dSet1 Is the Main H3K4 Di- and Tri-Methyltransferase Throughout *Drosophila* Development

Graham Hallson, Robert E. Hollebakken,[1] Taosui Li,[2] Monika Syrzycka, Inho Kim, Shawn Cotsworth, Kathleen A. Fitzpatrick, Donald A. R. Sinclair, and Barry M. Honda[3]

Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada

ABSTRACT In eukaryotes, the post-translational addition of methyl groups to histone H3 lysine 4 (H3K4) plays key roles in maintenance and establishment of appropriate gene expression patterns and chromatin states. We report here that an essential locus within chromosome 3L centric heterochromatin encodes the previously uncharacterized *Drosophila melanogaster* ortholog (*dSet1*, *CG40351)* of the Set1 H3K4 histone methyltransferase (HMT). Our results suggest that dSet1 acts as a "global" or general H3K4 di- and trimethyl HMT in *Drosophila*. Levels of H3K4 di- and trimethylation are significantly reduced in *dSet1* mutants during late larval and post-larval stages, but not in animals carrying mutations in genes encoding other well-characterized H3K4 HMTs such as *trr*, *trx*, and *ash1*. The latter results suggest that Trr, Trx, and Ash1 may play more specific roles in regulating key cellular targets and pathways and/or act as global H3K4 HMTs earlier in development. In yeast and mammalian cells, the HMT activity of Set1 proteins is mediated through an evolutionarily conserved protein complex known as Complex of Proteins Associated with Set1 (COMPASS). We present biochemical evidence that *dSet1* interacts with members of a putative *Drosophila* COMPASS complex and genetic evidence that these members are functionally required for H3K4 methylation. Taken together, our results suggest that dSet1 is responsible for the bulk of H3K4 di- and trimethylation throughout *Drosophila* development, thus providing a model system for better understanding the requirements for and functions of these modifications in metazoans.

POST-TRANSLATIONAL modification of histones can alter the local chromatin environment and affect the recruitment of transcriptional regulatory machinery (Ebert *et al.* 2006; Vermuelen *et al.* 2007; Choi and Howe 2009; Bannister and Kouzarides 2011). These modifications can play diverse roles in transcriptional activation or silencing (reviewed in Gartner *et al.* 2011), and cross talk between different activating and silencing modifications may fine-tune levels of transcription (reviewed in Munshi *et al.* 2009; Lee *et al.* 2010; Gartner *et al.* 2011).

doi: 10.1534/genetics.111.135863
Manuscript received July 13, 2011; accepted for publication October 22, 2011
Supporting information is available online at http://www.genetics.org/content/suppl/2011/11/02/genetics.111.135863.DC1.
[1]Present address: B.C. Center for Excellence in HIV/AIDS, St. Paul's Hospital, Vancouver, BC V6Z 1Y6, Canada.
[2]Present address: Department of Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.
[3]Corresponding author: Department of Molecular Biology and Biochemistry, Simon Fraser University, 8888 University Dr., Burnaby, BC V5A 1S6, Canada. E-mail: honda@sfu.ca

The post-translational addition of up to three methyl groups to histone H3 lysine 4 (H3K4) residues (H3K4me1, H3K4me2, and H3K4me3) correlates with active transcription (Bernstein *et al.* 2002; Santa-Rosa *et al.* 2002). H3K4 di- and trimethylation is often enriched at the promoter and 5′ coding regions of active genes, whereas H3K4 monomethylation is commonly found near the 3′ ends of active genes and within enhancer elements (Bernstein *et al.* 2002; Santa Rosa *et al.* 2002; Pokholok *et al.* 2005; Heintzman *et al.* 2007). Although the mechanisms of methyl-H3K4-mediated transcriptional activation are not fully elucidated, trimethyl-H3K4 is thought to act as a docking scaffold for the recruitment of the transcription pre-initiation complex and transcriptionally activating chromatin-remodelling complexes (Wysocka *et al.* 2006; Vermuelen *et al.* 2007).

In the budding yeast, *Saccharomyces cerevisiae*, all mono-, di-, and trimethylation of H3K4 is catalyzed by the Set1 enzyme, and the enzymatic activity of Set1 is modulated through a multi-subunit protein complex known as the Complex of Proteins Associated with Set1 (COMPASS) (Miller *et al.* 2001; Roguev *et al.* 2001; Nagy *et al.* 2002; Dehe

et al. 2005). COMPASS is evolutionarily conserved, with functional orthologs of Set1 acting as major H3K4 histone methyltransferases (HMTs) in metazoans (Lee and Skalnik 2005; Lee *et al.* 2007; Simonet *et al.* 2007).

Higher metazoans possess additional H3K4 methylases, the mixed lineage leukemia (MLL) class of proteins, which act through distinct complexes similar to COMPASS (reviewed by Eissenberg and Shilatifard 2010). The MLL proteins (MLL1–5) are required at limited but important subsets of gene targets, such as homeotic and hormone response genes (J. Lee *et al.* 2008; S. Lee 2008; Wang *et al.* 2009; reviewed by Ansari and Mandal 2010; Eissenberg and Shilatifard 2010). H3K4 methylases identified in *Drosophila melanogaster* to date include Trx (homologous to MLL1–2), Trr (homologous to MLL3–4), and Ash1 (Beisel *et al.* 2002; Byrd and Shearn 2003; Sedkov *et al.* 2003; Smith *et al.* 2004). Ash1 and Trx are members of the Trithorax group of proteins that antagonize Polycomb group-mediated gene silencing (Klymenko and Muller 2004). In addition, Trx methylates H3K4 at heat-shock loci upon induction and appears to be required for mediating stress responses to heat stimuli (Smith *et al.* 2004). Trr is recruited to and required for H3K4 methylation at gene targets responsive to the insect nuclear hormone ecdysone (Sedkov *et al.* 2003). Although these HMTs are known to catalyze H3K4 methylation and are widely believed to act as the main global H3K4 methylases in *Drosophila* (Beisel *et al.* 2002; Byrd and Shearn 2003; Sedkov *et al.* 2003; Smith *et al.* 2004), the functional roles of the *Drosophila* ortholog of Set1 (dSet1) have remained undefined, largely because its location within centric heterochromatin makes genetic and molecular analysis particularly challenging.

In our efforts to functionally annotate essential heterochromatic genes in *Drosophila* (see Fitzpatrick *et al.* 2005; Schulze *et al.* 2005; Hallson *et al.* 2008; Sinclair *et al.* 2009), we have linked *dSet1/CG40351*, the *Drosophila set1* ortholog, to an essential genetic locus previously known as *lethal 5* or *l(3L)h5*, residing in chromosome 3L centric heterochromatin (Marchant and Holm 1988; Fitzpatrick *et al.* 2005). Surprisingly, we find that dSet1, and not Trx, Trr, or Ash1, acts as the main global H3K4 di- and trimethylase throughout *Drosophila* development. We also provide genetic and molecular evidence that *Drosophila* orthologs of other COMPASS members are required for H3K4 methylation and physically interact with dSet1. Our findings establish a foundation for examining transcriptional regulatory mechanisms underlying this key post-translational modification.

Materials and Methods

Drosophila cultures/crosses

Cultures were maintained on standard yeast–cornmeal–molasses media, and crosses were performed at 25° unless otherwise stated. All background strains, *GAL4* drivers, RNAi lines, and stocks used for generating mosaics are publicly available [through the Bloomington, Kyoto, and Vienna Drosophila Rnai Center (VDRC) stock centers] and have been described previously (http://www.flybase.org and http://stockcenter.vdrc.at). Most mutations affecting and deficiencies removing the *lethal 5* region have been described previously (Marchant and Holm 1988; Fitzpatrick *et al.* 2005; Hallson *et al.* 2008).

The following strains were generous gifts*: G5 e/TM3 Sb* and *G12 e/TM3 Sb* (D. Deitcher); *Z480/TM3 Sb* (C. Zuker); trx^{Z11}*/TM3 Kr-GFP* and trx^{B11}*/TM3 Kr-GFP* (H. Brock); $ash1^{B1}$*/TM3 Sb* and $ash1^{B7}$*/TM3 Sb* (unpublished alleles from J. Kennison); trr^{1}*/FM7* (A. Mazo); and *UAS-ash2-3X HA/Cyo*; *MKRS/TM6* (M. Corominas). Unless noted, stocks listed have been described previously (Petruk *et al.* 2001; Sedkov *et al.* 2003; Koundakjian *et al.* 2004; Smith *et al.* 2004; Beltran *et al.* 2007; Fitzpatrick *et al.* 2005).

To isolate mutant animals at non-adult stages, mutants were balanced with the GFP-marked balancer chromosome, *TM3 Sb Ser Twi-GAL4 UAS-GFP*. *Trans*-heterozygous mutant animals were produced by crossing combinations of these GFP-balanced strains and were selected on the basis of the absence of GFP fluorescence. RNAi knockdown animals were generated in crosses using either VDRC RNAi lines (*dSet1*: VDRC #40682, #40683, and #10833; *wds*: VDRC #105371; *dRbbp5*: VDRC #106139; *dWdr82*: VDRC #25246; *hcf*: VDRC #46998; *trx*: VDRC #108122) or the Transgenic RNAi Resource Project (TRiP) RNAi line BL29563, which targets *trr*. Each of these RNAi lines was crossed to the *tub-GAL4/TM3 Sb Ser Twi-Gal4 UAS-GFP* driver line at 29° and RNAi knockdown animals were selected on the basis of the absence of GFP fluorescence. RNAi lines used in this study are not predicted to produce nonspecific 19-mers resulting in "off-target" effects (data not shown).

Single-embryo PCR

Homozygous embryos were selected from *lethal 5* mutation or deficiency strains as described previously (Hallson *et al.* 2008). To detect the presence or absence of the *CG40351* genomic sequence in various deficiencies, DNA from homozygous embryos was amplified using the primers SER2F (5′-AGGTTTGTTCATAATTGACACAGATGC-3′) and SER3R (5′-TCATACCTTTCCCATTACAGACTTTTG-3′). PCR conditions were as follows: 94° for 5 min, followed by 33 cycles of 94° for 30 sec, 60° for 40 sec, and 72° for 1 min. *Grip84* was also amplified to ensure DNA integrity as described previously (Hallson *et al.* 2008). Lesions associated with *lethal 5* alleles were identified as described in the supporting information, File S1.

