HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2017 November 01.

Published in final edited form as:

Nat Genet. 2017 November ; 49(11): 1647–1653. doi:10.1038/ng.3965.

Histone H3K4 monomethylation catalyzed by Trr and mammalian COMPASS-like proteins at enhancers is dispensable for development and viability

Ryan Rickels[1,6], **Hans-Martin Herz**[2,6,#], **Christie C Sze**[1], **Kaixiang Cao**[1], **Marc A Morgan**[1], **Clayton K Collings**[1], **Maria Gause**[3], **Yoh-hei Takahashi**[1], **Lu Wang**[1], **Emily J Rendleman**[1], **Stacy A Marshall**[1], **Annika Krueger**[2], **Elizabeth T Bartom**[1], **Andrea Piunti**[1], **Edwin Smith**[1], **Nebiyu A Abshiru**[4], **Neil L Kelleher**[4], **Dale Dorsett**[3], and **Ali Shilatifard**[1,2,3,5,*]

[1]Department of Biochemistry and Molecular Genetics, Northwestern University Feinberg School of Medicine, 320 E. Superior Street, Chicago, IL 60611, USA

[2]Stowers Institute for Medical Research, Kansas City, MO 64110, USA

[3]Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, 1100 South Grand Boulevard, Saint Louis, MO, 63104, USA

[4]Department of Chemistry, Northwestern University, Evanston, Illinois, 60208, USA

[5]Robert H. Lurie NCI Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, 320 E. Superior St., Chicago, Il 60611

Abstract

Histone H3 lysine 4 monomethylation (H3K4me1) is an evolutionarily conserved feature of enhancer chromatin catalyzed by the Trr/MLL3/4-COMPASS family[1–3]. Here we demonstrate that *Drosophila* embryos expressing catalytically deficient Trr-COMPASS eclose and develop to a productive adulthood. Parallel experiments with a Trr allele that augments enzyme product specificity reveal that conversion of H3K4me1 at enhancers to H3K4me2 and H3K4me3 is also compatible with life and results in minimal changes in gene expression. Similarly, loss of mammalian MLL3 and MLL4 catalytic SET domain in embryonic stem cells does not disrupt self-renewal capability of the ES cells. Trr catalytic mutant alleles manifest subtle developmental phenotypes when subjected to temperature stress or altered cohesin levels. Collectively, our

[*]Correspondence and proofs should be sent to the following address: Ali Shilatifard Department of Biochemistry and Molecular Genetics Northwestern University Feinberg School of Medicine Searle 6-512 320 E. Superior St. Chicago, Il 60611, ASH@Northwestern.edu.

[6]These authors contributed equally to this work.

[#]Current Address: Department of Cell and Molecular Biology, St. Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105, USA

Author Contributions

A.S. and H.M.H. conceived and initiated the project. R.R. performed *Drosophila* ChIP-seq studies and conducted genetic tests, stress response assays, and wrote the manuscript. C.C.S. and A.K. assisted with *Drosophila* experiments. Y.T. performed Trr/WRAD reconstitution and *in vitro* methyltransferase assays. K.C., M.A.M., generated CRISPR mESCs and L.W., A.P., C.C.S. assisted with mammalian studies. Next-generation sequencing data was analyzed by C.K.C. and E.T.B. while libraries were generated and sequenced by E.J.R. and S.A.M. Mass-spectrometry was performed and analyzed by N.A.A. and N.L.K. Cohesin genetic interaction data were contributed by M.G. and D.D. Critical feedback and advice was provided by E.S. throughout the course of this project.

Competing Financial Interests

The authors declare no competing financial interests.

findings suggest that metazoan development can occur in the context of Trr/COMPASS with H3K4me1 enzymatic deficiency, and points to a possible role for H3K4me1 on cis-regulatory elements in specific settings to fine-tune transcriptional regulation in response to environmental stress.

Transcriptional enhancers are cis-regulatory elements that potentiate transcriptional output even when separated from their cognate promoter by megabases of intervening DNA[4–7]. Enhancer-promoter communication is thought to occur through long-range looping mechanisms facilitated by cohesin complexes and other nuclear factors[8]. Enrichment of specific histone modifications, in particular H3K27-acetylation (H3K27ac) and H3K4me1, is an evolutionarily conserved feature of enhancer chromatin[9,10]. Previous work from our laboratory established the initial link between Trr/COMPASS in *Drosophila* and its mammalian homologues MLL3/4-COMPASS as the enhancer H3K4 monomethylases required for enhancer promoter communication during development[1–2]. Subsequent studies confirmed our original findings that MLL3/4-COMPASS are H3K4 monomethylases functioning on enhancers, and that they are essential for embryonic development in mammals[11,12]. Importantly, recent studies implicate defective MLL3/4 function in the pathogenesis of various forms of cancer[13]. However, the specific requirements for H3K4me1 catalytic activity during organismal development are largely undefined.

To investigate the role of Trr catalyzed H3K4me1 at cis-regulatory elements in *Drosophila*, we complemented the embryonic lethal *trr*[1] null allele[14] with various *trr* rescue transgenes. Using site-specific integration, we rescued *trr*[1] lethality using a 12kb transgenic genomic construct encompassing the *trr* locus and its associated regulatory elements. We also introduced specific amino acid substitutions into the Trr SET domain at highly conserved residues that disrupt Set1 methyltransferase activity in yeast (Figure 1B). A Trr cysteine-to-alanine (C2398A) mutation is catalytically deficient when reconstituted *in vitro* (Figure 1C), and cell lysates from C2398A (referred to as Trr-C/A) larval tissues demonstrate a large reduction in bulk H3K4me1 levels without affecting Trr, Trx, or dSet1 protein stability (Figure 1D). This closely resembles the effect of Trr-RNAi on H3K4me1 levels, and is consistent with our original findings that Trr predominantly catalyzes H3K4me1 in *Drosophila*[1].

We next tested the importance of Trr product-specificity by mutating the “F/Y switch” position, which converts Set1/COMPASS to a more efficient H3K4-di- and trimethylase, *in vitro* and *in vivo*[15,16]. Indeed, introducing this mutation at the corresponding residue (Y2383F, here on referred to as Trr-Y/F) in Trr’s SET domain shifts its enzymatic activity towards that of an H3K4-di and –trimethylase without affecting protein stability (Figures 1B–D).

Intriguingly, both the catalytic-deficient (Trr-C/A) and –hyperactive (Trr-Y/F) mutations rescue the recessive lethal *trr*[1] allele, producing fertile adults with normal life-span and no gross abnormalities (Figure 1E). Because loss of Trr results in embryonic lethality, we were intrigued that disrupting its catalytic activity produces no apparent phenotype, given the strong evolutionary conservation of the COMPASS SET domain (Figure 1B). Sanger-sequencing confirmed the intended mutations in both genomic DNA and mRNA

(Supplemental Figures 1A and 1B). We removed the *trr-rescue* transgene through genetic crosses and confirmed that *trr*[1] recessive lethality persists, thus ruling out the possibility that recombination occurred on the chromosome X to produce a functional Trr enzyme (Figure 1F). Two additional lethal alleles, *trr[C2375X]* and *trr[K662X]*[17], were also rescued in the same manner to confirm complementation is not unique to *trr*[1] (data not shown). These observations suggest an essential methylase-independent role for Trr in regulating enhancer-mediated processes critical for *Drosophila* development.

