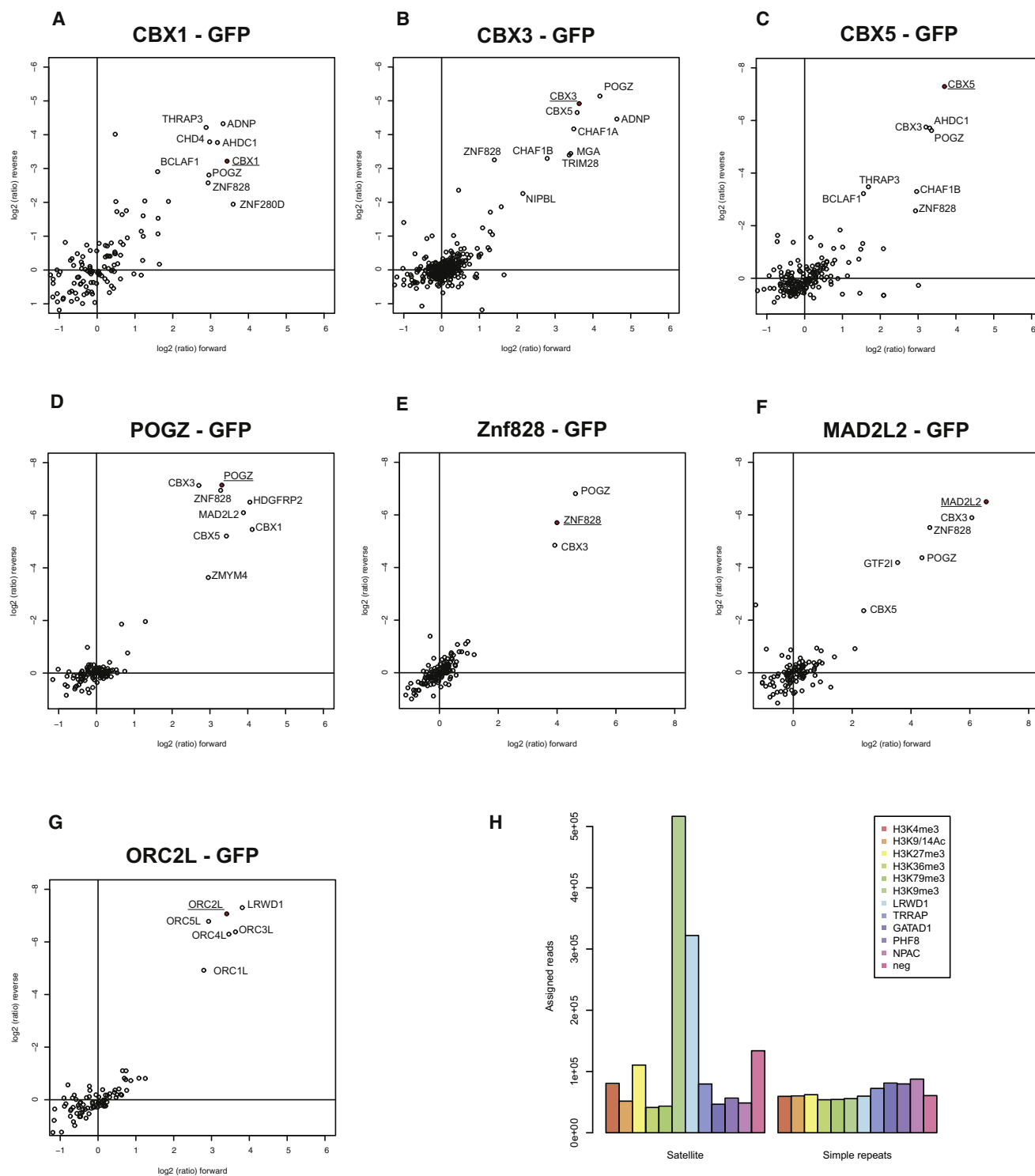
### Triple SILAC Pulldowns Reveal Differential Fine-Tuning of Trimethyl Lysine Binding

The five trimethyl lysine marks that we screened for interactors are flanked by numerous residues that can also be subjected to posttranslational modifications. These modifications could, either agonistically or antagonistically, affect trimethyl lysine binding. To study such potential interplay between different posttranslational modifications (PTMs) occurring in close proximity on the histone H3 tail, we applied triple pulldown experiments involving a combination of methylation and other PTM marks, in this case acetylations or phosphorylations (Vermeulen et al., 2007). In this approach, cells are grown in three different SILAC media, each containing different stable isotopic versions of lysine and arginine. These extracts, which are distinguishable

