

Genomic context-dependent histone H3K36 methylation by three *Drosophila* methyltransferases and implications for dedicated chromatin readers

Muhunden Jayakrishnan¹, Magdalena Havlová², Václav Veverka^{1,2,3}, Catherine Regnard^{1,*} and Peter B. Becker^{1,*}

¹Biomedical Center, Molecular Biology Division, Ludwig-Maximilians-Universität, Munich, Germany

²Institute of Organic Chemistry and Biochemistry (IOCB) of the Czech Academy of Sciences, Prague, Czech Republic

³Department of Cell Biology, Faculty of Science, Charles University, Prague, Czech Republic

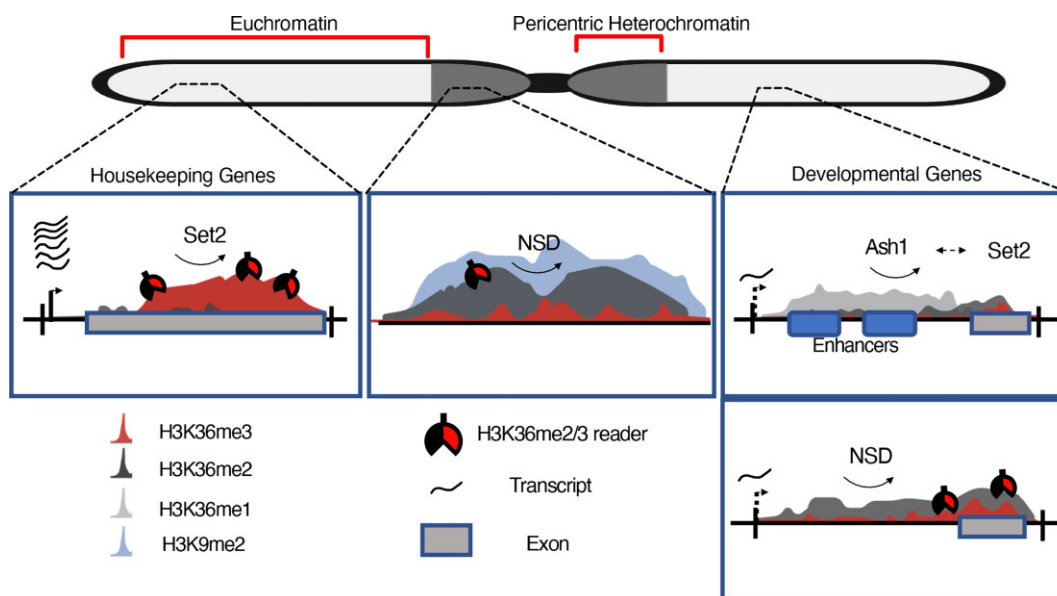
*To whom correspondence should be addressed. Tel: +49 89 2180 75427; Email: pbecker@bmc.med.lmu.de

Correspondence may also be addressed to Catherine Regnard. Tel: +49 89 2180 71602; Email: cregnard@bmc.med.lmu.de

Abstract

Methylation of histone H3 at lysine 36 (H3K36me3) marks active chromatin. The mark is interpreted by epigenetic readers that assist transcription and safeguard the integrity of the chromatin fiber. The chromodomain protein MSL3 binds H3K36me3 to target X-chromosomal genes in male *Drosophila* for dosage compensation. The PWWP-domain protein JASPer recruits the JIL1 kinase to active chromatin on all chromosomes. Unexpectedly, depletion of K36me3 had variable, locus-specific effects on the interactions of those readers. This observation motivated a systematic and comprehensive study of K36 methylation in a defined cellular model. Contrasting prevailing models, we found that K36me1, K36me2 and K36me3 each contribute to distinct chromatin states. A gene-centric view of the changing K36 methylation landscape upon depletion of the three methyltransferases Set2, NSD and Ash1 revealed local, context-specific methylation signatures. Set2 catalyzes K36me3 predominantly at transcriptionally active euchromatin. NSD places K36me2/3 at defined loci within pericentric heterochromatin and on weakly transcribed euchromatic genes. Ash1 deposits K36me1 at regions with enhancer signatures. The genome-wide mapping of MSL3 and JASPer suggested that they bind K36me2 in addition to K36me3, which was confirmed by direct affinity measurement. This dual specificity attracts the readers to a broader range of chromosomal locations and increases the robustness of their actions.

Graphical abstract



Introduction

Histone modifications highlight functional chromatin states. Particular chemical modifications inform about whether large

chromatin domains are permissive or repressive to transcription and demarcate regulatory sequences within these domains. Combinations of modifications, most prominently

Received: January 24, 2024. Revised: May 3, 2024. Editorial Decision: May 7, 2024. Accepted: May 28, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

acetylation and methylation of specific histone lysines, constitute epigenomic signatures that allow interpreting eukaryotic genomes (1–3).

A prominent example of a histone modification that is associated with a specific function is the trimethylation of lysine 36 of histone H3 (H3K36me₃; K36me₃ for short), which maps to bodies of transcribed genes. This mark correlates to active transcription in all eukaryotes from yeast to mammals. This methylation signals the ‘transcribed’ state of chromatin to localize a wide variety of processes that go along with active transcription, such as splicing (4,5), nucleosome turnover (6) deacetylation (7,8), DNA repair (9–11), RNA methylation (12), establishment of DNA methylation (13,14) and counteraction of facultative heterochromatin spreading (15,16), as well as regulation of *S*-adenosyl methionine flux (17,18). For example, K36me₃ recruits enzymes to reinstate the integrity of the nucleosome fiber during transcription and to repress cryptic, intragenic promoters in yeast (19). In human cells, H3K36me₃ can recruit DNA methyltransferases and decreased H3K36me₃ is associated with increased cryptic transcription during aging (20) as well as splicing defects (21). K36me₃ serves as an epigenetic mark that targets dosage compensation to X-linked genes in *Drosophila* males (22).

Drosophila has only two H3 variants and replacement of replication dependent H3.2 or the replication independent H3.3 with non-methylatable K36R mutants have allowed to study the role of K36 in development (23–25). H3.2K36 accounts for the majority of H3K36me₃, is essential for viability and its loss mostly affects post-transcriptional processing of mRNA (26). Intriguingly, the same study showed that splicing and transcription are only mildly affected, with observed defects frequently not restricted to K36me₃ regions (26), contrasting previous reports regarding involvement of K36me₃ in splicing (21) and repression of spurious transcription (19). In contrast, H3.3K36 accounts for ~20% of H3K36me₃ and supports longevity (27). The two variants may be redundant under certain contexts, for example during X-chromosome dosage compensation (28).

The effects of histone methylation are usually mediated by ‘reader’ proteins, i.e. proteins that contain domains that recognize and bind specifically to methylated lysines. As subunits of larger enzyme complexes, these reader proteins tether epigenetic regulators to specific chromatin regions (29). Reader domains for H3K36 methylation include PWWP (30,31), chromo (32,33) and tudor domains (34). We are interested in two *Drosophila* proteins for which K36me₃ seems to be the main determinant for genomic distribution: MSL3, the reader subunit of the male-specific-lethal dosage compensation complex (DCC) that boosts the transcription on the X chromosome in male flies (35) and JASPer, the targeting subunit of the JJ-complex responsible for interphase H3S10 phosphorylation which is also enriched on the male X-chromosome (36).

Lysines may be modified by one, two or three methyl groups. Whereas K36me₃ has been extensively studied, K36me_{1/2} is less understood. The prevailing view, mainly based on characterization of the distribution of the H3K36 methylation marks in yeast, was that K36me₃ corresponded to the true functional state, while K36me_{1/2} were mostly methylation intermediates.

Histone methylation profiles are established by dedicated histone methyltransferases (HMTs). In mammals, H3K36 HMTs are to some extent redundant. In vivo, SETD2 cat-

alyzes K36me₃, whereas NSD1-3 and ASH1L (and possibly other enzymes such as SETMAR or SETD3) contribute to K36me_{1/2} (17). In *Drosophila*, only three evolutionary conserved enzymes methylate K36, namely Set2 (alias dHypb), NSD (alias dMes-4) and Ash1. Set2 (orthologous to mammalian SETD2) is an interactor of the elongating RNA polymerase II and is thought to catalyze all K36me₃ (35,37). NSD (orthologous to mammalian NSD1/2/3) has been reported to localize to both eu- and heterochromatin and is known predominantly as a mono- and dimethyltransferase (37,38). Finally, Ash1 (orthologous to mammalian ASH1L) is described as dimethyltransferase whose role is developmentally restricted to a few hundred genomic loci to counteract Polycomb repression (39–41). Interestingly, some of the effects observed upon deletion of Set2 or other H3K36 methyltransferases have not been recapitulated by histone replacement by their non-methylatable counterparts. For example, while Set2 deletion leads to imbalanced gene expression of X-linked genes in male flies, the effect is absent in flies with H3.2K36R replacement (28). A similar study also showed that excessive Polycomb repression of *HOX* genes observed upon Ash1 loss is not recapitulated in K36R flies (25). These observations may hint at the possibility that some of the effects attributed to H3K36 methylation may be due to methylation of other non-histone proteins.

According to the prevailing model of the division of labor, NSD first ‘premodifies’ chromatin with K36me_{1/2}, which then serves as a substrate for Set2 to co-transcriptionally generate K36me₃, establishing a linear relationship between the two enzymes (37,42,43). Ash1 is believed to function independently of the Set2-NSD tandem (41,44). This simple view is currently being challenged by several independent observations. In *Drosophila*, K36me₃, me₂ and me₁ are differently distributed, implying distinct functions for each methylation state (2,28). Furthermore, the transcriptional and developmental consequences of NSD deletion are distinct from that of Set2 deficiency (25,28). For instance, Set2-null flies die during late larval stage, while NSD-null flies are viable and fertile (albeit with minor defects), arguing against an exclusive linear relationship of the two enzymes towards generating the trimethylated state (25,35,37). Interpretation of these organismal phenotypes is complicated by suspected cell type-specific and developmental effects.

To more systematically evaluate the function of the three *Drosophila* HMTs and to characterize the different patterns of the K36 methylation states at high resolution, we resorted to a defined tissue culture model. We mapped the individual K36 methylation states at high resolution using an enhanced ChIP protocol that combines the resolution of Micrococcal Nuclease digestion with solubilization of heterochromatin by mild shearing (36). We describe the consequences of systematic HMT depletion, individual and in combination, on the K36 methylation landscape and highlight differences in local HMT dependencies. In contrast to the prevailing model, we find that the three enzymes establish distinct methylation states at defined genomic regions.

We also monitored the distribution of the two K36me₃ reader proteins MSL3 and JASPer upon HMT depletion and showed that reduction of K36me₃ did not lead to a proportional reduction in reader’s binding. We propose that the robustness of the reader’s recruitment is at least in part due to their ability to bind both K36me₂ and K36me₃ combined with threshold effects and influence of the chromatin context.

Our systematic study uncovers an unexpected complexity of the three different H3K36 methylation states, which likely have different functions depending on the local chromatin context.

Materials and methods

Cell culture and RNAi

S2 and Kc167 cells were obtained from *Drosophila* Genomic Resource Center (DGRC) and were cultured in Schneider's *Drosophila* Medium (Thermo Fisher), supplemented with 10% heat-inactivated Fetal Bovine Serum (FBS, Sigma-Aldrich), 100 units/mL penicillin and 0.1 mg/mL streptomycin (Sigma-Aldrich) at 26°C. RNAi against target genes was performed either in large scale (for ChIP-seq) or small scale (Western Blots, Immunofluorescence staining, RNA extraction):

Large scale

1×10^6 S2 cells were seeded per well of a 6-well dish (Cellstar) (two wells per RNAi condition). Cells were washed once in serum-free medium followed by treatment with 10 µg dsRNA/well/RNAi target in 1 ml serum-free medium. Two different dsRNAs were used for each target to improve RNAi efficiency. Cells were incubated for 10 min at room temperature (RT) with slight agitation and further 50 min at 26°C. Two volumes of complete growth medium were added and cells were incubated for 5 days at 26°C. At day 5, cells were resuspended and counted. 2×10^7 cells per RNAi condition were transferred to a 75 cm² flask for a second round of dsRNA treatment (80 µg dsRNA/flask/RNAi target) in 8 ml serum-free media and incubated as mentioned above. At day 10, cells were counted and processed for ChIP-seq. For Ash1, to improve knockdown efficiency, three rounds of knockdowns were performed for 12 days as described by (45).

Small scale

3×10^5 S2 cells were seeded per well of a 12-well dish (Cellstar). Once adhered, cells were washed once in serum-free medium followed by treatment with 3–4 µg dsRNA/well/RNAi target in 300 µl serum-free medium. Cells were incubated for 10 min at RT with slight agitation and further 50 min at 26°C. Two volumes of complete growth medium were added and cells were incubated for 5 days at 26°C. At day 5, cells were resuspended and counted. 1.5×10^6 cells per RNAi condition were transferred to a 6-well dish for a second round of dsRNA treatment (10 µg dsRNA/flask/RNAi target) in 1 ml serum free media and incubated as mentioned above. At day 10, cells were counted and processed for Western Blot or Immunofluorescence.

dsRNA was generated by *in vitro* transcription (NEB T7 E2050s transcription kit) from PCR products generated by the following forward and reverse primers (separated by comma):

- *gst* RNAi: TTAATACGACTCACTATAGGGAGAATGTCCCCTATACTAGGTTA, TTAATACGACTCATATAGGGAGAACGCATCCAGGCACATTG; The sequence of *Schistosoma japonicum* GST was amplified from pGEX-4T-1 (GE Healthcare)
- *gfp* RNAi: TTAATACGACTCACTATAGGGTGTCTCAGGTAGTGGTTGTCG, TTAATACGACTCACTATAGGGCTGAAGTTCATCTGCACCA;

- Set2_#1 RNAi: TTAATACGACTCACTATAG GGAGAAAATCCTTGATTCCAAGCAA, TTAATACGACTCACTATAGGGAGAAGTGGTTTCTACATTTTCGT;
- Set2_#2 RNAi: TTAATACGACTCACTATAGGGAGACACGGCTTGAGATTGCTACA, TTAATACGACTCACTATAGGGAGACATGGACATGCTTTTGTGG;
- NSD_#1 RNAi: TTAATACGACTCACTATAGGGAGACGCAATTCTGAGCACGGACGCGCACTC, TTAATACGACTCACTATAGGGAGACGCTCTAGATGGACACACGCTGTTGTTGCTGTTT;
- NSD_#2 RNAi: TTAATACGACTCACTATAGGGAGACCTCCTCTGTGAGCATCGA, TTAATACGACTCACTATAGGGAGAACAACGTTTTTCGTACGTCTGG;
- Ash1_#1 RNAi: TTAATACGACTCACTATAGGGAGACTTTGTGGCCAGGACCAATCAA, TTAATACGACTCACTATAGGGAGACAGGCAAGGGATCGTGCTCGGT;
- Ash1_#2 RNAi: CTAATACGACTCACTATAGGGAGGACAGTCCATGGAGACCC, CTAATACGACTCACTATAGGGAGCAACACCCAGCAGCGTCC

D. virilis 79f7Dv3 cells (36) were cultured in Schneider's *Drosophila* Medium (Thermo Fisher), supplemented with 5% heat-inactivated FBS (Sigma-Aldrich), 100 units/ml penicillin and 0.1 mg/ml streptomycin (Sigma-Aldrich) at 26°C.

Immunofluorescence microscopy and analysis

Immunofluorescence microscopy analysis of S2 cells was performed as described in (46). Briefly, cells after RNAi were seeded on poly-L-lysine (Sigma #P8920, 0.01% (w/v) final concentration) coated coverslips and allowed to adhere for 1h. Coverslips were then fixed, permeabilized, blocked and incubated in primary antibodies overnight at appropriate dilutions (see table). Secondary antibody staining was performed the next day after which coverslips were incubated in DAPI and sealed onto glass slides with Vectashield mounting reagent.

Confocal images were acquired on a Leica TCS SP8 with a 63x/1.4NA oil-immersion objective. Image stacks were recorded at 100 Hz scan speed with a pixel size of 350 nm (or 50 nm for Figure S1B) and z-step size of 300 nm. Pinhole was set to 1 AU (580 nm reference wavelength). Fluorescence signals were recorded sequentially to avoid channel crosstalk. Further image processing and maximum intensity projections were done in Fiji (47). CellProfiler (48) was used to quantify histone modification immunostaining signals within DAPI defined regions (pipeline provided) and further plotted on R.