To test whether our Trr catalytic mutations specifically affect H3K4-methylation at *cis*-regulatory elements, we mapped the genomic distribution of these modifications using ChIP-seq in adult fly brain tissues (Figure 2A and Supplemental Figure 2A). Track examples of two representative genes, *kis* and *nvy*, show reduction of H3K4me1 at intronic and intergenic regions in both *trr-C/A* and *trr-Y/F* mutant brain tissues (Figure 2B). These changes overlap with sites of H3K27ac, a known mark of transcriptionally active chromatin, as well as binding of Trr and Lpt (a component of Trr/COMPASS[1]). Remarkably, these same sites exhibit conversion of H3K4me1 to H3K4me2/3 in *trr-Y/F* catalytic-hyperactive flies (Figure 2B and 2C).

To confirm that these effects occur genome-wide, we identified H3K27ac peaks and divided them into 5367 transcription start sites (TSS) and 1440 nonTSS sites. ChIP-seq reads were converted to Z-scores and displayed as density bar-plots centered on nonTSS H3K27ac peaks. At these 1440 putative enhancers, H3K4me1 is either diminished in the *trr-C/A* mutation, or converted to H3K4me2/3 in the *trr-Y/F* allele (Figures 2C and 2D, Supplemental Figure 2B). Interestingly, enhancer H3K27ac levels are modestly but reproducibly correlated with H3K4-methylation levels, decreasing in the *trr-C/A* and increasing slightly in the *trr-Y/F*. The effects caused by these *trr* catalytic mutations are specific to enhancers, and are not observed at TSS regions. By contrast, H3K4me1 in the *trr-C/A* is slightly increased at TSSs, consistent with Trr-RNAi experiments[1] (Supplemental Figure 2C). We performed ChIP-seq for both Trr and Lpt components of COMPASS in our transgenic flies and found that their genome-wide binding is not substantially affected (Supplemental Figure 2D). To confirm our findings in a different tissue-type, we repeated these experiments for H3K4me1/3 and H3K27ac using larval wing imaginal discs, and obtained similar results to brain tissues (Supplemental Figures 2E, 2F, and 2G), thus corroborating our previous findings that Trr functions predominantly at enhancers[1].

To explore the transcriptional consequences of altering H3K4 methylation at enhancers, we performed RNA-seq on both adult brains and larval wing imaginal discs from our transgenic flies (Figure 3A). Despite the alterations of enhancer histone methylation, gene expression in adult brain tissue is largely unaffected in the *trr-C/A* and *trr-Y/F* flies (Supplemental Figures 3A and 3B). This is predicted, as neither transgenic line demonstrates apparent abnormalities in their behavior or head morphology. In wing imaginal discs, gene-expression profiles for the two catalytic mutants are highly similar to *trr-WT* (>99%, Pearson correlation) (Figure 3B). By contrast, reducing total Trr levels by RNAi, using either *engrailed* or *T80* Gal4 drivers, elicits major changes in gene expression (Figure 3C) suggesting that there are H3K4 methylation-independent functions of Trr in developmental gene regulation. We performed unsupervised hierarchical clustering of the 3168

differentially expressed genes from our six datasets (adjusted p<0.01) and plotted Z-score transformed read counts for comparison. Control samples (LacZ-RNAi) and the three different *trr-rescue* alleles exhibit highly similar patterns of gene expression, whereas Trr-RNAi depleted samples form a separate cluster with dramatically altered RNA expression profiles (Figure 3C). These results support our hypothesis that the major functions of Trr in transcription regulation and organismal development are not strictly dependent on H3K4me1 methyltransferase activity.

By integrating ChIP-seq and wing disc RNA-seq data, we observed a link between subtle differences in gene expression (Figure 3C) and altered H3K4-methylation levels at nearby enhancers. Centering on total H3K27ac peaks (n=5908), we performed k-means clustering (k=3) (Figure 3D). Cluster #3 (bottom cluster) exhibits the strongest alteration in nonTSS enhancer H3K4-methylation (Figure 3D). Diminished enhancer H3K4me1 in in *trr-C/A* flies is associated with decreased expression of the nearest gene ($p = 1.22 \times 10^{-42}$, hypergeometric test), conversely *trr-Y/F* conversion of H3K4me1 to H3K4me2/3 is associated with increased expression of the nearest gene ($p = 6.25 \times 10^{-22}$, hypergeometric test) (Figure 3D). Despite these observed changes in gene expression, all three *trr*[1]-rescued fly lines produce adults with morphologically normal wings (Figure 1E), whereas both *trr-RNAi* lines die during pupation, prohibiting analysis of wing development.

After detailed analysis were we able to uncover subtle phenotypes in the three *trr-rescue* lines. When maintained at the elevated temperature of 29°C, *trr-C/A* flies display an additional L3/L4 cross-vein, and this phenotype disappears when *trr-C/A* is expressed in a *trr[+]* background (Figure 4A). Also, *trr-Y/F* females have much stronger pigmentation of their 7th abdominal segment, whereas the majority of *trr-C/A* females lack this pigmentation (Figure 4B). We also observe a mild bristle phenotype in the *trr-Y/F* line when maintained at 29°C. The occurrence of supernumerary thoracic macrochaetae (bristles) was significantly higher (p = 1.19 E –17, t-test with unequal variance) in the *trr-Y/F* than in *trr-C/A* flies (p = 0.36) (Supplemental Figure 4A). Remarkably, this phenotype also manifests in mutants of the H3K4me3-demethylase, *lid*[18,19]. Bulk levels of H3K4me3 are increased in *lid* mutants, similar to what we observe in *trr-Y/F* flies, suggesting this bristle phenotype results from increased H3K4me3, presumably through the Notch pathway. Together, these observations demonstrate a wide tolerance for differential H3K4-monomethylation at developmental enhancers in *Drosophila*, and suggest this modification is important for fine-tuning enhancer activity, especially under temperature stress.

In eukaryotes, the cohesin complex is essential for sister chromatid cohesion during mitosis and also plays a pivotal role in facilitating enhancer-promoter communication by bringing these elements into close physical proximity[20–24]. Nipped-B is responsible for loading cohesin onto chromatin, and over-expression of this gene in *Drosophila* causes abdominal segmentation defects in adult flies[25]. Interestingly, we detect abdominal segmentation defects in the *trr-Y/F* flies at low penetrance. If this is due to enhancer over-activation through hypermethylation, then Nipped-B overexpression in the *trr-Y/F* background should exacerbate this phenotype. Indeed, crossing *da*-GAL4:UAS-Nipped-B with *trr-Y/F* increases segmentation defect penetrance in a *trr[+]* background by ~12-fold relative to UAS-Nipped-B alone (Figure 4C). By contrast, *trr-C/A* does not dramatically increase the segmentation

defect penetrance of UAS-Nipped-B (Figure 4C). These results indicate that enhancer activity is sensitive to increased cohesin loading, and that local enhancer H3K4-monomethylation fine-tunes transcriptional output.

To test the conservation of enhancer H3K4me1 function between *Drosophila* Trr and mammalian MLL3 and MLL4 (encoded by *KMT2C* and *KMT2D*, respectively[1,2,26]), we used CRISPR/Cas9 to generate several mESC clones containing deletions of both the MLL3 and MLL4 SET domains (referred to as MLL3/4-ΔSET) (Figures 5A). These mutant cells exhibit bulk reductions in H3K4me1 (Figures 5B). In contrast to a recent report claiming MLL4 protein stability is dependent on its methyltransferase activity[27], we observe no protein stability defects in our MLL3/4-ΔSET clones (Figure 5C). In addition, while this manuscript was under revision, another study in mESCs used single amino acid catalytic-inactivating MLL3/4 point mutations, and also observed no effect on MLL3/4 protein stability[28]. We attribute these differences to the use of a potentially destabilizing triple-amino acid mutant versus the single point mutations or the deletion of the entire SET domain. Collectively, these results are in agreement that MLL3/4 regulates enhancer function largely through a methylation-independent mechanism[28].