by MS, are each incubated with a differently modified histone peptide (triple pulldown). Peptides appear as triplets in the MS spectra and a significant ratio between the first two peaks indicates specific binding to the H3K4me3 mark. The highest mass peak in the triplet originates from the eluate of the combinatorially modified peptide and its intensity compared with the eluate from the singly modified peptide (middle peak) indicates either agonistic or antagonistic binding or no effect. On genes that are actively being transcribed, H3K4me3 often co-occurs with acetylation of H3K9 and H3K14. A number of readers for H3K4me3 carry both a domain that recognizes H3K4me3 as well as one or multiple bromodomains, which bind to acetylated lysine residues. We therefore wondered whether these acetylations would function agonistically with H3K4me3 to bind H3K4me3 readers to the histone H3 tail. Consistent with our previous findings (Vermeulen et al., 2007), TFIID and BPTF bound more strongly to the H3K4me3 mark when it was flanked by acetylation on H3K9 and H3K14 acetylation (Figures S4A and S4B). In addition, we also observed—by quantitative proteomics and by western blotting—agonistic binding to the methylated and acetylated peptide for the SAGA complex (Sgf29 in Figures 6A and 6C). In contrast, recombinant Sgf29 does not display preferential H3K9,14Ac binding (Figure 6D), indicating that the observed effects in the triple pulldown are due to the agonistic binding effects of the Sgf29 double Tudor domain and the GCN5 bromodomain. Finally, we also observed agonistic binding of PHD finger protein 8 (PHF8) to H3K4me3 and H3K9,14 Ac (Figure 6B). PHF8 carries an H3K4me3-binding PHD finger (Horton et al., 2010), but it does not contain a bromodomain. Therefore, we hypothesize that this protein either carries an unidentified acetyl lysine binding motif, or interacts with an as-yet unidentified bromodomain-containing protein. These results indicate that agonistic H3K4me3 and H3K9,14Ac recognition occurs in several chromatin readers. The mechanisms are diverse; for example, a PHD finger domain can be combined with a bromodomain in one protein (BPTF), or in different subunits of the same complex (TAF3 PHD finger and TAF1 bromodomains in the TFIID complex). Moreover, a different recognition domain combination can be used (Tudor domain of Sgf29 with the bromodomain of GCN5 in the SAGA complex). Clearly, these chromatin readers have each evolved the ability to target combinatorially marked nucleosomes allowing regulation of specific subsets of genes.

To study potential antagonistic histone PTM crosstalk, we decided to focus on phosphorylations on the histone H3 tail. Phosphorylation of histone H3S10 results in the release of HP1 from chromatin during mitosis even though levels of H3K9me3 remain unchanged (Fischle et al., 2005). H3K27me3 is also flanked by a serine residue that can be phosphorylated (Winter et al., 2008). To investigate if these trimethylations co-occur with the respective adjacent phosphorylations, we analyzed our recent large-scale study of the proteome and the phosphoproteome of the cell cycle (Olsen et al., 2010). Indeed, we found the corresponding doubly modified peptides. Moreover cell cycle data indicates that they are specific for mitotic cells (Figures S4G–S4J). As shown in Figure 6E, H3S10 phosphorylation does not appear to drastically affect the binding of HP1 to H3K9me3. These results are in agreement with data reporting



