INTRODUCTION

Epigenetic posttranslational modifications on histones are highly dynamic; extensive changes of these histone states can affect recruitment of effector proteins that fundamentally shape gene expression programs underlying processes that govern development and disease (1–6). Specifically, methylation at histone lysine residues has received considerable attention due to the important involvement with transcriptional regulation. One such mark, histone H3 lysine 4 trimethylation (H3K4me3), is an evolutionarily conserved modification consistently found at transcription start sites (TSSs) and serves as a hallmark of active gene promoters (6–8). Studies have conferred that H3K4me3 interacts with chromatin remodelers (9, 10) and promotes recruitment of basic transcription factors to facilitate transcriptional activation (11, 12). Other functional implications ascribed to H3K4me3 include DNA damage repair response (13, 14), marking of bivalent genes coexpressed with the repressive H3K27me3 modification (15), and cell identity specification (16, 17). Thus, there is a consensus that H3K4me3 is critically linked to transcriptional output and cellular response.

Deposition of methylation on H3K4 is catalyzed by lysine methyltransferases (KMTs), namely, the family of COMPASS (COMplex of Proteins ASSociated with Set1) complexes, which is widely conserved from yeast to human (18–20). While Set1/COMPASS is solely responsible for implementing all three methylation patterns in yeast (19–21), our laboratory and others have described the division of labor in methylating H3K4 among the COMPASS members in higher organisms (21–28). In mammals, there are six Set1-related enzymes that reside in COMPASS-like complexes: Set1A, Set1B, Mll1, Mll2, Mll3, and Mll4 (18, 29, 30). Evolutionary expansion of the COMPASS family in higher metazoans denotes functional diversification of H3K4 methylation, showcasing the underlying complexity of epigenetic regulation (31). Mammalian Set1A/SET is a consensus that H3K4me3 is critically linked to transcriptional aberrations seen in cells with total protein (32, 33); Mll1/Mll2 catalyze H3K4me3 in a locus-specific manner (e.g., targeting Hox genes) (25, 27, 34–36); and Mll3/Mll4 are key H3K4 monomethyltransferases at enhancers (24, 29, 37). Recent studies have begun to reassess the biological significance of the catalytic activity as the primary function of COMPASS. For instance, the enzymatic SET domain of Set1A has been shown to be nonessential for embryonic stem cell (ESC) self-renewal (38), although deletion of the full-length protein impairs viability (39, 40). Likewise, inactivating mutations of the SET domain of Trr, the Drosophila homolog of Mll3/Mll4, does not result in clear developmental defects, and Mll3/Mll4 catalytically deficient ESCs have less transcriptional aberrations than seen in cells with total protein loss (41, 42). Our laboratory also determined Mll4/COMPASS could regulate enhancers without its enzymatic activity (43). These studies suggest that the COMPASS family function independent of H3K4 methylation is context dependent. However, deletion of any individual COMPASS member in mice results in embryonic or prenatal lethality with distinct phenotypes (37, 39, 44–46), signifying these proteins have partially nonredundant functions. Although research has provided insight into how each COMPASS member operates in unique contexts, mechanisms underlying their roles in cellular and developmental regulation remain elusive.

We previously reported that ablating the SET domain of Set1A (Set1ASET or Set1AΔSET) does not disrupt bulk H3K4me3 in ESCs (38), indicating the likelihood of other COMPASS members having functionally redundant roles to sustain global H3K4me3 in Set1AΔSET ESCs. Multiple lines of evidence thus far have suggested that both Set1A and its paralog Set1B contribute to genome-wide H3K4me3 deposition
Set1B to establish \( \Delta \))38\)), we used CRISPR-Cas9 to investigate the role of Set1B in H3K4 methylation (Fig. 1D). Because independent evidence suggests that Set1B binds to chromatin in ESCs, prompting us to determine the potential function of Set1B in ESCs. To explore this hypothesis, we deleted the first four exons of the Set1B gene in ESCs (Set1B KO) and examined the effect on H3K4 methylation. Our previous study demonstrated that deleting the C-terminal catalytic SET domain of Set1A does not disrupt global H3K4 methylation in ESCs, which suggested that other enzymes implement bulk H3K4 methylation in the absence of Set1A activity (38). Because Set1B is structurally homologous to Set1A (fig. S1A), we investigated whether Set1B could compensate for Set1A\(^{\text{SET}}\) function to sustain global H3K4 methylation in ESCs. Therefore, we generated Set1B knockout (Set1BKO) ESCs by deleting the first four exons of the Set1B genomic locus via CRISPR-Cas9 gene editing. Polymerase chain reaction (PCR) genotyping and RNA sequencing (RNA-seq) confirmed the deletion of the intended genomic region and Set1B transcript (fig. S1, B and C), and Western blotting substantiated the complete loss of the Set1B protein in ESCs (Fig. 1A). Ablating Set1B did not alter bulk H3K4 methylation as observed by Western blotting (Fig. 1B and fig. S1D). ChIP-seq [chromatin immunoprecipitation (ChIP) followed by high-throughput sequencing] analyses examined similar H3K4me3 levels in wild-type (WT) and Set1BKO ESCs (Fig. 1C). In addition, the expressions of key pluripotency factors Sox2, Nanog, and Oct4 remain invariant when comparing WT and Set1BKO cells (fig. S1E). These findings thus far agree with prior studies demonstrating that Set1B is not essential for ESC self-renewal and that Set1B removal does not affect bulk H3K4 methylation (39).