Quantitative PCR

For details, see File S1.

Nuclear extracts

All steps were performed at 4°. Nuclei were isolated as described previously (Shaffer *et al.* 1994). Nuclei were sonicated

for 5 min (30 sec on followed by 30 sec off) on medium power in a Diagenode Bioruptor in PIPES buffer [10 mM PIPES (pH 7.0), 300 mM sucrose, 3 mM MgCl, 1 mM EGTA, and 1× Protease Inhibitor Cocktail I (Calbiochem)] containing 400 mM NaCl and 0.5% Triton X-100 or HEPES high-salt buffer [20 mM HEPES (pH 7.5), 300 mM NaCl, 10% glycerol, 3 mM MgCl, 2mM DTT, and 0.1% Triton X-100] and then samples were centrifuged at 13,000 × *g* in a Heraeus Biofuge 13 tabletop centrifuge for 15 min and the pellet was discarded.

Western blots

Western blots were performed following standard procedures. The following primary antibodies and antibody concentrations were used: H3K4me3 (Active Motif, 39159; lot nos.: 573636, 961715) at 1/5000; H3K4me2 (Millipore, 07-030; lot nos.: DAM147603, DAM1724042) at 1/5000, H3K4me (Abcam, ab8895; lot nos.: 573636, 961715) at 1/5000; H3 (Abcam, ab1791) at 1/10,000; hemagluttinin (HA) tag (ABM, G036) at 1/2500; anti-FLAG (ABM, G0191) at 1/200; and anti-GST (ABM, G0181) at 1/10,000. Binding specificities of each lot of methyl-H3K4 antibodies have been tested by us and others: these data can be found in Figure S1 in File S1 and at the Antibody Validation Database (http://compbio.med.harvard.edu/antibodies/about) (Egelhofer *et al.* 2010). Quantitative Western blotting details are provided in File S1.

Binding assays

The cDNA sequences encoding the full-length dSet1 protein or the dSet1 carboxy terminus (amino acids 1329–1641; dSet1np) were cloned into the *CMV-HA* vector (Promega) in frame with an N terminus HA tag (described in detail in File S1). HEK293T cells cultured in 75-cm^2 flasks under standard conditions and grown to 50% confluence were transfected with 5 μg DNA and 50 μl of Polyfect (Qiagen) following the manufacturer's instructions. Thirty-six hours post transfection, nuclear extractions were performed as described above. These extracts were diluted fourfold in PIPES buffer (with a final concentration of 150 mM NaCl and 0.25% Triton X-100) and precleared overnight with GST attached to glutathione (GSH) sepharose. A total of 600 μl of the precleared extracts were incubated with 30 μl of GST, GST-Ash2, GST-Wds (see File S1 for cloning details), or nothing (mock) attached to GSH sepharose beads (GE Life Sciences) for 3 hr with gentle rotation at 4°. The beads were washed with 1 ml of PIPES buffer (with 0.25% Triton X-100 and 150 mM NaCl) for 5 min followed by liquid aspiration three times. Samples were run on SDS-PAGE gels, followed by Western blot analysis.

Immunoprecipitation

Nuclear protein extracts were prepared from adult flies of the genotype *ash2-3X HA/+*; *tub-Gal4/UAS-dSet1-2X FLAG* as described above. These extracts were diluted twofold in PIPES buffer (with 0.25% Triton X-100 and 150 mM NaCl) and precleared overnight with 40 μl of protein G agarose (PGA) beads (Genscript) at 4°. Anti-HA antibodies (Abcam, ab9110) were added at a 1:200 dilution to precleared extracts and extracts were rotated overnight followed by the addition of 30 μl of PGA beads (Genscript) for 30 min at 4°. The PGA beads were washed a total of three times for 5 min each with PIPES buffer (containing 0.25% Triton and 150 mM NaCl) at 4°. After washing, the PGA beads were boiled in Laemmli sample buffer and run on 5% SDS-PAGE gels, followed by Western blot analysis. A mock immunoprecipitation experiment was also performed as above without antibodies added.

Glycerol gradient centrifugation

Protein extracts (500 μl) from ~50 adult flies of the genotype *ash2-3X HA/+*; *tub-Gal4/UAS-dSet1-2X FLAG* were loaded onto a 16–40% glycerol step gradient (in HEPES high-salt buffer) with 1-ml steps in 3% increments in 16- × 102-mm polyallomer tubes (Beckman). Gradients were run at 30,000 × *g* for 18 hr at 4° in an SW32.1 swinging bucket rotor (Beckman). Fractions were run on SDS-PAGE gels, followed by Western blot analysis. Approximately 500-μl fractions were collected from top to bottom. Size markers (MWGF1000, Sigma-Aldrich) were run on parallel gradients and detected in fractions run on SDS-PAGE gels by Coomassie staining.

Immunostaining

Head structures from third instar larvae were isolated and fixed in 4% paraformaldehyde in PBST (1× PBS + 0.1% Triton X-100) and blocked in 1% BSA diluted in PBST, followed by incubation overnight at 4° with the primary antibodies rabbit anti-H3K4me3 (ActiveMotif) and mouse anti-phosphotyrosine-100 (Cell Signaling) diluted to 1:500 and 1:200, respectively, in PBT. Six 10-min washes with PBST were performed to remove unbound primary antibodies, followed by incubation with the secondary antibodies (goat anti-mouse DyeLight 694 1:500 (Thermo Scientific) and goat anti-rabbit Texas red 1:500 (Cell Signaling) for 2 hr and six 10-min washes with PBST. Imaginal discs were dissected and mounted on slides containing VectaShield mounting medium (Cell Signaling). Images were obtained on a Zeiss LSM140 spinning disc confocal microscope.

Results

dSet1 maps to the lethal 5/l(3L)h5 locus

dSet1 (*CG40351*), the *Drosophila* ortholog of Set1, has been difficult to characterize because it resides in centric heterochromatin on the left arm of *Drosophila* chromosome 3 (3L). To determine the location of *dSet1* along a genetic map of 3L centric heterochromatin (Schulze *et al.* 2001; Fitzpatrick *et al.* 2005; M. Syrzycka, unpublished data), we attempted to PCR-amplify the second and third exons of *dSet1* using, as template, DNA isolated from embryos carrying deletions of 3L centric heterochromatin (Figure 1A). In

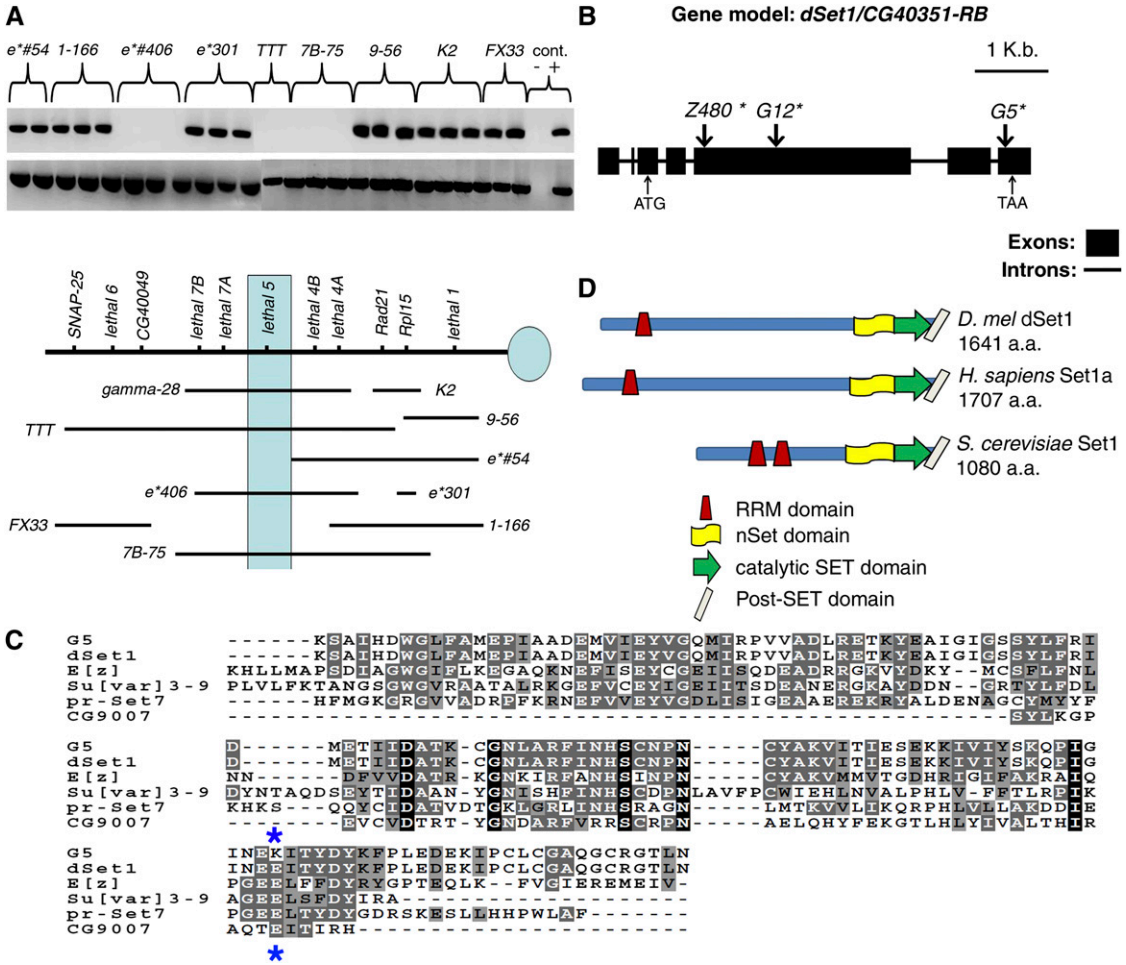

Figure 1 The *lethal 5* locus encodes the *Drosophila* ortholog of the Set1 H3K4 methylase (*dSet1*). (A) Using single-embryo PCR mapping, we demonstrate that *CG40351* is absent in deficiencies that remove the *lethal 5* and *lethal 4B* loci. (Top row gel lanes) *CG40351* amplicons. (Bottom row gel lanes) *Grip84* amplicons (+ve control). (B) Transcript model of *dSet1/CG40351-RB*. The position of the start codon (ATG) and stop codon (TAA) are indicated with small arrows. Locations of mutations are given below relative to the translational start site and are marked by large arrows along the gene model. Descriptions of mutations identified are the following: mutant G12 has a deletion of nt A1717, leading to a premature stop at amino acid 596; mutant Z480 has a T-to-A transversion at nt 548, resulting in premature termination at amino acid 183; and mutant G5 has a G-to-A transition at nt 4713, resulting in an E1613K missense mutation. (C) The alignment of diverse SET domains present in *Drosophila* demonstrates that the *G5* missense mutation changes a conserved residue (E1613, marked with asterisks) thought to be required for methylase activity (Wilson *et al.* 2002). (D) Structural comparisons indicate significant homology between dSet1 (CG40351), *Homo sapiens* Set1a, and *S. cerevisiae* Set1.