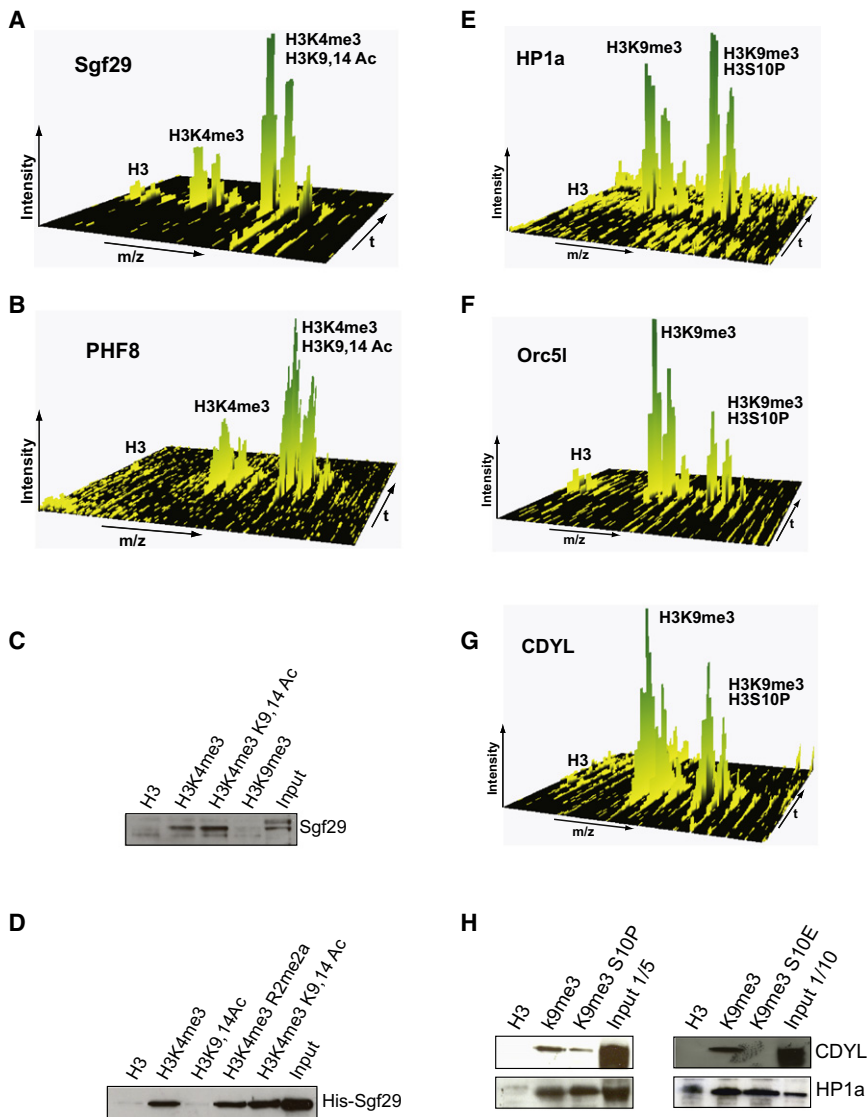


**Figure 5. GFP Pulldowns for Readers of the Repressive Histone Marks**

(A–G) GFP-fusion proteins expressed in SILAC-labeled HeLa cells were enriched on GFP-nanotrap beads. In each figure, the ratio of the identified proteins in the forward and reverse pulldown is plotted. Proteins interacting with the baits are indicated.

(H) The total number of ChIP-Seq reads present on either satellite repeats or simple repeats for the indicated proteins and histone marks is shown.

See also Table S2.



**Figure 6. Triple SILAC Pulldowns Revealing Histone Modification Crosstalk**

(A) Three-dimensional representation of the MS signal of an Sgf29 peptide identified in a triple pulldown SILAC experiment (the m/z scale is the x axis, the chromatographic retention time is the y axis, and the MS-signal is the z axis). Each group of signals represents the natural isotope pattern of the peptide. The relative intensities of the triplet peak of the Sgf29 peptide indicates the preference of binding to the modification states (unmethylated histone H3 peptide [left peak], H3K4me3 peptide [middle peak], and the double-modified H3K4me3/H3K9,14 Ac peptide [right peak]).

(B) Same as (A) for a PHF8 peptide identified in the same triple pulldown.

(C) Nuclear extracts derived from HeLa cells were incubated with the indicated histone peptides. The amount of Sgf29 protein bound to these peptides was determined by western blotting using an antibody against endogenous Sgf29.

(D) Bacterial lysates expressing recombinant His-tagged Sgf29 were incubated with the indicated peptides. Following incubation and washes, the amount of bound Sgf29 protein was determined by western blotting using an anti-His antibody. Note that Sgf29 does not bind to a peptide containing H3K9,14 acetylation and that the binding of Sgf29 to H3K4me3 is not affected by asymmetric dimethylation of H3R2.

(E–G) Three-dimensional representation of an HP1α (E), Orc5 (F), and CDYL (G) peptide identified in a triple pulldown SILAC experiment. The spectra show the MS-signal representing the relative binding of these peptides to the unmethylated histone H3 peptide (left peak), the H3K9me3 peptide (middle peak), and the double-modified H3K9me3/H3S10P peptide (right peak).

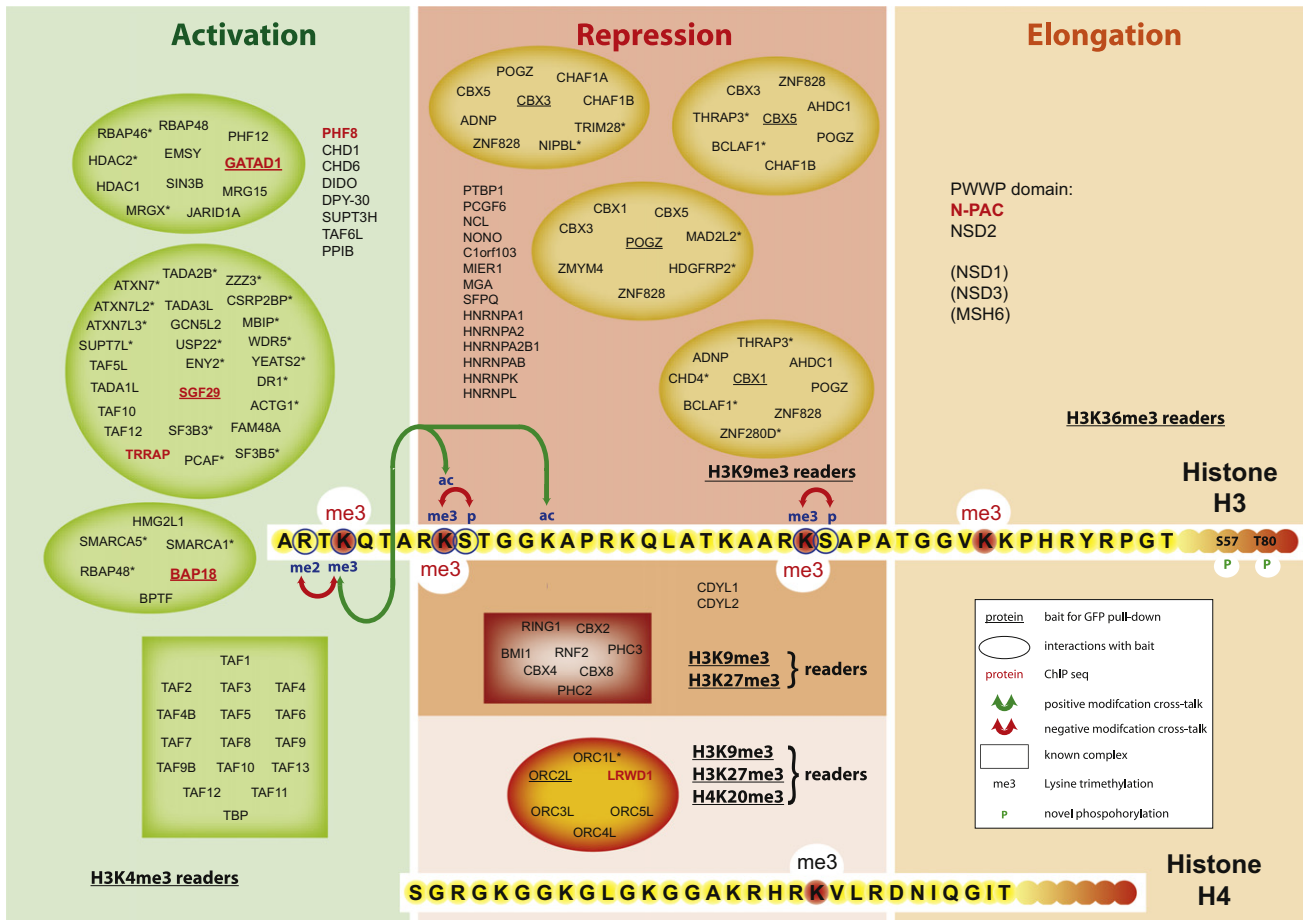
(H) Histone peptide pulldowns in HeLa nuclear extracts were performed with the indicated peptides. The amount of HP1α and CDYL binding to these peptides was determined by western blotting using an antibody against HP1α and CDYL. See also Figure S4.