Western blot

2–3 × 10⁶ RNAi-treated cells were pelleted and lysed in 1 × Laemmli Buffer at a concentration of 25 000 cells/µl. Samples were denatured 95°C for 10 mins. 6–8 µl of lysate per sample was electrophoresed on SDS ServaGel TGPrimer (14% gel for histone modifications, 8% gel for other proteins) for 1.5–2 h at 180 V. Proteins were transferred to AmershamTM ProtranTM 0.45 µM Nitrocellulose Blotting Membrane for 1.5 h at 300–400 mA in either regular transfer buffer (20% MeOH, 25 mM Tris, 192 mM Glycine) or high molecular weight transfer buffer for methyltransferases (10% MeOH, 0.037% SDS, 25 mM Tris, 192 mM Glycine). Membranes were blocked with 3% BSA for 1h at RT. The membrane was

incubated with primary antibody (see table) overnight at 4°C in 3% BSA PBS washed thrice with PBS-T (1x Phosphate-Buffered Saline, 0.1% Tween-20) and incubated with secondary antibody in PBS-T for 1 h in RT. Images were acquired and quantified using the LICOR Odyssey CLx. For all quantifications, we verified that the signal is linear in the range of loading used (not shown). However, due to blot-to-blot variability, the normalized signals should be interpreted as an estimate rather than an absolute quantification of protein levels.

RNA isolation, cDNA synthesis and RT-qPCR

Total RNA was extracted from 1.5×10^6 cells using the RNeasy Mini kit (Qiagen) according to the manufacturer's instructions. cDNA was synthesized from 500 ng of total RNA using Superscript III First Strand Synthesis System (Invitrogen, Cat. No 18080–051, random hexamer priming) and following standard protocol. cDNA was diluted 1:100, qPCR reaction was assembled using Fast SYBR Green Mastermix (Applied Biosystem, Cat. No 4385612) and ran on a Lightcycler 480 II (Roche) instrument. Primer efficiencies were calculated via serial dilutions. Primer sequences for Ash1 and 7sk (control) were obtained from (41) and (49), respectively.

Chromatin immunoprecipitation after MNase treatment (MNase ChIP-seq) with spike-in

ChIP-seq on MNase-digested chromatin and sonicated chromatin was performed as previously described (36). For spike-in ChIP-seq on MNase-digested chromatin in combination with mild sonication, S2 cells ($\sim 1.2 \times 10^8$ cells) after RNAi were harvested and cross-linked for 8 min by adding 1.1 ml $10 \times$ fixing solution (50 mM HEPES pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA and 10% (w/v) methanol-free formaldehyde) per 10 ml culture at RT. The reaction was stopped by adding glycine at 125 mM final concentration and incubating for 10 min on ice. Cells were washed twice in PBS before subsequent steps. For nuclei isolation, S2 cells were spiked with 5% (relative cell number) 79f7Dv3 fixed cells (processed as described for S2 cells), resuspended in PBS supplemented with 0.5% (v/v) Triton X-100 and complete EDTA-free Protease Inhibitor Cocktail (Roche) (PI), volume was adjusted to 7×10^7 cells/ml and cells incubated for 15 min at 4°C with end-over-end rotation. Nuclei were collected by centrifuging at 4°C for 10 min at $2000 \times g$ and washed once in PBS. For chromatin fragmentation, nuclei were spun down at 4°C for 10 min at $2000 \times g$, resuspended in RIPA (10 mM Tris/HCl pH 8.0, 140 mM NaCl, 1 mM EDTA, 1% (v/v) Triton-X 100, 0.1% (v/v) SDS, 0.1% (v/v) Sodium deoxycholate) supplemented with PI and 2 mM CaCl_2 at 7×10^7 cells/ml. These lysates were digested in 1 ml aliquots by adding 0.6 U MNase (Sigma Aldrich), resuspended in EX-50 (50 mM KCl, 10 mM HEPES pH 7.6, 1.5 mM MgCl_2 , 0.5 mM EGTA, 10% glycerol) at 0.6 U/ μl , and incubated at 37°C for 35 min with slight agitation. The reaction was stopped by adding 10 mM EGTA and placing on ice. Digested chromatin was mildly sheared further with Covaris AFA S220 using 12×12 tubes at 50 W peak incident power, 20% duty factor and 200 cycles per burst for 8 min at 5°C. We noticed that this step results in a more uniform solubilization of chromatin across the entire genome compared

to protocols that use either MNase-only (50) or shearing-only (41,51), based on comparison of respective chromatin input coverages, and thus is more suited for quantitative comparisons between eu- and heterochromatic ChIP signals. Subsequent steps were performed as described in (36). Libraries were prepared with NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, E7645) and analyzed with either 2100 Bioanalyzer or TapeStation systems (Agilent). Libraries were sequenced on NextSeq1000 (Illumina) instrument yielding typically 20–25 million 150 bp paired-end reads per sample.

CUT&RUN

CUT&RUN was performed as previously described (52) with the Rb α -H3K36me3 antibody diluted 1/500.

Library preparation

For ChIP-seq samples, libraries were prepared using NEBNext Ultra II DNA Library with a starting ChIP DNA amount of 3–6 ng according to manufacturer's instructions. All libraries were sequenced on an Illumina NextSeq1000 sequencer at the Laboratory of Functional Genomic Analysis (LAFUGA, Gene Center Munich, LMU). About 20 million paired-end reads were sequenced per sample for each of the ChIP samples.

Antibodies

Antibody	Western	IF	ChIP	Supplier/reference
Rb α -H3K36me1	1:5000	1:800	2 μl	Abcam-ab9048
Rb α -H3K36me2	1:5000	1:800	2 μl	Abcam-ab9049
Rb α -H3K36me3	1:5000	1:800	2 μl	Abcam-ab9050
Rb α -H4K16ac			2 μl	Millipore 07-329
Ms α -H3K9me2			2 μl	Abcam-ab1220
Rb α -H3K27me3			2 μl	Millipore 07-449
Rb α -H2AV	1:500			Börner and Becker (53)
Gp α -MSL3			3 μl	Albig <i>et al.</i> (36)
Gp α -JASPer			3 μl	Albig <i>et al.</i> (36)
Rb α -NSD	1:1000		5 μl	Bell <i>et al.</i> (37)
Ms α -NSD	1:50	1:10		Bell <i>et al.</i> (37)
Ms α -Set2	1:50			Bell <i>et al.</i> (37)
RBP1-S2ph (ePol)			2 μl	Abcam-ab5095
Lamin	1:1000			Gift from H Saumweber

Recombinant protein expression and purification

JASPer full length sequence was cloned into pET derived T7 driven vector (ampicillin resistance) behind His₆ affinity tag and TEV (Tobacco Etch Virus) protease cleavage site. After the cleavage, a cloning artefact of five amino acids (SNAAS) remained at the N-terminus of JASPer.

Chemical competent NiCo21 pRARE2 (DE3) *Escherichia coli* (NEB C2529H with additional plasmid coding rare codons and Chloramphenicol resistance) were transformed by heat shock with JASPer plasmid and cultivated on LB agar medium (100 $\mu\text{g/ml}$ ampicillin, 25 $\mu\text{g/ml}$ chloramphenicol and 1% glucose) overnight in 37°C. Colonies were resuspended in LB broth (Sigma) supplemented with 100 $\mu\text{g/ml}$ ampicillin, 25 $\mu\text{g/ml}$ chloramphenicol and 0.5% glycerol to $\text{OD}_{600} = 0.05$. Bacterial cultures were cultivated in LEX Epiphyte 3 bioreactor at 37°C until $\text{OD}_{600} = 0.8$ and then transferred to 18°C. Protein expression was induced with 400 μM IPTG (isopropyl- β -D-thiogalactopyranoside). The cells were

harvested after 20 h expression by centrifugation (5000 g, 4°C, 20 min), were stored in -80°C.

Bacterial cells were lysated by Emulsiflex C3 in lysis buffer (25 mM Tris-HCl pH 7.5, 1 M NaCl, 2 mM β -mercaptoethanol, 10 μ M EDTA) with protease inhibitor cocktail tablet (EDTA free) and DNase I. The lysate was cleared by centrifugation (25 000 g, 4°C, 30 min) and loaded on Ni-NTA-resin (Sigma). The bound protein (with His₆ tag) was then eluted by imidazole buffer (25 mM Tris-HCl pH 7.5, 1 M NaCl, 2 mM β -mercaptoethanol, 250 mM imidazole). The purest fractions were dialyzed in presence of TEV protease to lysis buffer overnight. JASPer without His₆ tag was present in the flow-through fraction in the second Ni affinity chromatography. JASPer sample was concentrated using an Amicon Ultra centrifugal filtration unit and loaded onto size exclusion chromatography Superdex 200 10/300 GL in buffer 25 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM TCEP, 10 μ M EDTA). The final concentration of the purified protein was determined on a NanoDrop spectrophotometer, and its purity was evaluated by SDS-PAGE stained by Coomassie Brilliant Blue.

Nucleosomes were prepared as described in (31).

MicroScale Thermophoresis (MST)

JASPer protein (1-475) was dialysed overnight to MST buffer without Tween (20 mM HEPES pH 7.5, 50 mM NaCl, 1 mM TCEP). The stock of exchanged protein was divided in to equal aliquots and stored in -80°C. After thawing, each aliquot was centrifuged (20 000 g, 4°C, 10 min), and the protein concentration was re-measured. JASPer at constant concentration 0.4 μ M was mixed in 16-step dilution series (2:1 ratio, 0.5 nM and reaching up to 16 μ M) with reconstituted nucleosomes (147 bp H3K36me0/2/3). The samples were adjusted to final volume of 30 μ l MST buffer (20 mM HEPES pH 7.5, 50 mM NaCl, 1 mM TCEP, 0.05% Tween-20). All reaction tubes were centrifuged and incubated for 15 min at room temperature. Mixtures were loaded into Monolith NT.LabelFree premium capillaries and measured at 20% LED Power and 40% MST Power on a NanoTemper Monolith NT.LabelFree device at 24°C. Each mixture was prepared twice, where each replicate was measured in technical triplicates. Values were fitted in the GraphPad Prism using a built-in equation for specific binding with Hill slope.

Data analysis

Read processing

ChIP-seq Sequence reads were Demultiplexed by JE demultiplexer (54) using the barcodes from the Illumina Index read files. Demultiplexed files were aligned to either *D. melanogaster* reference genome (BDGP6, release 104) or independently to *D. virilis* genome (droVir3, Feb 2006) using Bowtie2 (55) version 2.28.0 (parameter '-end-to-end -very-sensitive -no-unal -no-mixed -no-discordant -X 500 -I 10') and filtered for quality using samtools 1.6 (56) with a MAPQ score cutoff of -q 10. For transposons, a custom genome containing repetitive regions was used as before (36) at the alignment step. Tag directories and input-normalized coverage files were generated using Homer (57) with the parameter -totalReads set to the number of reads mapped to *D. virilis* genome for spike-in normalization. Input-normalized, scaled

D. melanogaster coverage per base pair files were visualized using the Integrative Genomics Viewer (58). Replicate coverages were first analyzed independently to confirm similarities in HMT-dependency patterns after which they were averaged for subsequent analyses. Resizing of coverages to fixed window sizes of mean signal was performed with bedops and bedmap (59).

Published RNA-seq reads were obtained from respective sources using sra prefetch and processed according to (36). Single-end parameters were used for read alignment by STAR for Huang *et al.* (41) and paired-end settings for Depierre *et al.* (51).

ChIP-seq peak calling and annotation

Broad domains of modified H3K36me1/2/3, H327me3 and H3K9me2 were called using MACS2 v2.1.2 (60) bdgpeakcall function using parameters -l 1000 (-c 3 for K36me1/2/3; -c 0.8 for K27me3; -c 2.0 for K9me2). Manipulation (filtering, merging etc.) of peak sets was performed with BEDTOOLS2 v2.28.0. Genomic annotation of peaks was done using HOMER annotatePeaks.pl script.

Data analysis and plotting

Data analysis was conducted in R (61) (R Core Team, 2020) using tidyverse libraries. ChromoMaps were generated using R chromoMap package (62). Clustered heatmaps were made using R package 'ComplexHeatmap' (63). Chromatin State Annotations were derived from modENCODE (2). Gene annotations were obtained from FlyBase GTF annotations BDGP6 release 104. Only genes associated with a unique FlyBase (FBgn) ID ($N = 17.8k$) were used for subsequent analyses. Detailed explanation for certain analyses can be found within provided R scripts.

Differential binding analysis

To identify significant differential K36me1/2/3 regions across RNAi conditions, csaw v1.24.3 (64) was used. Reads for each unique experimental sample were counted into sliding windows of 250 bp across the entire genome. Windows were filtered using filterWindowsControl function using the corresponding Input profiles as controls. Data were binned into larger 10-kbp bins and subsequently used to calculate normalization factors using normFactors function. Differential binding was assessed for using the quasi-likelihood (QL) framework in the edgeR package v3.32.1 with robust = TRUE for glmQLFit. The design matrix was constructed using a layout specifying the RNAi treatment as well as the experimental batch. Proximal tested windows were merged into regions of maximum 3 kb by clusterWindows with a cluster level FDR target of 0.05.

Z-score transformation

Z-score data transformation for average signal within 5-kb genomic windows (Figure 3) or genes (Figure 6) was performed similar to previously described method (50). Average feature coverage instead of RPKM was used in calculations. A consensus peak set representing all genomic regions containing at least one K36 modification in any RNAi condition was further used to filter windows before Z-score representation in Figure 3.

External datasets

Dataset	Source
Ash1 ChIP (S2 cells)	Huang <i>et al.</i> (41)
H3K36me2 ChIP in Ash1 RNAi (S2 cells)	Huang <i>et al.</i> (41)
RNAseq in Ash1 RNAi (S2 cells)	Huang <i>et al.</i> (41)
Chromatin States	Kharchenko <i>et al.</i> (2)
H3K36me3 modENCODE (S2 cells)	modENCODE (GSE20785)
H3K36me3 Native-ChIP (Fly)	Chaouch <i>et al.</i> (50)
H3K36me3 modified MNase ChIP (S2 cells)	Albig <i>et al.</i> (36)
Su(var)3–7	modENCODE (GSE23487)
H3K9me2	modENCODE(GSE20792)
H4K20me1	modENCODE(GSE27743)
H3K27ac	modENCODE(GSE20779)
H3K4me1	modENCODE(GSE32826)
Nurf	modENCODE(GSE20829)
H2B-ub	modENCODE(GSE20773)
H3K79me3	modENCODE(GSE45090)
Kdm4A	modENCODE(GSE32839)
Kdm2	modENCODE(GSE45061)
MRG15	modENCODE(GSE25367)
Beaf-32	modENCODE(GSE20760)
SMC3	modENCODE(GSE45054)
CTCF	modENCODE(GSE32750)
H3K27me3	modENCODE(GSE20781)
H3.3	Henikoff <i>et al.</i> (65)
SNR1	Hendy <i>et al.</i> (66)
ISWI	Hendy <i>et al.</i> (66)
RNAseq in NSD/Set2 RNAi (S2 cells)	Depierre <i>et al.</i> (51)
WT S2 cells RNAseq	Albig <i>et al.</i> (36)
Fly HMT KO RNAseq	Lindehell <i>et al.</i> (28)
FlyAtlas Tissue Expression Data	Chintapalli <i>et al.</i> (67)
H3K36me2/3 in HMT KO (C3H10T cells)	Weinberg <i>et al.</i> (13)
RNAseq of WT C3H10T cells	Weinberg <i>et al.</i> (13)
H3K36me3 in HMT dTAG cells (mESC)	Sun <i>et al.</i> (68)

Genomic coordinates of datasets previously aligned to genome version dm3 were transformed to newer dm6 using liftOver tool (69).

Results

H3K36 methylation states and methyltransferases mark distinct chromatin types in S2 cells

To study the genomic distribution of the histone H3K36 methylation states (and their corresponding HMTs, we generated high-resolution MNase chromatin immunoprecipitation (ChIP-seq) profiles in male *Drosophila* S2 cells. Our optimized protocol includes the solubilization of chromatin through mild shearing, yielding significantly more signals at mappable heterochromatin compared to other methods (Supplementary Figure S1A).

We mapped K36me1/2/3 using highly specific antibodies (70), the HMT NSD and the elongating RNA polymerase II (ePol) as a proxy for Set2 (for which no ChIP-grade antibodies exist). A recently published Ash1 ChIP profile (41) was included. We also mapped two *bona fide* K36me3 binders: JASPer, a PWWP-domain reader that colocalizes with K36me3 genome-wide and MSL3, a K36me3 chromodomain reader restricted to the X chromosome (36). The browser views (Figure 1A) reveal the distinct distributions of consid-

ered features, in particular for K36me1/2 and 3. To visualize the differences in the distribution of the K36 methylation states as well as the associated factors at a chromosome-wide level, we generated chromoMaps representing average ChIP signals in 10-kbp bins superimposed on the chromosomes (Figure 1B, Supplementary Figure S1B). Tracks depicting annotated genes as well as K9me2 peaks highlighting the gene-poor mappable pericentric heterochromatic regions (PCH) were included for reference. Of note, the PCH region is not prominently represented on the X chromosome because it is poorly mapped in the dm6 reference genome. K36me1 is largely euchromatic, and only partly overlaps with K36me2/3. K36me2 is heterochromatic, but also present in stretches of euchromatin. Surprisingly, K36me3 is distributed rather uniformly between euchromatin and PCH, where it is not restricted to genes (Supplementary Figure S1A). Set2 (inferred from ePol) and Ash1 are largely enriched at euchromatin, while NSD showed a clear heterochromatin enrichment but is also found in euchromatic regions. The two K36me3 reader proteins JASPer and MSL3 correlate best with K36me3 as already described (Figure 1A). Of note, the chromoMaps of the X chromosome for MSL3 and JASPer predominantly illustrate the X-chromosomal enrichment due to the association to dosage compensation.