Our laboratory has recently established that Set1B, unlike the rest of its COMPASS family relatives, resides predominantly in the cytoplasm (47); however, we find that a small fraction of Set1B still binds to chromatin in ESCs, prompting us to determine the potential role of Set1B in H3K4 methylation (Fig. 1D). Because independent abrogation of Set1B or the catalytic SET domain of Set1A did not change overall H3K4 methylation levels (38), we used CRISPR-Cas9 to establish Set1BKO-Set1A\(^{\text{SET}}\) ESCs. Successful generation of homozygous double-mutant ESCs was verified by PCR genotyping and RNA-seq (fig. S1F and Fig. 1E). Expressions of pluripotency markers Oct4, Nanog, and Sox2 are comparable between WT and Set1BKO-Set1A\(^{\text{SET}}\) ESCs at the significance level of \( P < 0.01 \) (fig. S1H). We assessed H3K4 methylation in the double-mutant cells compared with that in WT ESCs and found that while bulk H3K4me1/me2 levels did not differ, H3K4me3 was somewhat decreased in Set1BKO-Set1A\(^{\text{SET}}\) ESCs by Western blotting (fig. S1G and Fig. 1F). The reduction in H3K4me3 levels in the double mutant compared with WT ESCs was also consistently detected by ChIP-seq (Fig. 1G). The diminished H3K4me3 in the combined deletions of the SET domain of Set1A and knockout of Set1B, compared with either mutation alone, suggests functional redundancy between the two Set1 paralogs in regulating H3K4 methylation in ESCs.

Several studies have implicated a role for H3K4me3 in transcriptional activation by RNA polymerase II (RNAPII) (11, 12). Given the noticeable decrease in H3K4me3 in Set1BKO-Set1A\(^{\text{SET}}\) compared with WT cells, we used ChIP-seq to evaluate RNAPII occupancy genome-wide. Reduced RNAPII II is evident at sites with H3K4me3 loss, and accompanying reduction in gene expression can be seen in representative track examples (Fig. 2A). Such decrease in H3K4me3 and RNAPII II is not due to the loss of Set1A SET domain or Set1B alone (38), further supporting their functionally redundant roles in H3K4me3 implementation and transcription regulation at these regions (fig. S2, A and B). Genome-wide analyses in the Set1BKO-Set1A\(^{\text{SET}}\) double-mutant cells revealed that H3K4me3 is lost across regions proximal to TSSs, with a corresponding decrease in RNAPII II occupancy compared with WT ESCs (fig. S2C). Partitioning of these H3K4me3 peaks using K-means clustering revealed that H3K4me3 loss in the double mutant occurs in the second and third clusters, where Set1A binding is the strongest (Fig. 2, B and C). Both clusters also exhibit somewhat diminished RNAPII II levels in Set1BKO-Set1A\(^{\text{SET}}\) compared with WT cells (Fig. 2C). The accompanying RNA-seq data demonstrate moderate differences in gene expression pattern, with a stronger decrease in expression pertinent to cluster 3, which has a relatively more severe H3K4me3 loss (Fig. 2C). Cluster 3 also contains genes that are more highly expressed in WT ESCs (fig. S2D) and is especially enriched for housekeeping factors such as ribosomal-related proteins as shown by gene ontology (GO) annotation (fig. S2E). Collectively, these data signify that H3K4me3 at promoters of more highly expressed genes are more perturbed by the combinatorial deletions of Set1B and the SET domain of Set1A in ESCs. These findings are consistent with previous studies reporting that H3K4me3 at genes with higher expression are more sensitive to loss of CXXC1, a key subunit of the Set1/COMPASS complexes (48, 49).

We previously found that Mll2, another member of the COMPASS family of H3K4 methyltransferases, also catalyzes H3K4me3 in ESCs (25); therefore, we deleted the SET domain of Mll2 using CRISPR-Cas9 (fig. S2F). Through Western blotting, we noted that disrupting the SET domain of Mll2 adversely affects the protein’s stability (fig. S2G), indicating that Mll2 protein level in Mll2\(^{\text{SET}}\) ESCs is quite comparable to Mll2KO cells, which were previously generated in our laboratory (34). Histone H3K4me3 ChIP-seq analysis in Mll2\(^{\text{SET}}\) ESCs revealed loss of H3K4me3 at sites distinct from those seen in Set1BKO-Set1A\(^{\text{SET}}\) when compared with WT cells (fig. S2, C and H). Particularly, the greatest H3K4me3 reduction in Mll2\(^{\text{SET}}\) cells is confined mainly to cluster 1 peaks, where there is also a corresponding decrease in RNAPII II level and overall expression of genes nearest to these sites (Fig. 2, D and E). Cluster 1 contains the TSSs of genes that are typically less transcriptionally active and are linked to function in proper development (fig. S2, D and E), concordant with previous findings that Mll2/COMPASS implements H3K4me3 primarily at specific loci such as bivalent genes (25). Analysis of H3K4me3 ChIP-seq in ESCs without Mll1 (50), also known to deposit H3K4me3 in mammals, showed comparable decreases of global H3K4me3 between Mll1KO
and WT cells (fig. S2I), signifying that Mll1/COMPASS is not a crucial regulator of bulk H3K4me3 in ESCs.

The K-means clustering of H3K4me3 in Fig. 2 effectively divided peaks by their width, which also correlated with gene expression, such that wider H3K4me3 coincided with higher nearest gene expression, and narrower H3K4me3 concurred with lower nearest gene expression (Fig. 2, B to E, and fig. S2D), suggesting an underlying biological significance to H3K4me3 breadth. Multiple studies have recently explored such functional implications, reporting a positive relationship between H3K4me3 breadth and gene expression (17, 51) and attributing the role of breadth to determining specific cell identity during development and disease (16, 52, 53). We noted the predilection of Set1/COMPASS mutant cells to present H3K4me3 loss in clusters 2 and 3, which have greater H3K4me3 breadth, while mutant Mll2/COMPASS primarily affected cluster 1 with the narrowest breadth (Fig. 2, B to E). To investigate the function of H3K4me3 breadth further, we classified broad and narrow WT peaks, annotated to the nearest TSS, and sorted these regions from wide to narrow. We retrieved 500 sites with either the broadest or narrowest peaks and examined H3K4me3 levels in WT, Set1B KO–Set1A △SET, and Mll2 △SET ESCs. In comparison to WT, Set1B KO–Set1A △SET cells exhibit abated H3K4me3 levels at genes with mainly narrower peaks (Fig. 3A). These unequivocal differences of altered ChIP density by peak width in the two mutant cell lines are further illustrated quantitatively in composite profiles (Fig. 3B) and representative track examples (Fig. 3C). GO term analyses indicate that genes with broader H3K4me3 between double-mutant and WT cells is shown on the right. Two biological replicates of each genotype were analyzed.
Fig. 2. H3K4me3 is implemented by Set1/COMPASS at more transcriptionally active promoters, while Mll2 catalyzes H3K4me3 at lowly expressed genes. (A) Genome browser track examples showing ChIP-seq tracks of H3K4me3 and RNAP II, with corresponding RNA-seq tracks for Gpt2 (left) and Hexb (right). (B) TSS-proximal H3K4me3 (13,486) peaks were identified and partitioned into three groups via K-means clustering, and corresponding Set1A (left), H3K4me3 (middle), and RNAP II (right) levels were plotted for WT and double-mutant ESCs. (C) Log2FC changes in H3K4me3 (left) and RNAP II (middle) occupancy, as well as in RNA-seq (right), were determined in Set1BKO-Set1A−SET relative to WT cells for the three clusters of peaks identified and ordered in (B). (D) Mll2 (left), H3K4me3 (middle), and RNAP II (right) occupancy levels were plotted for WT and Mll2−SET ESCs at the clustered 13,486 H3K4me3 peaks identified in (B). (E) Log2FC changes in H3K4me3 (left) and RNAP II (middle) occupancy, as well as in RNA-seq (right), were assessed in Mll2−SET compared with WT ESCs for the three clusters of peaks determined and ordered in (B).
enriched for biological processes that are relatively more important for proper maintenance of stem cell identity (fig. S3A), supporting previous studies associating H3K4me3 breadth with cell identity specification (16, 17). In addition, narrow peaks are enriched for processes that are more pertinent to neuronal development (fig. S3A), consistent with published studies showing Mll2 implementing H3K4me3 at development-related genes in ESCs (25, 54, 55).