stabilization of HP1 binding by H3S10 phosphorylation (Mateescu et al., 2004). Indicating that our assay can indeed reveal antagonistic effects, we observed that CDYL as well as the ORC complex subunits do show reduced H3K9me3 binding in combination with H3S10 phosphorylation (Figures 6F and 6G). These experiments were further confirmed by western blotting, also making use of a phosphomimetic peptide where H3S10 was mutated to glutamic acid (Figure 6H). Similarly, H3S28 phosphorylation destabilizes the binding of CDYL and ORC complex subunits to H3K27me3, whereas this phosphorylation only mildly affects the binding of Polycomb group proteins (Figures S4C–S4F). Taken together, these results suggest that phosphorylations on the N-terminal tails of histones selectively affect the binding of proteins to adjacent modified lysines residues. Such so-called phospho-methyl switches are quite common on core histones (Fischle et al., 2003a). We have also identified H3S57 and H3T80 as phosphorylation sites on histone H3 (for H3S57P and H3T80, Figures S4K and S4L), both of which

are adjacent to modified lysine residues. Thus, almost all of the modified lysine residues on histone H3 can be flanked by phosphorylated residues. An important function of these phosphorylation sites could be the differential regulation of protein binding to neighboring methylated or acetylated lysines in specific cellular situations and for specific genes.

## DISCUSSION

Here, we have characterized the association of chromatin readers with histone trimethyl-lysine modifications by a combination of three technologies. The major findings from our integrated approach are visualized and summarized in Figure 7. High-accuracy, quantitative proteomics based on SILAC identified known and previously unknown binders to each of the chromatin marks. Plotting SILAC ratios from forward and reverse experiments grouped distinct protein clusters together, representing functional complexes. To investigate these complexes, we turned



**Figure 7. Visualization of the Histone Trimethyl-Lysine Interactome**

Proteins interacting with the five trimethyl lysine marks are indicated. Encircled are proteins that were additionally identified in GFP pull-down experiments; baits in these pull-downs are underlined. Proteins in those circles marked with an asterisk were not identified as interactors in the peptide pull-downs. Proteins clustered in rectangles were identified in the peptide pull-downs and were previously shown to interact with each other (TFIID for H3K4me3 and PRC1 for H3K9me3 and H3K27me3). For proteins that are color coded red *in vivo* verification by ChIP-Seq is also provided. The arrows and associated labels indicate histone modification crosstalk investigated in this study. In the globular part of histone H3, two identified histone phosphorylations (H3S57P and H3T80), are indicated.

to the recently developed BAC-transgeneOmics technology (Poser et al., 2008), which allowed rapid generation of stable cell lines containing the entire gene of interest fused to GFP in its endogenous context. Therefore, this technology provides a generic ‘‘handle’’ for the members of chromatin reader complexes while maintaining endogenous control. We used these cell lines in a next round of SILAC-based quantitative interaction screens to establish physical interactions between the chromatin readers. Furthermore, the GFP-tag was utilized for chromatin immunoprecipitation followed by next generation DNA sequencing to localize the readers on the genome. The synergistic use of these three approaches enabled us to create data sets and reagents that provide a resource for researchers interested in epigenetic questions. While shown here for histone modifications, our approach can be extended to posttranslational modifications on other chromatin-associated proteins and to other cellular systems such as stem cells. Illustrating the usefulness of this resource, we were able to dissect several

mechanisms of chromatin reader associations with their chromatin marks starting from basic interaction data.