We called peaks to estimate the proportion of genome decorated by each K36me mark. K36me2 and K36me3 occupy a similar proportion of the genome (~20%) while K36me1 is concentrated more locally (~12%). Correlating the coverage to estimated absolute modification numbers derived from late-stage embryos (71) suggests that K36me1 is the most densely placed mark, followed by K36me2 and then K36me3 (Supplementary Figure S1C).

To examine the distribution of the marks and the associated factors at a higher resolution, we next correlated ChIP enrichments to each of the nine functional Chromatin States as defined by the modENCODE consortium (Figure 1C). We included various modENCODE ChIP datasets as well as more recently published profiles generated from published data to highlight features distinguishing the 9 States. This revealed several interesting patterns confirming the chromoMap observations. K36me3 is strongly enriched within euchromatic States 1 and 2 representing active promoters and transcribed chromatin, but also shows a mild heterochromatin enrichment (State 7). K36me2 marks similar states but is much more abundant at heterochromatin. In contrast, K36me1 is almost exclusively present at enhancer-like Chromatin States 3 and 4.

Correlating the HMT occupancies to the different K36 methylation states showed that Ash1 is most strongly enriched with K36me1 at enhancers (State 3). NSD is distributed between eu- and heterochromatin (States 1 and 7). Set2 (ePol) is rather restricted to euchromatin at transcribed regions (State 1 and 2) and enhancers (State 3) but rather not correlating with the mild K36me3 enrichment in heterochromatin (State 7). The dual distribution of NSD was confirmed by immunofluorescence confocal microscopy (Supplementary Figure S1D). In addition to a diffuse signal in euchromatin, intense speckles adjacent to (but not overlapping) DAPI-dense chromocenters were observed. Both populations were sensitive to RNA interference against NSD, confirming the specificity of the antibody.

We then intersected the peaks of K36me1/2/3 with genomic features considering the strongest peaks (top 10 percent) to highlight differences (Figure 1D). K36me1 is strongly

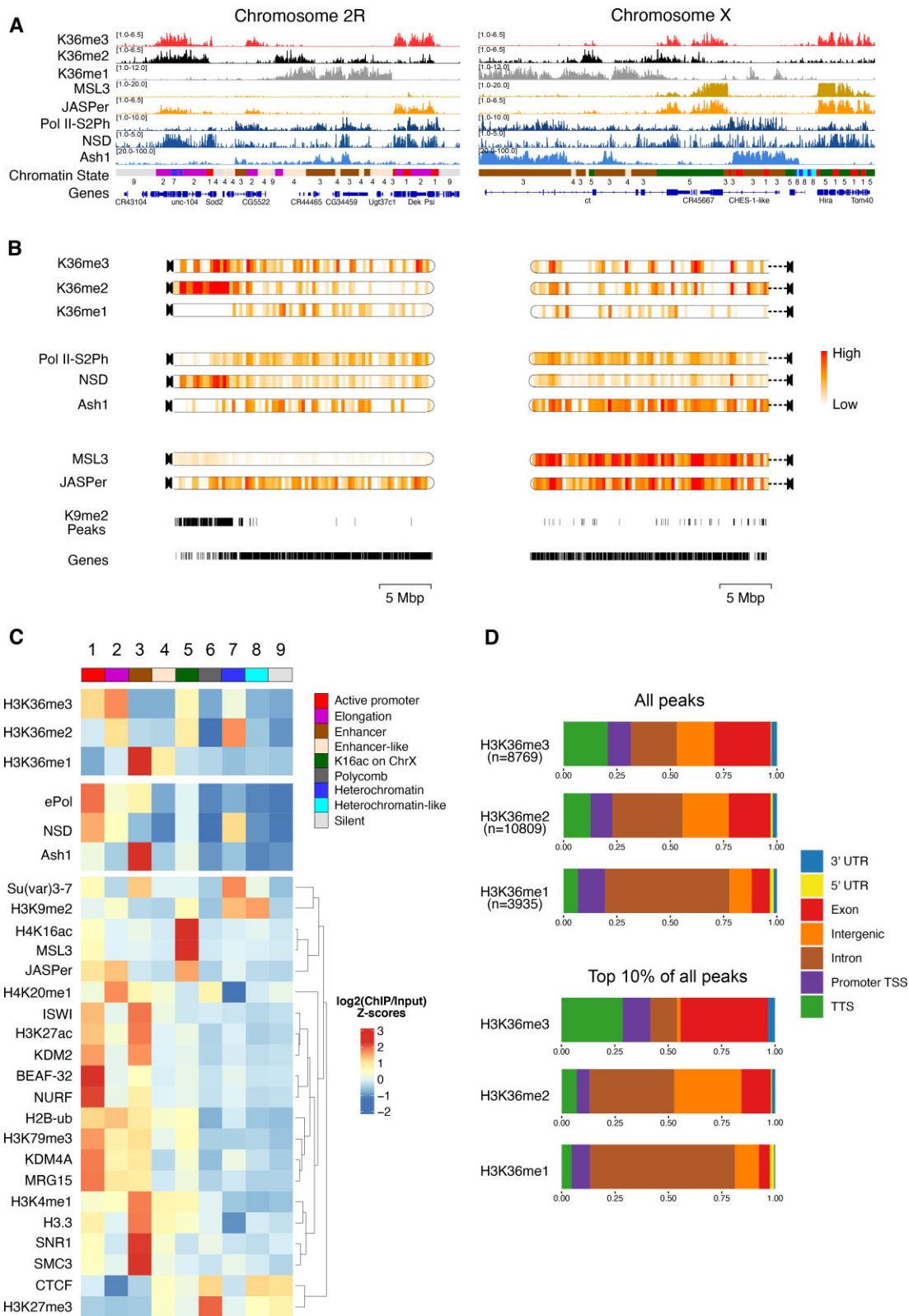


Figure 1. H3K36 modifications and HMTs mark distinct chromatin states in male *Drosophila* cells. **(A)** Genome browser profiles for representative *Drosophila* chromosomes 2R and X of spike-normalized MNase-ChIP of K36me1/2/3, the K36me3 reader proteins JASPer and MSL3 and the HMTs NSD and Ash1. ‘ePol’ refers to the signal generated by an antibody against RBP1-S2ph, as a proxy for RNA-polymerase-S2ph-interacting Set2. The Ash1 profile was taken from (41). The 9-state ChromHMM (modENCODE) is color-coded and explained in (C). **(B)** ChromoMaps representing steady-state enrichment of K36me1/2/3, HMTs and K36me3 readers in 10-kbp genomic bins for chromosomes 2R and X. Scale bars are different for each chromoMap to facilitate visualization of changes. The individual scale values for each state and chromosome are shown in Supplementary Figure S1. Tracks representing H3K9me2 peaks and annotated genes serve as a reference for mappable pericentric heterochromatin domains and transcribed chromatin, respectively. **(C)** Chromatin State enrichment (9-state ChromHMM, (2)) for K36me1/2/3, HMTs and K36me3 readers. Published ChIP-seq/CUT&RUN profiles of histone modifications and other chromatin proteins were hierarchically clustered to highlight differences between chromatin states. **(D)** Genomic features marked by either all K36me1/2/3 domains (top) or filtered for the strongest 10% (bottom).

enriched in introns, in agreement with its localization to enhancers. Strong K36me3 signals accumulate at TTS and exons, whereas K36me2 marks mostly genic features in particular introns but also some intergenic regions. In conclusion, different H3K36 methylation states label distinct types of chromatin and genomic features. NSD stands out due to its association with both eu- and heterochromatin.

Contribution of methyltransferases to different H3K36 methylation states

To determine the contribution of HMTs to the three K36 methylation states, we depleted each HMT individually as well as in combination by RNA interference (RNAi) and followed the levels of the HMTs and K36me1/2/3 by quantitative Western blotting. The depletion of Set2 and NSD was efficient (>90% and ~80%, respectively, [Supplementary Figure S2A](#)). Because the Ash1 antibody performed poorly in Western blotting, we estimated Ash1 mRNA levels by RT-qPCR too, which revealed a 75% reduction upon RNAi ([Supplementary Figure S2B](#)).

We first followed bulk changes in K36 methylation upon HMT RNAi (Figure 2A, representative blots in [Supplementary Figure S2C](#)). Loss of Set2 led to ~75% reduction of western blot signal for K36me3, while K36me2 appeared unaffected. Depletion of NSD strongly reduced Western blot signal for K36me2 (~80%), but led to only modest reduction of K36me3 signal (~40%) in agreement with previous reports (25,43). These observations are at odds with the idea that NSD only generates intermediates for Set2-dependent trimethylation. We also observed similar trends in female Kc cells, ruling out cell- or sex-specific effects ([Supplementary Figure S2D, E](#)).

Interestingly, K36me1 was rather unaffected in RNAi of Set2 and/or NSD, arguing against the idea that NSD is the major K36me1 methyltransferase. Partial depletion of Ash1 did not affect K36me3 levels, moderately decreased K36me2 by about 20%, but led to very prominent reduction of K36me1 (50%) as indicated by their corresponding Western blot signals. This reduction may be an underestimation, since the previously mentioned modENCODE study (70) reported that certain lots of K36me1 antibody have mild reactivity towards unmodified K36 in western blot assays.

We confirmed the Western blot trends by immunofluorescence microscopy (Figure 2B), quantifying the K36me1/2/3 signals upon HMT depletion (Figure 2C, [Supplementary Figure S2F](#)). K36me1 was exclusively decreased upon Ash1 RNAi. K36me2 was decreased reproducibly only in NSD RNAi while an Ash1-dependent decrease was observed in only one replicate ([Supplementary Figure S2F](#)). K36me3 was reduced in a Set2-dependent manner. The immunofluorescence staining did not reflect the reduction in K36me3 upon NSD RNAi found in western blot, possibly due to the sensitivity limit on one hand and to the local distribution of the changes on the other hand (see below).

These results suggest that K36me1/2 do not simply serve as intermediates of trimethylation. They also reveal a complex interplay between the three HMTs, where each individual enzyme predominantly catalyzes one K36 methylation state, and only partially contributes to the others. Given the different distributions of K36me1/2/3 and HMT's genome wide (Figure 1A–C), the differences may also be restricted to some genomic locations. We thus expect different phenotypes upon HMT loss of function mutants in flies. To explore this possibility, we reanalyzed published transcriptome data of NSD, Set2 and Ash1 knockout mutant fly brain (28), focusing on gene-level correlations. Indeed, the significantly deregulated genes poorly correlate between individual HMT mutants ([Supplementary Figure S3](#)), especially between Set2 and NSD, in support of the hypothesis of distinct roles of Set2 and NSD in gene regulation.

types upon HMT loss of function mutants in flies. To explore this possibility, we reanalyzed published transcriptome data of NSD, Set2 and Ash1 knockout mutant fly brain (28), focusing on gene-level correlations. Indeed, the significantly deregulated genes poorly correlate between individual HMT mutants ([Supplementary Figure S3](#)), especially between Set2 and NSD, in support of the hypothesis of distinct roles of Set2 and NSD in gene regulation.

Chromosome locus-specific H3K36 methylation changes upon HMT depletions

We next explored the genome-wide changes in K36 methylation after depletion of HMTs (Figure 3). Z-scores representing the direction and magnitude of change in K36me ChIP signal within 5-kb genomic bins were calculated for each HMT RNAi relative to the control (Figure 3A). To identify genomic regions that show significantly altered K36me1/2/3 upon RNAi, we adopted a statistical approach using the csaw pipeline (see Materials and methods). The differentially marked regions obtained from this analysis were further visualized using high-resolution 2-kb chromoMaps to provide information regarding spatial patterns of methylation changes in each RNAi condition. For conciseness, only chromosome 3L is shown (all chromosomes are shown in [Supplementary Figure S4A–C](#)). Profiles representing the respective K36me levels in unperturbed S2 cells above each chromoMap aid the interpretation of the relative changes while the gene annotation and H3K9me2 peak tracks illustrate local chromatin environment (Figure 3B–D).

The chromoMaps for K36me3 show that depletion of Set2 leads to a reduction of predominantly euchromatic K36me3 as previously described (35,37), while depletion of NSD leads to loss of K36me3 at gene-poor pericentric heterochromatin and a subset of euchromatic bins (distinct from Set2-dependent regions as well as euchromatic K9me2 nanodomains) (Figure 3B). A few heterochromatic bins, which were marked by NSD-dependent trimethylation, were identified as significantly increasing in the Set2 knockdown condition.

To directly compare the differential effect of HMT depletion at higher resolution we visualized pairwise correlations at genomic bins that contain at least one K36 modification as Z-score scatter plots (see principle in Figure 3A). Off-center dots indicate the bins affected by Set2 or NSD or by both depletions while bins centered around (0,0) represent unchanging bins which are typically 'background' regions of low K36me3 signal (but are likely decorated by K36me1/2) (Figure 3B, C). K36me3 is reduced in a large fraction of genomic regions upon depletion of Set2 only (top-left quadrant), consistent with its role as a major trimethyl-transferase. Conversely, NSD RNAi leads to decrease of K36me3 in distinct genomic regions (bottom-right quadrant) which are not affected upon Set2 depletion, likely corresponding to PCH regions (and few euchromatic bins). Interestingly, depletion of either enzyme led to a mild gain of K36me3 at genomic regions methylated by the other enzyme (indicated by the positive shift from the horizontal or vertical axis, respectively), a phenomenon we term the 'see-saw effect'. For example, in the absence of Set2 more K36me3 is found at pericentric heterochromatin, where NSD is strongly enriched (Figure 1). The see-saw effect can also be observed in several other instances. We assured that it was not due to a systematic bias in library preparation by verifying ChIP signals by qPCR at several diagnostic loci (not shown)

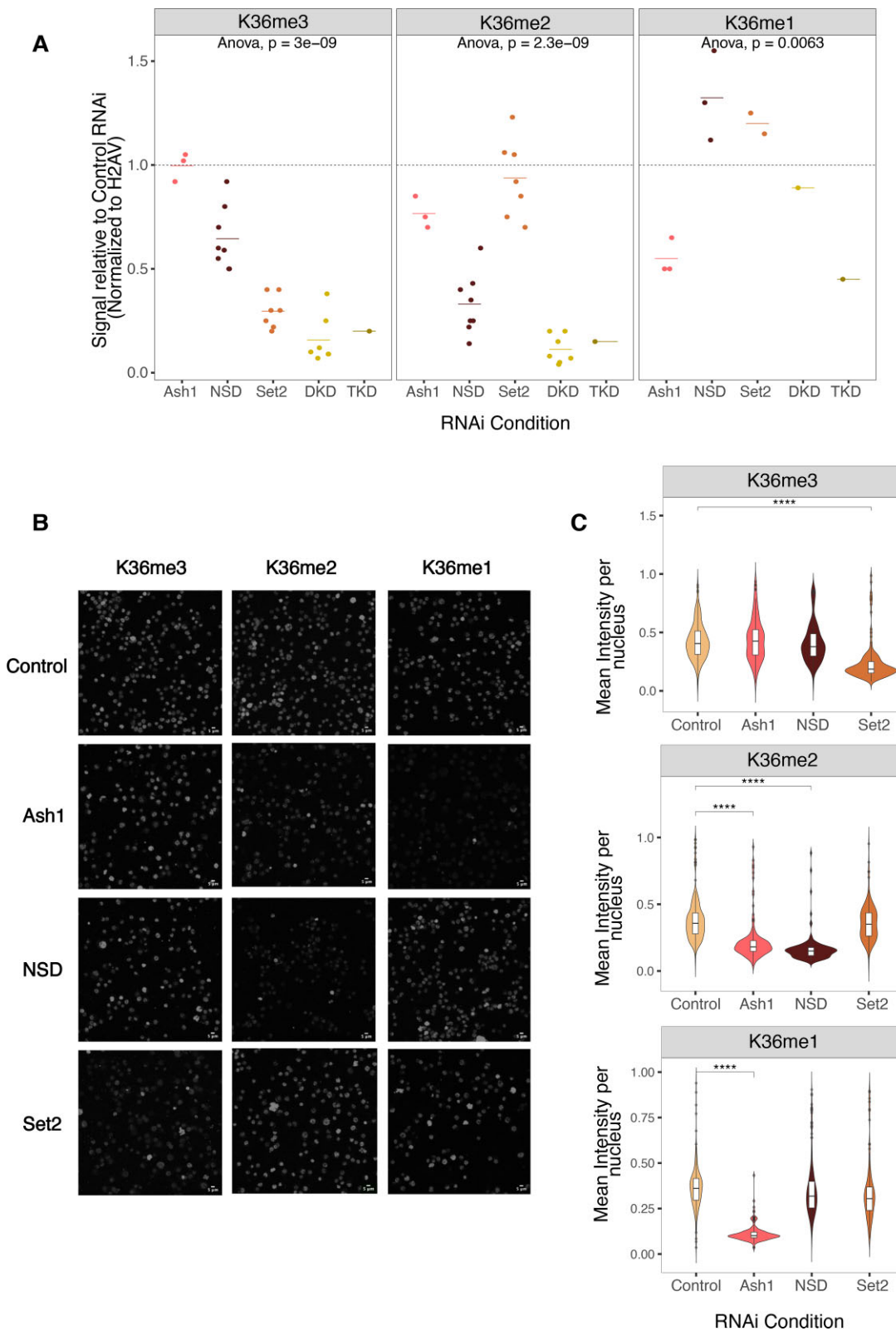


Figure 2. Effect of HMT depletion on H3K36 methylation bulk levels reveals distinct HMT dependencies in male cells. **(A)** K36me1/2/3 levels in whole cell extracts from S2 cells were determined by quantitative Western blotting using specific antibodies. Cells were treated with RNAi against Ash1, NSD, Set2, NSD + Set2 (DKD) or NSD + Set2 + Ash1(TKD). An irrelevant GST RNAi served as control. Representative blots are shown in [Supplementary Figure S2C](#); source data for all blots can be found in Zenodo Repository (see Data Availability). Values were normalized to histone H2AV signals on the same membrane and are represented as fraction relative to GST RNAi, which was run on same blot. Each dot represents an independent biological replicate. Calculated ANOVA p-values (null hypothesis: difference between means = 0) are presented for each antibody. **(B)** Representative immunofluorescence microscopy (IFM) images for α -H3K36me1/2/3 in S2 cells treated with RNAi against GST (Control), Ash1, NSD or Set2. The scale bar is 5 μ m. **(C)** Quantification of IFM images ($n = \sim 500$ nuclei from 1st biological replicate, shown in Figure 2B). ANOVA followed by post-hoc Tukey HSD was performed to identify groups with significantly different mean relative to GST RNAi (**** denotes a significance value of $P < 0.001$).