To further investigate the consequence of H3K4me3 peak width, the top and bottom breadth quintiles were evaluated. For the top 20% H3K4me3 peaks, the Set1BKO-Set1A\textsuperscript{ASET} mutant featured a significantly drastic decrease in H3K4me3 coverage compared with WT, while the Mll2\textsuperscript{ASET} mutant exhibited comparable H3K4me3 occupancy relative to WT (Fig. 3D). In contrast, Mll2\textsuperscript{ASET} ESCs showed a more significant reduction in H3K4me3 for the 20% narrowest H3K4me3 peaks compared with WT, while the Set1BKO-Set1A\textsuperscript{ASET} mutant is much less altered for H3K4me3 at these narrow peak regions (Fig. 3D). Furthermore, when assessing peak width of differential H3K4me3 TSS-proximal peaks for each mutant, Set1/COMPASS mainly affected wider peaks, while Mll2\textsuperscript{ASET} mutant primarily affected narrower peaks (fig. S3B). By examining the levels of RNAPII at H3K4me3-proximal TSSs in Set1BKO-Set1A\textsuperscript{ASET} and the Mll2\textsuperscript{ASET} mutants compared with WT, we observe decreased RNAPII levels at broader peaks for the Set1BKO-Set1A\textsuperscript{ASET} mutant and reduced RNAPII levels at narrower peaks for the Mll2\textsuperscript{ASET} mutant (Fig. 3E). On the basis of these findings, we speculate the following: (i) transcription of cell identity genes, which are marked by broad H3K4me3 (16), is regulated by Set1/COMPASS and, thus, illuminates the importance of Set1/COMPASS in maintaining stem

Fig. 3. H3K4me3 peak breadth is coordinately controlled by Set1 and Mll2/COMPASS. (A) Left: Heatmaps of H3K4me3 occupancy in WT, Set1BKO-Set1A\textsuperscript{ASET}, and Mll2\textsuperscript{ASET} ESCs at 500 genes with the broadest (top) and 500 genes with the narrowest (bottom) H3K4me3 peaks in WT. TSSs are ordered by descending peak width of their nearest H3K4me3 peak. Right: Log2FC in H3K4me3 levels comparing either Set1BKO-Set1A\textsuperscript{ASET} or Mll2\textsuperscript{ASET} relative to WT cells for the same ordered set of most broad (top) and narrow (bottom) peaks. (B) Profiles of log2FC of average H3K4me3 signal density (RPM) at the loci with either 500 most broad (top) or narrow (bottom) H3K4me3 peaks in Set1BKO-Set1A\textsuperscript{ASET} or Mll2\textsuperscript{ASET} compared with WT ESCs. (C) Example genome tracks of broad (right) and narrow (left) H3K4me3 peaks in WT, Set1BKO-Set1A\textsuperscript{ASET}, and Mll2\textsuperscript{ASET} ESCs. (D) Box plot contrasting normalized H3K4me3 ChIP-seq coverage for the top 20% broadest H3K4me3 peaks (n = 2626) and top 20% narrowest H3K4me3 peaks (n = 2627) in WT, Set1BKO-Set1A\textsuperscript{ASET} and Mll2\textsuperscript{ASET} ESCs. Coverage was normalized to peak width. P values (indicated by lowercase letters) were determined using the Wilcoxon rank sum test with continuity correction (paired). Calculated P values are as follows: a, P = 0.01901; b, P < 6.173 × 10\textsuperscript{-282}; c, P = 6.173 × 10\textsuperscript{-282}; d, P = 1.7914 × 10\textsuperscript{-74}. (E) Log2FC heatmaps of RNAPII levels at H3K4me3-proximal TSSs for each indicated mutant ESC line relative to WT cells for the top 20% broadest (top) or narrowest (bottom) H3K4me3 peaks initially defined in (D).
cell viability (38, 39), and (ii) the observed phenomenon of Mll2 affecting narrow H3K4me3 peaks, which mark developmental genes, supports the critical role of Mll2 in development (34, 56, 57). There are several plausible explanations for the differential recruitment and activity of these COMPASS methyltransferases at distinct groups of loci. Studies have shown that Set1, the yeast homolog to mammalian Set1A and Set1B, is recruited to chromatin through its association with elongating RNAP II (58–60). One study has recently proposed that repeated passaging of elongation complexes containing Set1/COMPASS contributes to the widening or broadening of H3K4me3 levels, which correlate with increased transcription frequency (61). While this model has yet to be demonstrated for Mll2/COMPASS, it is possible that Mll2 is recruited to chromatin by other factors that include Menin, which also resides in the Mll/COMPASS complexes, and LEDGF (27, 62–65). In addition, different DNA-binding specificities of the CXXC motif found in the CXXC1 protein, a key subunit in the Set1/COMPASS complex, and in Mll2 have been reported to contribute to selective targeting of various COMPASS complexes to their respective genomic loci (66–68). In sum, these findings demonstrate that the difference in H3K4me3 patterns established by Set1A/COMPASS and Mll2/COMPASS is indicative of functional significance in stem cells.