One such example is the human SAGA complex, all identified members of which clustered tightly in the two-dimensional interaction plot (Figure 1B). SAGA is a highly conserved complex, which plays key roles in the activation of transcription of RNA polymerase II target genes. However, the mechanisms of activation are not completely understood. In yeast, it has been suggested that CHD1 links the complex to H3K4me3 (Pray-Grant et al., 2005). However, this association is controversial as it has been reported that yeast CHD1 does not bind to H3K4me3 (Sims et al., 2005). While we identified human CHD1 as a specific binder to this mark, it did not co-cluster with the SAGA subunits in our H3K4me3 peptide pull-downs. Furthermore, we were not able to identify CHD1 as an interactor of the SAGA subunit Sgf29 in a GFP pull-down. Instead, starting with the observation that Sgf29, which we identified as a H3K4me3 interactor, has a double Tudor domain (Lee and Workman, 2007) and given

the fact that Tudor domains can bind methylated lysines (Huang et al., 2006), we established by biochemical and biophysical means that the double Tudor domain of Sgf29 forms the direct molecular link between SAGA and H3K4me3. This binding mode is likely conserved down to yeast, which has a homolog of Sgf29 that also contains a double Tudor domain (Figure 2A). Such conservation is not universal as it is not the case for association of TFIID with H3K4me3. This interaction is mediated by the PHD-finger domain of human TAF3, but yeast TAF3 lacks the PHD-finger domain (Vermeulen et al., 2007).

Bioinformatic analysis of the interactors of the activating H3K36me3 mark revealed that four of the most prominent specific interactors shared the same domain. This PWWP domain is part of the Tudor domain "Royal family" of domains (Maurer-Stroh et al., 2003) and therefore almost certainly mediates direct binding to H3K36me3. In agreement with this, deletion analysis revealed that the PWWP domain of N-PAC is essential for its interaction with H3K36me3 (Figure 4E).

In the interactome of the repressive marks we identified, in addition to expected heterochromatin associated proteins, several other proteins. Interaction studies with BAC GFP-fusion proteins uncovered many interactions with members of the HP1 family. This HP1 family and associated proteins represent a large portion of the H3K9me3 interactome and establish the HP1 proteins as interaction hubs in mediating repressive gene functions. Interestingly, several HP1 interactors contain zinc finger domains (such as POGZ and Znf828), which may serve to recruit HP1 to specific sites in the genome.

The origin recognition complex (ORC) has a key function in replication firing. It is known to localize to heterochromatic regions (Prasanth et al., 2004) and it interacted with all three repressive marks. LRWD1 grouped with the ORC complex members in the two-dimensional interaction plots. Pulldowns with an Orc2L BAC-GFP cell line demonstrated that LRWD1 is indeed an ORC complex subunit and ChIP-Seq experiments established that it co-enriches with H3K9me3 on satellite repeats. The WD40 repeat domain of LRWD1 may mediate the interaction of the ORC complex with the repressive marks as it was recently shown that the WD40 repeats of the Polycomb group protein EED directly binds to H3K27me3 (Margueron et al., 2009).

A triple-encoding variant of the SILAC peptide pulldown allowed us to directly address the question of agonistic and antagonistic binding to combinatorial histone modifications. These experiments recapitulated several known combinatorial interactions, such as the agonistic effects between H3K4me3 and nearby acetylations. The general conclusion from these experiments is that the trimethyl marks constitute the major docking sites for chromatin readers and that other nearby modifications fine-tune these primary interactions by augmenting or destabilizing specific interactions. For example, our data show that H3S10 phosphorylation destabilizes the ORC complex and CDYL binding to H3K9me3, whereas HP1 binding does not appear to be affected. Consistent with this paradigm, we have not been able to determine specific interactors with peptides bearing only the ancillary modifications. This is unlikely to be an artifact due to pulldowns with synthetic peptides because similar results are obtained when performing pulldowns

with entire immobilized nucleosomes carrying particular epigenetic marks (T. Bartke, M.V., M.M., and T. Kouzarides, unpublished data). In this context, mass spectrometry can also contribute by identifying and quantifying the combinatorially modified peptides *in vivo*, as shown for several examples here.

A striking finding that emerges from our integrative investigation into the nature of the relationship between histone marks and their readers is the degree of overlap between the known biological functions of the marks and the biological functions of their associated readers (Figure 7). Histone modifications are usually studied by techniques such as ChIP, ChIP-Seq, or immunofluorescence that associate them with particular genes or nuclear processes. The same holds true for transcription factors or other chromatin regulators. By its nature, our strategy combines investigation of chromatin marks and transcriptional regulators and is thereby uniquely suited as an integrative tool for the investigation of epigenetic regulation of gene expression.