this way, we determined that *dSet1* resides in the vicinity of the *lethal 5/l(3)h5*, *lethal 7A/l(3)h7A*, and *lethal 7B/l(3)h7B* loci (Figure 1A). Further PCR mapping of *dSet1* using small chromosomal aberrations that genetically complement *lethal 7A* and *7B* alleles, but not *lethal 5* alleles, narrowed the candidacy of *dSet1* to *lethal 5* (data not shown).

lethal 5 mutations contain corresponding lesions in dSet1

To determine if EMS-induced *lethal 5* mutant alleles result from mutations in *dSet1*, we PCR-amplified and sequenced the *dSet1* genic region from homozygous *lethal 5* mutant embryos. Figure 1B shows a correspondence between *lethal 5* mutant alleles and molecular lesions in *dSet1*. The mutations identified were predicted to strongly reduce dSet1 function (described in the legend to Figure 1B).

lethal 5 mutants are rescued with a dSet1 cDNA transgene

Most *lethal 5* mutants hemizygous with *Df(3L)γ28* (which completely removes *dSet1*) are predominantly lethal at the pupal stage with low levels of late L3 larval lethality (<20%); we obtained similar patterns of lethality by driving *dSet1*-targeting RNAi constructs ubiquitously (data not shown) (Table 1). We were able to rescue lethality associated with complete loss of dSet1 function [amorphic *G12* allele over the deficiency *Df(3L)γ28*] by expressing a *dSet1* cDNA transgene (Table 2). This result unequivocally links phenotypic defects resulting from *lethal 5* mutations to molecular defects in the *dSet1* coding sequence.

Table 1 RNAi knockdown of Drosophila COMPASS members results in lethality

RNAi target	VDRC ID	Total flies	*dsRNA + GAL4* flies expected	*dsRNA + GAL4* flies observed	Predominant lethal stage
CG40351	40682 (ch2)	202	101	0	Pupal[a]
CG40351	40683 (ch2)	158	79	0	Pupal[a]
CG40351	10833 (ch3)	189	63	0	Pupal[a]
wds	105371 (ch2)	343	172	0	L1–L3 larval[b]
CG5585/dRbbp5	106139 (ch2)	247	148	0	Pupal
CG17293/dWdr82	25246 (ch2)	287	144	0	Pupal
Ash2+Dcr2	7141 (ch2)	230	115	0	Pupal[c]
dHcf1	46999 (ch3)	108	54	9	Semilethal/pupal[d]

RNAi-mediated depletion of transcripts encoding *Drosophila* COMPASS homologs was accomplished by driving specific double-strand RNA targeting sequences with constitutively expressed *GAL4* (*UAS-dsRNA/UAS-dsRNA* × *tub-GAL4/TM3 Sb* or *UAS-dsRNAi/TM3 Sb* × *tub-GAL4/TM3 Sb* performed at 29°). ch, chromosome.
[a] *dSet1* mutants display similar lethal phases.
[b] Based on analysis of wds mutant phenotypes described in Shannon *et al.* (1972).
[c] Based on analysis of ash2 mutant phenotypes described in Adamson and Shearn (1996).
[d] Survivors display phenotypic defects of varying penetrance including held-out wings and mild wing vein and tergite defects.

Global levels of H3K4me2 and H3K4me3 are reduced in dSet1 mutants

The predicted dSet1 protein appears structurally orthologous to the Set1 class of H3K4 methylases present in yeast and mammals (Figure 1D) (Roguev *et al.* 2001; Lee and Skalnik 2005; Lee *et al.* 2007). To determine if dSet1 functions similarly to Set1 in other organisms, we quantified levels of methylated H3K4 in nuclear extracts from wild-type (*OreR*) or amorphic *dSet1* mutant animals (*G12/γ28*) using antibodies specific for each methylated form of H3K4 (details of our methodology and an example of this approach are provided, as supporting information and Figure S2 respectively, in File S1). Our results suggest that dSet1 acts as a major H3K4 di- and trimethyltransferase: relative to wild-type levels, we observed reductions of di- and trimethyl H3K4, but not monomethyl H3K4, in *dSet1* mutant animals (Figure 2). Reductions of di- and trimethyl H3K4 in *dSet1* mutants relative to wild-type levels appear to increase over developmental time (Figure 2). We have confirmed the requirement of dSet1 for H3K4 di- and trimethylation by using other *dSet1* mutant allele combinations (Figure S3 in File S1) and by observing similar decreases after knocking down *dSet1* via RNAi (VDRC line #40683).

Other Drosophila COMPASS members are required for bulk H3K4 methylation

Since, in other eukaryotes, the Set1 class of proteins act through conserved multimeric protein complexes (Miller *et al.* 2001; Roguev *et al.* 2001; Nagy *et al.* 2002; Lee and Skalnik 2005; Lee *et al.* 2007), we predicted that a similar complex exists in *Drosophila*. RNAi knockdown of several genes encoding proteins orthologous to members of the human and yeast COMPASS complexes demonstrates that putative *Drosophila* COMPASS members are essential (Table 1). Loss of function via mutation or RNAi-induced depletion of several putative COMPASS members (Ash2, Wds, dRbbp5, and dWdr82) results in dramatic reductions of bulk H3K4 methylation levels (Figure 3). In addition, *hcf* knockdown results in slightly decreased H3K4 trimethylation—to ~60% of wild-type levels on the basis of scanned band intensities (Figure 3, quantitation not shown). These results implicate *Drosophila* COMPASS members as key players required for global H3K4 methylation.

dSet1 is a member of a Drosophila COMPASS complex

To demonstrate that dSet1 is present in a high-molecular-weight complex similar to yeast and mammalian COMPASS, we used glycerol gradients to fractionate high-molecular-weight species from nuclear extracts isolated from adult flies overexpressing tagged forms of Ash2 and dSet1. In this way, we isolated tagged dSet1 (predicted molecular weight ~190 kDa) in a complex with a molecular mass of ~667 kDa (fraction 13, Figure 4A), correlating closely with the additive mass of *Drosophila* COMPASS members (~610 kDa). We noted a significant presence of dSet1 in a low-molecular-weight form (~146 kDa, fraction 5), likely corresponding to unbound dSet1 fractionating slightly more slowly than expected. A large proportion of tagged Ash2 (molecular weight ~67 kDa) copurifies in fraction 13 (~667 kDa) with dSet1, suggesting that these two proteins are indeed present together in a complex. Ash2 also appears in other fractions, although to a lesser extent, which can be explained by the presence of multiple complexes containing Ash2 in flies,

Table 2 *lethal 5* alleles placed over a deficiency [*Df(3L)γ28*]

Genotype of rescued flies	Total flies scored	Rescued flies
w^{1118}/UAS-CG40351-4F-X; Act5C-Gal4/+; G5 e/Df(3L)γ28	197	25
w^{1118}/UAS-CG40351-4F-X; Act5C-Gal4/+; G12 e/Df(3L)γ28	221	30
w^{1118}/UAS-CG40351-4F-X; Act5C-Gal4/+; Z480/Df(3L)γ28	128	10

lethal 5 alleles were rescued with a *dSet1 cDNA* transgene (*UAS-CG40351-4F-X*) driven by GAL4 mimicking the pattern of *Act5C* expression The progeny reported were generated by crossing *w^{1118}/w^{1118}; Act5C-Gal4/Cyo; Df(3L)γ28/TM6 Hu Tb* flies to flies of the genotype *UAS-CG40351-4F-X/Y; +/+; G5 e/TM6 Hu Tb*; *UAS-CG40351-4F-X/Y; +/+; G12 e/TM6 Hu Tb* or *UAS-CG40351-4F-X/Y; +/+; Z480 e/TM6 Hu Tb*.

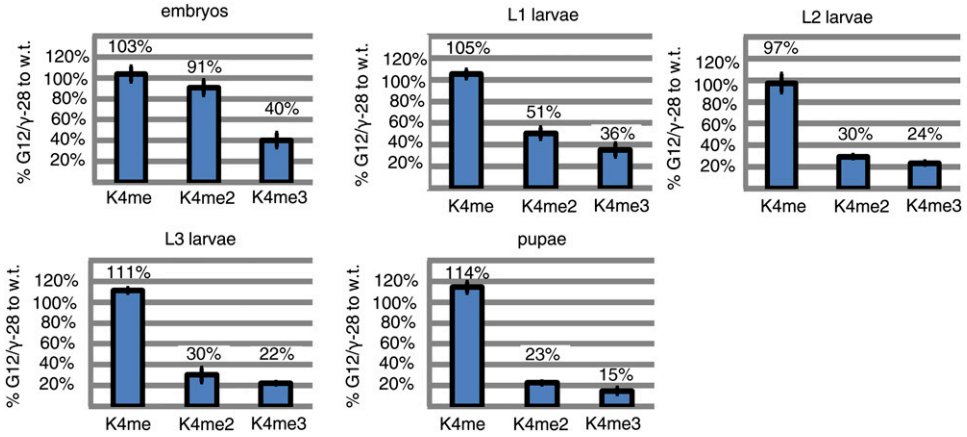

Figure 2 *dSet1* mutations result in dramatic losses of di- and trimethylated H3K4 at multiple developmental time points. H3K4 methylation levels from nuclear extracts were detected on Western blots using H3K4 methyl-specific antibodies. Extracts were normalized to total histone H3 present. The H3K4 methylation levels measured from wild-type (*OreR*) and *dSet1* mutants (*G12/γ-28*) were compared across a range of developmental time points: embryonic extracts (14–20 hr after egg lay), L1 extracts (34–40 hr post egg lay), L2 larval extracts (56–64 hr post egg lay), L3 larval extracts (104–110 hr post egg lay), and pupal extracts (7–8 days post egg lay).

yeast, and mammals (Papoulas *et al.* 1998; Roguev *et al.* 2003; Secombe *et al.* 2007; Ansari and Mandal 2010; Eissenberg and Shilatifard 2010).