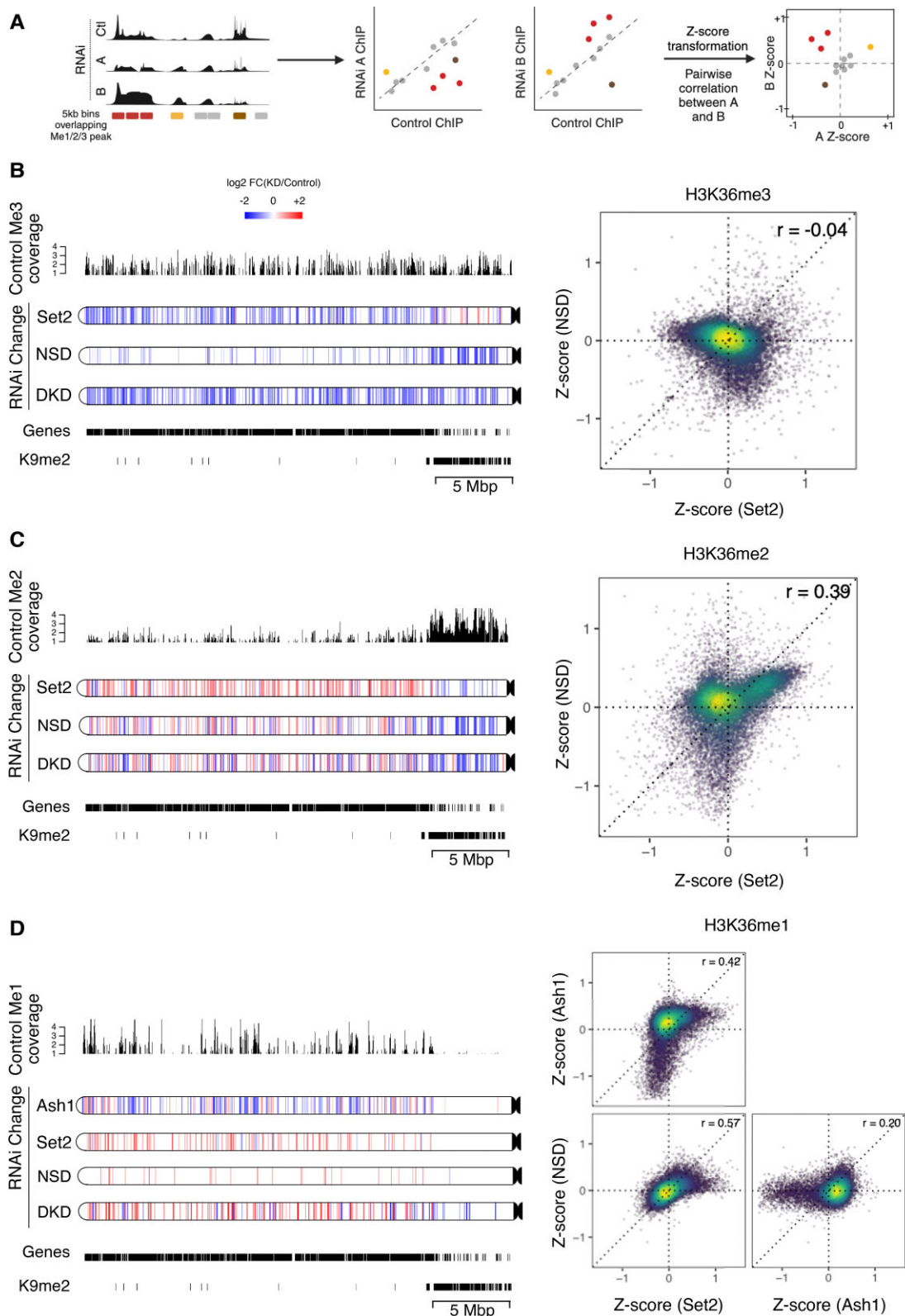


Figure 3. HMTs act on largely distinct genomic regions. **(A)** Schematic depicting Z-score analysis workflow. Pairwise Z-score scatter plots representing correlations among changes in ChIP signal upon RNAi of HMTs with respect to control RNAi in genome-wide 5 kb bins. Only bins overlapping at least one of K36me1/2/3 peaks in any RNAi condition were included. A negative Z-score indicates a reduction relative to control while a positive Z-score indicates an increase. Schematic was generated with BioRender.com. **(B–D)** Pairwise Z-score scatter plots for K36me3/2/1 upon RNAi of indicated factor (right). The color overlaid on the scatter indicates the local density of points. Pearson correlations are provided for each pair. Corresponding chromoMaps representing regions of significant difference in K36me3/2/1 signal for indicated RNAi conditions as derived from csaw analysis for chromosome 3L at 2-kbp resolution (left). The color of the regions (as indicated by the common scale in B) represent log₂-transformed value of number of normalized reads in RNAi condition relative to control condition. Control K36me3/2/1 signal overlaid above chromoMaps aid interpretation of relative changes. K9me2 peaks and gene annotations provided for reference.

and that spike-in normalization does not change the results (see Methods). However, we emphasize that this mild reciprocal increase is often not identified as statistically significant by csaw, potentially due to limited detection power with three biological replicates. Nevertheless, this observation may suggest that the loss of one HMT somehow increases the activity of the remaining HMTs.

The contributions of NSD and Set2 to K36me2 are more complex (Figure 3C), most likely because K36me2 can arise from methylation (me0/1 → me2) as well as from demethylation (me3 → me2) reactions. The chromoMaps illustrate that K36me2 at heterochromatin and at certain euchromatic bins are dependent on NSD (compare NSD RNAi in Figure 3B, C) as also observed for K36me3 above. The corresponding bins where K36me2 is reduced in an NSD-dependent manner are in the bottom half of the scatterplot (Figure 3C). Conversely, K36me2 at heterochromatin was largely unaffected upon Set2 RNAi, although a mild decrease can be observed at a few bins which also display increased K36me3, hinting at the possibility that NSD is locally more active at converting K36me2 to K36me3 in the absence of Set2.

Further, an increase in K36me2 at many euchromatic bins can be observed upon NSD RNAi, despite unchanged K36me3, suggesting augmented K36me2-deposition by Set2. In contrast, the increased K36me2 levels upon depletion of Set2 are accompanied by corresponding drops in K36me3 levels and may thus be due to demethylation of K36me3 (compare Set2 RNAi in Figure 3B, C). These bins, which show increased K36me2 in both Set2 and NSD RNAi, are represented in the top right quadrant of the corresponding scatter plot (Figure 3C).

Lastly, chromoMaps and scatter plots highlight a strong loss of K36me1 upon depletion of Ash1 at loci with a strong enrichment of K36me1 in control cells confirming that Ash1 is the major K36me1-transferase in S2 cells (Figure 3D). Depletions of Set2 or NSD are rather associated with an increased K36me1 in distinct bins which have very little K36me1 in control cells, and likely result from demethylation of K36me2/3.

Taken together, our results suggest that the three methyltransferases establish the genomic K36 methylation signatures according to locus-specific rules. Individual HMTs may dominate the methylation state in certain spatial domains but may also indirectly affect each other's activities exemplified by the see-saw effect which is significant at only few bins after csaw analysis.

A gene-centric view of the K36 methylation landscapes

The observation of distinct contributions of the three relevant HMTs to locus-specific methylation signatures prompted a more gene-centric analysis. K36 methylations are predominantly found on transcribed chromatin (Figure 1D). We selected around 10,500 genes enriched for at least one K36 methylation state in any one condition. The signals over the gene bodies were averaged and the resulting profiles clustered across all RNAi conditions. This yielded 12 clusters that once again illustrate the complexity of dynamic K36 methylation patterns across the genome. To facilitate a first interpretation of broad biological patterns while avoiding overinterpretation, we manually merged the clusters into three super-clusters

(supercluster I = clusters 1–4, ochre; supercluster II = clusters 5–8, brown; supercluster III = clusters 9–11, pink; Figure 4A). The minor supercluster IV (cluster 12, grey) contains a few hundred genes with very low K36me signals and was excluded from all subsequent analyses. The genes within the superclusters I to III are strongly correlated to specific K36 methylation profiles in steady state and tend to show consistent changes upon HMT depletion. Genome browser profiles for representative genes in each supercluster are shown in Figure 4B. Interestingly, these clusters correlate with distinct chromatin factors which may provide hints about their function/regulation (Supplementary Figure S5A).

Supercluster I genes depend on Set2 for K36 methylation

The ochre supercluster I consists of genes which have the highest average K36me3 signal and very low me1/2 in control conditions (Figure 4B, C). At these genes K36me3 strictly depends on Set2. The reduction in K36me3 in absence of Set2 is accompanied by an increase in K36me2 and, most prominently in clusters 3 and 4, an increase in K36me1. Variable effects may be due to heterogeneity in locus-specific histone demethylase (HDM) activity (see below and Figure 6). We also observe a mild general increase in K36me2/3 upon depletion of NSD, reminiscent of the see-saw effect mentioned earlier and exemplified by the gene CG6701 (Figure 4A, B).

The combined depletion of both Set2 and NSD (DKD) results in patterns very similar to Set2 KD alone, although a some genes (particularly in cluster 4 and in cluster 1) show a slightly stronger reduction of K36me3 compared to Set2 knockdown alone. This might be due to mild contribution of NSD towards K36me3, which is not detectable in the individual NSD knockdown, because of compensatory Set2 activity (see-saw effect). Regardless, these observations show that Set2 is highly efficient in generating K36me3 with almost no observable intermediate K36me1/2 and mostly does not rely on prior methylation by NSD.

NSD is responsible for H3K36me2/3 at euchromatic and heterochromatic genes in supercluster II

Supercluster II comprises of 1710 genes with high K36me2, moderate K36me3, and variable K36me1 levels in unperturbed cells. Only a minor fraction of these genes (136) was located in pericentric heterochromatin and carried the heterochromatic K9me2 mark (Figure 4B, Supplementary Figure S6A). The abundant K36me2 is strongly reduced upon depletion of NSD, but unaffected upon depletion of Set2. The NSD-dependent reduction in K36me2 is in most cases accompanied by a reduction in K36me3, but no detectable changes for K36me1. Evidently, NSD is the major HMT responsible for K36me2/3 in this cluster of genes, of which 90% reside in euchromatin. For further downstream analyses (below), we distinguish euchromatic genes (supercluster SC-IIA) from heterochromatic NSD-dependent genes (supercluster SC-IIB) for clarity.

Ash1 is responsible for H3K36me1 in supercluster III

Supercluster III genes stand out because they are strongly marked by K36me1, which is markedly reduced by depletion of Ash1, but not affected by either Set2 or NSD depletion. This shows that K36me1 is a state in itself and not only a methylation intermediate (Figure 4A).

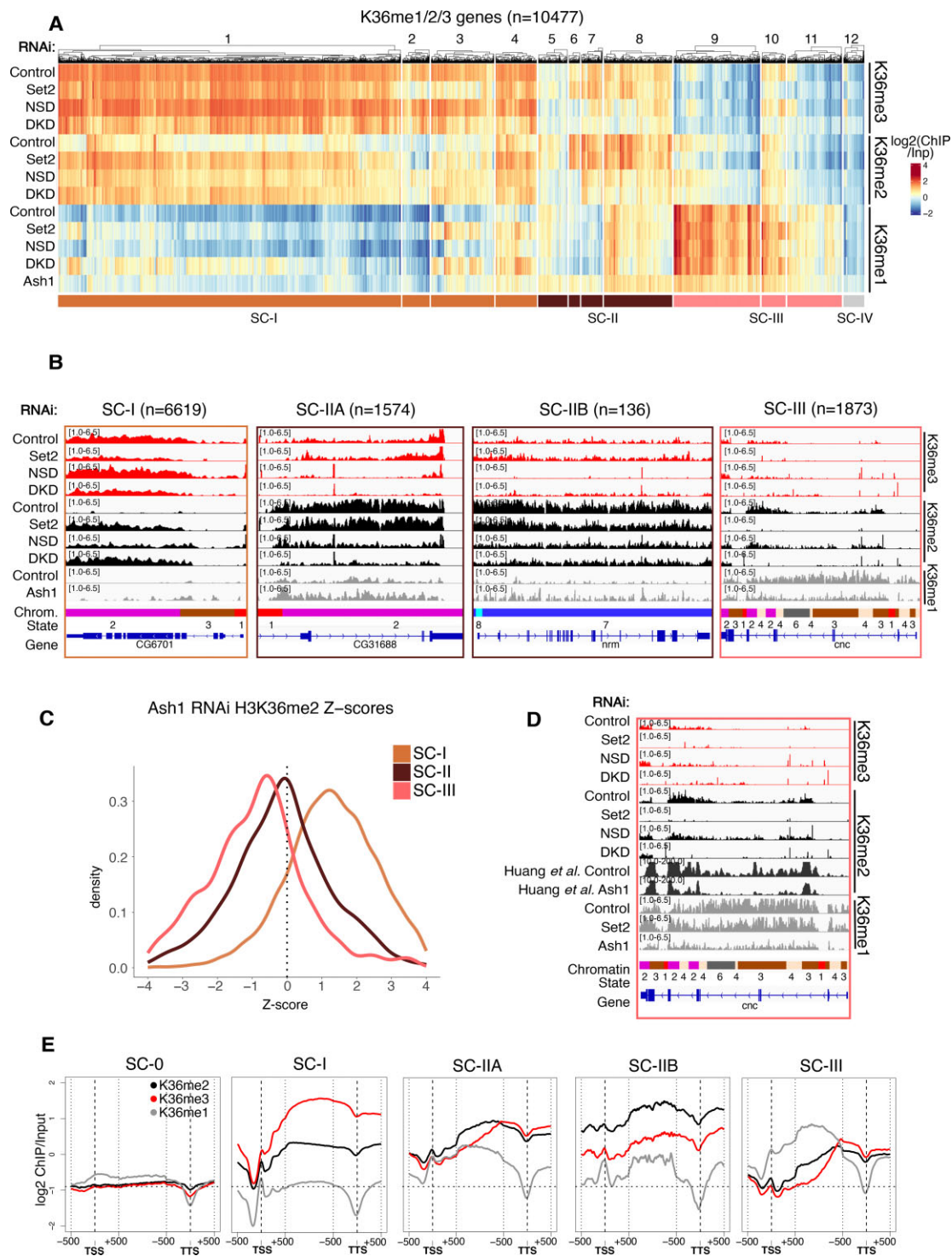


Figure 4. A gene-centric view of the K36 methylation landscapes. **(A)** Clustered heatmaps of gene body-averaged ChIP signal for K36me1/2/3 indicated on the right and RNAi condition indicated on the left. Only genes overlapping at least one of K36me1/2/3 peaks in any RNAi condition ($n = 10\,477$) were used for clustering. Clusters are numbered 1–12 as indicated above the heatmap. The track below the main heatmap represents the manual grouping of clusters to define Superclusters (SC) which display similar patterns, resulting in ochre SC-I ($n = 6619$), brown SC-II ($n = 1710$) and pink SC-III ($n = 1873$). **(B)** Genome browser profiles of representative genes from each supercluster along with the number of grouped genes. Supercluster II was further classified into Euchromatic (IIA) or Heterochromatic (IIB) based on overlap with H3K9me2 peaks. RNAi condition and immunoprecipitation target indicated on the left and right, respectively; 9-state Chromatin States (as in Figure 1C) serve as reference. Representative genes shown for SC-I, IIA, IIB and III lie within clusters 1, 8, 7 and 9 respectively. **(C)** Density plot of Z-scores representing change in K36me2 signal upon Ash1 RNAi (41) for superclusters defined in Figure 4A. **(D)** Genome browser profiles of supercluster 3 representative gene highlighting effect of Ash1 RNAi on K36me2. RNAi condition and immunoprecipitation target are indicated on the left and right, respectively; Chromatin States (as in Figure 1C) shown for reference. **(E)** Cumulative plots for each cluster representing gene body relative distributions of K36me1/2/3. 1 kb regions centered around TSS and TTS are unscaled, while the rest of the gene body was scaled to 500 bins. SC-0 genes lack any detectable H3K36 methylation and serve as a reference for zero signal/baseline. This is also represented by the horizontal dotted line in all individual plots. Only genes of minimum length of 1500 bp were included in the analysis.