We noted that Set1A and Mll2 are both highly enriched at H3K4me3 TSS-proximal peaks depicted in the K-means clustered heatmaps, indicating colocalization of Set1A and Mll2 on chromatin (Fig. 2, B and D). In addition, at annotated TSS sites, there is >98% overlap in Mll2 and Set1A binding regions (fig. S4A). This colocalization could be explained by the fact that both CXXC1, a key subunit in the Set1A/COMPASS complex as discussed earlier, and Mll2 harbor a CXXC motif that recognizes and binds to unmethylated CpG-containing DNA (34). It is therefore possible that Mll2 could compensate for Set1/COMPASS loss in H3K4me3 deposition. Consequently, we depleted Mll2 using a short hairpin RNA (shRNA) in Set1BKO-Set1A\textsuperscript{ASSET} lines and compared to WT ESCs (fig. S4B). Knockdown of Mll2 resulted in a marked decrease in H3K4me3 in Set1BKO-Set1A\textsuperscript{ASSET} ESCs (Fig. 4, A and B). To investigate this observation in further detail, we therefore generated a triple-mutant cell line, where we deleted the SET domain of Mll2 in Set1BKO-Set1A\textsuperscript{ASSET} ESCs by CRISPR-Cas9 (fig. S4, C and D). Simultaneously, we removed Mll1 in Set1BKO-Set1A\textsuperscript{ASSET} ESCs using previously reported guide RNAs (gRNAs) for targeting Mll1 (50), in the event that Mll1 may manifest a compensatory role in implementing H3K4me3 under conditions of Set1/COMPASS mutation (fig. S4, C and D). We were able to successfully retrieve homozygous triple-mutant ESCs harboring the intended Set1BKO-Set1A\textsuperscript{ASSET}, Mll2\textsuperscript{ASSET} and Set1BKO-Set1A\textsuperscript{ASSET}, Mll1KO mutations, as validated by PCR genotyping and RNA-seq (fig. S4, C and D). However, we noted that the Set1BKO-Set1A\textsuperscript{ASSET}, Mll2\textsuperscript{ASSET} cells proliferate more slowly than WT, Set1BKO-Set1A\textsuperscript{ASSET}, and Mll2\textsuperscript{ASSET} cells (data not shown). To ascertain if the appended deletion of either SET domain of Mll2 or of Mll affected H3K4 methylation, we performed Western blotting and ChIP-seq analyses to evaluate H3K4me3 levels. Consistent with the effects seen in knocking down Mll2 in the double-mutant ESCs (Fig. 4, A and B), bulk H3K4me3, including the level at TSS-proximal regions, is substantially lowered in the Set1BKO-Set1A\textsuperscript{ASSET}, Mll2\textsuperscript{ASSET} cells (fig. S4, E to G). However, no additive perturbation effect on H3K4me3 levels was observed in the Set1BKO-Set1A\textsuperscript{ASSET}, Mll1KO triple mutant relative to the double mutant (fig. S4, E to G), affirming the minimal contribution to H3K4me3 deposition by Mll1 in ESCs.

Extending our analyses by evaluating H3K4me3 changes in the previously defined clusters, we observe a robust and synergistic decrease in H3K4me3 occupancy in all three clusters in the Set1BKO-Set1A\textsuperscript{ASSET}, Mll2\textsuperscript{ASSET} triple-mutant cells (Fig. 4C), evincing that Mll2/COMPASS is functionally redundant to Set1/COMPASS in sustaining global H3K4me3 levels in ESCs. When we analyzed H3K4me3 changes at Set1- versus Mll2-dependent H3K4me3 regions in Set1BKO-Set1A\textsuperscript{ASSET}, Mll2\textsuperscript{ASSET}, and Set1BKO-Set1A\textsuperscript{ASSET}, Mll2\textsuperscript{ASSET} compared with WT, we noted the following: The significant decrease in H3K4me3 in the triple mutant compared with that in the double mutant at Set1-dependent H3K4 methylation is unquestionably greater than the significant reduction in H3K4me3 in the triple mutant compared with that in the Mll2\textsuperscript{ASSET} at Mll2-dependent sites (fig. S4, H and I). This supports the interpretation that while Mll2 is functionally redundant to Set1 at Set1-controlled regions, Set1 does not appear to reciprocate redundancy to Mll2 at Mll2-controlled sites, suggesting a unidirectional compensatory relationship between Mll2 and Set1. The added mutation of Mll2\textsuperscript{ASSET} in Set1BKO-Set1A\textsuperscript{ASSET} ESCs also further perturbed H3K4me3 breadth, such that the reduction in H3K4me3 density is seen at both genes with broad or narrow breadth, with a more radical decrease at broader peaks in the triple mutant (Fig. 4, D to G, and fig. S4J; also compare Fig. 3A to Fig. 4E). These data pinpoint the role of Mll2/COMPASS to be a key compensator for Set1/COMPASS function in bolstering overall H3K4me3 level and breadth in ESCs.

On the basis of global analysis, we noted an additional decrease in RNAP II at broader and narrower H3K4me3 peaks for the Set1BKO-Set1A\textsuperscript{ASSET}, Mll2\textsuperscript{ASSET} versus Set1BKO-Set1A\textsuperscript{ASSET} cells (Fig. 4H). The severe loss of H3K4me3 in the Set1BKO-Set1A\textsuperscript{ASSET}, Mll2\textsuperscript{ASSET} mutant at these regions correlates with a significant decrease in the transcription of cell identity genes compared with WT and Set1BKO-Set1A\textsuperscript{ASSET} cells and as a result may contribute to the proliferation defect in these triple-mutant cells. Thus, in retrospect, the less marked decrease in RNAP II levels in the double mutant at these broader peak regions depicted in Fig. 3E may indicate that the remaining H3K4me3 at these loci is adequate to maintain expression of such cell identity genes to the levels that cell proliferation is not impaired. Furthermore, we previously demonstrated that proper ESC differentiation requires the catalytic activity of Set1A (38), which is a key contributor to the formation of broad H3K4me3. Therefore, we deduce that loss of H3K4me3 at cell identity genes in Set1-mutant cells has a direct effect on cellular differentiation. Because self-renewal and differentiation are hallmarks of pluripotent stem cells, the loss of differentiation potential in Set1-mutant ESCs reflects the impairment of their identity. Together, these data signify that H3K4me3 levels at broad domains play an instructive role in maintaining ESC identity.