To obtain more direct evidence that dSet1 and Ash2 interact within COMPASS, we immunoprecipitated tagged Ash2 from whole-fly nuclear extracts and, by doing so, coprecipitated dSet1 (Figure 4B). We also performed *in vitro* binding assays: HA-tagged dSet1 or a truncated version of dSet1 containing the catalytic carboxy terminus (aa 1329–1641) were pulled down by GST-tagged Ash2 or Wds attached to glutathione-conjugated beads (Figure 5). These results indicate that dSet1 directly interacts with Ash2 and Wds and that the catalytic carboxy-terminal region of dSet1 is sufficient for facilitating these interactions (Figure 5).

Trx, Trr, and Ash1 are not late-acting global H3K4 methylases

Given the surprising result that dSet1 is required for most H3K4 di- and trimethylation from the larval stages onward, we asked whether Trr, Trx, and Ash1 are also required for bulk H3K4 methylation at these same developmental stages. Levels of H3K4 methylation observed in *trx* mutant (trx^{B11}/trx^{Z11}) or RNAi knockdown larvae were similar to wild-type levels; a similar result was obtained with *ash1* mutant ($ash1^{B1}/ash1^{B7}$) (Figure 6, A and B). Our findings are consistent with those of Srinivasan *et al.* (2008) who demonstrate that H3K4 trimethylation patterns in third instar salivary glands are not significantly altered in the presence of *ash1* and *trx* mutations (Srinivasan *et al.* 2008).

Since *trr* mutants arrest during embryogenesis, we were unable to compare levels of methyl-H3K4 between wild-type and *trr* mutant whole larvae. Instead, we generated clonal patches homozygous for the trr^1 null allele in third instar larval haltere discs. We noted no differences in H3K4 trimethylation levels between trr^1/trr^1 haltere disc cells (marked by an absence of GFP) and cells containing functional Trr (marked by the presence of GFP) (Figure 6C), which is in stark contrast to the clear reductions of H3K4me3 that we observed in *dSet1* mutant imaginal discs (File S1, Figure S4). Levels of H3K4 methylation present in *trr* RNAi knockdown larvae were also comparable to wild-type amounts (Figure 6B). Taken together, our results indicate Trx, Trr,

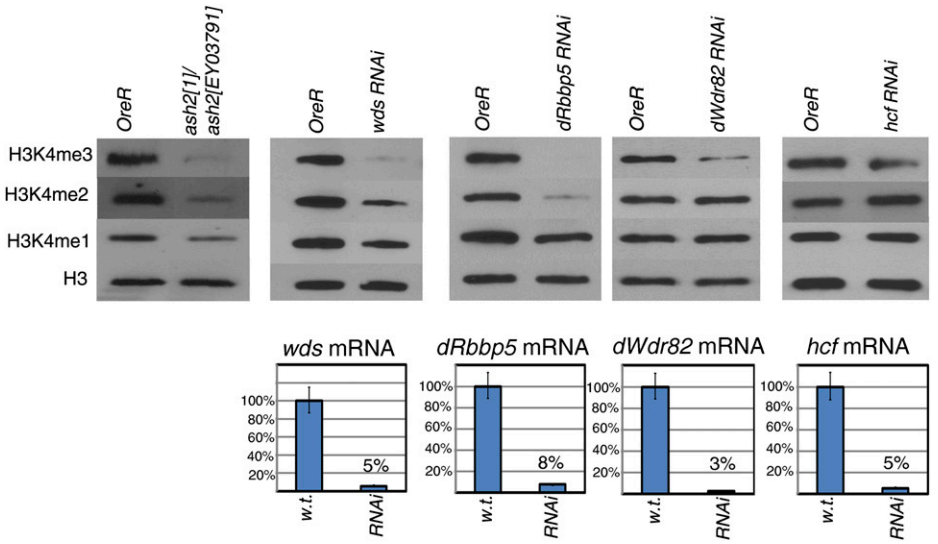

Figure 3 The loss of function or depletion of *Drosophila* COMPASS members results in reduced global H3K4 methylation. A comparison of global H3K4 methylation levels in L3 larvae containing a *trans*-heterozygous combination of *ash2* mutant alleles ($ash2^1/ash2^{EY03971}$) or expressing double-strand RNA targeting *wds*, *dRbbp5*, *dWdr82*, and *hcf*, relative to wild-type (*OreR*) larvae. Levels of mRNA knockdown in RNAi experiments are represented as a percentage of wild-type levels and presented as bar graphs below.

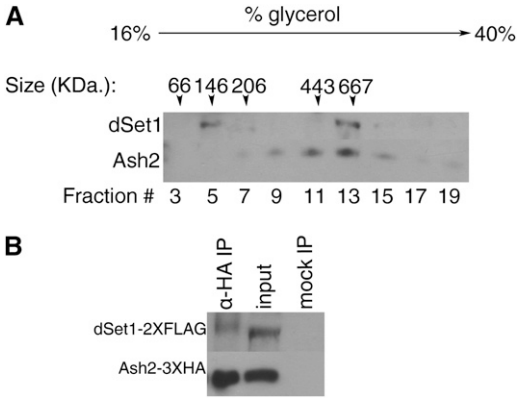

Figure 4 dSet1 is a component of a macromolecular complex. (A) Size fractionation of adult nuclear extracts containing dSet1-2X FLAG and Ash2-3X HA on a 16–40% glycerol gradient. Fractions were recovered from top to bottom (left to right) and detected with anti-HA and anti-FLAG antibodies. The numbers listed above lanes indicate sizes (in kDa) of markers that were isolated at the respective fraction number when run on a parallel gradient. Note that the peak of dSet1-2X FLAG recovery occurs in a high-molecular-weight fraction (fraction 13, 667 kDa) and that significant levels of ash2-3X HA are also recovered in this fraction. (B) Anti-HA immunoprecipitation of Ash2-3X HA from adult nuclear extracts isolated from the genotype *UAS-ash2-3X HA/+; tub-Gal4/UAS-dSet1-2X FLAG*. dSet1 and Ash2 were detected with anti-FLAG and anti-HA antibodies, respectively. Note that dSet1 coprecipitates with Ash2.

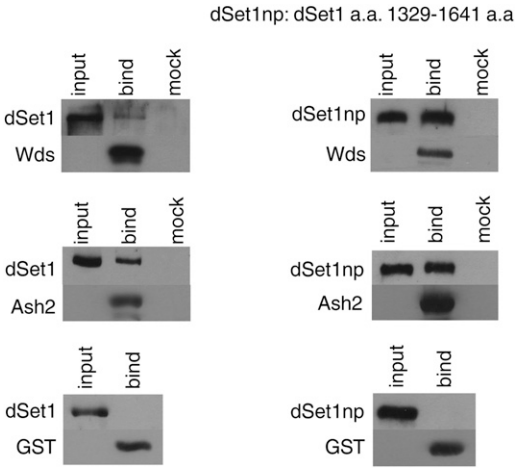

Figure 5 dSet1 physically interacts with the *Drosophila* COMPASS members Wds and Ash2. To identify binding partners of dSet1, *in vitro* binding assays were performed by incubating a solution containing HA-tagged versions of dSet1 or a fragment containing the catalytic carboxy terminus of dSet1 (HA-dSet1np amino acids 1329–1641) ("input" lanes) with GST, GST-Wds, or GST-Ash2 immobilized on glutathione sepharose beads ("bind" lanes) or unbound beads ("mock" lanes), followed by washes to remove nonspecific interactors. Bound products were detected on Western blots using anti-HA and anti-GST antibodies.

or Ash1 are not required for bulk H3K4 methylation at the larval stages.

Discussion

Our results indicate that dSet1 acts as the main global H3K4 methylase throughout *Drosophila* development and is required for completion of late developmental stages. Although developmental roles of Set1 have been reported in *Caenorhabditis elegans* (Simonet *et al.* 2007), to our knowledge, this is the first report that a Set1 ortholog is essential for the somatic development of a multicellular organism.

We cannot rule out a persisting maternal contribution of dSet1 mRNA or protein to the catalysis of bulk H3K4me2/me3 during early developmental stages. Indeed, RNA sequencing data available on FlyBase (http://www.flybase.org) indicate that dSet1 is present in embryos aged 0–2 hr post egg lay, suggesting significant maternal loading of *dSet1* transcripts (Daines *et al.* 2011). However, knocking down this maternal contribution during embryogenesis by expressing *dSet1* RNAi in a *dSet1* mutant background only slightly increases lethality at the L3 larval stage relative to *dSet1* mutants (data not shown). It is possible that, in addition to dSet1, other H3K4 methylases are responsible for "early" bulk H3K4 methylation. Consistent with this, mutations in *trr* result in major losses of H3K4me2 and H3K4me3 levels during *Drosophila* embryogenesis (Sedkov *et al.* 2003). Moreover, bulk H3K4 dimethylation during *C. elegans* embryogenesis depends mostly on ASH-2, and not on SET-2 (the *C. elegans* dSet1 ortholog), suggesting that other H3K4 HMT players are involved (Xiao *et al.* 2011).

Trx and Trr, while apparently not required for bulk H3K4 methylation (Figure 6, A–C), may be important for transcriptional regulation of a subset of specific gene targets later in development. This would be consistent with their proposed role in human cells, where the Trr and Trx homologs MLL1-2 and MLL3-4 are thought to methylate H3K4 at a limited number of non-overlapping gene targets (Wang *et al.* 2009; Ansari and Mandal 2010; Eissenberg and Shilatifard 2010). We will not discuss the roles of Ash1 further, as recent findings suggest that the predominant function of Ash1 is the catalysis of H3K36, and not H3K4 methylation (Tanaka *et al.* 2007; An *et al.* 2011).