## EXPERIMENTAL PROCEDURES

### Recombinant Protein Expression and ITC Calorimetry

Full-length Sgf29 constructs were expressed with an N-terminal His-tag and a maltose binding protein (MBP) domain using expression plasmid pETM44 (Novagen). For histone peptide pulldown experiments, crude induced bacterial lysates were used as described (Vermeulen et al., 2007). His-tag westerns were performed using a penta-His antibody (QIAGEN). For isothermal calorimetry (ITC) experiments the Sgf29 protein was enriched using Ni NTA beads after which the protein was further purified on a Superdex 200 column. ITC measurements were performed on a VP-ITC Microcal calorimeter (Microcal, Northampton, MA) at 25°C. During titration, 7  $\mu$ l of H3K4me3 peptide (aa 1–17) at a concentration of 300  $\mu$ M was injected into a solution of 25  $\mu$ M Sgf29 protein.

### GFP Pulldowns

Generation of the BACs with GFP-fusion constructs was done as described (Poser et al., 2008). Nuclear extracts from BAC-GFP-tagged or wild-type HeLa cells were SILAC labeled with heavy lysine (Isotec, Sigma). For CBX3, no BAC was available and SILAC-labeled HeLa cells were transfected with plasmid pBCHGN-CBX3 (Addgene). GFP nanotrap beads (Chromotek) were used to precipitate GFP-tagged proteins from these lysates. Approximately 500–1000  $\mu$ g of nuclear extract was used per pulldown in a buffer containing 300 mM NaCl, 0.25% NP40, 0.5 mM DDT, 20 mM HEPES KOH (pH 7.9), and protease inhibitors. Following incubation and washes with the same buffer, beads from both pulldowns were combined, proteins were eluted with acidic glycine (0.1 M [pH 2.0]) and digested overnight with LysC (Wako Biochemicals, Japan) using the FASP protocol (Wisniewski et al., 2009) prior to LC/MS-MS analysis.

### Mass Spectrometry of Proteins

Gel lanes representing each pulldown were cut into eight equally sized slices as described (Vermeulen et al., 2007). Peptide identification was performed on an LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific, Germany) essentially as described (Olsen et al., 2004). Full-scan MS spectra were acquired with a resolution of 60,000 in the Orbitrap analyzer. For every full scan, the five most intense ions were fragmented in the linear ion trap. Raw data were processed and analyzed using the MaxQuant software (version 1.0.12.33) and searched with the Mascot search engine against a human IPI database 3.52 as described (Butter et al., 2010). Phosphopeptide enrichment of core histones and MS analysis of these were performed as described (Hurd et al., 2009).

### Deposition of MS-Related Data

Mass spectrometric data for peptide pulldowns and GFP pulldowns, consisting of raw data files, unfiltered "proteingroups" tables, and identified

peptides, can be accessed at the TRANCHE repository (<https://proteomecommons.org/>) under the name “Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers.”

### Chromatin Immunoprecipitation and Deep Sequencing

ChIP experiments were performed using  $3.3 \times 10^6$  cells per ChIP according to standard protocols (Denissov et al., 2007), with two minor modifications. Crosslinking of the cells was done on the culture plates for 20 min, while ChIP'ed DNA was purified by Qiaquick PCR purification Kit (QIAGEN cat. no. 28106). ChIP enrichment levels were analyzed by qPCR using specific primers (available upon request) for quality control. ChIP-Seq samples were prepared and analyzed according to the manufacturer (Illumina). Enriched regions were identified by FindPeaks (Fejes et al., 2008). Table S4 summarizes the ChIP-Seq output. For the repeat analysis of the H3K9me3 and LRWD1 ChIP-Seq profiles, mappings were performed by maq aligner (Li et al., 2008). For further information about the ChIP-Seq methods and data analysis see Extended Experimental Procedures. All ChIP-Seq data are present in the NCBI GEO SuperSeries GSE20303.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, four figures, and five tables and can be found with this article online at doi:10.1016/j.cell.2010.08.020.

### ACKNOWLEDGMENTS

Eva Janssen-Megens, Kees-Jan François, and Yan Tan helped with the sequencing, the MPI core facility expressed and purified recombinant proteins, and Dr. F. Tashiro kindly provided Sgf29 antibody. This work was supported by the Max-Planck Society for the Advancement of Science, HEROIC, an Integrated Project funded by the European Union under the 6th Framework Program (LSHG-CT-2005-018883) and by Interaction Proteome (LSHG-CT-2003-505520). M.V. received a fellowship of the Dutch Cancer Society (KWF/NKB) and a grant from the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research.

Received: February 17, 2010

Revised: June 9, 2010

Accepted: August 13, 2010

Published: September 16, 2010

### REFERENCES

Butter, F., Kappei, D., Buchholz, F., Vermeulen, M., and Mann, M. (2010). A domesticated transposon mediates the effects of a single-nucleotide polymorphism responsible for enhanced muscle growth. *EMBO Rep.* *11*, 305–311.

Cloos, P.A., Christensen, J., Agger, K., and Helin, K. (2008). Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. *Genes Dev.* *22*, 1115–1140.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* *26*, 1367–1372.

