The tet methylcytosine dioxygenase 2 (TET2) enzyme catalyzes the conversion of the modified DNA base 5-methylcytosine to 5-hydroxymethylcytosine. TET2 is frequently mutated or dysregulated in multiple human cancers, and loss of TET2 is associated with changes in DNA methylation patterns. Here, using newly developed TET2-specific antibodies and the estrogen response as a model system for studying the regulation of gene expression, we demonstrate that endogenous TET2 occupies active enhancers and facilitates the proper recruitment of estrogen receptor α (ERα). Knockout of TET2 by CRISPR-CAS9 leads to a global increase of DNA methylation at enhancers, resulting in attenuation of the estrogen response. We further identified a positive feedback loop between TET2 and ERα, which further requires MLL3 COMPASS at these enhancers. Together, this study reveals an epigenetic axis coordinating a transcriptional program through enhancer activation via DNA demethylation.

INTRODUCTION

The epigenetic pathways that regulate the methylation of cytosine bases in DNA, modifications of histone amino acids, and positioning of nucleosomes are commonly found to be dysregulated in human diseases, especially in cancer. Failure to maintain these epigenetic marks can result in changes in the expression pattern of oncogenes and tumor suppressors, which leads to the development and progression of cancer (1–3). The tet methylcytosine dioxygenase enzymes (TET1, TET2, and TET3) were demonstrated to oxidize 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) (4, 5). The 5hmC and its oxidized derivatives are subsequently replaced with an unmodified cytosine by base excision repair to achieve demethylation (6). Different from the TET1 and TET3 genes, the human TET2 gene lacks the CpG DNA binding CXXC domain. The TET2 gene is the product of a gene fission event during vertebrate evolution, whereby a separate protein is encoded by the new CXXC4 gene, with the CXXC4 protein regulating the stability of the catalytic domain—containing TET2 protein (7). In human hematopoietic cells, loss of TET2 can lead to DNA hypermethylation at up to 25% of active transcriptional enhancer elements (8). In mouse embryonic stem cells, depletion of TET2 causes extensive loss of 5hmC at enhancers, leading to enhancer hypermethylation and reduction of enhancer activity (9). These studies suggested that TET2 plays a nonredundant, critical role in maintaining enhancer demethylation and activity. However, because of a lack of commercially available TET2 antibodies suitable for chromatim immunoprecipitation sequencing (ChIP-seq), whether TET2 directly or indirectly regulates enhancer DNA methylation and how TET2 is recruited to specific loci remain under debate (8, 10).

In tumors, aberrant DNA methylation can lead to silencing of tumor suppressor genes and thus can contribute to malignancy (11). Decreased expression of the TET genes and the resulting loss of 5hmC have been found in multiple human cancers, including leukemia, breast cancer, colorectal cancer, melanoma, and lung cancer (12–14), suggesting that this gene family has a critical role in regulating DNA methylation for the maintenance of normal cellular function (15, 16).

The MLL3 and MLL4 branches of the COMPASS family of histone H3K4 methyltransferases have been shown to be epigenetic regulators of enhancer function (17, 18). Mutation or loss of expression of MLL3/4 COMPASS has been identified in multiple human cancers (19–22). Genome-wide studies revealed that MLL3 COMPASS implements monomethylation of histone H3 lysine 4 (H3K4me1) at enhancers and positively regulates tumor suppressor gene expression (23, 24). In our current study, we report that TET2 binds to active enhancers and can facilitate the proper recruitment and function of transcription factors, as exemplified by estrogen receptor α (ERα) in response to estrogen signaling. We further report a positive feedback between TET2 and ERα, which is directly mediated by MLL3 activity at enhancers, that underlies the proper regulation of ERα gene expression.