By integrating our datasets with those of Wysocka and colleagues[28], we observe significant reductions in levels of H3K4me1/2 and H3K27ac at MLL3/4-bound enhancers, consistent with multiple ChIP-seq studies in MLL3/4 double-KO mESCs[28,29] (Figures 5D and 5E). K-means clustering (k=2) of the ChIP-seq data reveals that sites of MLL3/4-dependent H3K4-methylation are weakly associated with reductions in nearby gene expression, similar to results shown in Figure 3D (Figure 5F). Interestingly, the effects on nearby transcription are more evident in MLL3/4-ΔSET versus MLL3/4-point mutants, but less severe than the double-KO cells, indicating the MLL3/4 point-mutants may retain some residual catalytic-activity (Figure 5F). Additionally, MLL3/4-ΔSET mESCs stain positive for alkaline phosphatase activity, whereas MLL3/4 double-KO cells do not[29], thus the consequences of disrupting MLL3/4 methyltransferase activity are likely less severe than removing the entire gene product (Figure 5G). These results agree with our findings in *Drosophila* and point to a vital methylase-independent function for the Trr/MLL3/MLL4 COMPASS family *in vivo*.

Methods & Materials

Fly Stocks

Genomic *trr* rescue flies were generated using *pattB* plasmid for site-specific integration on 3R (89E11) and injections performed by BestGene (strain 9744). Transgenic flies were crossed to *trr*[1] and then made homozygous for the *trr*[1] allele and *trr-rescue* construct. UAS-Nipped-B was overexpressed using a *da*-Gal4 driver line. *T80*-Gal4 (1878), *en*-Gal4 (33557), UAS-*trr*-RNAi (29563) were purchased from Bloomington Drosophila Stock Center.

HMTase analysis of the reconstituted Trr/hWRAD complex *in vitro*

Trr complex was reconstituted by co-transfecting Sf9 cells with the cocktail of five baculoviruses expressing Trr-WinSET, hRBBP5, hASH2L, hWDR5, and hDPY30,

respectively, and purified with M2 agarose resin (sigma-aldrich). All five components are N-terminally FLAG-tagged. HMTase assay was performed in 20μL of 50mM Tris-HCl pH8.8, 20mM KCl, 5mM MgCl2, 0.5mM DTT at 37°C for 1hr with 1μg of recombinant histone H3 (NEB), 200μM of S-adenosyl-methionine, and near equal amounts of Trr + hWRAD.

Chromatin Immunoprecipitation (ChIP)

ChIP-seq experiments in adult brains were performed as follows. Approximately 10,000 adult flies (4–6 days after eclosion) were flash-frozen in a 50mL tube, vortexed to decapitate heads, and separated by passing through 710 μm and 425 μm sieves. Heads were cross-linked for 10 minutes while homogenizing with a 50mL dounce (loose pestle) in Buffer A1 (15mM HEPES pH 7.5, 15mM NaCl, 60mM KCl, 4mM $MgCl_2$, 0.5% Triton X-100, 0.5mM DTT) plus 2% paraformaldehyde and protease inhibitor (Sigma P8340) added fresh. Glycine was added to final concentration of 225mM to quench fixative and the mixture was then passed through a 70 μm cell-strainer to remove debris. Brain cells were pelleted for 5 minutes at 2000 × g, 4°C, and washed once with Orlando/Paro buffer (10mM Tris pH7.5, 10mM EDTA, 0.5mM EGTA, 0.25% Triton X-100, 0.5mM DTT, protease inhibitors) and pelleted again at 2000 × g. Final cell pellet was resuspended in sonication buffer (10mM Tris pH 8.0, 1mM EDTA, 0.1% SDS, protease inhibitor) and chromatin was fragmented to between 200–600 bp using a Covaris E220 bath sonicator. Chromatin was incubated with antibodies overnight at 4°C, immunoprecipitated the following day by incubating with Protein A/G agarose beads (SantaCruz), washed six times in RIPA buffer (25mM Tris pH7.5, 140mM NaCl, 1% Triton X-100, 1mM EDTA, 0.1% SDS, 0.1% Na-deoxycholate, 0.5mM DTT), and eluted (0.1M $NaHCO_3$, 1% SDS). After cross-link reversal and Proteinase K digestion overnight at 65°C, DNA was purified using Qiagen PCR purification spin-columns.

Larval wing imaginal disc ChIP-seq (~100 discs/ChIP) performed similar to previous reports[30], except chromatin was sheared using a Covaris E220 bath sonicator.

Qiagen RNeasy kits were used for all RNA-purification.

Antibodies

Antibodies recognizing H3K4me1, H3K4me2, H3K4me3, Trr, Lpt, Trx, dSet1, MLL3, and MLL4 were generated in the Shilatifard lab[1,31] and anti-H3K27ac was purchased from CST (D5E4, Rb mAb #8173).

Genome Editing by CRISPR/Cas9

Plasmid Cloning—CRISPR sgRNA constructs were cloned into pX330. pX330-U6-Chimeric_BB-CBh-hSpCas9 was a gift from Feng Zhang (Addgene plasmid # 42230)[32]. The following oligos were heat-denatured, annealed by gradual cooling and cloned into the BbsI site of pX330:

Generation of CRISPR mutations—Mouse v6.5 ES cells were cultured in 2i media plus LIF and electroporated with 25ug of each sgRNA construct and 10ug of CAG-EGFP-IRES-Puro (a kind gift from the Hitoshi Niwa lab). One day post electroporation, cells were

selected for 24 hours with 1ug/ml Puromycin. Cells were allowed to recover for 4 days and then plated at low density (1,000 cells per 10cm dish) and allowed to form colonies. Single colonies were picked into 96-well plates and allowed to grow to confluence at which point plates were split in half for freezing and lysis. Cells were lysed in DNA extraction buffer (10mM Tris, pH 8.5, 50mM KCl, 1.5mM $MgCl_2$, 0.45% NP-40, 0.45% Tween-20, 0.5 mg/ml proteinase K) and incubated overnight at 55°C. Proteinase K was inactivated by heating to 95°C for 12 minutes and DNA was analyzed by PCR.

NGS Data Processing—RNA-seq and ChIP-seq samples were sequenced with the Illumina NextSeq technology, and output data were processed with the bcl2fastq software tool. Sequence quality was assessed using FastQC v 0.11.2 (Andrew 2010), and quality trimming was done using the FASTX toolkit. RNA-seq and ChIP-seq reads were aligned to the mm9 and dm3 genomes using TopHat v2.0.9 and Bowtie v0.12.9, respectively, and only uniquely mapped reads with a two-mismatch threshold were considered for downstream analysis. Gene annotations from Ensembl 67 were used for mouse cells, and gene annotations from Ensembl 70 were used for Drosophila cells. Output bam files were converted into bigwig track files to display coverage throughout the genome (in RPM) using the GenomicRanges package[33].

RNA-seq Analysis—Gene count tables were used as input for edgeR 3.0.8[34]. Genes with Benjamini-Hochburg adjusted p-values less than 0.01 were considered to be differentially expressed. Heatmaps displaying gene expression levels transformed into Z-scores were generated using the pheatmap R package, and the rows (genes) and/or columns (samples) in these heatmaps were subjected to unsuperivised hierarchical clustering.