In summary, we present findings delineating the functionally redundant roles of the COMPASS family members in H3K4me3 implementation in mammalian stem cells. By investigating H3K4me3 enrichment and breadth in a series of ESC lines harboring combounding mutations of COMPASS enzymes relative to WT cells, we report that COMPASS members Set1A, Set1B, and Mll2 coordinate regulation of H3K4me3 level and peak breadth across the mammalian genome. We also establish that Mll2 plays an important compensatory role in sustaining global H3K4me3 level and breadth in the absence of Set1/COMPASS. This study enhances our current knowledge, shedding new light on the extraordinary ability of our cells to adapt to contextual changes for sustainment. Moreover, with numerous studies indicating the role of COMPASS in disease pathogenesis—Set1A/Set1B...
Fig. 4. Mll2 is functionally redundant to Set1/COMPASS in sustaining global H3K4me3 level and breadth. (A) Heatmaps of H3K4me3 occupancy in WT and Set1BKO-Set1A$^{LSET}$ ESCs upon control (shNT) and Mll2 (shMll2) knockdowns. Profiles are plotted in decreasing H3K4me3 peak width. Two replicates of each genotype were analyzed. (B) Log2FC heatmaps comparing differences in H3K4me3 levels in WT-shMll2, Set1BKO-Set1A$^{LSET}$-shNT, or Set1BKO-Set1A$^{LSET}$-shMll2 and WT-shNT conditions for the same ordered peaks as in (A). (C) Heatmaps of clustered H3K4me3 occupancy (left) and corresponding log2FC (right) comparing Set1BKO-Set1A$^{LSET}$-Mll2$^{(Sh)}$ relative to WT cells at the 13,486 H3K4me3 TSS-proximal peaks initially determined in Fig. 2B. (D) Track examples of broad (top) and narrow (bottom) H3K4me3 peaks in WT, Set1BKO-Set1A$^{LSET}$, Mll2$^{(Sh)}$, and Set1BKO-Set1A$^{LSET}$-Mll2$^{(Sh)}$ ESCs. (E) Heatmaps of H3K4me3 occupancy (left) and corresponding log2FC (right) at genes with the 500 most broad (top) and 500 most narrow (bottom) H3K4me3 peaks in Set1BKO-Set1A$^{LSET}$-Mll2$^{(Sh)}$ and WT ESCs. Broad and narrow peaks were initially determined in Fig. 3A. (F) Average log2FC H3K4me3 signal density plots (RPM) at genes with the 500 most broad (top) and 500 most narrow (bottom) H3K4me3 peaks in Set1BKO-Set1A$^{LSET}$, Mll2$^{(Sh)}$, and Set1BKO-Set1A$^{LSET}$-Mll2$^{(Sh)}$ cells. Normalized coverage, peak number per group, and P values (indicated by lower case letters) were calculated as in Fig. 3D. Calculated P values are as follows: a, c, d, e, f, and I, P < 6.173 × 10$^{-282}$; b, P = 0.01901; g, P = 6.1730 × 10$^{-282}$; h, P = 1.7914 × 10$^{-219}$; i, P = 5.0143 × 10$^{-54}$; j, P = 4.0960 × 10$^{-204}$; k, P = 9.0609 × 10$^{-204}$; l, P = 0.0017; m, P = 6.9636 × 10$^{-38}$; n, P = 0.4571.
in cancer (69, 70), Set1A in schizophrenia (71, 72), and Mll2 in childhood-onset dystonia (73, 74)—our work offers insight into discovering potential targets for future therapy against these relevant diseases through characterization and assessment of the broad/narrow H3K4me3 epigenetic signature.

**MATERIALS AND METHODS**

**Murine ESC culture, CRISPR-Cas9–mediated gene editing, and shRNA knockdown**

V6.5 ESCs were cultured in N2B27 media supplemented with MEK inhibitor and GSK3 inhibitor and LIF (2i/LIF) as previously described (38). Vectors expressing Cas9 and CRISPR single guide RNAs (sgRNAs) were generated by annealing and cloning oligos encoding the desired sgRNA sequences into the pX459 plasmid following a previously published protocol (75). Subsequent ESC transfection was performed as follows: 5 million ESCs were electroporated with the CRISPR sgRNA targeting vectors using Lonza’s Nucleofector2b (following the manufacturer’s instructions). One day after transfection, ESCs were selected with puromycin (2 μg/ml) (Life Technologies) for 2 days and grown in 2i/LIF ESC medium for 10 days until clone picking. gRNA sequences used in this study are listed as follows: Set1BKO, cctagatccgccgaaagcc (left) and ctacgccgctctctcctg (right); Mll2Δ, ggctcagaagttctcaag (left) and tgcccctggttgacagat (right); and Set1BKO-Set1AΔ, ggctcagaaggtctcaag (left) and gtaaggctgcaagtttg (right). gRNA sequences to generate Mll1 were previously reported (50). The following primers were used to genotype clones: Set1B—gaaaggttcacaaagcag (WT forward), acctgcttcagaaaaa (WT reverse), tcacggttaactcag (mutant forward), and cccagtcatctactatcagt (mutant reverse); Mll2—gaggtgcagggctcagcag (forward) and ggaatctgtgacacactac (reverse); Set1BKO-Set1A—, tcacaggttgtctcaagaa (forward) and cctagatccgccgaaagcc (reverse), with primers previously reported (38). Knockdown was performed as described in Cao et al. (43). The lentiviral construct against Mll2 was reported previously (25).