RESULTS

TET2 is a transcriptional coactivator in the estrogen response

Mutation or depletion of the TET methylcytosine dioxygenase enzymes (TET1, TET2, and TET3) was previously demonstrated to be pervasive in tumors and proposed to be a cancer hallmark (13). To begin to investigate the role of TET proteins in human breast cancer, we first determined the correlation between the mRNA level of TET1 to TET3 and breast cancer patient survival. We found a significant correlation between TET2 expression and patient survival, with lower levels of TET2 predicting shortened patient life span (fig. S1A).

To assess the function of TET2 in breast cancer cell lines, we used CRISPR-CAS9 technology to knock out the TET2 gene in both ERα− and ERα+ cell lines. Exon 3, which contains the start codon for TET2, was chosen as the target for CRISPR (fig. S1B). Polymerase chain reaction (PCR) amplicons within and outside exon 3 were used to confirm successful knockout (KO) of TET2 in both MCF7 (ERα−) and CAL51 (ERα+) cell lines (fig. S1, B and C). Western
blotting (Fig. 1A) and RNA sequencing (RNA-seq; fig. S1D) were further used to ensure that no protein product of the TET2 gene was made in the KO cells, while ERα protein levels were not altered by TET2-KO (Fig. 1A).

We found that loss of TET2 does not significantly affect ERα− breast cancer cell growth (Fig. 1B, left) or alter cell morphology (fig. S1E); however, loss of TET2 in MCF7 (ERα+) cells results in a significant growth defect in regular medium (Fig. 1B, right). Accordingly, cell morphology of MCF7-TET2-KO cells is distinct from that of parental cells (Fig. 1C). On the basis of this observation, we hypothesized that the estrogen response and estrogen-dependent growth of MCF7 cells may be affected by TET2 depletion. To determine the sensitivity of MCF7-WT (wild-type) and MCF7-TET2-KO cells to estrogen stimulation, we grew cells in phenol red−free medium supplemented with charcoal-stripped FBS for 4 days before treatment with DMSO or 10 nM 17β-estradiol (E2). TET2-KO cells exhibited estrogen-dependent

**Fig. 1. TET2 is a coactivator of ERα.** (A) TET2 and ERα protein levels were determined by Western blotting in TET2-WT and TET2-KO CAL51 cells (ERα−) and MCF7 cells (ERα+). HSP90 was used as an internal control (n = 3). (B) Cells (2 × 10⁴) of each cell line were seeded in six-well plates, and the cell growth ability of TET2-WT and TET2-KO cells was determined by cell counting at the indicated days. Data are means ± SD; n = 3 independent experiments. **P < 0.01, two-tailed unpaired Student's t test. (C) Representative images of the cell morphology of MCF7-TET2-WT and MCF7-TET2-KO clones (n = 3). (D) MCF7-TET2-WT and MCF7-TET2-KO cells were maintained in phenol red−free medium containing 5% charcoal-stripped fetal bovine serum (FBS) for 4 days. Subsequently, 5 × 10³ cells were seeded in six-well plates in the presence of dimethyl sulfoxide (DMSO) or E2 at 10 nM, and cell growth was determined by cell counting at the indicated days. Data are means ± SD; n = 3 independent experiments. **P < 0.01, two-tailed unpaired Student's t test. (E) Schematic of the RNA-seq experimental workflow using MCF7-TET2-WT or MCF7-TET2-KO cells treated with either DMSO or E2 (10 nM) for 4 hours (n = 2). (F and G) Venn diagrams (F) and heat maps (G) showing the overlap of genes induced by E2 between TET2-WT and TET2-KO cells. Log2FC, log2 fold change. (H) Representative RNA-seq tracks of genes differentially induced by E2 in TET2-WT and TET2-KO cells (n = 2).
cell growth defects (Fig. 1D). To determine whether the estrogen-controlled gene expression profile is also altered in TET2-depleted cells, we performed RNA-seq with parental MCF7 cells and TET2-KO MCF7 cells treated with either DMSO or E₂ (10 nM) for 4 hours (Fig. 1E). In parental MCF7 cells, treatment with E₂ resulted in an increased expression of 803 genes and a decreased expression of 205 genes, while in TET2-KO cells, 481 genes are induced and 222 genes are suppressed by E₂ treatment (Fig. 1F). In general, TET2 loss was associated with an impaired estrogen response, as seen by gene expression analysis (Fig. 1, G and H).