ChIP-seq Analysis—Peaks were called with the MACS v1.4.2 software[35] using default parameters. H3K27ac peaks were separated into TSS and nonTSS groups based on whether or not they overlapped regions within 500 bp of a TSS. Density bar plots were generated using unpublished perl and R scripts written by Yaping Lui, which incorporated some UCSC genome browser tools. For the density bar plots, mean RPM values for each sample were computed along the genome in 10-bp bins, input was subtracted, and ChIP-seq RPM values were transformed into Z-scores by subtracting the mean RPM value across the genome and dividing by the standard deviation of the genome-wide mean RPM value. Subsequently, these Z-scores were aligned to H3K27ac nonTSS and TSS peaks. Metaplots and heatmaps and were generated using ngsplot[36]. Metaplots show log fold changes relative to input, and the heatmaps show log fold changes relative to wild type. K-means clustering was also performed using ngsplot, and nearest-gene log fold changes in gene expression (from the RNA-seq edgeR ouput) corresponding to the clustered peaks in the heatmaps were determined using in house scripts and visualized with Java TreeView[37].

Statistical Analysis—For statistical analyses, R and Microsoft Excel were used. Appropriate statistical tests were used for all data where a statistical analysis was reported. An *F* test was performed to determine whether different groups had the same variance or not. For ChIP-seq Z-score analysis, *t* tests were calculated between the area-under-the-curve

values. *P* values <0.05 were considered to be statistically significant. All experiments were conducted in unblinded conditions.

Data Availability Statement—RNA-seq and ChIP-seq raw data are available in the Gene Expression Omnibus (GEO) database: GEO-GSE95781

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

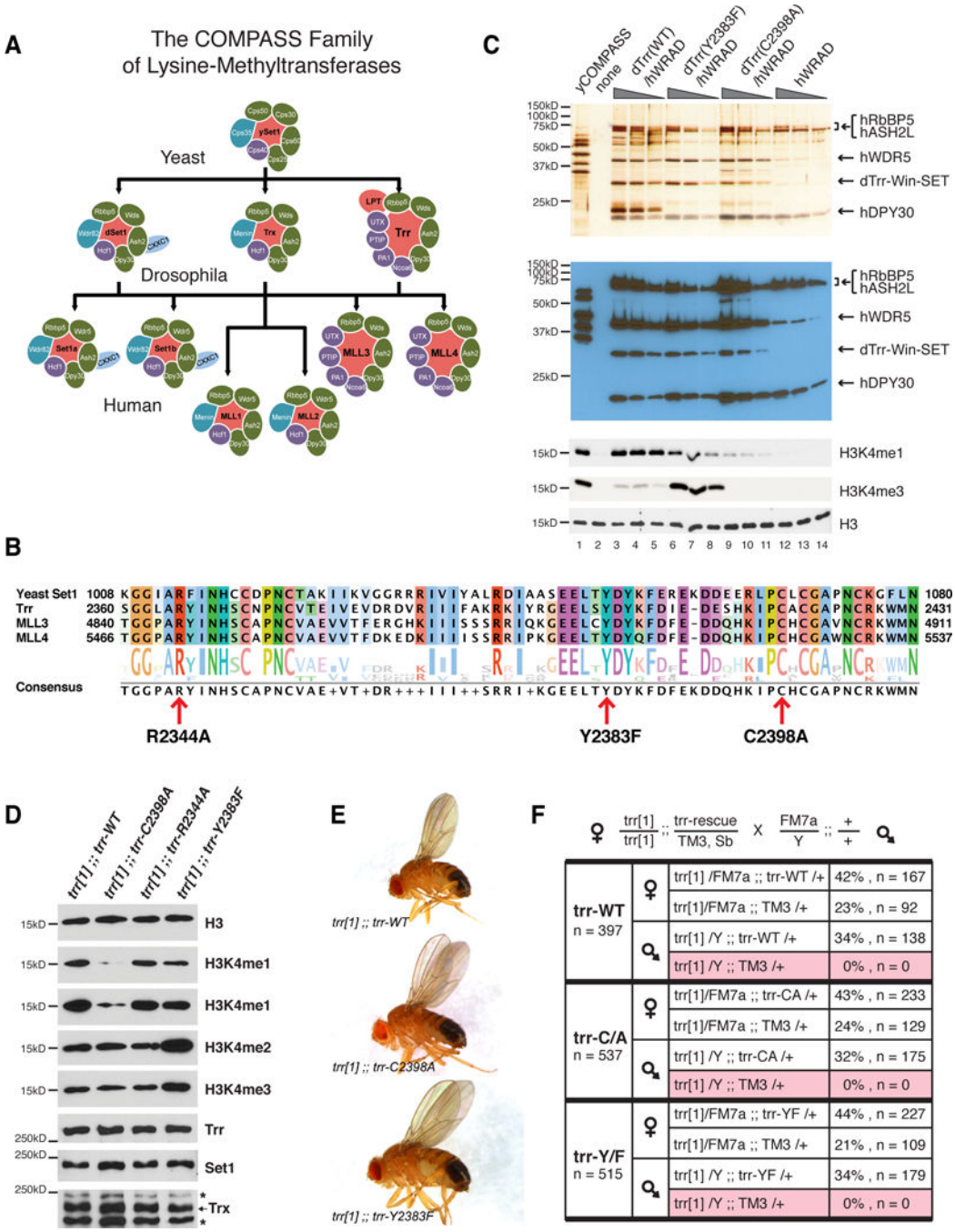
We thank members of the Shilatifard laboratory for their critical feedback during preparation of the manuscript and L. Shilatifard for editorial assistance. R.R. is supported by the NIH/NCI under the Ruth L. Kirschstein National Research Service Award (F31CA213928). A.P. is an EMBO postdoctoral fellow (ALTF 372-2015), and his work in the Shilatifard laboratory is supported by AIRC and Marie Curie Actions–People–COFUND. H.-M.H. is supported by the NIH (R00CA181506). E.R.S. is supported by the NIH (R50CA211428). C.C.S. is supported in part by an NIH/NCI training grant (T32CA09560). Studies on Drosophila Trr and the mammalian MLL3 and MLL4 COMPASS-like family in the Shilatifard laboratory are supported in part by a generous Outstanding Investigator Award from the National Institute of Health to A.S. (R35CA197569).

References

1. Herz HM, et al. Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. Genes & Development. 2012; 26:2604–2620. [PubMed: 23166019]
2. Hu D, et al. The MLL3/MLL4 Branches of the COMPASS Family Function as Major Histone H3K4 Monomethylases at Enhancers. Molecular and Cellular Biology. 2013; 33:4745–4754. [PubMed: 24081332]
3. Shilatifard A. The COMPASS Family of Histone H3K4 Methylases: Mechanisms of Regulation in Development and Disease Pathogenesis. Annu Rev Biochem. 2012; 81:65–95. [PubMed: 22663077]
4. Banerji J, Rusconi S, Schaffner W. Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell. 1981; 27:299–308. [PubMed: 6277502]
5. Fukaya T, Lim B, Levine M. Enhancer Control of Transcriptional Bursting. Cell. 2016; 166:358–68. [PubMed: 27293191]
6. Levine M, Cattoglio C, Tjian R. Looping back to leap forward: transcription enters a new era. Cell. 2014; 157:13–25. [PubMed: 24679523]
7. Smith E, Shilatifard A. Enhancer biology and enhanceropathies. Nat Struct Mol Biol. 2014; 21:210–9. [PubMed: 24599251]
8. Dorsett D. Distant liaisons: long-range enhancer-promoter interactions in Drosophila. Curr Opin Genet Dev. 1999; 9:505–14. [PubMed: 10508687]
9. Creyghton MP, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A. 2010; 107:21931–6. [PubMed: 21106759]
10. Heintzman ND, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nature Genetics. 2007; 39:311–318. [PubMed: 17277777]
11. Lee JE, et al. H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. Elife. 2013; 2:e01503. [PubMed: 24368734]
12. Piunti A, Shilatifard A. Epigenetic balance of gene expression by Polycomb and COMPASS families. Science. 2016; 352:aad9780. [PubMed: 27257261]
13. Lawrence MS, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature. 2014; 505:495–501. [PubMed: 24390350]
14. Sedkov Y, et al. Molecular genetic analysis of the Drosophila trithorax-related gene which encodes a novel SET domain protein. Mech Dev. 1999; 82:171–9. [PubMed: 10354481]