**Antibodies**

The following antibodies were generated in-house: anti-Set1B (47), anti-Mll2, anti-Set1A, anti-H3, anti-H3K4me1, anti-H3K4me2, and anti-H3K4me3. Other antibodies used in this study were H3 (CST #1B1B2), H3K4me3 (CST #9727), Rpb1 (CST #D8L4Y), HSP90 (Santa Cruz #7947), and Rbpb5 (Bethyl Laboratories #A300-109A). Set1A ChIP-seq data in Fig. 2B were from the Sze et al. (38).

**RNA-seq, ChIP-seq, and next-generation sequencing data processing**

RNA was extracted using the RNeasy mini kit following Qiagen’s instructions, and RNA-seq libraries were prepared using the TruSeq Stranded Total RNA Preparation Kit (Illumina). ChIP was performed as previously described (38, 76). ChIP-seq libraries were prepared using the KAPA HTT library preparation kit (Kapa Biosystems). Libraries were single read sequenced on the Illumina NextSeq 500 Sequencing System, and raw BCL output files were processes using bcl2fastq (Illumina, version 2.17.1.14) before quality trimming using Trimmomatic (77). Trimmed ChIP-seq and RNA-seq reads were aligned to the mouse genome [University of California at Santa Cruz (UCSC) mm9] using Bowtie v1.1.2 (78) and TopHat v2.1.0 (79), respectively. Only uniquely mapped reads satisfying the two-mismatch maximum threshold within the entire length of the gene were considered for ensuing analyses. Mapped ChIP-seq reads were extended to 150 base pairs toward the interior of the sequenced fragment, and raw read counts were normalized to total reads per million (RPKM). Output BAM files were converted into bigwig coverage plots to generate UCSC genome browser tracks. In the case of RNA-seq, exonic reads were assigned to genes from Ensembl release 67 using the python package HTSeq-0.6.0 (80). At least two biological replicates were performed for RNA-seq and ChIP-seq under each experimental condition.

**ChIP-seq and RNA-seq analyses**

For ChIP-seq, peak calling for H3K4me3 was performed using MACS v1.4.2, with default parameters. Heatmaps and metaplots were generated using deepTools 2.0 v3.1.1 (81). Occupancy levels in RPM were observed over peak regions as denoted and were either ranked by peak width or partitioned by K-means clustering. For the classification of broad and narrow H3K4me3 peaks in Figs. 3 and 4, we overlapped broad peaks called using SICER v1.1 (82) and narrow TSS-proximal peaks called using MACS, retaining the peaks called using SICER that overlapped with narrow TSS-proximal peaks. Peaks were mapped to a TSS using the GREAT algorithm (83), where TSS-proximal peaks were required to be located within 5 kb from the identified TSS. To determine the H3K4me3 levels for breadth analyses shown in Figs. 3D and 4(G and H), and figs. S3B and S4J, the BEDTools map function was used to quantify the coverage over peak regions, and custom R scripts were used for data plotting. The differential H3K4me3 peaks shown in figs. S3B and S4J for the respective mutant lines were defined as having a log2 fold change (log2FC) of H3K4me3 peak coverage of <−1 relative to WT peak coverage. To identify Set1-dependent H3K4me3 and Mll2-dependent H3K4me3 peaks in fig. S4 (H and I), H3K4me3 peaks were first called from at least two WT biological replicates using MACS, and overlapping regions determined using BEDTools intersect were chosen for subsequent differential binding analysis. To define sites with differential binding between WT and the indicated mutant line, BAM files of H3K4me3 ChIP-seq from WT and mutant samples were converted to a bed file using BEDTools v.2.29.1. BEDTools coverage was then used to quantify read counts within the initially identified overlapping H3K4me3 regions in WT and mutant samples. The edgeR v3.12.1 (84) package from Bioconductor, the quantified counts were normalized to counts per million, and differential peaks were then identified by fitting a negative binomial generalized log linear model to the normalized counts data. Statistically significantly decreased peaks in the mutant cell lines were identified using Benjamini-Hochberg adjusted P values <0.01 and log2FC <−1. Set1- or Mll2-dependent H3K4me3 peaks were determined by overlapping the respective decreased H3K4me3 regions with Set1A or Mll2 binding. For RNA-seq, for differential gene expression, gene count tables from HTSeq were used as input for edgeR v3.12.1 (84), and genes with Benjamini-Hochberg adjusted P values <0.01 were regarded as differentially expressed and used for downstream GO functional analysis with Metascape (85). Log2FC heatmaps of nearest gene expression relevant to the clustered peaks were determined using custom scripts and visualized using Java TreeView.

### SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/26/eaaz4764/DC1

---

**View/request a protocol for this paper from Bio-protocol.**
REFERENCES AND NOTES

1. R. Jaenisch, A. Bird, Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. Nat. Genet. 33 (Suppl.), 245–254 (2003).

2. E. I. Campos, D. Reinberg, Histone: Annotating chromatin. Annu. Rev. Genet. 43, 559–599 (2009).

3. A. J. Bannister, T. Kouzarides, Regulation of chromatin by histone modifications. Cell Res. 21, 381–391 (2011).

4. L. Wang, A. Shilatifard, UTX mutations in human cancer. Cancer Cell 35, 168–176 (2019).

5. J. E. Audia, R. M. Campbell, Histone modifications and cancer. Cold Spring Harb. Perspect. Biol. 8, a019521 (2016).

6. A. Piunti, A. Shilatifard, Epigenetic balance of gene expression by polycomb and COMPASS families. Science 352, and9780 (2016).

7. H. Santos-Rosa, R. Schneider, A. J. Bannister, J. Sherriff, B. E. Bernstein, N. C. Emre, S. L. Schreiber, J. Mellor, T. Kouzarides, Active genes are tri-methylated at K4 of histone H3. Nature 419, 407–411 (2002).

8. B. E. Bernstein, M. Kamal, K. Lindblad-Thor, S. Bekiranov, W. W. Pijnappel, M. Wilm, R. Aasland, A. F. Stewart, The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. EMBO J. 20, 7137–7148 (2001).

9. P. Wang, C. Lin, E. R. Smith, H. Guo, B. W. Sanderson, M. Wu, M. Gogol, T. Alexander, C. Seidel, L. M. Wiedemann, K. Ge, R. Krumlauf, A. Shilatifard, 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, 6074–6085 (2009).