We have shown that dSet1 interacts within the *Drosophila* COMPASS and have demonstrated the requirement of other *Drosophila* COMPASS members for H3K4 methylation, which now places us in a position to dissect the functional roles of individual COMPASS members. Upon *hcf* and *dWdr82* RNAi knockdown, only levels of H3K4me3 are reduced, an effect differing from that associated with loss of dSet1 function and suggesting specialized roles for Hcf and dWdr82 within the COMPASS (Figure 3). A nearly complete loss of H3K4me2 and H3K4me3 in *ash2* mutants and *dRbbp5* and *wds* RNAi knockdown animals suggests that these members are critical for COMPASS function; a loss of H3K4 monomethylation in these same animals is an effect not observed in *dSet1* mutants and suggests distinct roles of dRbbp5, Wds, and Ash2 aside from their roles within the COMPASS.

Although Set1 has been reported to target H3K4me1 in *Saccharomyces cerevisiae* (Santa-Rosa *et al.* 2002), there have been no reports indicating that Set1 plays a similar

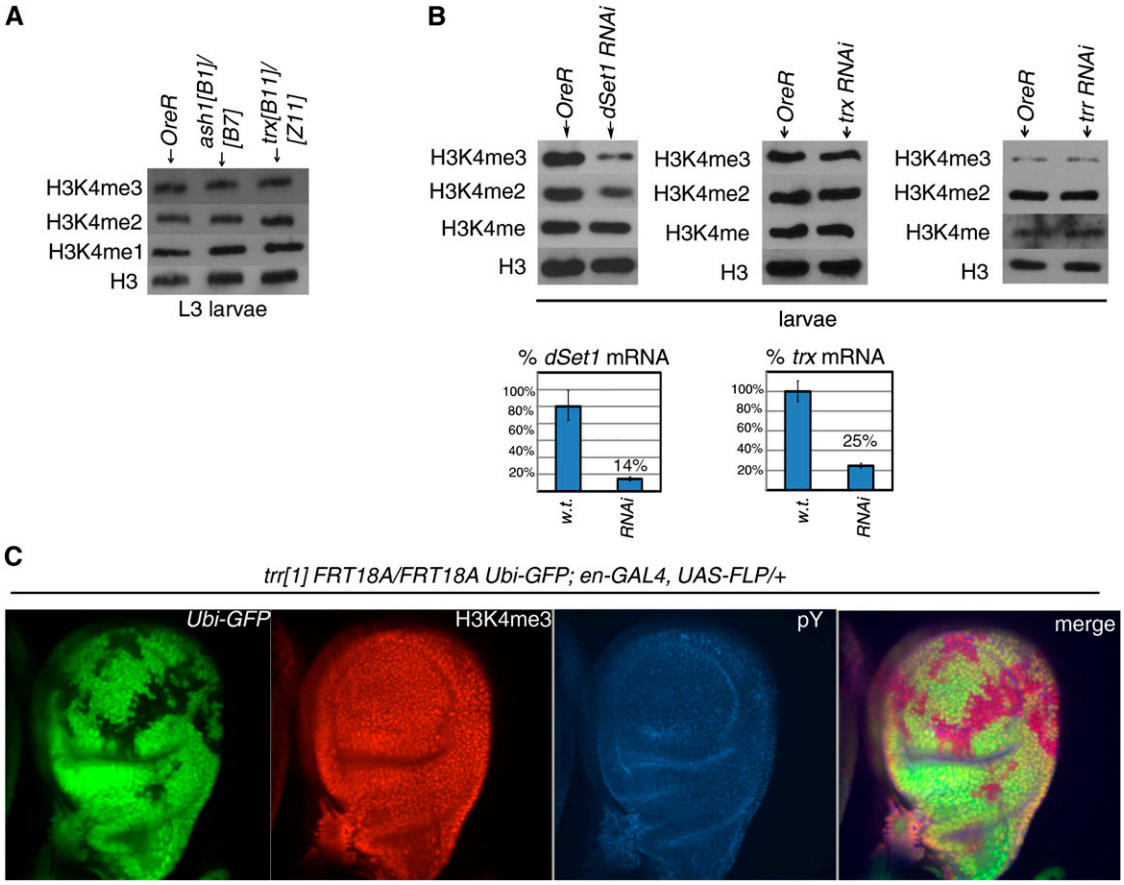

Figure 6 *trr*, *ash1*, and *trx* mutations do not significantly alter global H3K4 methylation levels. (A) Global H3K4 levels are unaltered in *trx* and *ash1* mutant animals. Comparison of H3K4 methylation profiles of *OreR*, $ash1^{B1}/ash1^{B7}$, and trx^{B11}/trx^{Z11} larvae. (B) Comparison of global H3K4 methylation levels between *OreR* larvae and larvae expressing double-strand RNA targeting (left to right) of *dSet1*, *trx*, and *trr*. Levels of *dSet1* and *trx* mRNA knockdown in RNAi experiments are represented as a percentage of wild-type levels and presented as bar graphs below. We had technical problems in quantifying the extent of knockdown in the *trr* RNAi line, but Mohan *et al.* (2011) have verified that this line shows significantly reduced levels of Trr protein. (C) H3K4 methylation defects are not observed in trr^1/trr^1 mutant haltere disc clones (marked by an absence of GFP). Clones were generated in L3 larvae of the genotype *trr*^{1} *FRT18A/Ubi-GFP FRT18A*; *UAS-FLP*, *en-Gal4*/+. Note the similar amounts of H3K4 trimethylation present (red) in trr^1/trr^1 clones *vs.* surrounding GFP-positive tissue. (Left to right) *Ubi-GFP* (green); H3K4me3 staining (red); phosphotyrosine (pY) cell-surface staining (blue); and merge.

role in metazoans, and we did not observe reductions in monomethyl H3K4 in *dSet1* mutants. Our results also rule out the individual contributions of Ash1, Trr, or Trx to general H3K4 monomethylation. It has been reported that the human MLL/COMPASS complex subunits WDR5, Rbbp5, ASH2L, and DPY-30 form a complex (known as WRAD) that monomethylates recombinant histone H3 at lysine 4 *in vitro* (Patel *et al.* 2009, 2011). The involvement of a *Drosophila* form of WRAD in H3K4me1 seems plausible as *ash2* mutations as well as *wds* and *dRbbp5* RNAi knockdown (Figure 3) result in decreased levels of H3K4me1. Alternatively, bulk H3K4 monomethylation may be catalyzed or targeted by an as-yet-uncharacterized H3K4 HMT complex containing these members or by combinatorial effects of Trr- and Trx-containing complexes.

In summary, our results indicate that dSet1 acts through the COMPASS to promote global H3K4 di- and trimethylation and appears to be indispensable during *Drosophila* development. While we were preparing this report, Ardehali *et al.* (2011) reported that dSet1 is associated within a COMPASS complex and is responsible for the majority of H3K4 methylation in *Drosophila* S2 tissue culture cells. Mohan *et al.* (2011) have also recently reported on the central role of dSet1; moreover, they have comprehensively characterized all three COMPASS complexes in *Drosophila* containing dSet1, Trx, or Trr and associated proteins (Mohan *et al.* 2011). Our work provides a complementary analysis of dSet1 function in the context of whole-organism development and includes data on functional roles for other COMPASS members. These results lay the groundwork for studying mechanisms and functional roles of H3K4 methylation by the COMPASS and other HMTs in metazoans.

Acknowledgments

We thank Zhuo (Bill) Li, Jessica Tom, Sabrina Rayworth, Mary Chan, Michelle Higginson, Amanda Berscht, Rachel Chan, Sara Brown, Linda Yang, and Arthur Lee for technical

assistance. We also thank Nicholas Harden for helpful discussions and Hugh Brock for critical insights prior to and during manuscript preparation, as well as two anonymous referees for valuable feedback. In addition, we thank the Bloomington Stock Center and the TRiP at Harvard Medical School (National Institutes of Health/National Institute of General Medical Sciences R01-GM084947) for providing fly stocks used in this study. This work was supported by a Discovery grant from the National Sciences and Engineering Research Council (NSERC) Canada to Barry Honda and NSERC postgraduate scholarships to Graham Hallson and Monika Syrzycka.

Literature Cited

Adamson, A. L. and A. Shearn, 1996 Molecular genetic analysis of Drosophila ash2, a member of the trithorax group required for imaginal disc pattern formation. Genetics 144: 633–651.

An, S., K. J. Yeo, Y. H. Jeon, and J. J. Song, 2011 Crystal structure of the human histone methyltransferase ASH1L catalytic domain and its implication for the regulatory mechanism. J. Biol. Chem. 286: 8360–8374.

Ansari, K. I., and S. S. Mandal, 2010 Mixed lineage leukemia: roles in gene expression, hormone signalling and mRNA processing. FEBS J. 277: 1790–1804.

Ardehali, M. B., A. Mei, K. L. Zobeck, M. Caron, J. T. Lis *et al.*, 2011 Drosophila Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription. EMBO J. 30: 2817–2828.

Bannister, A. J., and T. Kouzarides, 2011 Regulation of chromatin by histone modifications. Cell Res. 3(2): 381–395.

Beisel, C., A. Imhof, J. Greene, E. Kremmer, and F. Sauer, 2002 Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1. Nature 419: 857–862.

Beltran, S., M. Angulo, M. Pignatelli, F. Serras, and M. Corominas, 2004 Functional dissection of the *ash2* and *ash1* transcriptomes provides insights into the transcriptional basis of wing phenotypes and reveals conserved protein interactions. Genome Biol. 8(4): R67.

Bernstein, B. E., E. L. Humphrey, R. L. Erlich, R. Schneider, P. Bouman *et al.*, 2002 Methylation of histone H3 Lys4 in coding regions of active genes. Proc. Natl. Acad. Sci. USA 99(13): 8695–8700.

Byrd, K. N., and A. Shearn, 2003 ASH1, a Drosophila trithorax group protein, is required for methylation of lysine 4 residues on histone H3. Proc. Natl. Acad. Sci. USA 100(20): 11535–11540.