15. Takahashi YH, et al. Regulation of H3K4 trimethylation via Cps40 (Spp1) of COMPASS is monoubiquitination independent: implication for a Phe/Tyr switch by the catalytic domain of Set1. Mol Cell Biol. 2009; 29:3478–86. [PubMed: 19398585]
16. Collins RE, et al. In vitro and in vivo analyses of a Phe/Tyr switch controlling product specificity of histone lysine methyltransferases. J Biol Chem. 2005; 280:5563–70. [PubMed: 15590646]
17. Kanda H, Nguyen A, Chen L, Okano H, Hariharan IK. The Drosophila ortholog of MLL3 and MLL4, trithorax related, functions as a negative regulator of tissue growth. Mol Cell Biol. 2013; 33:1702–10. [PubMed: 23459941]
18. Eissenberg JC, et al. The trithorax-group gene in Drosophila little imaginal discs encodes a trimethylated histone H3 Lys4 demethylase. Nat Struct Mol Biol. 2007; 14:344–6. [PubMed: 17351630]
19. Gildea JJ, Lopez R, Shearn A. A screen for new trithorax group genes identified little imaginal discs, the Drosophila melanogaster homologue of human retinoblastoma binding protein 2. Genetics. 2000; 156:645–63. [PubMed: 11014813]
20. Kagey MH, et al. Mediator and cohesin connect gene expression and chromatin architecture. Nature. 2010; 467:430–5. [PubMed: 20720539]
21. Michaelis C, Ciosk R, Nasmyth K. Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. Cell. 1997; 91:35–45. [PubMed: 9335333]
22. Phillips-Cremins JE, et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell. 2013; 153:1281–95. [PubMed: 23706625]
23. Rollins RA, Korom M, Aulner N, Martens A, Dorsett D. Drosophila nipped-B protein supports sister chromatid cohesion and opposes the stromalin/Scc3 cohesion factor to facilitate long-range activation of the cut gene. Mol Cell Biol. 2004; 24:3100–11. [PubMed: 15060134]
24. Rollins RA, Morcillo P, Dorsett D. Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. Genetics. 1999; 152:577–93. [PubMed: 10353901]
25. Gause M, Misulovin Z, Bilyeu A, Dorsett D. Dosage-sensitive regulation of cohesin chromosome binding and dynamics by Nipped-B, Pds5, and Wapl. Mol Cell Biol. 2010; 30:4940–51. [PubMed: 20696838]
26. Mohan M, et al. The COMPASS Family of H3K4 Methylases in Drosophila. Molecular and Cellular Biology. 2011; 31:4310–4318. [PubMed: 21875999]
27. Jang Y, Wang C, Zhuang L, Liu C, Ge K. H3K4 Methyltransferase Activity Is Required for MLL4 Protein Stability. J Mol Biol. 2016
28. Dorighi KM, et al. Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation. Molecular Cell. 2017
29. Wang C, et al. Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. Proc Natl Acad Sci U S A. 2016; 113:11871–11876. [PubMed: 27698142]
30. Papp B, Muller J. Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. Genes Dev. 2006; 20:2041–54. [PubMed: 16882982]
31. Rickels R, et al. An Evolutionary Conserved Epigenetic Mark of Polycomb Response Elements Implemented by Trx/MLL/COMPASS. Mol Cell. 2016; 63:318–28. [PubMed: 27447986]
32. Cong L, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013; 339:819–23. [PubMed: 23287718]
33. Lawrence M, et al. Software for computing and annotating genomic ranges. PLoS Comput Biol. 2013; 9:e1003118. [PubMed: 23950696]
34. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2009; 26:139–140. [PubMed: 19910308]
35. Zhang Y, et al. Model-based Analysis of ChIP-Seq (MACS). Genome Biol. 2008; 9:R137. [PubMed: 18798982]
36. Shen L, Shao N, Liu X, Nestler E. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. BMC Genomics. 2014; 15:284. [PubMed: 24735413]

37. Saldanha AJ. Java Treeview–extensible visualization of microarray data. Bioinformatics. 2004; 20:3246–8. [PubMed: 15180930]



trr-WT n = 397	♀	trr[1] /FM7a ;; trr-WT /+	42% , n = 167
		trr[1]/FM7a ;; TM3 /+	23% , n = 92
	♂	trr[1] /Y ;; trr-WT /+	34% , n = 138
		trr[1] /Y ;; TM3 /+	0% , n = 0
trr-C/A n = 537	♀	trr[1]/FM7a ;; trr-CA /+	43% , n = 233
		trr[1]/FM7a ;; TM3 /+	24% , n = 129
	♂	trr[1] /Y ;; trr-CA /+	32% , n = 175
		trr[1] /Y ;; TM3 /+	0% , n = 0
trr-Y/F n = 515	♀	trr[1]/FM7a ;; trr-YF /+	44% , n = 227
		trr[1]/FM7a ;; TM3 /+	21% , n = 109
	♂	trr[1] /Y ;; trr-YF /+	34% , n = 179
		trr[1] /Y ;; TM3 /+	0% , n = 0

Figure 1. The catalytic activity of Trr is dispensable for viability

(A) The COMPASS family of lysine-methyltransferases is conserved from yeast to humans. Common subunits are shown in green, shared subunits in purple, and complex-specific subunits in blue.

(B) Alignment of the SET domains from Trr, MLL3, MLL4 and yeast Set1. Single amino acid substitutions were introduced into the SET domain of Trr at the positions highlighted by red arrows. Alignment plots were created in Jalview.

(C) *Drosophila* Trr SET domain was reconstituted using Baculovirus in SF9 cells along with flag-tagged human COMPASS core subunits (WRAD): WDR5, RbBP5, ASH2L, and DPY30. Silver-staining and anti-flag western blots show the purified components, and western blots for H3K4me1 and H3K4me3 demonstrate diminished catalytic activity of the C2398A mutation and hyperactivating effect of Y2383F on Trr activity *in vitro*. Uncropped images in Supplemental Figure 5. Image representative of at least two independent experiments.

(D) Western blot of lysates from wing imaginal discs. The catalytically deficient *trr-C2398A* construct is not able to restore H3K4me1 (lane 2, panels 2–3, short and long exposures) while the hyperactive *trr-Y2383F* construct exhibits increased levels in H3K4me2 (lane 4, panel 4) and H3K4me3 (lane 4, panel 5). Lane 1: *trr[1] ;; trr-WT*. Lane 2: *trr[1] ;; trr-C2398A*. Lane 3: *trr[1] ;; trr-R2344A*. Lane 4: *trr[1] ;; trr-Y2383F*. Panel 1: α-H3. Panel 2: α-H3K4me1, short exposure. Panel 3: α-H3K4me1, long exposure. Panel 4: α-H3K4me2. Panel 5: α-H3K4me3. Panel 6: α-Trr. Panel 7: α-Set1. Panel 8: α-Trx. The arrow marks the Trx band and the asterisks indicate unspecific bands. The R2344A mutation showed no effect on bulk H3K4me1 levels and was not included in subsequent experiments. Uncropped images in Supplemental Figure 5. Image representative of at least two independent experiments.