A subset of supercluster III genes are decorated by moderate K36me2 and low K36me3 (cluster 9) or moderate K36me2 and K36me3 (cluster 10). This K36me2/3-marking appears particularly Set2-dependent, suggesting that Set2 can add methyl groups processively to pre-methylated K36 (K36me1 and/or K36me2) in this particular context. Upon loss of Set2, these sites are demethylated, but K36me1 is maintained by Ash1. This response to Set2 depletion is markedly different from supercluster I where a widespread increase in K36me2 is observed. We used a previously published dataset (41) to show that K36me2 of genes within the supercluster III was indeed also dependent on Ash1 (Figure 4C, D). The simplest model explaining this co-dependence is that Set2 locally uses K36me1 placed by Ash1 to generate K36me2/3.

Remarkably, the above-mentioned see-saw effect is also visible upon Ash1 depletion on Supercluster-I genes which clearly gain K36me2 but not on supercluster-II genes (Figure 4C). However, a widespread (though mild) increase of K36me1 detected in supercluster II genes (Figure 4A, B) suggests that NSD-dependent *de novo* K36 methylation is increased in the absence of Ash1. Some Set2-dependent genes (in particular clusters 3 and 4) also experience a K36me1 increase upon Ash1 depletion. These genes probably overlap the K36me1-devoid bins that significantly gain K36me1 upon Ash1 RNAi in the previously shown chromoMap (Figure 3D) suggesting that Set2 and/or NSD contribute to K36me1 in certain contexts. Evidently, the see-saw effect is a rather locus-specific phenomenon (see Discussion).

Context-dependent distribution of H3K36 methylation states along gene bodies

H3K36me3 is frequently referred to as a histone mark enriched towards 3' ends of genes. Anecdotal genome browser observations (Figure 4B) suggest that the intragenic location of K36 methylation states differs. To investigate this more systematically, we generated cumulative profiles of K36 methylation states along scaled gene bodies between transcription start sites (TSS) and transcription termination sites (TTS), separately for each supercluster (Figure 4E). Unmethylated genes (SC-0) which were not included in the clustering serve as an internal negative control.

Genes in supercluster I show a strong 3' bias for K36me3 but not K36me2, consistent with previous reports (72). K36me2 is broadly distributed throughout the bodies of the euchromatic genes in SC-IIA, whereas K36me1 drops towards the TTS and K36me3 increases sharply towards their 3' ends. Such a modulation of K36me3 was not observed for the heterochromatic cluster II genes (SC-IIB), where the gene body signals are similar to those upstream of TSS and downstream of TTS. This observation fits the genome browser views (Supplementary Figure S1A) that show heterochromatic genes embedded in larger K36me2/3 domains.

Finally, at supercluster III genes, K36me1 broadly marks the gene bodies, whereas K36me2/3 are enriched towards the 3' ends of transcription units. Notably, the intragenic distribution of the dimethyl mark is more similar to K36me3 rather than K36me1, further supporting the idea that most K36me2/3 at these genes is deposited by Set2 and not directly by Ash1 (Figure 4A).

The observation of distinct profiles of K36 methylation along the transcription unit supports the concept of context-dependent methylation patterns.

Genes of superclusters I, II and III are enriched in ePol, NSD and Ash1, respectively

The distinct distributions of K36 methylation marks at genes may be explained by targeting of the corresponding HMTs. If that was the case, we should find the HMT enriched at sites where K36 methylation is sensitive to its depletion. HMTs are often found enriched at distinct loci within larger domains of K36 methylation. To avoid averaging HMT ChIP coverage over the entire gene body, we selected the binding maxima for each HMT and calculated average signals within 2-kb windows around these maxima. The intensity distribution of the HMTs within the superclusters was visualized as violin plots (Figure 5A). A set of 3000 randomly sampled unmethylated genes (SC-0) served as an internal negative control.

Set2 (defined by the presence of ePol) was strongly enriched on genes with Set2-dependent K36me2/3, most prominently supercluster I genes, followed by supercluster III genes (likely contributed by clusters 9 and 10) which show Set2-dependent K36me2 and K36me2/3, respectively. Expectedly, NSD enrichment was stronger at heterochromatic genes (SC-IIB) compared to NSD-dependent euchromatic genes (SC-IIA). NSD was only weakly enriched at supercluster III genes. Set2-dependent supercluster I genes showed comparable levels of NSD to SC-IIA genes suggesting that NSD may be recruited to SC-I genes although its contribution to K36me at those genes is subtle. Lastly, analyzing the published Ash1 profile (41), we confirm that Ash1 was enriched the most in supercluster III. Mild enrichment within superclusters I and II-A may be due to overlapping domains (Figure 1A). The Ash1 enrichment observed within SC-IIB heterochromatic genes is probably contributed by non-specific signals which remain in Ash1 RNAi condition (data not shown).

Overall, these observations suggest that the three HMTs are the dominant drivers of K36 methylation at their highlighted gene clusters, but may also contribute to K36 methylation in other contexts.

Housekeeping functions are marked by Set2 while NSD and Ash1 target differentially expressed genes

The context for K36 methylation signatures (Figure 4E) may relate to gene-specific features like transcriptional activity, length and intron content. We visualized those features for genes in each supercluster as boxplots (Figure 5B). Genes in supercluster I were most highly transcribed and transcription activity correlated strongly with occupancy of the elongating RNA polymerase II (ePol) (Figure 5A, left). Genes in superclusters II and III are less expressed in agreement with lower ePol occupancy, tend to be longer and contain more introns. Given the specific patterns of steady state transcription in the different superclusters, we explored whether perturbation of individual HMTs leads to altered transcription within their respective superclusters (Supplementary Figure S7A-E). Interestingly, the majority of the genes do not change in expression upon RNAi of the corresponding HMT, suggesting that K36 methylation does not have a major influence on bulk transcript levels. Supplementary Figure S7 presents a reanalysis and discussion of published transcriptome data.

Our observations above show that genes of superclusters II and III are frequently long and intron-rich. Genes of this type are often differentially expressed across cell types, in contrast to house-keeping genes. To evaluate the degree of differential gene expression, we classified genes according

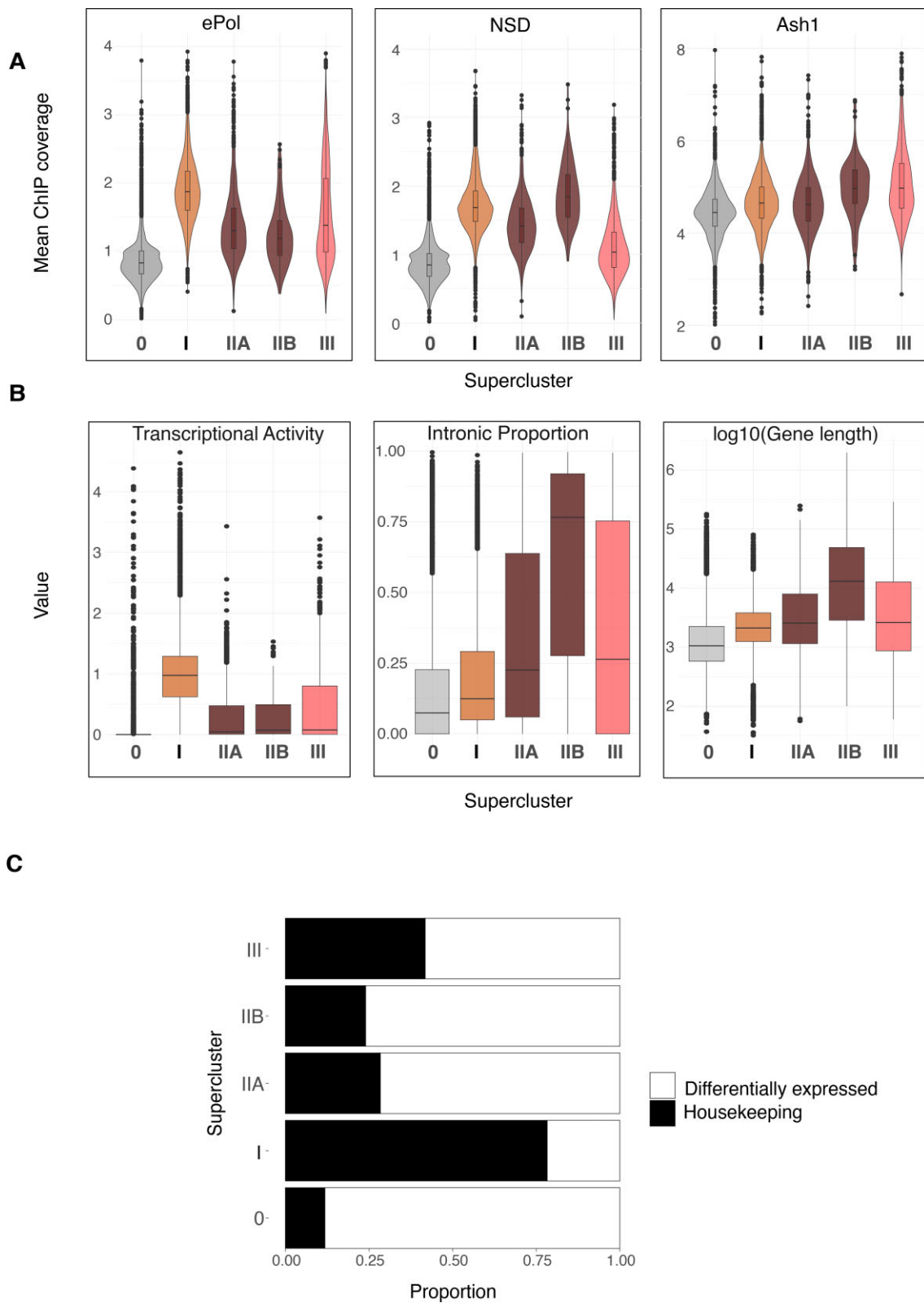


Figure 5. Gene clusters defined by HMT methylation patterns are correlated to different genic features. **(A)** Violin plots representing average ChIP signal for 2 kb windows around HMT gene body peaks for NSD, ePol and Ash1 within superclusters defined in Figure 4A. Supercluster 0 represents randomly sampled genes ($n = 3000$) lacking any detectable K36 methylation and serves as reference for zero signal/baseline. **(B)** Boxplots showing the proportion of introns (sum of length of introns/total gene length), transcriptional activity (denoted by \log_{10} -transformed RNA-seq TPM values) and $\log_{10}(\text{gene length})$ for gene superclusters. **(C)** Proportion of tissue-specific/invariant genes for gene superclusters based on FlyAtlas Expression Data for 25 tissues.

to their FlyAtlas expression profiles for 25 different tissues (67), as either ‘tissue-invariant’ or ‘differentially expressed’ (Supplementary Figure S8A) and calculated the proportion of these two classes across the superclusters. This revealed that genes in superclusters II and III were predominantly differentially expressed, while supercluster I genes were ubiquitously transcribed (‘housekeeping genes’) (Figure 5C). Tissue-specific genes are frequently subject to silencing by the Polycomb complex. Since K36 methylation has been linked to counteracting Polycomb-mediated K27me3 (41,51), we explored whether the K27me3 landscape is perturbed upon individual K36 HMT RNAi (Supplementary Figure S9A-F). While the majority of K27me3 domains remain unchanged upon HMT RNAi, we do observe several *de novo* K27me3 domains appearing within genes that lose K36 methylation bolstering the idea that K36 methylation contributes to restricting K27me3 spreading.

A further determinant of HMT targeting may relate to intron/exon content. K36me3 maps predominantly to exons, perhaps due to a slowing of ePol (and associated Set2) during splicing (4). Other studies also documented K36 methylation on introns (3). Anecdotal genome browser views (Figure 4B) hinted that K36me1/2/3 may mark introns and exons differentially in the different superclusters. We therefore examined the exon/intron distribution of steady state K36 methylation (Supplementary Figure S10A). We found K36me3 strongly enriched in exons versus introns in Set2-dependent supercluster I genes, whereas the bias is not observed for K36me2. The exon bias for K36me3 is substantial but also much less pronounced for K36me2 in the NSD-dependent SC-IIA genes and could be contributed by some genes that depend on both, Set2 and NSD (a certain fraction of cluster 7 genes). Interestingly, the bias is not lost upon Set2 depletion (Supplementary Figure S10B), suggesting that NSD may have a bias as well, despite no known interaction with ePol. Since a much weaker bias is observed for SC-IIB genes, exon selection is not an intrinsic feature of NSD but rather depends on the chromatin context (Supplementary Figure S10A).

Supplemental Figure S10B shows how the exon/intron distribution and bias shifts upon depletion of the HMTs and is revealing complex patterns that may be explored by future analyses.

The division of labor between HMTs for establishment of K36 methylation patterns genome-wide and the correlation to various genic features appears to be conserved during evolution. Similar SETD2-dependent and NSD1/2-dependent gene clusters were observed in two independent mouse data sets (Supplementary Figure S11A).

K36 methylation at transposable elements

Some transposable (TEs) are marked by H3K36me3 in *Drosophila* (36), and K36 HMTs were suggested to regulate TE expression (73). We mapped K36me2/3 at various repeat families across RNAi conditions and clustered the profiles (Supplementary Figure S12A). Roughly 50% of TE families were not methylated in any condition. Among the families that were exclusively methylated by Set2, telomeric TE stand out, which are defined by high K36me3/low K36me2. Many transposon families that are characterized by NSD-dependent high K36me2 and moderate K36me3 levels are probably embedded in pericentric heterochromatin (74) as highlighted in the chromoMaps of Figure 3. These transposons also show in-

creased K36me3/decreased K36me2 upon Set2 RNAi, probably resulting from the see-saw effect involving both NSD and Set2.

H3K36me reader binding is determined by density and turnover of K36me3 and K36me2

Changes in K36 methylation patterns upon HMT depletion provide a unique opportunity to observe and compare the dependency of reader proteins on these marks. We explored the redistribution of exemplary reader proteins JASPer (36) and MSL3 (ChrX-specific) (22,35) upon HMT depletion for supercluster I and II genes, which comprise of almost all reader binding events (Figure 6A). The bulk levels of JASPer and MSL3 (inferred from other MSL complex proteins) were not affected upon Set2 or NSD depletion (Supplementary Figure S13A).

Comparing the heat maps of JASPer binding upon HMT depletion (Figure 6A) revealed that JASPer follows K36me3 at some genes, in particular within clusters 2 and 3, but much less so in clusters 1 and 4. However, these latter sites are characterized by strong gains of K36me2 upon Set2 depletion. Upon depletion of NSD, JASPer binding is lost when both K36me2 and K36me3 are lost, most prominently in clusters 6, 7 and 8. Genomic profiles of representative genes in clusters 1, 3 and 6 illustrate the different scenarios inferred from the heat map (Figure 6B). These observations led us to hypothesize that JASPer may be attracted to both K36me3 and K36me2.