Choi, J. K., and L. J. Howe, 2009 Histone acetylation: Truth of consequences? Biochem. Cell Biol. 87(1): 139–150.

Daines, B., H. Wang, L. Wang, Y. Li, Y. Han *et al.*, 2011 The Drosophila melanogaster transcriptome by paired-end RNA sequencing. Genome Res. 21(2): 315–324.

Dehe, P.-M., M. Pamblanco, P. Luciano, R. Lebrun, D. Moiner *et al.*, 2005 Histone H3 lysine 4 mono-methylation does not require ubiquination of histone H2B. J. Mol. Biol. 353(3): 477–484.

Ebert, A., S. Lein, G. Schotta, and G. Reuter, 2006 Histone modification and the control of heterochromatic gene silencing in Drosophila. Chromosome Res. 14(4): 377–392.

Egelhofer, T. A., A. Minoda, S. Klugman, K. Lee, P. Kolasinka-Zwierz *et al.*, 2010 An assessment of histone-modification antibody quality. Nat. Struct. Mol. Biol. 18: 91–93.

Eissenberg, J. C., and A. Shilatifard, 2010 Histone H3 lysine 4 (H3K4) methylation in development and differentiation. Dev. Biol. 339(2): 240–249.

Fitzpatrick, K. A., D. A. Sinclair, S. R. Schulze, M. Syrzycka, and B. M. Honda, 2005 A genetic and molecular profile of third chromosome centric heterochromatin in *Drosophila melanogaster*. Genome 48(4): 571–584.

Gartner, K. E., C. D. Allis, and B. D. Strahl, 2011 OPERating ON chromatin, a colorful landscape where context matters. J. Mol. Biol. 409(1): 36–46.

Hallson, G., M. Syrzycka, S. A. Beck, J. A. Kennison, D. Dorsett *et al.*, 2008 The Drosophila cohesion subunit Rad21 is a trithorax group (trxG) protein. Proc. Natl. Acad. Sci. USA 105: 12405–12410.

Heintzman, N. D., R. K. Stuart, G. Hon, Y. Fu, C. W. Ching *et al.*, 2007 Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. **39**(3): 311–318.

Kylmenko, T., and J. Muller, 2004 The histone methyltransferases trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. EMBO Rep **5**(4): 373–377.

Koundakjian, E. J.Mulle, J., D. M. Cowan, R. W. Hardy, and A. H. Becker 2004 The Zucker Collection: a resource for the analysis of autosomal gene function in *Drosophila melanogaster*. Genetics 167: 203–206.

Lee, J., P. K. Saha, Q. H. Yang, S. Lee, J. Y. Park *et al.*, 2008 Targeted inactivation of MLL3 histone H3-Lys4 methyltransferase activity in the mouse reveals vital roles for MLL3 in adipogenesis. Proc. Natl. Acad. Sci. USA 105(49): 19229–19234.

Lee, J. H., and D. G. Skalnik, 2005 CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex. J. Biol. Chem. 280(50): 41725–41731.

Lee, J. H., C. M. Tate, J. S. You, and D. G. Skalnik, 2007 Identification and characterization of the human Set1B histone H3-Lys4 methyltransferase complex. J. Biol. Chem. 282 (18): 13419–13428.

Lee, J. S., E. Smith, and A. Schilatifard, 2010 The language of histone crosstalk. Cell 142(5): 682–685.

Lee, S., J. Lee, S. K. Lee, and J. W. Lee, 2008 Activating signal cointegrator-2 is an essential adaptor to recruit histone H3 lysine 4 methyltransferases MLL3 and *MLL4 to liver X* receptors. Mol. Endocrinol. 22(6): 1312–1319.

Marchant, G. E., and D. G. Holm, 1988 Genetic analysis of the heterochromatin of chromosome 3 in *Drosophila melanogaster*. II. Vital loci identified through EMS mutagenesis. Genetics 120: 519–532.

Miller, T., N. J. Krogan, J. Dover, H. Erdjument-Bromage, P. Tempst *et al.*, 2001 COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. Proc. Natl. Acad. Sci. USA 98(23): 12902–12907.

Mohan, M., H. M. Herz, E. R. Smith, Y. Zhang, J. Jackson *et al.*, 2011 The COMPASS family of H3K4 methylases in Drosophila. Mol. Cell. Biol. 31: 4310–4318.

Munshi, A., G. Shafi, N. Aliya, and A. Jyothy, 2009 Histone modifications dictate specific biological readouts. J. Genet. Genomics 36(2): 75–88.

Nagy, P. L., J. Griesenbeck, R. D. Kornberg, and M. L. Cleary, 2002 A trithorax-group complex purified from *Saccharomyces cerevisiae* is required for methylation of histone H3. Proc. Natl. Acad. Sci. USA 99(1): 90–94.

Papoulas, O., and S. J. Beek, S. J. Moseley, C. M. McCallum, M. Sarte *et al.*, 1998 The Drosophila trithorax group proteins BRM, ASH1 and ASH2 are subunits of distinct protein complexes. Development 125: 3955–3966.

Patel, A., V. Dharmarjan, V. E. Vought, and M. S. Cosgrove, 2009 On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-(MLL1) complex. J. Biol. Chem. 284(36): 24242–24256.

Patel, A., V. E. Vought, V. Dharmarjan, and M. S. Cosgrove, 2011 A novel non-SET domain multi-subunit methyltransferase required for the sequential nucleosomal histone H3 methylation by the mixed lineage leukemia protein-1 (MLL1) core complex. J. Biol. Chem. 286(5): 3359–3369.

Petruk, S., Y. Sedkov, S. Smith, S. Tillim, V. Kraevski *et al.*, 2001 Trithorax and dCBP acting in a complex to maintain expression of a homeotic gene. Science 294: 1331–1334.

Pokholok, D. K., C. T. Harbison, S. Levine, M. Cole, N. M. Hannett *et al.*, 2005 Genome-wide Map of Nucleosome Acetylation and Methylation in Yeast. Cell 122(4): 2005.

Roguev, A., D. Schaft, A. Shevchenko, W. W. Pijnappel, M. Wilm *et al.*, 2001 The *Sacchromyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone H3 lysine 4. EMBO J. 20(24): 7137–7148.

Roguev, A., D. Schaft, A. Shevchenko, R. Aasland, A. Shevchenko *et al.*, 2003 High conservation of the Set1/Rad6 axis of histone 3 lysine 4 methylation in budding and fission yeasts. J. Biol. Chem. 278(10): 8487–8493.

Santa-Rosa, H., R. Schneider, A. J. Bannister, J. Sherriff, B. F. Bernstein *et al.*, 2002 Active genes are tri-methylated at K4 of histone H3. Nature 419(6905): 407–411.

Schulze, S. R., D. A. R. Sinclair, E. Silva, K. A. Fitzpatrick, M. Singh *et al.*, 2001 Essential genes in proximal 3L heterochromatin in *Drosophila melanogaster*. Mol. Gen. Genet. 264(6): 782–789.

Schulze, S. R., D. A. Sinclair, K. A. Fitzpatrick, and B. M. Honda, 2005 A genetic and molecular characterization of two proximal heterochromatic genes on chromosome 3 of *Drosophila melanogaster*. Genetics 169: 2165–2177.

Secombe, J., L. Li, L. Carlos, and R. N. Eisenman, 2007 The Trithorax group protein Lid is a trimethyl histone H3K4 demethylase required for dMyc-induced cell growth. Genes Dev. 21(5): 537–551.

Sedkov, Y., E. Cho, S. Petruk, L. Cherba, S. T. Smith *et al.*, 2003 Methylation at lysine 4 of histone H3 in ecdysone-dependent development of Drosophila. Nature 426(6962): 78–83.

Shaffer, C. D., J. M. Wuller, and S. C. R. Elgin, 1994 Preparation of *Drosophila* nuclei, pp. 185–189 in *Drosophila melanogaster: Practical Uses in Cell and Molecular Biology*, edited by L. S. B. Goldstein and E. A. Fyrberg. Academic Press, San Diego.

Shannon, M. P., T. C. Kaufman, M. W. Shen, and B. H. Judd, 1972 Lethality patterns and morphology of selected lethal and semi-lethal mutations in the zeste-white region of *Drosophila melanogaster*. Genetics 72: 615–638.

Simonet, T., R. Dulermo, S. Schott, and F. Palladino, 2007 Antagonistic functions of SET-2/SET1 and HPL/HP1 proteins in C. elegans. Dev. Biol. 312(1): 367–383.

Sinclair, D. A. R., M. Syrzycka, M. S. Macauley, T. Rastgardani, I. Komljenovic *et al.*, 2009 *Drosophila* O-GlcNac transferase (OGT) is encoded by the *Polycomb* group (PcG) gene, *super sex combs* (sxc). Proc. Natl. Acad. Sci. USA 106: 13427–13432.

Smith, S. T., and S. Petruk., Y. Sedkov, E. Cho, S. Tillub *et al.*, 2004 Modulation of heat shock gene expression by TAC1 chromatin modifying complex. Nat. Cell Biol. 6: 162–167.

Srinivasan, S., K. M. Dorighi, and J. W. Tamkun, 2008 Drosophila Kismet regulates histone H3 lysine 27 methylation and early elongation by RNA Polymerase II. PLoS Genet. 4(10): e1000217.

Tanaka, Y., Z. Katagiri, K. Kawahasi, D. Kioussis, and S. Kitajima, 2007 Trithorax-group protein ASH1 methylates histone H3 lysine 36. Gene 396: 161–168.

Vermuelen, M., K. W. Mulder, S. Denissov, W. W. Pijnappel, F. M. Van Schaik *et al.*, 2007 Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell 131(1): 58–69.

Wang, P., C. Lin, E. R. Smith, H. Guo, B. W. Anderson *et al.*, 2009 Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. Mol. Cell. Biol. 29(22): 6074–6085.

Wilson, J. R., C. Jing, P. A. Walker, S. R. Martin, S. A. Howell *et al.*, 2002 Crystal structure and functional analysis of the histone methyltransferase SET7/9. Cell 111(1): 105–115.