(E) Lethality of the *trr* null allele *trr*[1] can be rescued by various genomic *trr* constructs containing the *trr* gene and putative regulatory regions. Only male rescued flies are shown. From top to bottom: *trr-WT*, *trr-C2398A* (inactive), and *trr-Y2383F* (hyperactive).

(F) Genetic analyses confirm the *trr*[1] allele remains hemizygous lethal in the absence of either of the three *trr-rescue* transgenes. Lack of *trr*[1] males are highlighted in pink.

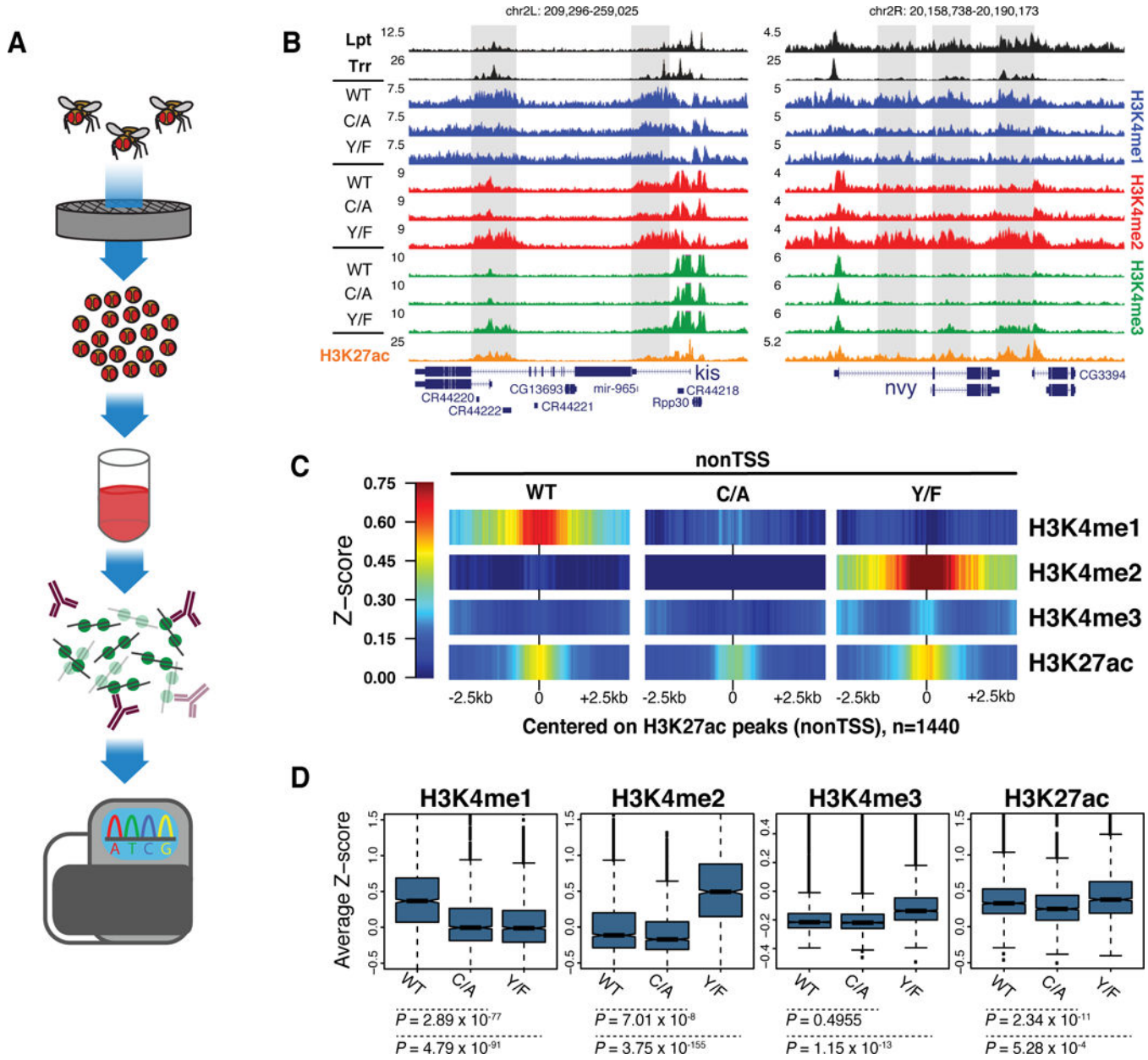


Figure 2. Enhancer chromatin modification is specifically affected by Trr catalytic mutations

(A) Graphical depiction of ChIP-seq experimental workflow using adult fly brain tissues, as described in the methods section.

(B) Representative track examples highlighting changes (gray boxes) in enhancer H3K4-methylation. Note the reductions in H3K4me1 in the *trr-C/A* (inactive), and the conversion of H3K4me1 to H3K4me2/3 in the *trr-Y/F* (hyperactive) specifically at H3K27ac-marked nonTSS sites. Image representative of two independent experiments.

(C) Density bar-plots comparing H3K4-methylation Z-scores at 1440 active enhancers. Plots are centered on H3K27ac (nonTSS) peaks, +/− 2.5kb. Note the reductions in H3K4me1 in the *trr-C/A* and conversion of H3K4me1 to H3K4me2/3 in the hyperactive *trr-Y/F*. The Z-score scale for H3K4me3 is −0.25 to 0.5, and for H3K27ac is 0 to 1.5.

(D) Box plots quantifying the H3K4-methylation changes shown in (C). Upper/lower boundaries represent the interquartile range (IQR). Upper whiskers signify the third quartile (Q3) + 1.5*IQR. Lower whiskers signify Q1 – 1.5*IQR. Dots represent outliers. P-values were calculated by conducting t-tests.