10. M. Wu, P. F. Wang, J. S. Lee, S. Martin-Brown, L. Florens, M. Washburn, A. Shilatifard, Molecular regulation of H3K4 methylation by Wdr62, a component of human Set1/COMPASS. Mol. Cell. Biol. 28, 7337–7344 (2008).

11. D. Hu, X. Gao, M. A. Morgan, H. M. Herz, E. R. Smith, A. Shilatifard, The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. Mol. Cell 33, 4745–4754 (2013).

12. M. Mohan, H. M. Herz, E. R. Smith, Y. Zhang, J. Jackson, M. P. Washburn, L. Florens, J. C. Eisenberg, A. Shilatifard, The COMPASS family of H3K4 methylases in Drosophila. Mol. Cell. Biol. 31, 4310–4318 (2011).

13. C. C. Sze, A. Shilatifard, MLL3/MLL4/COMPASS family on epigenetic regulation of enhancer function and cancer. Cold Spring Harb. Perspect. Med. 6, a026427 (2016).

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

15. Z. Tang, W. Y. Chen, M. Shimada, U. T. Nguyen, J. Kim, X. J. Sun, T. Sengoku, R. K. McGinty, J. P. Fernandez, T. W. Muir, R. G. Roeder, SET1 and p300 act synergistically, through coupled histone modifications, in transcriptional activation by p53. Cell 154, 297–310 (2013).

16. D. Hu, X. Gao, K. Cao, M. A. Morgan, G. Mas, E. R. Smith, A. G. Volk, E. T. Bartom, J. D. crispino, I. Di croce, A. Shilatifard, Not all H3K4 methylations are created equal: MLL3/COMPASS dependency in primordial germ cell specification. Mol. Cell 65, 460–475, e6 (2017).

17. P. Ernst, M. Mabon, A. J. Davidson, L. I. Zon, S. J. Karsaner, My dMll-dependent hox program drives hematopoietic progenitor expansion. Curr. Biol. 14, 2063–2069 (2004).

18. R. Rickels, D. Hu, C. K. Collins, A. R. Woodfin, A. Piunti, M. Mohan, H. M. Herz, E. Kvon, A. Shilatifard, An evolutionary conserved epigenetic mark of polycluster response elements implemented by Trx/MLL/COMPASS. Mol. Cell 63, 318–328 (2016).

19. J. E. Lee, C. Wang, S. Xu, Y. W. Cho, L. Wang, X. Feng, A. Baldridge, V. Sartorelli, L. Zhuang, W. Peng, K. Ge, H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation, elife 2, e01993 (2013).

20. C. C. Sze, K. Cao, C. K. Collins, S. A. Marshall, E. J. Rendleman, P. A. Ozark, F. X. Chen, M. A. Morgan, L. Wang, A. Shilatifard, Histone H3K4 methylation-dependent and -independent functions of Set1A/COMPASS in embryonic stem cell self-renewal and differentiation. Genes Dev. 31, 1732–1737 (2017).

21. A. S. Bledau, K. Schmidt, K. Neumann, U. Hill, G. Scotta, A. Gupta, D. C. Torres, J. fu, A. Kranz, A. F. Stewart, K. Anastasiadis, The H3K4 methyltransferase Set1a is first required at the epiblast stage, whereas Set1b becomes essential after gastrulation. Development 141, 1022–1034 (2014).

22. L. Fang, J. Zhang, H. Zhang, X. Yang, X. Jin, L. Zhang, D. G. Skalnik, Y. Jin, Y. Zhang, X. Huang, J. Li, J. W. hong, H3K4 methyltransferase Set1a is a key Oct4 coactivator essential for generation of Oct4 positive inner cell mass. Stem Cells 34, 565–580 (2016).

23. R. Rickels, H. M. Herz, C. C. Sze, K. Cao, M. A. Morgan, C. K. Collins, M. Gause, Y. Y. Takahashi, L. Wang, E. J. Rendleman, S. A. Marshall, A. Krueger, E. T. Bartom, A. Piunti, E. R. Smith, N. A. Abshiu, N. L. Kelleher, D. Dorsett, A. Shilatifard, Histone H3K4 monomethylation catalyzed by Tr and mammalian COMPASS-like proteins at enhancers is dispensable for development and viability. Nat. Genet. 49, 1647–1653 (2017).

24. K. M. Dorighi, T. Swigut, T. Henriques, N. V. Bhanu, B. S. Scruuggs, N. Nady, C. D. Still II, B. A. Garcia, K. Adamel, J. Wysocka, Mll3 and Mll4 facilitate enhancer RNA synthesis and transcription from promoters independently of H3K4 monomethylation. Mol. Cell 66, 568–576.e4 (2017).

25. K. Cao, C. K. Collins, M. A. Morgan, S. A. Marshall, E. J. Rendleman, P. A. Ozark, E. R. Smith, A. Shilatifard, An Mll4/COMPASS-Lsd1 epigenetic axis governs enhancer function and pluripotency transition in embryonic stem cells. Sci. Adv. 4, eaap8747 (2018).

26. B. D. Yu, A. L. Hess, S. E. Horning, G. A. Brown, S. J. Karsmeyer, Altered Hox expression and segmentation identity in Mll mutant mice. Nature 378, 505–508 (1995).

27. Y. Hagi, K. Deguchi, A. Aono, Y. Tani, T. kishimoto, T. kornori, Growth disturbance in fetal liver hematopoiesis of Mll mutant mice. Blood 92, 108–117 (1998).

28. Sze et al., Sci. Adv. 2020; 6: eaaz4764 24 June 2020 9 of 11
D. A. Brown, V. Di Cerbo, A. Feldmann, J. Ahn, S. Ito, N. P. Blackledge, M. Nakayama, T. Clouaire, S. Webb, P. Skene, R. Illingworth, A. Kerr, R. Andrews, J. H. Lee, D. Skalnik, K. Cao, C. K. Collings, S. A. Marshall, M. A. Morgan, E. J. Rendleman, L. Wang, C. C. Sze, J. A. Dahl, C. Wang, W. Liu, J. Li, J. Li, J. Chen, Y. Zhao, H. Gao, H. Wang, Y. Zhang, C. W. Hanna, A. Taudt, J. Huang, L. Gahurova, A. Kranz, S. Andrews, W. Dean, A. F. Stewart, L. M. Soares, P. C. He, Y. Chun, H. Suh, T. Kim, S. Buratowski, Determinants of Histone H3K4 methylation in oocytes.