Wysocka, J., T. Swigut, H. Xiao, T. A. Milne, S. J. Kwon *et al.*, 2006 A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature 442(7098): 86–90.

Xiao, Y., C. Bedet, V. J. Robert, T. Simonet, S. Dunkelbarger *et al.*, 2011 Caenorhabditis elegans chromatin-associated proteins SET-2 and ASH-2 are differentially required for histone H3 Lys 4 methylation in embryos and adult germ cells. Proc. Natl. Acad. Sci. USA 108(20): 8305–8310.

Communicating editor: J. A. Birchler

GENETICS

Supporting Information
http://www.genetics.org/content/suppl/2011/11/02/genetics.111.135863.DC1

dSet1 Is the Main H3K4 Di- and Tri-Methyltransferase Throughout *Drosophila* Development

Graham Hallson, Robert E. Hollebakken, Taosui Li, Monika Syrzycka, Inho Kim, Shawn Cotsworth, Kathleen A. Fitzpatrick, Donald A. R. Sinclair, and Barry M. Honda

DOI: 10.1534/genetics.111.135863

File S1

Supporting Materials and Methods

Gene specific PCR:

DNA was isolated from homozygous mutant embryos, and PCR was performed with primers designed based upon the *CG40351* genomic sequence available on Flybase (www.flybase.org). PCR conditions were as follows: 94° for 5 minutes followed by 33 cycles of 94° for 30 s, annealing (varying temperatures listed in Table S1) for 40s and 72° for 1 min. PCR fragments were directly sequenced with the same primers used to amplify each fragment. Resultant sequences were compared to the published *D. melanogaster* genomic sequence available on Flybase (www.flybase.org). Mutations were identified by the presence of mismatches when compared to the reference sequence.

TABLE S1. Sequences and annealing temperatures of *CG40351* PCR primers pairs.

Primer pair used:	Sequences:	Tanneal:
SER1F/	*5′-TTAGCCATAGTCTGAACTAGTTATTATTTACG-3′*	60°
SER2R	*5′-TTGACCAGGTTGATATTCCGAAC-3′*	
SER2F/	*5′-AGGTTTGTTCATAATTGACACAGATGC-3′*	62°
SER3R	*5′-TCATACCTTTCCCATTACAGACTTTTG-3′*	
SER3F/	*5′-ACTTGGGCATTGCACGTATAGTG-3′*	62°
SER4R	*CTACTTATTGTATTCGAAGAACGTCTAACATC*	
SER4F/	*5′-GGATCGCGGAAGTCTTAAGTATTGTAG-3′*	63°
47936M2R	*5′-CCCATAGCCTGAAGCATGGTAATAAAAAGG-3′*	
SER4BF/	*5′-TATCCAGACCAAGGAAGGTGCAG-3′*	64°
SER4CR	*5′-CATTAGAATCCTCCTTATCAGAGTCGGAGC-3′*	
SEQ1F/	*5′-CATCGTTTCGGCGTATAAGAAAACATCC-3′*	65°
SER4ER	*5′-CTCGTTCCATTTGAGCCATGTACTC-3′*	
SER4EF/	*5′-GTGTGATATGTATAAAACTCGTGATC-3′*	56°
SEQ1R	*5′-TACTTATGCTTAGCCTTTTCTCGAACATC-3′*	
SER6F/	*5′-GGTTAACTCGACTAAATCATTCTCTCTTG-3′*	61°
SER6BR	*5′-GTATTAGTTAAGTGTACCCCGACATCC-3′*	
SER6F/	*5′-ATGTAAGTAAGCTGGATGTTGACATTATG-3′*	63°
SER6R	*5′-GAACAGCACCTCTGACCTCGTTG-3′*	
SER4DF/	*5′-TGAATACGATAGAATTTACAGCGACTC-3′*	61°
SEQ1R	*5′-TACTTATGCTTAGCCTTTTCTCGAACATC-3′*	

Cloning of inducible *UAS-dSet1* cDNA transgenes:

Since there was no available cDNA clone containing the full length dSet1 cDNA sequence, we separately amplified and cloned fragments from the DGRC cDNA clones *RE47936* and *LD15202*, which contain incomplete but overlapping regions of the full length *dSet1* cDNA sequence. Regions of the *dSet1* coding sequence contained within *RE47936* and *LD15202* were amplified with *SER2R*/*SER4DR* and *SER4CF*/*SER6BR* respectively (primer sequences listed in Table S1) and each respective PCR product was separately cloned into *pTZ57R/T* using the InsTAclone PCR Cloning Kit (Fermentas). Each fragment was cleaved from the *pTZ57R/T* vector, followed by cleavage of each fragment at an overlapping *Xho*I site, ligated together using T4 DNA ligase (Invitrogen) as per manufacturer's instructions and finally into the *pTZ57R/T* vector. Sequence errors contained within the *dSet1* sequence were corrected with the Quikchange Site Directed Mutagenesis Kit (Stratagene) following manufacturer's instructions so that the full length *dSet1* ORF matched the sequence for *CG40351* (gi: 281366741,68-5018). This corrected *dSet1* coding sequence was cut from *pTZ57R/T* using flanking *Xba*I sites and ligated into *Xba*I digested *pUAST (*BRAND and PERRIMON 1993*).* The completed rescue construct (*pUAST-dSet1*) was fully sequenced with the primers listed in Table S1 prior to microinjection by Bestgene Inc..

A sequence encoding a 2X FLAG™ tag was PCR amplified without template DNA using the overlapping primers *2XFLAGpUASTF: 5'-TTTGCGGCCGCGACTACAAGGATGACGATGACAAAGACTACAAGGATGACGATGACAAA- 3'* and *2XFLAGpUASTR: TTTCTCGAGTTATTTGTCATCGTCCATCCTTGTAGTCTTTGTCATCGTCATCCTTGTAGTC-3'*. PCR conditions were as follows: 94° for 5 minutes followed by 20 cycles of 94° for 30s, 60° for 40 s and 72° for 30 s. The 2X FLAG fragment was digested with *Not*I and *Xba*I and ligated into *pUAST* (*pUAST-2X FLAG*). The *dSet1* cDNA sequence was amplified from 20 ng of *pUAST-dSet1* template using a standard PCR recipe with the primers *SETNotIF: 5'-TTTGCGGCCGCTCATGCAGGACGTTCGG-3'*and *SET2XFLAGNotIR: 5'-TTTGCGGCCGCGTTAAGTGTACCCCGACATCC-3'* containing engineered *Not*I restriction sites. PCR conditions were: 95° for 5 min., followed by 33 cycles of 95° for 30 s, 55°C for 40 s and 72° for 5 min. This amplicon and *pUAST-2X FLAG* were digested with *Not*I and ligated together with T4 DNA ligase (Invitrogen). The completed *pUAST-dSet1-2X FLAG* was fully sequenced with primers listed in Table S1 prior to microinjection by Bestgene Inc..

Cloning of GST fusion proteins:

The full length *wds* and *ash2* cDNA sequences were amplified from the Drosophila Gold Collection clones *RE31658* and *LD31689* respectively, with the primers *wdsGSTF: 5'-TTGAATTCGCCAAGGAGCATAAGCAGAAT-3'* and *wdsGSTR: 5'-TTGTCGACAGCTGGATACCAGCCACTCTAT-3'*, and *ash2GSTF: 5'- TTGAATTCATCACTGCAATGGAGGACAG-3'* and *ash2GSTR: 5'-TTCTCGAGCCAAAACTCATGTGTATGACGA-3'*. PCR conditions were 94° for 5 minutes followed by 33 cycles of 94° 30 seconds, 58° for 40 seconds, followed by 72° for 2 minutes. *wds* and *ash2* cDNA amplicons were digested with *Eco*RI/*Sal*I and *Eco*RI/*Xho*I

respectively (sites engineered in each primer set), and were cloned into *pGEX-4T1* in frame with a 5' sequence coding for GST. The *CG17293* cDNA sequence was subcloned from *pEG202* into *pGEX-4T-1* using the restriction sites previously described.

Quantitative PCR:

RNA was extracted from *OreR* animals and animals expressing RNAi targeting Drosophila COMPASS members using TRIzol reagent (Invitrogen) followed by chloroform extraction and isopropanol precipitation using standard procedures. 300 ng of RNA from each stage tested was treated with 1 U of DNAse I for 1 hour, heated at 65° for 15 minutes and converted into cDNA using the iScript Select cDNA synthesis kit (Biorad). Primers were designed for *rp49* and genes coding for COMPASS orthologues and are listed in Table S1. Reaction mixtures were prepared in triplicate as follows: 2 µL cDNA, 1 µL primer/probe mix, 7 µL of ddH2O and 10 µL of KAPA Sybr Fast Mastermix (KAPA). PCR conditions were 40 cycles of 94° for 3 seconds, followed by 60° for 30 seconds, followed by 72° for 1 minute. Primer pairs used are listed in Table S2. Primer efficiencies were estimated from the slope of standard curves generated using dilutions of adult male cDNA. Transcript levels in RNAi knockdown experiments were calculated relative to wild-type using the ΔΔCt method and using *rp49* as a reference gene to normalize differences in input cDNA levels (PFFAFL 2001).

TABLE S2. qPCR primers used to validate RNAi experiments.

Gene target:	Sequences:
dWdr82	*5'*- GCA CGG TGA ACT CGA AGA AG *-3'*
	5'- TCG TGC AGG CTG AGA TAG *-3'*
wds	*5'*- AGC TCC TCC GCT GAT AAA CTA ATC *-3'*
	5'- ACC TTG AGC GTC TTG TCA TCA C *-3'*
hcf	*5'*- TTC TTA ACC CAA CCG GAC CG *-3'*
	5'- CGC ATC CGT TCG GTA CAT CA *-3'*
dRbbp5	*5'*- TGA GCG GAT GCG TTT CAT C *-3'*
	5'- GGA ACC GCA TGA CTA TGA GGA *-3'*
trx	*5'- TCG ACT ATG GAT CGG ATC AG-3'*
	5'-GAT GGG CTG CTT GCG ATG TCT-3'
dSet1	*5'*- CGT TCG GAA TAT CAA CCT GGT C *-3'*
	5'- CGT AAC GAT AGA GTC TGG TAC CAC *-3'*
rp49	5'-TGC TAA GCT GTC GCA CAA ATG GCG -3'
	5'-CAT GTG GCG GGT GCG CTT GTT C -3'

Methyl-H3K4 antibody specific tests:

0.5 µg and 0.05 µg of unmodified H3 peptides and H3 peptides containing mono-, di- and trimethylation at lysine 4 (Epigenetek) were spotted onto nitrocellulose membranes and dried for 30 minutes. Immunoblotting was then performed using standard procedures. Antibody dilutions are indicated in the main text.