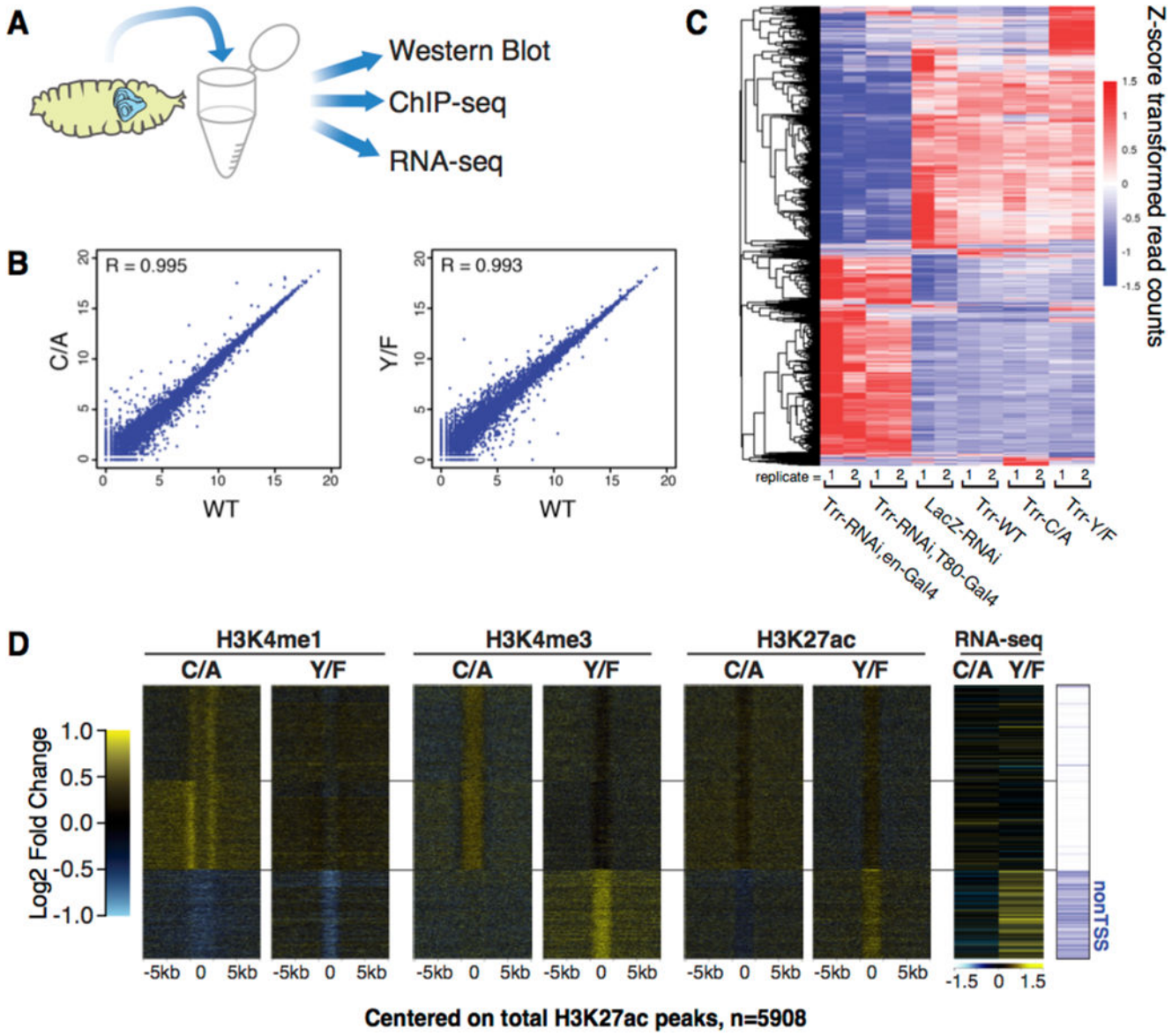


Figure 3. Gene expression is only modestly affected by Trr-dependent enhancer methylation

(A) Graphical depiction of 3rd instar larval wing imaginal discs and experiments performed with these tissues.

(B) Gene expression analysis shows >99% correlation (Pearson) between *trr*[1] flies rescued with *trr-WT* and flies rescued with either *trr-C/A* (inactive) or *trr-Y/F* (hyperactive).

(C) *trr*-RNAi is expressed in wing discs using separate Gal4 drivers, *en* and *T80*, and gene expression profiles are compared with three *trr[1]*-rescue lines. Hierarchical clustering of differentially expressed genes segregates *trr*-RNAi samples from *trr* catalytic mutants. Two independent biological replicates from each sample are shown in the heatmap displaying the total number of differentially expressed genes among these 6 samples. Note the similarities among *trr*[1]-rescue lines versus *trr*-RNAi.

(D) K-means clustering (k=3) of H3K4me1/3 and H3K27ac ChIP-seq data displayed as $\log_2$ fold change heat maps centered on total H3K27ac peaks (+/− 5kb). Cluster 3 (bottom) is

over-enriched for nonTSS sites ($p=3.61 \times 10^{-404}$, hypergeometric test) most affected by Trr catalytic mutations. RNA-seq $\log_2$ fold changes show the effects of enhancer chromatin modifications on nearby gene expression. NonTSS sites are depicted as blue in the far-right panel, whereas TSS are white. Wing disc ChIP-seq experiments were performed once.

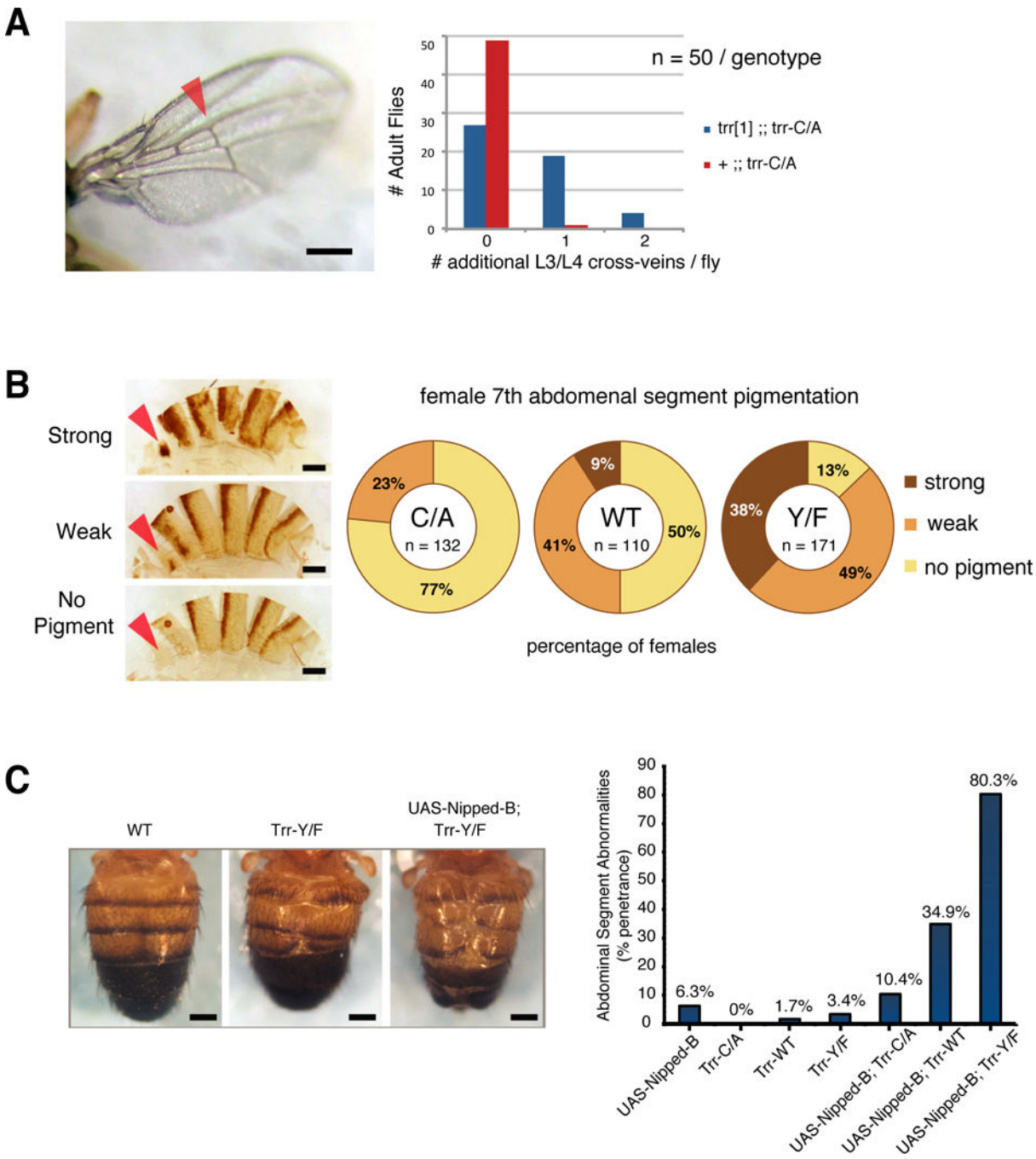


Figure 4. Subtle phenotypes of *trr* catalytic mutants

(A) Histograms quantitate increased frequency of additional L3/L4 cross-vein phenotype in *trr[1] ;; trr-C/A* versus +;; *trr-C/A* when grown at 29°C. The x-axis shows whether one or both wings display the additional cross-vein. 50 adult flies were scored for each genotype. Scale bar = 0.5mm.