To explore this possibility further, we calculated Z-scores to quantify and compared the magnitude and direction of changes in reader binding and K36me2/3 signals upon HMT depletion in all genes of clusters 1, 3 and 6. At cluster 1 genes (Figure 6C, left panel), JASPer binding is unperturbed upon Set2 depletion. A significant reduction of K36me3 may be compensated by a corresponding strong increase of K36me2. In contrast, at cluster 3 genes, strong loss of K36me3 upon Set2 depletion is accompanied by only a mild gain of K36me2 and correlates with a loss of JASPer binding (Figure 6C, middle panel). In the third scenario, highlighted by cluster 6 genes, both K36me2/3 are eliminated in a coordinated fashion upon NSD knockdown resulting in reduction of JASPer binding (Figure 6C, right panels). These findings support the hypothesis that JASPer binds to both K36me2 and K36me3.

In order to generalize our observations to supercluster I and IIA genes, we visualized the change in JASPer binding (as Z-score) in response to Set2 depletion as a function of the absolute K36me3 ChIP signal (Supplementary Figure S13B). The scatter plot revealed that reduction in JASPer binding ($Z < 0$) predominantly occurs when residual K36me3 levels fall below a threshold of 1.35. Below this threshold, decreasing K36me3 values result in a proportional JASPer reduction whereas JASPer continues to bind to ‘resistant’ genes that maintain higher levels of K36me3. Those resistant genes also acquire more K36me2 upon depletion of Set2, possibly because the turnover of both marks is linked (see below). Likewise, the same threshold of 1.35 is observed for the loss JASPer binding at SC-IIA genes upon NSD depletion (Supplementary Figure S13C). This 1.35 threshold is marked by a dotted line in all genome browser views (Figure 6B).

In light of the recent biochemical observation that the PWWP domain of LEDGF, the human ortholog of JASPer, can bind K36me3 and K36me2 with similar affinities (31), we measured the affinity of recombinant JASPer to K36me0/2/3

mononucleosomes by microscale thermophoresis (MST) (Figure 6D). JASPer was found to have a slight preference for K36me2 ($K_D = 0.32 \pm 0.02 \mu\text{M}$) over K36me3 mononucleosomes ($K_D = 0.60 \pm 0.04 \mu\text{M}$), whereas the affinity for unmodified nucleosomes was much lower ($K_D = 4.00 \pm 1.54 \mu\text{M}$).

We explored if the binding principles inferred from JASPer profiles also applied to the chromodomain protein MSL3. Because MSL3 binding is restricted to genes on the X chromosome, it cannot be displayed along with the reference heat map that clusters all genes. Nevertheless, representative X-chromosomal genes of the reference clusters 1, 3 and 6 show a similar behaviour and threshold effect for MSL3 (Supplementary Figure S13D, E). The corresponding density plots also show that MSL3 binding follows the general pattern of JASPer (Supplementary Figure S13F). At cluster 1 genes, MSL3 binding is unperturbed, whereas at clusters 3 and 6 MSL3 binding is lost along with K36 methylation in Set2 and NSD knockdowns, respectively. Because MSL3 binding profiles mirror that of JASPer, we hypothesize that MSL3 also binds to both K36me3 and K36me2. Altogether, MSL3 and JASPer leave chromatin when the combined density of K36me2 and K36me3 drop below a critical threshold.

The question remained why some genes in clusters 1 and 4 maintain relatively high levels of combined K36me2/3 despite significant depletion of Set2 during 7- or 10-day RNAi protocols (the duration of RNAi did not affect the general outcome, Supplementary Figure S14A), while genes in other clusters lose most K36me2/3 methylation. We consider that the residual H3K36me2/3 at specific loci after HMT depletion depends much on genomic context, which may differ in terms of transcription-associated nucleosome turnover and methylation/demethylation dynamics. In line with these considerations, we found more elongating RNA polymerase II (ePol) at cluster 2/3 genes that have much lower combined levels of K36me2/3 ('sensitive genes') compared to 'resistant' genes in clusters 1/4 (Supplementary Figure S14B) which may be linked to turnover of K36me2/3-bearing nucleosomes. Furthermore, we found the putative K36 demethylase KDM2 (75), but not KDM4A, is enriched at 'sensitive' clusters 2/3 genes. These findings suggest that robustness in K36me reader binding is determined by a complex interplay between dedicated HMTs and counteracting turnover mechanisms.

Discussion

A variety of epigenetic processes have been linked to the methylation of histone H3 at lysine 36 in different organisms, mostly considering K36me2/3. Many of those processes influence the transcriptional output (5). Although K36 methylation is generally considered as an active mark, it is relevant for silencing by DNA methylation (13,14) and for silencing at facultative heterochromatin by H3K27me3 (15,16). In addition, K36me3 has been linked to genome stability (for review, see (76)) and HMTs establishing H3K36 methylation as well as K36me readers have been linked to cancer (77–79).

In mammals, HMTs that methylate K36 and corresponding reader proteins are numerous, suggesting partially redundant functions. To more easily uncover fundamental principles, we use the *Drosophila* model to explore the role of the three relevant HMTs Set2, NSD and Ash1 in the genome-wide distribution of the three K36 methylation states. We also determine

the genomic locations of two dedicated readers proteins: the chromodomain MSL3, a subunit of the male-specific dosage compensation complex (DCC) (35) and the PWWP protein JASPer, which tethers JIL1 kinase to active chromatin, where it reinforces the active state by H3S10 phosphorylation (36). Importantly and in contrast to other methods, our optimized ChIP protocol enables improved mapping of K36 methylation and readers within constitutive heterochromatin (Supplemental Figure S1A), which are often poorly solubilized in traditional ChIP protocols (80,81) highlighting that they are not strictly active chromatin marks.

The mapping of each individual K36 methylation state in unperturbed conditions revealed complex patterns of distributions. Yet, the full complexity of the profiles only became apparent when the changes upon single or combined HMT depletion were scored in a gene-centric manner and the resulting heat maps subjected to unsupervised clustering. We hypothesize that the 11 clusters that emerged correspond to gene classes with shared properties and chromatin context. The grouping into superclusters served to outline broad features without getting lost in details. The heterogeneity of the loss-of-function phenotypes revealed local, context-dependent activities of the HMTs. Interestingly, a recent preprint adopting a very similar approach involving single and combinatorial knockouts of murine H3K36 methyltransferases echo our observations regarding complexity in establishment of the K36 methylation landscape (82).

Ash1 establishes H3K36me1 domains overlapping enhancers

Monomethylation is the most abundant methylation state of K36 in *Drosophila* (71), but not well studied in other organisms. We found that K36me1 represents a defined chromatin state in regions with enhancers features, often in introns of differentially expressed genes. These sites cover about 15% of the genome and may correspond to the most densely K36-methylated domains, according to the relative abundance of the three methylation states (71). Remarkably, those sites bear very little K36me2/me3 (see exception below) and K36me1 appears often rather complementary to me2/me3 signals. Evidently, K36me1 is not only a methylation intermediate towards K36me2/3, but may have its own function at enhancers, which needs to be further explored. To our knowledge, readers that selectively recognize K36me1 are not known.

Most of K36me1 is placed by Ash1, an HMT that so far is mostly described as a K36 dimethylase important for restricting Polycomb repression in flies and mammals (15,83). At other sites, such as a fraction of supercluster III genes, Set2 converts K36me1 placed by Ash1 into K36me2/3. The function of this locus- and context-specific methylation pathway at a large fraction of genic K36me1 enhancer domains remains to be explored.

H3K36 methylation by NSD and Set2 represent distinct pathways

Contrary to our expectations, in steady state conditions in S2 cells, NSD and Set2 act largely independently of each other at distinct chromatin domains and different genes. Set2 catalyzes K36me3 predominantly within intron-poor, highly transcribed housekeeping genes, while NSD catalyzes

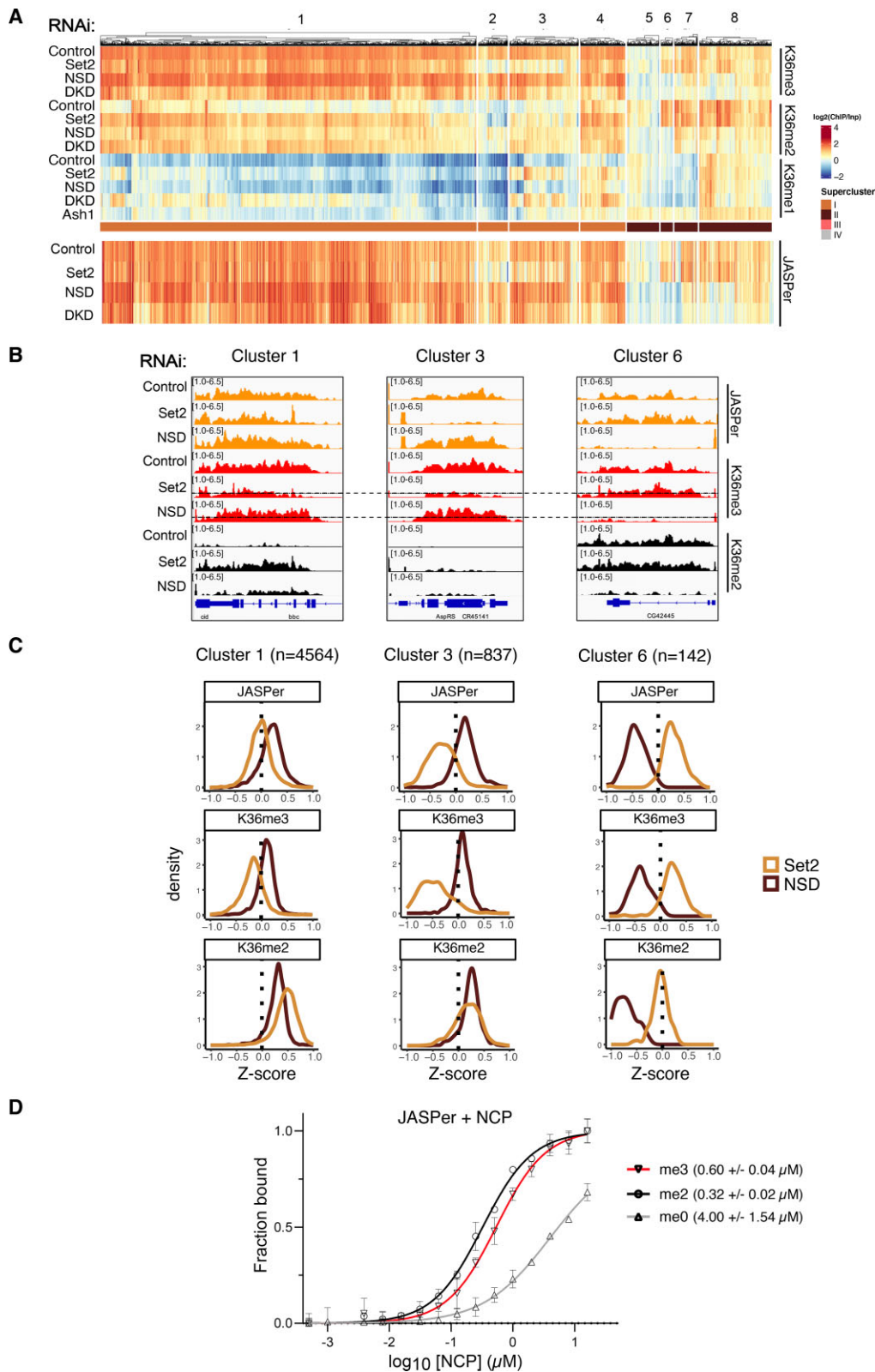


Figure 6. Robust binding of K36me3 readers above threshold K36me3 density. **(A)** Heatmap of gene body average ChIP signal for JASPer in indicated RNAi condition ordered according to heatmap from Figure 4A. Note that supercluster III was excluded as it has very few genes bound by JASPer. **(B)** Genome browser profiles for three distinct genomic loci highlighting differential response of JASPer to Set2 RNAi and NSD RNAi within indicated clusters. RNAi condition and immunoprecipitation target indicated on left and right, respectively. The dotted line in K36me3 ChIP in Set2/NSD RNAi represents the threshold derived from scatter plot in [Supplementary Figure S10B](#). (Note that genome browser profiles display non-log transformed ChIP values. The calculated threshold of 1.35 corresponds to ~ 2.5 in non-log transformed units which is marked by the dotted line.) **(C)** Z-score density plots representing direction and magnitude of change in gene body average ChIP signal for K36me2/3 along with reader JASPer for indicated representative clusters. **(D)** Equilibrium binding between recombinant JASPer and unmodified, di- or trimethylated nucleosomes, respectively, determined using microscale thermophoresis (MST). Error bars represent the standard deviation from the mean values obtained from $n = 2$ experiments. Calculated dissociation constants are indicated within brackets.

K36me2/3 in a heterochromatic environment as well as at numerous euchromatic, weakly transcribed genes with many introns and cell-specific expression profiles. Our data suggests that, in general, both NSD and Set2 can methylate nucleosomes *de novo* without prior ‘pioneering’ methylation by another enzyme. Set2 rarely adds a methyl group to K36me2 generated by NSD, in contrast to earlier interpretations of *in vitro* studies (37). Furthermore, we find that NSD is not only a K36 dimethylase but is also capable of generating K36me3 in *Drosophila*, which resonates with a recent report in mice (84). These results are in line with recent studies documenting distinct loss-of-function phenotypes of Set2 and NSD in flies (25,28). Moreover, these notions are corroborated by our reanalysis of independent mouse datasets, where we find that the mammalian K36 HMTs contribute to separate methylation profiles at distinct gene classes.

Targeting, regulation and function of H3K36 methyltransferases

Unsupervised clustering of the K36 methylation changes at genes upon HMT depletion revealed a clear functional partitioning of the three HMT activities, which is, however, not exclusive. Each HMT appears to be most strongly enriched at loci, where depletion studies reveal their activity (Set2 in supercluster I genes, NSD at supercluster II and Ash1 at supercluster III). However, they can also be found at other sites, where they are apparently not active. For example, NSD is detected at supercluster I genes where K36 methylation largely depends on Set2. Vice versa, Set2 can be mapped to supercluster II genes, where most K36me depends on NSD. These observations suggest that the functions of K36 HMTs can be regulated at two different steps, their targeting and the activation of enzymatic activity. This is reminiscent of yeast Set2, which is targeted to transcribed genes by interaction of the conserved SRI domain with the elongating RNA polymerase II. The SRI domain directly contributes to activation since its deletion abolishes K36me3 and reduces H3K36me2 activity *in vitro*. Further facilitation of Set2 methylation activity is achieved in presence of monoubiquitinated Lys123 of H2B on nucleosome substrates (for review see (85)).

While Set2 is recruited to transcribed chromatin by interaction with the transcription machinery, it is less clear how NSD is targeted. It has been reported to be recruited to Beaf-32 bound active promoters (43), but we found no systematic association between Beaf-32 peaks and promoters of SC-II NSD-dependent genes (not shown). The molecular basis of NSD localization to heterochromatin was hypothesized to depend on HP1 (38), however preliminary experiments (unpublished observations, and personal communication from D. Atkinson and T. Jenuwein) as well as another recent study (86) argue for an HP1-independent targeting mechanism. Sun *et al.* described a catalytic-activity independent role for the mouse NSD1 in promoting enhancer function by localizing to active enhancers (68). In contrast to these observations, we do not observe NSD at H3K27ac/H3K4me1 domains. For mammalian NSD enzymes, an autoinhibitory state is relieved by binding to nucleosomes (87) to trigger activation, while NSD2 has been shown to be inhibited by H1 in nucleosomes (88). However, one conserved feature of all NSD orthologs is the presence of two PWWP domains, which could potentially modulate the activity of NSD and be implicated in a maintenance function. K36me3 is a relatively rare mark (71), thus it

is also conceivable that NSD acts as a ‘maintenance’ methyltransferase by associating with a K36me2/3 mark. Such a scenario may serve to increase the local density of K36me3, to place K36me2 in the vicinity of K36me3 ‘seeds’ or to preserve the signature of active chromatin at poorly transcribed genes. Altogether, the targeting, activation and molecular function of *Drosophila* NSD, potentially through different mechanisms at euchromatin and heterochromatin, remains to be addressed in future studies. The function reported for mammalian NSD to establish faithful DNA methylation (13,89) will not be conserved in *Drosophila* which lacks DNA methylation-related machinery.

The determinants of genomic distribution of Ash1 have not been elucidated. Although we do not see extensive H3K36me2/3 dependent on Ash1 that is not dependent on Set2, it has been shown that *Drosophila* Ash1 in particular is activated by MRG15 (41,44).

While our study thoroughly documents the changes in K36 methylation occurring upon depletion of K36 HMTs, the consequences of altered K36 methylation is less clear. We investigated two phenomena commonly discussed in the context of K36 methylation, namely transcriptional regulation and counteracting the spreading of polycomb-dependent K27me3. In agreement with previous studies (26), we found that changes of K36 modification and transcription upon HMT depletion were poorly correlated, precluding a causal link between K36 methylation and transcription. We are aware of potential indirect effects due to HMT-dependent methylation of non-histone proteins, such as methylation of alpha-tubulin by SETD2 (90). Any effect of K36 methylation on transcription is likely to be mediated by the nature (and combination) of functionally diverse K36me readers bound at a particular gene locus and may involve RNA processing (26). Further, our preliminary observations also suggest that K36 HMTs, especially NSD and Ash1, independently counteract K27me3 spreading at some developmental genes. However, this effect is restricted to a fraction of all genes carrying K36 methylation, suggesting that counteracting K27me3 is not the sole function of K36 methylation.