Denissov, S., van Driel, M., Voit, R., Hekkelman, M., Hulsen, T., Hernandez, N., Grummt, I., Wehrens, R., and Stunnenberg, H. (2007). Identification of novel functional TBP-binding sites and general factor repertoires. *EMBO J.* *26*, 944–954.

Fejes, A.P., Robertson, G., Bilenky, M., Varhol, R., Bainbridge, M., and Jones, S.J. (2008). FindPeaks 3.1: a tool for identifying areas of enrichment from massively parallel short-read sequencing technology. *Bioinformatics* *24*, 1729–1730.

Fischle, W., Wang, Y., and Allis, C.D. (2003a). Binary switches and modification cassettes in histone biology and beyond. *Nature* *425*, 475–479.

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003b). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev.* *17*, 1870–1881.

Fischle, W., Tseng, B.S., Dormann, H.L., Ueberheide, B.M., Garcia, B.A., Shabanowitz, J., Hunt, D.F., Funabiki, H., and Allis, C.D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* *438*, 1116–1122.

Fischle, W., Franz, H., Jacobs, S.A., Allis, C.D., and Khorasanizadeh, S. (2008). Specificity of the chromodomain Y chromosome family of chromodomains for lysine-methylated ARK(S/T) motifs. *J. Biol. Chem.* *283*, 19626–19635.

Franz, H., Mosch, K., Soeroes, S., Urlaub, H., and Fischle, W. (2009). Multimerization and H3K9me3 binding are required for CDYL1b heterochromatin association. *J. Biol. Chem.* *284*, 35049–35059.

Garcia, B.A., Shabanowitz, J., and Hunt, D.F. (2007). Characterization of histones and their post-translational modifications by mass spectrometry. *Curr. Opin. Chem. Biol.* *11*, 66–73.

Horton, J.R., Upadhyay, A.K., Qi, H.H., Zhang, X., Shi, Y., and Cheng, X. (2010). Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. *Nat. Struct. Mol. Biol.* *17*, 38–43.

Huang, Y., Fang, J., Bedford, M.T., Zhang, Y., and Xu, R.M. (2006). Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science* *312*, 748–751.

Hubner, N.C., Bird, A.W., Cox, J., Spletstoesser, B., Bandilla, P., Poser, I., Hyman, A., and Mann, M. (2010). Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J. Cell Biol.* *189*, 739–754.

Hughes-Davies, L., Huntsman, D., Ruas, M., Fuks, F., Bye, J., Chin, S.F., Milner, J., Brown, L.A., Hsu, F., Gilks, B., et al. (2003). EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. *Cell* *115*, 523–535.

Huisinga, K.L., and Pugh, B.F. (2004). A genome-wide housekeeping role for TFIIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol. Cell* *13*, 573–585.

Hurd, P.J., Bannister, A.J., Halls, K., Dawson, M.A., Vermeulen, M., Olsen, J.V., Ismail, H., Somers, J., Mann, M., Owen-Hughes, T., et al. (2009). Phosphorylation of histone H3 Thr-45 is linked to apoptosis. *J. Biol. Chem.* *284*, 16575–16583.

Ikura, T., Ogryzko, V.V., Grigoriev, M., Groisman, R., Wang, J., Horikoshi, M., Scully, R., Qin, J., and Nakatani, Y. (2000). Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis. *Cell* *102*, 463–473.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. *Science* *293*, 1074–1080.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* *128*, 693–705.

Lechner, M.S., Schultz, D.C., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd. (2005). The mammalian heterochromatin protein 1 binds diverse nuclear proteins through a common motif that targets the chromoshadow domain. *Biochem. Biophys. Res. Commun.* *337*, 929–937.

Lee, K.K., and Workman, J.L. (2007). Histone acetyltransferase complexes: one size doesn't fit all. *Nat. Rev. Mol. Cell Biol.* *8*, 284–295.

Lee, N., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2009). The H3K4 Demethylase Lid Associates with and Inhibits Histone Deacetylase Rpd3. *Mol. Cell. Biol.* *29*, 1401–1410.

Li, H., Ruan, J., and Durbin, R. (2008). Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res.* *18*, 1851–1858.

Mandel, S., Rechavi, G., and Gozes, I. (2007). Activity-dependent neuroprotective protein (ADNP) differentially interacts with chromatin to regulate genes essential for embryogenesis. *Dev. Biol.* *303*, 814–824.

Margueron, R., Justin, N., Ohno, K., Sharpe, M.L., Son, J., Drury, W.J., 3rd, Voigt, P., Martin, S.R., Taylor, W.R., De Marco, V., et al. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* *461*, 762–767.