(B) Female 7th abdominal segment pigmentation is modulated in a Trr-methylation-dependent manner. Note, increased frequency of strong pigmentation in *trr-Y/F*, and

decreased pigmentation in *trr-C/A* compared with *trr-W/T*. Red arrowhead designates 7th abdominal segment. Number of flies scored is listed for each genotype. Scale bar = 0.25mm. (C) *trr-Y/F* flies display subtle abdominal segmentation defects, which are exacerbated by Nipped-B over-expression in the *trr[+]* background. Note the roughly 12-fold increase in penetrance when Nipped-B levels are increased in the *trr[+];;trr-Y/F* background. Number of flies counted per genotype is: 267, 177, 223, 192, 113, 103, and 89, with respect to their order shown in the graph. Scale bar = 0.25mm.

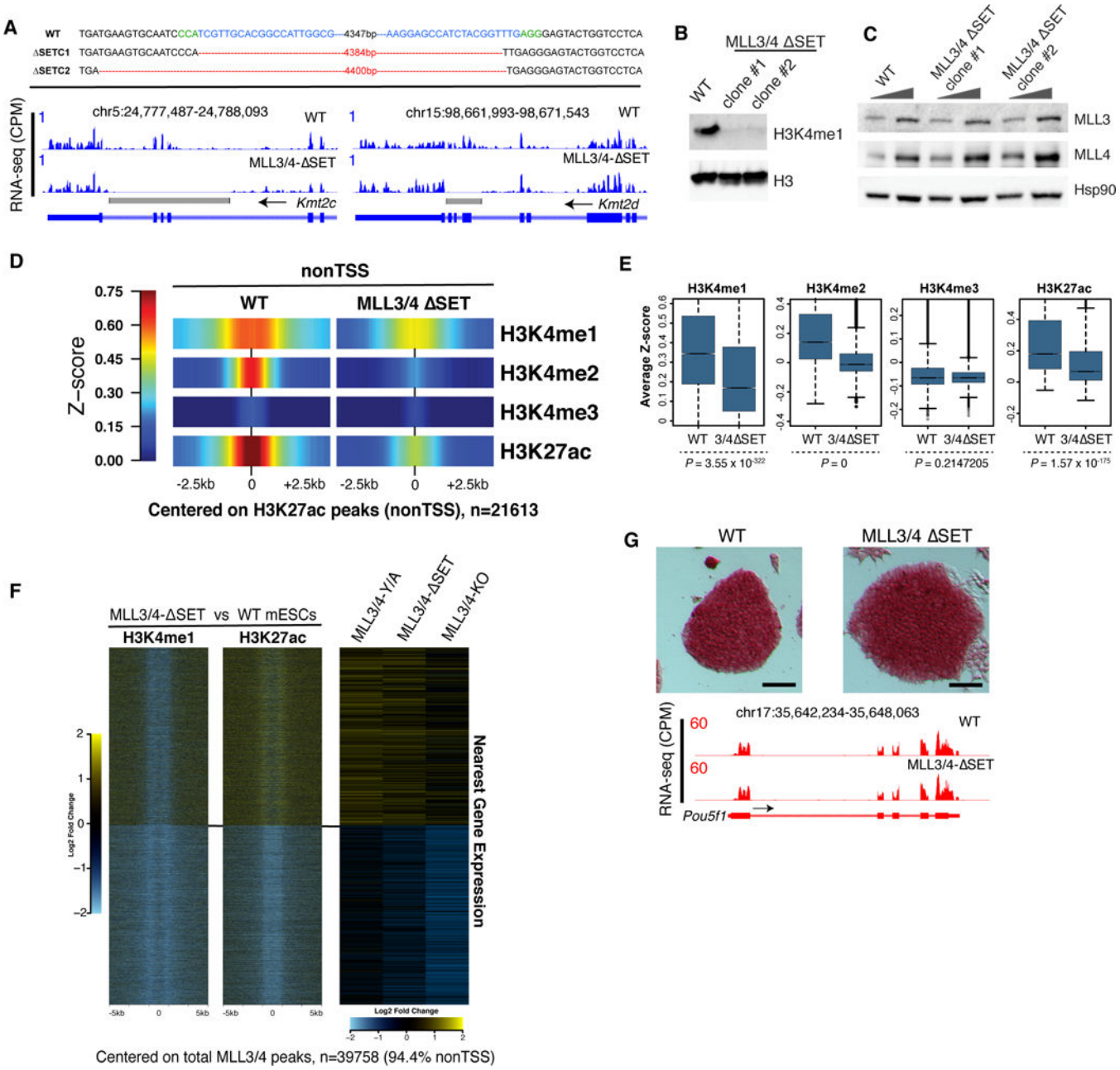


Figure 5. MLL3/4-COMPASS catalytic activity is not required for enhancer function in mouse ES cells

(A) DNA sequence and track examples from RNA-sequencing showing the CRISPR deleted regions to generate double MLL3/4-ΔSET homozygous mESC lines. Deletion of these three 3′ exons of *KMT4C* and *KMT4D* create truncations shortly after the WDR5-interacting motif.

(B) Western blots showing decreased H3K4me1 in both MLL3/4-ΔSET clones. Uncropped images in Supplemental Figure 6B.

(C) Western blots showing our SET domain deletions do not reduce MLL3/4 protein stability. Uncropped images in Supplemental Figure 6C. Image representative of at least two independent experiments.

(D) Density bar-plots of Z-score transformed ChIP-seq reads are centered on nonTSS H3K27ac peaks (n=21613) and show decreased enhancer H3K4me1/2 and H3K27ac in MLL3/4-ΔSET cells.

(E) Box plots quantifying changes in histone modifications displayed in (D). Upper/lower boundaries represent the interquartile range (IQR). Upper whiskers signify the third quartile (Q3) + 1.5*IQR. Lower whiskers signify Q1 – 1.5*IQR. Dots represent outliers. P-values were calculated by conducting t-tests. Results are representative of two independent experiments.

(F) K-means clustered heatmaps (k=2) centered on total MLL3/4 peaks from Wysocka and colleagues[28] (n=39758, 94.4% enriched for nonTSS) display log2 fold-changes of H3K4me1 and H3K27ac in MLL3/4-ΔSET versus WT mESCs. The accompanying heatmap shows log2 fold changes in nearby gene expression, comparing the published MLL3/4 point-mutation (MLL3/4-Y/A) and MLL3/4-KO[28] with our MLL3/4-ΔSET cells provides evidence that deletion of the SET domain is less consequential than a double-KO. Cluster #2 shows 831 significantly (adjusted p-value < 0.01) down-regulated genes in MLL3/4-ΔSET cells versus 3419 down-regulated in MLL3/4-KO cells. Average log2 fold gene-expression changes between 2 independent replicates are shown.

(G) MLL3/4-ΔSET mESCs retain their self-renewal capacity, express pluripotency factors such as Pou5f1, and stain positive for alkaline phosphatase activity, contrary to what is reported for MLL3/4-double-KO cells[29]. Scale-bar = 100 microns.

Tables

MLL3-ΔSET guides:	Sense: cgccaatggccgtgcaacga
	Anti-sense: aaggagccatctacggtttg
MLL4-ΔSET guides:	Sense: AGGCGAGGGGCCCCGATTGA
	Antisense: CAGCTTAAATTCCGGCCTTG