Indirect communication between HMTs: the see-saw effect

In several instances we observed that the depletion of one HMT led to an increase of the relevant mark at sites where another HMT is predominantly active, a phenomenon we termed see-saw effect. While the effect did not withstand rigorous statistical testing at our limited experimental power, it is nevertheless noteworthy. For example, Set2 depletion leads to an increase of K36me3 at pericentric heterochromatin, where NSD is strongly enriched. Similar effects can be observed across two different mouse data sets (13,68). Such an effect may be explained by competition for a shared substrate, such as SAM. Indirect communication between HMTs of this kind may lead to an underestimation of co-dependencies between Set2 and NSD, particularly at genes where they colocalize, if the activity of one is increased when the other is depleted. However, it also highlights a possible mechanism that ensures robustness of H3K36me2/3 distribution upon metabolic perturbation. Interestingly, recent studies in mammals have linked H3K36me3 to the methionine cycle and vitamin B12 levels, which determine SAM availability, in reprogramming and cell differentiation (91,92).

Robust readout of H3K36 methylation by readers bearing chromo- and PWWP-domains

At transcribed chromatin, the function of K36 methylation is mediated by dedicated reader proteins, of which we considered two, MSL3 and JASPer, which associate with methylated K36 with different domains. Both proteins behave similarly in our experiments. For MSL3, our systematic analysis extends previous studies (35,42), which either did not involve high-throughput sequencing or utilized Set2-deficient fly larvae where K36me3 is completely absent. Our RNAi approach provides a spectrum of intermediate K36me2/3 levels, which we exploit to obtain quantitative insights into reader binding.

We show that JASPer binds nucleosomes with K36me2 and K36me3 with similar affinity, as described for its mammalian ortholog LEDGF (31). This may explain why JASPer stays associated with genes that significantly lose K36me3, but retain K36me2. The ability of readers to recognize both states may ensure that transcribed chromatin is reliably bound upon changes. Furthermore, the fact that K36me3 is enriched at exons, while K36me2 is more prevalent at introns may assure that all aspects of transcribed chromatin are equally well addressed.

It appears that the chromatin interactions of both, the chromodomain reader MSL3 and the PWWP reader JASPer, depend on critical concentrations of K36me3/K36me2 regardless of which enzyme catalyzes the methylation. One way of rationalizing the observation is that the combined capacity of the K36me3/2 marks normally exceeds the requirements for productive tethering of the reader proteins, conferring robustness in reader binding against fluctuations during local turnover of histone or methyl-mark. While we formally determined a threshold value for K36me3, we note that this value represents a proxy for all potentially covarying chromatin modifications (exemplified by K36me2), which may contribute to reader binding in our model system. Loci where reader binding appears sensitive to Set2 or NSD depletion may have a higher turnover of H3K36me2/3 and/or a less efficient maintenance of K36 methylation. Corresponding physiological conditions may include changes in metabolic conditions (methionine or vitamin B12 availability, see above) for Set2-dependent housekeeping genes, and expression changes occurring during differentiation processes for NSD-dependent differentially expressed genes.

Methylation read-out at heterochromatin

We found that K36me3 partitions between eu- and heterochromatin. The presence of K36me3, a mark globally associated with active chromatin, at H3K9-methylated heterochromatin has also been observed in other studies (3,84) and been described in human datasets (93,94). Such overlap of active with more repressive chromatin features may not simply be due to population heterogeneity since engineered dual readers validated the existence of such dual domains (84,95) and have been proposed to bookmark poised enhancers genome-wide in mouse (84). In *Drosophila*, the co-occurrence of those marks is mostly observed at genes embedded at pericentric heterochromatin, which also recruit JASPer and MSL3 readers. However, the relationship between K36me3/2 at PCH and reader binding is complicated as we also observe many instances of heterochromatic K36me2 domains which fail to recruit JASPer. Another heterochromatin

modification/component may hinder binding to K36me2 locally. We thus speculate that the selective recruitment to K36me2/3 *in vivo* may be modulated in the chromatin context.

Limitations of our study

Our systematic and comprehensive approach necessitated to focus on a simple cellular model. The fundamental principles we uncovered are likely to be modulated in cell-specific ways in the fly organism, where the expression of HMTs and KDMs vary spatiotemporally. For instance, NSD is maternally deposited and present during early development while Set2 appears later during zygotic genome activation.

The RNAi approach results in strong reduction of target protein levels, but does not eliminate it entirely. For non-processive methyltransferases, the efficiency with which the three methyl groups are added to the same lysine may get progressively lower. Upon RNA interference there may still be enough enzyme to add two methyl groups but not enough to bring it to me3. In ChIP-seq, this may look like the knockdown leads to the ‘substitution’ of H3K36me3 by H3K36me2. This may be an alternative explanation to the observed appearance of K36me2 within supercluster I genes which we interpreted as differences in turnover of K36me3.

Due to a lack of ChIP-grade antibody, we did not directly map Set2, but infer the genomic Set2 localization from the profiles of elongating RNA pol II.

Our antibody-based approach is blind to the H3 variant that bears the K36me mark and it is likely that the replication-independent histone H3.3 variant contributes to the observed ChIP signals. H3.3 is enriched at active chromatin, transposons and the hyperactive male X chromosome in *Drosophila* (96). Since H3.3 accounts for about 20% of all H3 and is placed where nucleosomes turn over, it is possible that a large fraction of the K36me marks we score reside on the variant (97).

The genomic annotation of introns and exons are based on whole-fly RNA datasets. It is possible that the long introns methylated by NSD might contain unannotated tissue-specific exons. Alternatively, K36me2/3 at these long introns may mark cell-specific regulatory domains (e.g. enhancers).

Data availability

The raw sequencing files in fastq format and the summarized genome browser tracks in bigwig format are available in the GEO database GSE253391. Custom code for analysis of data is available on Zenodo (<https://doi.org/10.5281/zenodo.10894660>). Western blots and Immunofluorescence images used for quantifications in Figure 2 and Supplementary Figure S2 are available on Zenodo (<https://doi.org/10.5281/zenodo.10514405>).

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

We thank A. Campos-Sparr for technical assistance, S. Krause for conducting pilot experiments and R. Shetty for setting up

CUT&RUN for H3K36me3. We are grateful to C. Wirbelauer and D. Schübeler from the FMI in Basel for sharing their HMT antibodies and A. Thomae from the Bioimaging Unit of the Biomedical Center, LMU for advice on microscopy. We thank T. Straub and T. Schauer from the BMC Bioinformatics Unit for access to the computational cluster, advice on bioinformatics and helpful scripts. We thank Yuri Schwartz for suggestions that improved the manuscript. We thank S. Krebs of LMU LAFUGA for Illumina sequencing. We thank V. Lux and E. Koutná from IOCB Prague for their help with the recombinant protein production and nucleosome reconstitutions. We thank the members of the Becker lab for sharing reagents and comments on the manuscript. Graphical abstract and schematics in the manuscript were created on BioRender.com.

Author contributions: M.J., C.R. and P.B.B. conceived the study. M.J. carried out all experiments and performed the data analysis, except for MST experiments. M.H. purified recombinant JASPer, assembled modified nucleosomes and obtained MST measurements supervised by V.V. C.R. established collaboration with V.V. M.J., C.R. and P.B.B. interpreted the data and wrote the manuscript with input from all authors.

Funding

M.J. is member of the International Max-Planck Research School ‘Molecules of Life’ and the Integrated Research Training Group of the DFG-funded Collaborative Research Center ‘Chromatin Dynamics’; German Research Council (DFG) [BE1140/11-1 to P.B.B., RE4742/1-1 to C.R.]. Funding for open access charge: German Research Council (DFG) [BE1140/11-1 to P.B.B., RE4742/1-1 to C.R.].

Conflict of interest statement

None declared.

References

- Gorkin,D.U., Barozzi,I., Zhao,Y., Zhang,Y., Huang,H., Lee,A.Y., Li,B., Chiou,J., Wildberg,A., Ding,B., *et al.* (2020) An atlas of dynamic chromatin landscapes in mouse fetal development. *Nature*, **583**, 744–751.
- Kharchenko,P.V., Alekseyenko,A.A., Schwartz,Y.B., Minoda,A., Riddle,N.C., Ernst,J., Sabo,P.J., Larschan,E., Gorchakov,A.A., Gu,T., *et al.* (2011) Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature*, **471**, 480–485.
- Delandre,C., McMullen,J.P.D., Paulsen,J., Collas,P. and Marshall,O.J. (2022) Eight principal chromatin states functionally segregate the fly genome into developmental and housekeeping roles. bioRxiv doi: <https://doi.org/10.1101/2022.10.30.514435>, 01 November 2022, preprint: not peer reviewed.
- Luco,R.F., Allo,M., Schor,I.E., Kornblihtt,A.R. and Misteli,T. (2011) Epigenetics in alternative pre-mRNA splicing. *Cell*, **144**, 16–26.
- Sorenson,M.R., Jha,D.K., Ucles,S.A., Flood,D.M., Strahl,B.D., Stevens,S.W. and Kress,T.L. (2016) Histone H3K36 methylation regulates pre-mRNA splicing in *Saccharomyces cerevisiae*. *RNA Biol.*, **13**, 412–426.
- LeRoy,G., Oksuz,O., Descostes,N., Aoi,Y., Ganai,R.A., Kara,H.O., Yu,J.R., Lee,C.H., Stafford,J., Shilatifard,A., *et al.* (2019) LEDGF and HDGF2 relieve the nucleosome-induced barrier to transcription in differentiated cells. *Sci. Adv.*, **5**, eaay3068.
- Kim,J.H., Lee,B.B., Oh,Y.M., Zhu,C., Steinmetz,L.M., Lee,Y., Kim,W.K., Lee,S.B., Buratowski,S. and Kim,T. (2016) Modulation of mRNA and lncRNA expression dynamics by the Set2-Rpd3S pathway. *Nat. Commun.*, **7**, 13534.
- Martin,B.J., McBurney,K.L., Maltby,V.E., Jensen,K.N., Brind’Amour,J. and Howe,L.J. (2017) Histone H3K4 and H3K36 methylation independently recruit the NuA3 Histone acetyltransferase in *saccharomyces cerevisiae*. *Genetics*, **205**, 1113–1123.
- Jha,D.K. and Strahl,B.D. (2014) An RNA polymerase II-coupled function for histone H3K36 methylation in checkpoint activation and DSB repair. *Nat. Commun.*, **5**, 3965.
- Carvalho,S., Vitor,A.C., Sridhara,S.C., Martins,F.B., Raposo,A.C., Desterro,J.M., Ferreira,J. and de Almeida,S.F. (2014) SETD2 is required for DNA double-strand break repair and activation of the p53-mediated checkpoint. *eLife*, **3**, e02482.
- Daugaard,M., Baude,A., Fugger,K., Povlsen,L.K., Beck,H., Sorensen,C.S., Petersen,N.H., Sorensen,P.H., Lukas,C., Bartek,J., *et al.* (2012) LEDGF (p75) promotes DNA-end resection and homologous recombination. *Nat. Struct. Mol. Biol.*, **19**, 803–810.
- Huang,H., Weng,H., Zhou,K., Wu,T., Zhao,B.S., Sun,M., Chen,Z., Deng,X., Xiao,G., Auer,F., *et al.* (2019) Histone H3 trimethylation at lysine 36 guides m(6)A RNA modification co-transcriptionally. *Nature*, **567**, 414–419.
- Weinberg,D.N., Papillon-Cavanagh,S., Chen,H., Yue,Y., Chen,X., Rajagopalan,K.N., Horth,C., McGuire,J.T., Xu,X., Nikbakht,H., *et al.* (2019) The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape. *Nature*, **573**, 281–286.
- Dukatz,M., Holzer,K., Choudalakis,M., Emperle,M., Lungu,C., Bashtrykov,P. and Jeltsch,A. (2019) H3K36me2/3 Binding and DNA Binding of the DNA methyltransferase DNMT3A PWWP domain both contribute to its chromatin interaction. *J. Mol. Biol.*, **431**, 5063–5074.
- Schmitges,F.W., Prusty,A.B., Faty,M., Stutzer,A., Lingaraju,G.M., Aiwazian,J., Sack,R., Hess,D., Li,L., Zhou,S., *et al.* (2011) Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol. Cell*, **42**, 330–341.
- Alabert,C., Loos,C., Voelker-Albert,M., Graziano,S., Forne,I., Reveron-Gomez,N., Schuh,L., Hasenauer,J., Marr,C., Imhof,A., *et al.* (2020) Domain model explains propagation dynamics and stability of histone H3K27 and H3K36 methylation landscapes. *Cell Rep.*, **30**, 1223–1234.
- Wagner,E.J. and Carpenter,P.B. (2012) Understanding the language of Lys36 methylation at histone H3. *Nat. Rev. Mol. Cell Biol.*, **13**, 115–126.
- Ye,C., Sutter,B.M., Wang,Y., Kuang,Z. and Tu,B.P. (2017) A metabolic function for phospholipid and histone methylation. *Mol. Cell*, **66**, 180–193.
- Carrozza,M.J., Li,B., Florens,L., Suganuma,T., Swanson,S.K., Lee,K.K., Shia,W.J., Anderson,S., Yates,J., Washburn,M.P., *et al.* (2005) Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell*, **123**, 581–592.
- McCauley,B.S., Sun,L., Yu,R., Lee,M., Liu,H., Leeman,D.S., Huang,Y., Webb,A.E. and Dang,W. (2021) Altered chromatin states drive cryptic transcription in aging mammalian stem cells. *Nat. Aging*, **1**, 684–697.
- Luco,R.F., Pan,Q., Tominaga,K., Blencowe,B.J., Pereira-Smith,O.M. and Misteli,T. (2010) Regulation of alternative splicing by histone modifications. *Science*, **327**, 996–1000.
- Sural,T.H., Peng,S., Li,B., Workman,J.L., Park,P.J. and Kuroda,M.I. (2008) The MSL3 chromodomain directs a key targeting step for dosage compensation of the *Drosophila melanogaster* X chromosome. *Nat. Struct. Mol. Biol.*, **15**, 1318–1325.
- McKay,D.J., Klusza,S., Penke,T.J., Meers,M.P., Curry,K.P., McDaniel,S.L., Malek,P.Y., Cooper,S.W., Tatomer,D.C., Lieb,J.D., *et al.* (2015) Interrogating the function of metazoan histones using engineered gene clusters. *Dev. Cell*, **32**, 373–386.
- Salzler,H.R., Vandadi,V., McMichael,B.D., Brown,J.C., Boerma,S.A., Leatham-Jensen,M.P., Adams,K.M., Meers,M.P.,