- Martens, J.H., O'Sullivan, R.J., Braunschweig, U., Opravil, S., Radolf, M., Steinlein, P., and Jenuwein, T. (2005). The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.* *24*, 800–812.
- Mateescu, B., England, P., Halgand, F., Yaniv, M., and Muchardt, C. (2004). Tethering of HP1 proteins to chromatin is relieved by phosphoacetylation of histone H3. *EMBO Rep.* *5*, 490–496.
- Maurer-Stroh, S., Dickens, N.J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F., and Ponting, C.P. (2003). The Tudor domain “Royal Family”: Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends Biochem. Sci.* *28*, 69–74.
- Moshkin, Y.M., Kan, T.W., Goodfellow, H., Bezstarosti, K., Maeda, R.K., Pilyugin, M., Karch, F., Bray, S.J., Demmers, J.A., and Verrijzer, C.P. (2009). Histone chaperones ASF1 and NAP1 differentially modulate removal of active histone marks by LID-RPD3 complexes during NOTCH silencing. *Mol. Cell* *35*, 782–793.
- Nagy, Z., and Tora, L. (2007). Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* *26*, 5341–5357.
- Nagy, Z., Riss, A., Fujiyama, S., Krebs, A., Orpinell, M., Jansen, P., Cohen, A., Stunnenberg, H.G., Kato, S., and Tora, L. (2010). The metazoan ATAC and SAGA coactivator HAT complexes regulate different sets of inducible target genes. *Cell. Mol. Life Sci.* *67*, 611–628.
- Olsen, J.V., Ong, S.E., and Mann, M. (2004). Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics* *3*, 608–614.
- Olsen, J.V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M.L., Jensen, L.J., Gnad, F., Cox, J., Jensen, T.S., Nigg, E.A., et al. (2010). Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci. Signal.* *3*, ra3.
- Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* *1*, 376–386.
- Poser, I., Sarov, M., Hutchins, J.R., Heriche, J.K., Toyoda, Y., Pozniakovsky, A., Weigl, D., Nitzsche, A., Hegemann, B., Bird, A.W., et al. (2008). BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals. *Nat. Methods* *5*, 409–415.
- Prasanth, S.G., Prasanth, K.V., Siddiqui, K., Spector, D.L., and Stillman, B. (2004). Human Orc2 localizes to centrosomes, centromeres and heterochromatin during chromosome inheritance. *EMBO J.* *23*, 2651–2663.
- Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., 3rd, and Grant, P.A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* *433*, 434–438.
- Ringrose, L., Ehret, H., and Paro, R. (2004). Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of polycomb complexes. *Mol. Cell* *16*, 641–653.
- Shilatifard, A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu. Rev. Biochem.* *75*, 243–269.
- Sims, R.J., 3rd, Chen, C.F., Santos-Rosa, H., Kouzarides, T., Patel, S.S., and Reinberg, D. (2005). Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *J. Biol. Chem.* *280*, 41789–41792.
- Squatrito, M., Gorrini, C., and Amati, B. (2006). Tip60 in DNA damage response and growth control: many tricks in one HAT. *Trends Cell Biol.* *16*, 433–442.
- Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* *14*, 1025–1040.
- Tyteca, S., Vandromme, M., Legube, G., Chevillard-Briet, M., and Trouche, D. (2006). Tip60 and p400 are both required for UV-induced apoptosis but play antagonistic roles in cell cycle progression. *EMBO J.* *25*, 1680–1689.
- van Oevelen, C., Wang, J., Asp, P., Yan, Q., Kaelin, W.G., Jr., Kluger, Y., and Dynlacht, B.D. (2008). A role for mammalian Sin3 in permanent gene silencing. *Mol. Cell* *32*, 359–370.
- Vermeulen, M., and Selbach, M. (2009). Quantitative proteomics: a tool to assess cell differentiation. *Curr. Opin. Cell Biol.* *21*, 761–766.
- Vermeulen, M., Mulder, K.W., Denisov, S., Pijnappel, W.W., van Schaik, F.M., Varier, R.A., Baltissen, M.P., Stunnenberg, H.G., Mann, M., and Timmers, H.T. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* *131*, 58–69.
- Vermeulen, M., Hubner, N.C., and Mann, M. (2008). High confidence determination of specific protein-protein interactions using quantitative mass spectrometry. *Curr. Opin. Biotechnol.* *19*, 331–337.
- Vezzoli, A., Bonadies, N., Allen, M.D., Freund, S.M., Santiveri, C.M., Kvinlaug, B.T., Huntly, B.J., Gottgens, B., and Bycroft, M. (2010). Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1. *Nat. Struct. Mol. Biol.* *17*, 617–619.
- Wang, G.G., Song, J., Wang, Z., Dormann, H.L., Casadio, F., Li, H., Luo, J.L., Patel, D.J., and Allis, C.D. (2009a). Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. *Nature* *459*, 847–851.
- Wang, Y., Reddy, B., Thompson, J., Wang, H., Noma, K., Yates, J.R., III, and Jia, S. (2009b). Regulation of Set9-mediated H4K20 methylation by a PWWP domain protein. *Mol. Cell* *33*, 428–437.
- Wang, Z., Zang, C., Cui, K., Schones, D.E., Barski, A., Peng, W., and Zhao, K. (2009c). Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* *138*, 1019–1031.
- Winter, S., Fischle, W., and Seiser, C. (2008). Modulation of 14-3-3 interaction with phosphorylated histone H3 by combinatorial modification patterns. *Cell Cycle* *7*, 1336–1342.
- Wisniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. *Nat. Methods* *6*, 359–362.
- Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., et al. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* *442*, 86–90.
- Zhang, Y., Buchholz, F., Muirers, J.P., and Stewart, A.F. (1998). A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.* *20*, 123–128.