M. M. Matzuk, MLL2 is required in oocytes for bulk histone 3 lysine 4 trimethylation on bivalent promoters during early mammalian development. Nature 537, 553–557 (2016).

S. Denissenov, H. Hofemeister, M. Marks, A. Kranz, G. Ciotta, S. Singh, K. Anastassiadis, H. G. Stunnenberg, A. F. Stewart, Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, where MLL1 is redundant. Development 141, 526–537 (2014).

G. Mas, E. Blanco, C. Ballaré, M. Sansó, Y. G. Spill, D. Hu, Y. Aoi, F. Dílyi, A. Shilatifard, M. A. Martí-Renom, L. Di Croce, Promoter bivalency favors an open chromatin architecture in embryonic stem cells. Nat. Genet. 50, 1452–1462 (2018).

C. V. Andreu-Vieyra, R. Chen, J. E. Agno, S. Glaser, K. Anastassiadis, A. F. Stewart, M. M. Matzuk, MLL2 is required in oocytes for bulk histone 3 lysine 4 trimethylation and transcriptional silencing. PLOS Biol. 8, e1000453 (2010).

C. W. Hanna, A. Taudt, J. Huang, L. Gahurova, A. Kranz, S. Andrews, W. Dean, A. F. Stewart, M. Colomé-Tatché, G. Kelsey, MLL2 conveys transcription-independent H3K4 trimethylation in oocytes. Nat. Struct. Mol. Biol. 25, 73–82 (2018).

N. J. Krogan, J. Dover, A. Wood, J. Schneider, J. Heidt, M. A. Boontang, K. Engdahl, O. W. Ryan, A. Golshani, M. Johnstone, F. Greenblatt, A. Shilatifard, The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: Linking transcriptional elongation to histone methylation. Mol. Cell. 71, 721–729 (2003).

J. H. Lee, D. G. Skalnik, Wdr82 is a C-terminal domain-binding protein that recruits the Set1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. Mol. Cell. Biol. 28, 609–618 (2008).

H. H. Ng, F. Robert, R. A. Young, K. Struhl, Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol. Cell 11, 709–719 (2003).

L. M. Soares, P. C. He, Y. Chun, H. Suh, T. Kim, S. Buratowski, Determinants of Histone H3K4 methylation complex and its functional role in MLL-AF9 leukemia. Nat. Struct. Mol. Biol. 17, 62–68 (2010).

A. Yokoyama, M. L. Cleary, Menin critically links MLL proteins with LEDGF on cancer-susceptible pathways. Nat. Rev. Cancer 19, W160–W165 (2019).

C. Ang, D. E. Schones, C. Zeng, K. Cui, K. Zhao, W. Peng, A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Nat. Biotechnol. 28, 729–748 (2010).

A. M. Bolger, M. Lohse, B. Usadel, Trimomatic: A flexible trimmer for illumina sequencing data. Bioinformatics 30, 2114–2120 (2014).

B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25 (2009).

C. Trapnell, L. Pachter, S. L. Salzberg, TopHat: Discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105–1111 (2009).

S. Anders, P. T. Pyl, W. Huber, HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169 (2015).

F. Ramirez, D. P. Ryan, B. Grünving, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dündar, S. Anders, P. T. Pyl, W. Huber, HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169 (2015).

N. Tripathi, M. O. Pohlf, Y. Zhou, A. Rodriguez-Frandsen, G. W. D. Steen, H. M. Moulton, P. De Jesus, J. Che, L. C. Muller, E. Yáñez, D. Andemmen, L. Pache, M. Bannasch, R. A. Albrecht, M. G. Gonzalez, Q. Nguyen, A. Brass, S. Elledge, M. White, S. Shapira, N. Hacohen, A. Karlas, T. F. Meyer, M. Shales, A. Gatorano, J. R. Johnson, G. J. Johnson, The Human CXXC domain database. Nat. Rev. Genet. 20, 95–101 (2019).

A. García-Sastre, S. K. Chanda, Meta- and orthogonal integration of influenza “OMICs” data defines a role for UBR4 in virus budding. Cell Host Microbe 18, 723–735 (2015).
Acknowledgments: We express gratitude to members of the Shilatifard laboratory for constructive criticism toward this manuscript. We thank E. R. Smith for the critical reading and editing of this manuscript. We thank M. Iwanaszko for insightful feedback on statistical tests and bioinformatics analyses in this manuscript. We are thankful to K. Chen of The Methodist Hospital Research Institute in Houston, Texas, for instructive advice on performing breadth analyses. We also thank B. K. Cenik for inspiring scientific quip.

Funding: C.C.S. was supported, in part, by NIH/NCI Kirschstein-NRSA F31CA228149 and is currently supported, in part, by NIH/NCI Predoctoral to Postdoctoral Transition Award F99CA234945. K.C. is supported, in part, by NIH/NICHD Pathway to Independence Award K99HD094906. Studies in the Shilatifard laboratory regarding the role of the COMPASS family in development and cancer are supported by the NCI’s Outstanding Investigator Award R35CA197569.

Author contributions: C.C.S. and A.S. conceived and initiated the project. C.C.S., M.U., L.W., C.A.R., S.A.M., E.J.R., D.Z., and D.D. performed the experiments. C.C.S., P.A.O., K.C., F.X.C., and S.D. analyzed the data. C.C.S. and A.S. wrote the manuscript.

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Next-generation sequencing data have been deposited at Gene Expression Omnibus database under accession number GSE152595. Additional data related to this paper may be requested from the authors.

Submitted 12 September 2019
Accepted 8 May 2020
Published 24 June 2020
10.1126/sciadv.aaz4764

Citation: C. C. Sze, P. A. Ozark, K. Cao, M. Ugarenko, S. Das, L. Wang, S. A. Marshall, E. J. Rendleman, C. A. Ryan, D. Zha, D. Douillet, F. X. Chen, A. Shilatifard, Coordinated regulation of cellular identity–associated H3K4me3 breadth by the COMPASS family. Sci. Adv. 6, eaaz4764 (2020).