- Simon, J.M., Duronio, R.J., *et al.* (2023) Distinct roles for canonical and variant histone H3 lysine-36 in Polycomb silencing. *Sci. Adv.*, **9**, eadf2451.
25. Dorafshan, E., Kahn, T.G., Glotov, A., Savitsky, M., Walther, M., Reuter, G. and Schwartz, Y.B. (2019) Ash1 counteracts polycomb repression independent of histone H3 lysine 36 methylation. *EMBO Rep.*, **20**, e46762.
 26. Meers, M.P., Henriques, T., Lavender, C.A., McKay, D.J., Strahl, B.D., Duronio, R.J., Adelman, K. and Matera, A.G. (2017) Histone gene replacement reveals a post-transcriptional role for H3K36 in maintaining metazoan transcriptome fidelity. *eLife*, **6**, e23249.
 27. Brown, J.C., McMichael, B.D., Vandadi, V., Mukherjee, A., Salzler, H.R. and Matera, A.G. (2024) Lysine-36 of drosophila histone H3.3 supports adult longevity. *G3 (Bethesda)*, **14**, jkae030.
 28. Lindehell, H., Glotov, A., Dorafshan, E., Schwartz, Y.B. and Larsson, J. (2021) The role of H3K36 methylation and associated methyltransferases in chromosome-specific gene regulation. *Sci. Adv.*, **7**, eabh4390.
 29. Soshnev, A.A., Josefowicz, S.Z. and Allis, C.D. (2016) Greater than the sum of parts: complexity of the dynamic epigenome. *Mol. Cell*, **62**, 681–694.
 30. Wang, H., Farnung, L., Dienemann, C. and Cramer, P. (2020) Structure of H3K36-methylated nucleosome-PWWP complex reveals multivalent cross-gyre binding. *Nat. Struct. Mol. Biol.*, **27**, 8–13.
 31. Koutna, E., Lux, V., Kouba, T., Skerlova, J., Novacek, J., Srb, P., Hexnerova, R., Svachova, H., Kukacka, Z., Novak, P., *et al.* (2023) Multivalency of nucleosome recognition by LEDGF. *Nucleic Acids Res.*, **51**, 10011–10025.
 32. Zhang, P., Du, J., Sun, B., Dong, X., Xu, G., Zhou, J., Huang, Q., Liu, Q., Hao, Q. and Ding, J. (2006) Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3. *Nucleic Acids Res.*, **34**, 6621–6628.
 33. Li, W., Cui, H., Lu, Z. and Wang, H. (2023) Structure of histone deacetylase complex Rpd3S bound to nucleosome. *Nat. Struct. Mol. Biol.*, **30**, 1893–1901.
 34. Musselman, C.A., Gibson, M.D., Hartwick, E.W., North, J.A., Gatchalian, J., Poirier, M.G. and Kutateladze, T.G. (2013) Binding of PHF1 Tudor to H3K36me3 enhances nucleosome accessibility. *Nat. Commun.*, **4**, 2969.
 35. Larschan, E., Alekseyenko, A.A., Gortchakov, A.A., Peng, S., Li, B., Yang, P., Workman, J.L., Park, P.J. and Kuroda, M.I. (2007) MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. *Mol. Cell*, **28**, 121–133.
 36. Albig, C., Wang, C., Dann, G.P., Wojcik, F., Schauer, T., Krause, S., Maenner, S., Cai, W., Li, Y., Girton, J., *et al.* (2019) JASPer controls interphase histone H3S10 phosphorylation by chromosomal kinase JIL-1 in Drosophila. *Nat. Commun.*, **10**, 5343.
 37. Bell, O., Wirbelauer, C., Hild, M., Scharf, A.N., Schwaiger, M., MacAlpine, D.M., Zilbermann, F., van Leeuwen, F., Bell, S.P., Imhof, A., *et al.* (2007) Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in Drosophila. *EMBO J.*, **26**, 4974–4984.
 38. Alekseyenko, A.A., Gorchakov, A.A., Zee, B.M., Fuchs, S.M., Kharchenko, P.V. and Kuroda, M.I. (2014) Heterochromatin-associated interactions of Drosophila HP1a with dADD1, HIP1, and repetitive RNAs. *Genes Dev.*, **28**, 1445–1460.
 39. Klymenko, T. and Muller, J. (2004) The histone methyltransferases trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep.*, **5**, 373–377.
 40. Tanaka, Y., Katagiri, Z., Kawahashi, K., Kioussis, D. and Kitajima, S. (2007) Trithorax-group protein ASH1 methylates histone H3 lysine 36. *Gene*, **397**, 161–168.
 41. Huang, C., Yang, F., Zhang, Z., Zhang, J., Cai, G., Li, L., Zheng, Y., Chen, S., Xi, R. and Zhu, B. (2017) Mrg15 stimulates Ash1 H3K36 methyltransferase activity and facilitates Ash1 trithorax group protein function in Drosophila. *Nat. Commun.*, **8**, 1649.
 42. Bell, O., Conrad, T., Kind, J., Wirbelauer, C., Akhtar, A. and Schubeler, D. (2008) Transcription-coupled methylation of histone H3 at lysine 36 regulates dosage compensation by enhancing recruitment of the MSL complex in Drosophila melanogaster. *Mol. Cell. Biol.*, **28**, 3401–3409.
 43. Lhoumaud, P., Hennion, M., Gamot, A., Cuddapah, S., Queille, S., Liang, J., Micas, G., Morillon, P., Urbach, S., Bouchez, O., *et al.* (2014) Insulators recruit histone methyltransferase dMes4 to regulate chromatin of flanking genes. *EMBO J.*, **33**, 1599–1613.
 44. Schmahling, S., Meiler, A., Lee, Y., Mohammed, A., Finkl, K., Tauscher, K., Israel, L., Wirth, M., Philippou-Massier, J., Blum, H., *et al.* (2018) Regulation and function of H3K36 di-methylation by the trithorax-group protein complex AMC. *Development*, **145**, dev163808.
 45. Schwartz, Y.B., Kahn, T.G., Stenberg, P., Ohno, K., Bourgon, R. and Pirrotta, V. (2010) Alternative epigenetic chromatin states of polycomb target genes. *PLoS Genet.*, **6**, e1000805.
 46. Villa, R., Jagtap, P.K.A., Thomae, A.W., Campos Sparr, A., Forne, I., Hennig, J., Straub, T. and Becker, P.B. (2021) Divergent evolution toward sex chromosome-specific gene regulation in Drosophila. *Genes Dev.*, **35**, 1055–1070.
 47. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods*, **9**, 676–682.
 48. Stirling, D.R., Swain-Bowden, M.J., Lucas, A.M., Carpenter, A.E., Cimini, B.A. and Goodman, A. (2021) CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinf.*, **22**, 433.
 49. Muller, M., Schauer, T., Krause, S., Villa, R., Thomae, A.W. and Becker, P.B. (2020) Two-step mechanism for selective incorporation of lncRNA into a chromatin modifier. *Nucleic Acids Res.*, **48**, 7483–7501.
 50. Chaouch, A., Berlandi, J., Chen, C.C.L., Frey, F., Badini, S., Harutyunyan, A.S., Chen, X., Krug, B., Hebert, S., Jeibmann, A., *et al.* (2021) Histone H3.3 K27M and K36M mutations de-repress transposable elements through perturbation of antagonistic chromatin marks. *Mol. Cell*, **81**, 4876–4890.
 51. Depierre, D., Perrois, C., Schickele, N., Lhoumaud, P., Abdi-Galab, M., Fosseprez, O., Heurteau, A., Margueron, R. and Cuvier, O. (2023) Chromatin in 3D distinguishes dMes-4/NSD and Hypb/dSet2 in protecting genes from H3K27me3 silencing. *Life Sci. Alliance*, **6**, e202302038.
 52. Eggers, N., Gkountromichos, F., Krause, S., Campos-Sparr, A. and Becker, P.B. (2023) Physical interaction between MSL2 and CLAMP assures direct cooperativity and prevents competition at composite binding sites. *Nucleic Acids Res.*, **51**, 9039–9054.
 53. Borner, K. and Becker, P.B. (2016) Splice variants of the SWR1-type nucleosome remodeling factor Domino have distinct functions during Drosophila melanogaster oogenesis. *Development*, **143**, 3154–3167.
 54. Girardot, C., Scholtalbers, J., Sauer, S., Su, S.Y. and Furlong, E.E. (2016) Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. *BMC Bioinf.*, **17**, 419.
 55. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
 56. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Subgroup, G.P.D.P. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
 57. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H. and Glass, C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, **38**, 576–589.
 58. Robinson, J.T., Thorvaldsdottir, H., Turner, D. and Mesirov, J.P. (2023) igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). *Bioinformatics*, **39**, btac830.
 59. Quinlan, A.R. (2014) BEDTools: the Swiss-Army tool for genome feature analysis. *Curr. Protoc. Bioinformatics*, **47**, 11.12.1–11.12.34.

60. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
61. R Core Team (2020) R: a language and environment for statistical computing. In: *R Foundation for Statistical Computing*. Vienna, Austria.
62. Anand, L. and Rodriguez Lopez, C.M. (2022) ChromoMap: an R package for interactive visualization of multi-omics data and annotation of chromosomes. *BMC Bioinf.*, **23**, 33.
63. Gu, Z., Eils, R. and Schlesner, M. (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, **32**, 2847–2849.
64. Lun, A.T. and Smyth, G.K. (2016) csaw: a bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. *Nucleic Acids Res.*, **44**, e45.
65. Henikoff, S., Henikoff, J.G., Sakai, A., Loeb, G.B. and Ahmad, K. (2009) Genome-wide profiling of salt fractions maps physical properties of chromatin. *Genome Res.*, **19**, 460–469.
66. Hendy, O., Serebreni, L., Bergauer, K., Muerdter, F., Huber, L., Nemcko, F. and Stark, A. (2022) Developmental and housekeeping transcriptional programs in *Drosophila* require distinct chromatin remodelers. *Mol. Cell*, **82**, 3598–3612.
67. Chintapalli, V.R., Wang, J. and Dow, J.A. (2007) Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nat. Genet.*, **39**, 715–720.
68. Sun, Z., Lin, Y., Islam, M.T., Koche, R., Hedehus, L., Liu, D., Huang, C., Vierbuchen, T., Sawyers, C.L. and Helin, K. (2023) Chromatin regulation of transcriptional enhancers and cell fate by the Sotos syndrome gene NSD1. *Mol. Cell*, **83**, 2398–2416.
69. Hinrichs, A.S., Karolchik, D., Baertsch, R., Barber, G.P., Bejerano, G., Clawson, H., Diekhans, M., Furey, T.S., Harte, R.A., Hsu, F., *et al.* (2006) The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res.*, **34**, D590–D598.
70. Egelhofer, T.A., Minoda, A., Klugman, S., Lee, K., Kolasinska-Zwiercz, P., Alekseyenko, A.A., Cheung, M.S., Day, D.S., Gadel, S., Gorchakov, A.A., *et al.* (2011) An assessment of histone-modification antibody quality. *Nat. Struct. Mol. Biol.*, **18**, 91–93.
71. Bonnet, J., Lindeboom, R.G.H., Pokrovsky, D., Stricker, G., Celik, M.H., Rupp, R.A.W., Gagneur, J., Vermeulen, M., Imhof, A. and Muller, J. (2019) Quantification of proteins and histone marks in *Drosophila* embryos reveals stoichiometric relationships impacting chromatin regulation. *Dev. Cell*, **51**, 632–644.
72. de Almeida, S.F., Grosso, A.R., Koch, F., Fenouil, R., Carvalho, S., Andrade, J., Levezinho, H., Gut, M., Eick, D., Gut, I., *et al.* (2011) Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. *Nat. Struct. Mol. Biol.*, **18**, 977–983.
73. Lindehell, H., Schwartz, Y.B. and Larsson, J. (2023) Methylation of lysine 36 on histone H3 is required to control transposon activities in somatic cells. *Life Sci. Alliance*, **6**, e202201832.
74. Kaminker, J.S., Bergman, C.M., Kronmiller, B., Carlson, J., Svirskas, R., Patel, S., Frise, E., Wheeler, D.A., Lewis, S.E., Rubin, G.M., *et al.* (2002) The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol.*, **3**, RESEARCH0084.
75. Holowatyj, A., Yang, Z.Q. and Pile, L.A. (2015) Histone lysine demethylases in *Drosophila melanogaster*. *Fly (Austin)*, **9**, 36–44.
76. Sharda, A. and Humphrey, T.C. (2022) The role of histone H3K36me3 writers, readers and erasers in maintaining genome stability. *DNA Repair (Amst.)*, **119**, 103407.
77. He, J., Xu, T., Zhao, F., Guo, J. and Hu, Q. (2023) SETD2-H3K36ME3: an important bridge between the environment and tumors. *Front. Genet.*, **14**, 1204463.
78. Vougiouklakis, T., Hamamoto, R., Nakamura, Y. and Saloura, V. (2015) The NSD family of protein methyltransferases in human cancer. *Epigenomics*, **7**, 863–874.
79. Zhang, Y., Guo, W., Feng, Y., Yang, L., Lin, H., Zhou, P., Zhao, K., Jiang, L., Yao, B. and Feng, N. (2023) Identification of the H3K36me3 reader LEDGF/p75 in the pancancer landscape and functional exploration in clear cell renal cell carcinoma. *Comput. Struct. Biotechnol. J.*, **21**, 4134–4148.
80. Nicetto, D., Donahue, G., Jain, T., Peng, T., Sidoli, S., Sheng, L., Montavon, T., Becker, J.S., Grindheim, J.M., Blahnik, K., *et al.* (2019) H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification. *Science*, **363**, 294–297.
81. Becker, J.S., McCarthy, R.L., Sidoli, S., Donahue, G., Kaeding, K.E., He, Z., Lin, S., Garcia, B.A. and Zaret, K.S. (2017) Genomic and proteomic resolution of heterochromatin and its restriction of alternate fate genes. *Mol. Cell*, **68**, 1023–1037.
82. Shipman, G.A., Padilla, R., Horth, C., Hu, B., Bareke, E., Vitorino, F.N., Gongora, J.M., Garcia, B.A., Lu, C. and Majewski, J. (2023) Systematic perturbations of SETD2, NSD1, NSD2, NSD3 and ASH1L reveals their distinct contributions to H3K36 methylation. bioRxiv doi: <https://doi.org/10.1101/2023.09.27.559313>, 18 October 2023, preprint: not peer reviewed.
83. Dorighi, K.M. and Tamkun, J.W. (2013) The trithorax group proteins kismet and ASH1 promote H3K36 dimethylation to counteract polycomb group repression in *Drosophila*. *Development*, **140**, 4182–4192.
84. Barral, A., Pozo, G., Ducrot, L., Papadopoulos, G.L., Sauzet, S., Oldfield, A.J., Cavalli, G. and Dejardin, J. (2022) SETDB1/NSD-dependent H3K9me3/H3K36me3 dual heterochromatin maintains gene expression profiles by bookmarking poised enhancers. *Mol. Cell*, **82**, 816–832.
85. Molenaar, T.M. and van Leeuwen, F. (2022) SETD2: from chromatin modifier to multipronged regulator of the genome and beyond. *Cell. Mol. Life Sci.*, **79**, 346.
86. Saha, P., Sowpati, D.T., Soujanya, M., Srivastava, I. and Mishra, R.K. (2020) Interplay of pericentromeric genome organization and chromatin landscape regulates the expression of *Drosophila melanogaster* heterochromatic genes. *Epigenetics Chromatin*, **13**, 41.
87. Li, W., Tian, W., Yuan, G., Deng, P., Sengupta, D., Cheng, Z., Cao, Y., Ren, J., Qin, Y., Zhou, Y., *et al.* (2021) Molecular basis of nucleosomal H3K36 methylation by NSD methyltransferases. *Nature*, **590**, 498–503.
88. Willcockson, M.A., Heaton, S.E., Weiss, C.N., Bartholdy, B.A., Botbol, Y., Mishra, L.N., Sidhwani, D.S., Wilson, T.J., Pinto, H.B., Maron, M.I., *et al.* (2021) H1 histones control the epigenetic landscape by local chromatin compaction. *Nature*, **589**, 293–298.
89. Hamagami, N., Wu, D.Y., Clemens, A.W., Nettles, S.A., Li, A. and Gabel, H.W. (2023) NSD1 deposits histone H3 lysine 36 dimethylation to pattern non-CG DNA methylation in neurons. *Mol. Cell*, **83**, 1412–1428.
90. Park, J.Y., Powell, R.T., Tripathi, D.N., Dere, R., Ho, T.H., Blasius, T.L., Chiang, Y.C., Davis, I.J., Fahey, C.C., Hacker, K.E., *et al.* (2016) Dual chromatin and cytoskeletal remodeling by SETD2. *Cell*, **166**, 950–962.
91. Kovatcheva, M., Melendez, E., Chondronasiou, D., Pietroccola, F., Bernad, R., Caballe, A., Junza, A., Capellades, J., Holguin-Horcajo, A., Prats, N., *et al.* (2023) Vitamin B(12) is a limiting factor for induced cellular plasticity and tissue repair. *Nat. Metab.*, **5**, 1911–1930.
92. Sun, Y., Ramesh, V., Wei, F. and Locasale, J.W. (2023) Methionine availability influences essential H3K36me3 dynamics during cell differentiation. bioRxiv doi: <https://doi.org/10.1101/2023.11.22.568331>, 22 November 2023, preprint: not peer reviewed.
93. Corless, S., Singh, N.-P., Benabdalla, N.S., Böhm, J., Simon, A.M., Dolejš, V., Anders, S., Banito, A. and Earhardt, S. (2023) The bromodomain inhibitor JQ1 is a molecular glue targeting centromeres. bioRxiv doi:

- <https://doi.org/10.1101/2023.03.15.532673>, 15 March 2023, preprint: not peer reviewed.
94. Gershman,A., Sauria,M.E.G., Guitart,X., Vollger,M.R., Hook,P.W., Hoyt,S.J., Jain,M., Shumate,A., Razaghi,R., Koren,S., *et al.* (2022) Epigenetic patterns in a complete human genome. *Science*, **376**, eabj5089.
95. Villasenor,R., Pfaendler,R., Ambrosi,C., Butz,S., Giuliani,S., Bryan,E., Sheahan,T.W., Gable,A.L., Schmolka,N., Manzo,M., *et al.* (2020) ChromID identifies the protein interactome at chromatin marks. *Nat. Biotechnol.*, **38**, 728–736.
96. Mito,Y., Henikoff,J.G. and Henikoff,S. (2005) Genome-scale profiling of histone H3.3 replacement patterns. *Nat. Genet.*, **37**, 1090–1097.
97. McKittrick,E., Gafken,P.R., Ahmad,K. and Henikoff,S. (2004) Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 1525–1530.