Quantitative Western blotting:

Two-fold dilutions series of wild-type and *dSet1* mutant (*G12*/*γ28*) nuclear extracts were run on SDS-PAGE gels and transferred to nitrocellulose membranes using standard procedures. Immunoblotting was then performed using standard procedures with antibody dilutions indicated in the text. Detection was performed using Supersignal Pico or Supersignal Femto (Thermo Scientific) and exposure to film. Exposed films were scanned and integrated density (pixel area multipled by baseline subtracted intensity) of bands were calculated within an appropriate exposure range, and bands from *dSet1* mutant lanes were compared with bands from wild-type lanes exposed to similar levels. These integrated densities were then normalized according to total H3 loaded. Levels of mutant to wild-type methyl-H3K4 were calculated based on the normalized integrated density of the band corresponding to *dSet1* mutant extracts divided by that corresponding to the normalized wild-type extract multiplied by the dilution factor. All blots were performed in duplicate and error is reported as the standard error of the mean (SEM).

SUPPORTING REFERENCES

BRAND, A. H. and N. PERRIMON, 1993 Targeted gene expression as a means of altering cell fates generating dominant phenotypes. *Development* **118**: 401-415.

PFAFFL, M. W., 2001 A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**(9): e45.

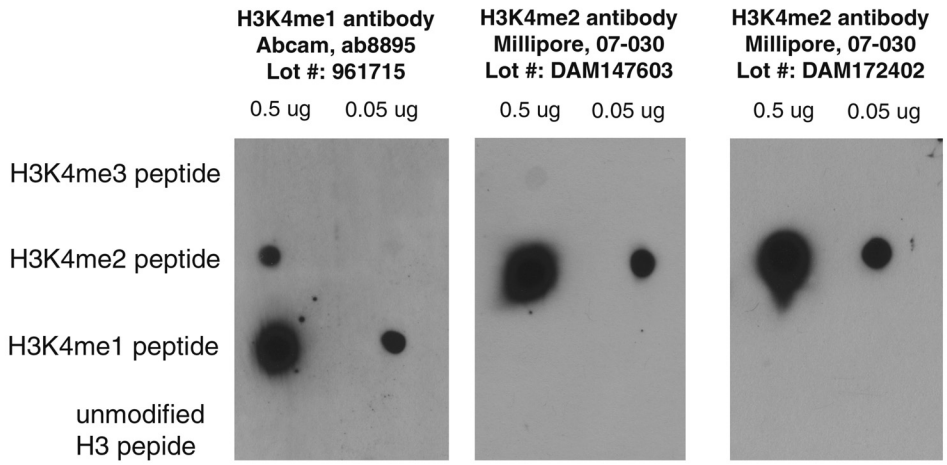

Figure S1 Methyl-H3K4 antibody specificity tests.

.

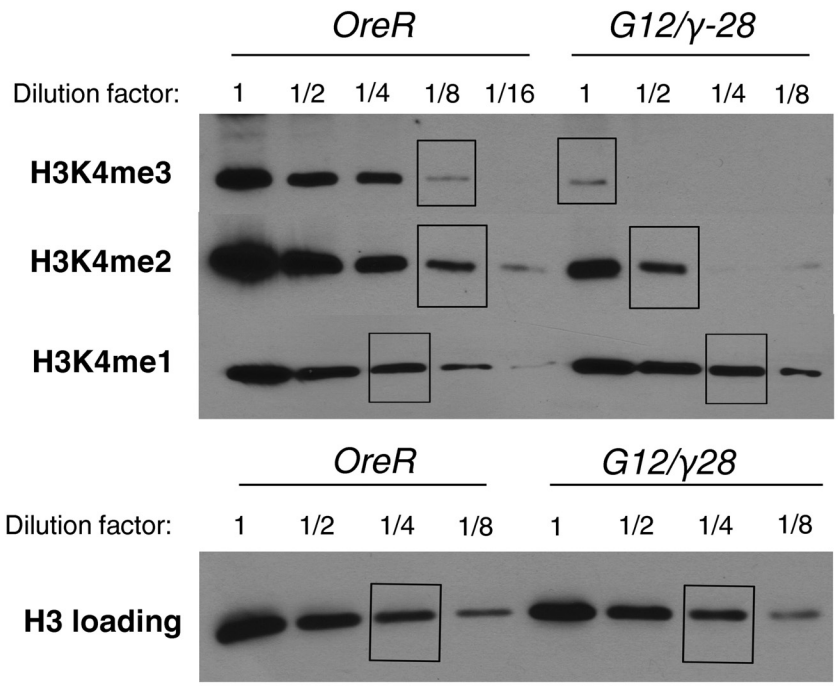

Figure S2 An example of our quantitative approach for calculating methyl H3K4 levels in *dSet1* mutants (*G12/γ28*) relative to wild-type (*OreR*) levels. This example contains dilution series of pupal extracts. Boxed bands correspond to bands used for quantification purposes.

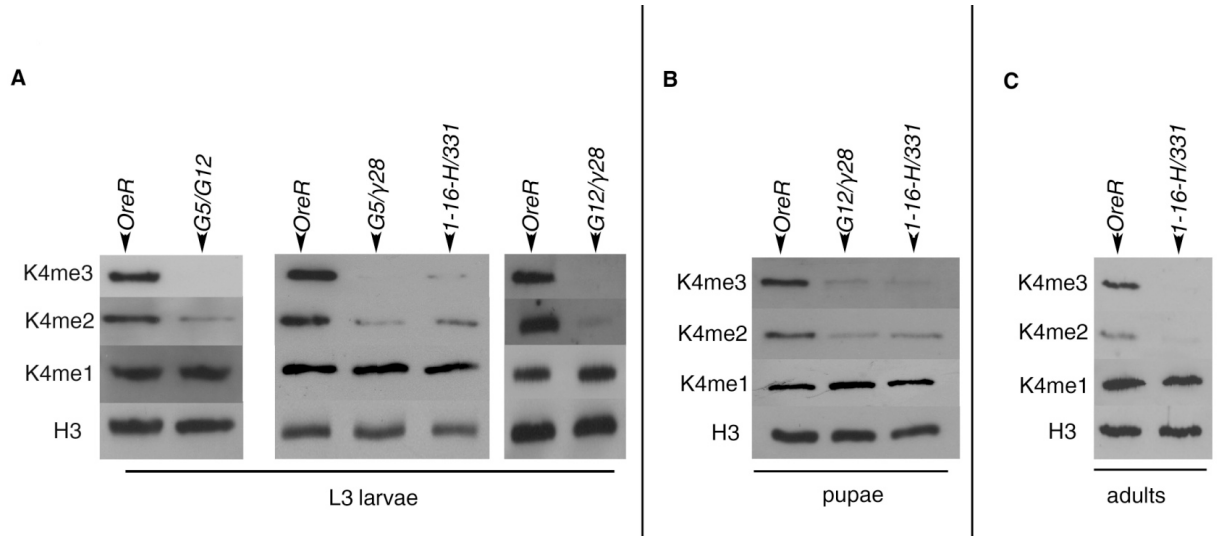

Figure S3 Comparisons of total H3K4 levels between wild-type (*OreR*) and *dSet1* mutant transheterozygous (A) third instar larvae (B) pupae and (C) adults.

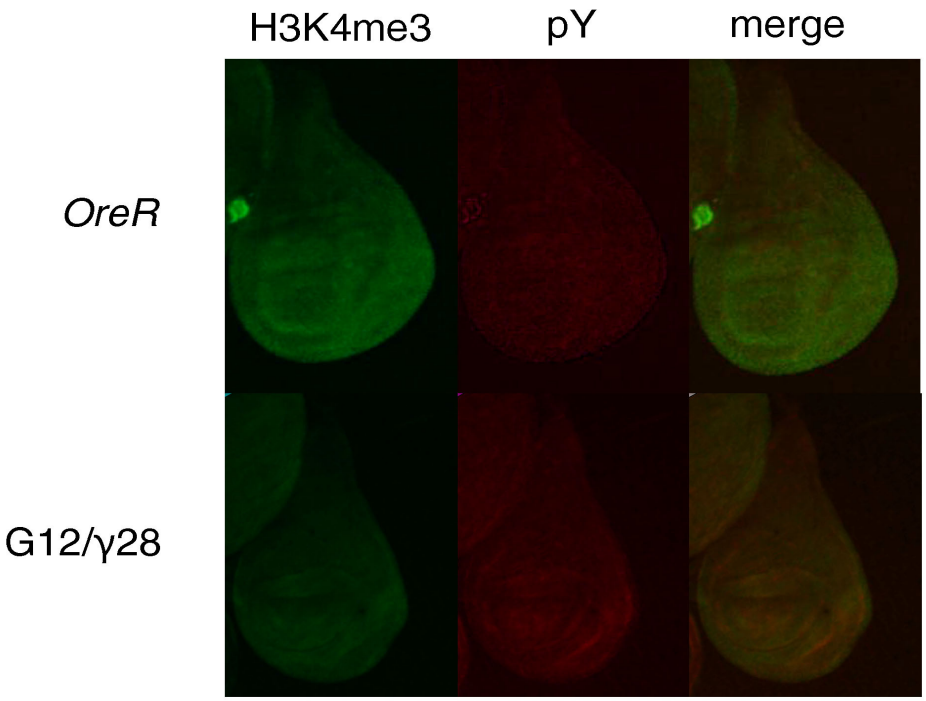

Figure S4 *dSet1* mutations result in reduced H3K4 trimethylation in halterediscs. Left to right: H3K4me3 (green), phosphotyrosine (red), merge. Immunostaining was performed as described in the main text using FITC anti-rabbit and Texas-Red anti-mouse secondaries (Cell Signalling) for detection.