**TET2 mediates proper recruitment of ERα to enhancers**

To determine how loss of TET2 affects E₂-dependent gene expression, we performed ChIP-seq of ERα to identify sites where the recruitment of ERα is altered in response to E₂ treatment. Consistent with previous studies, we found that most ERα was recruited to non-TSS (transcription start site) regions after E₂ treatment (Fig. 2A (right) and B (left), and fig. S2B). E₂-dependent induction of genes nearest to these non-TSS ERα peaks is also attenuated in TET2-KO cells (Fig. 2B, right). These data suggested that loss of TET2 affects proper recruitment of ERα to active enhancers and thereby attenuates the estrogen response.

Because there are no reliable commercially available TET2-specific ChIP-seq grade antibodies, and to determine whether TET2 directly binds to ERα active enhancers, we generated polyclonal antibodies toward the N terminus of human TET2 protein (fig. S2C). To validate the specificity of our antibodies, we performed Western blotting and ChIP-seq in nontargeting and TET2 short hairpin RNA (shRNA) cells and observed decreased signal by both assays (fig. S2, D and E). We further validated our antibodies in TET2-WT and TET2-KO MCF7 cells and found that the TET2 peaks detected by the antibodies are reduced in TET2-KO cells (fig. S2F).

To determine whether TET2 binds to ERα-bound active enhancers, we sorted TET2 peaks according to ERα-clustered peaks in Fig. 2A and found enrichment for TET2 in cluster 1 with the active enhancer peaks (Fig. 2C). Furthermore, loss of TET2 impairs the proper and full recruitment of ERα at TET2-bound active enhancers (Fig. 2D). In these cells, there are 773 ERα active enhancer peaks that are co-occupied by TET2 (Fig. 2E), and pathway analysis by GREAT (Genomic Regions Enrichment of Annotations Tool) analysis demonstrated that these loci are associated with genes for hormone stimulus (fig. S2G). We further compared the occupancy of H3K4me1, H3K4me3, H3K27me3, and H3K27Ac at the 773 ERα/TET2 co-bound peaks with 1869 ERα-only peaks. We found that H3K4me1 and H3K27Ac levels are significantly higher at TET2-ERα co-bound loci, we performed modified reduced representation bisulfite sequencing (mRRBS) with TET2-WT and TET2-KO MCF7 cells. We found that loss of TET2 does not lead to increased DNA methylation at ERα-bound loci, performed modified reduced representation bisulfite sequencing (mRRBS) with TET2-WT and TET2-KO MCF7 cells. We found that loss of TET2 does not lead to increased DNA methylation globally (fig. S2H), active promoters (fig. S1, I and J), or at ERα-only-bound TSS peaks (Fig. 2G). However, as can be seen in two different TET2-KO clones, loss of TET2 leads to an increase of CpG methylation at the ERα–non-TSS peaks, especially at cluster 1 loci (Fig. 2, H to J, and fig. S2K).

**Positive feedback loop between TET2 and ERα regulates the estrogen response**

Our RNA-seq analysis identified TET2 as an E₂-inducible gene (Fig. 3A). Real-time PCR confirms that TET2, but not TET1 or TET3, is induced by E₂ treatment (Fig. 3B). Accordingly, Western blotting of TET2 protein levels at different time points of E₂ treatment demonstrated a steady increase in TET2 protein levels over a 24-hour period (Fig. 3C). These data suggested that there is a positive feedback loop between TET2 and ERα. To understand the role of ERα in regulating TET2 expression, we performed H3K4me3, H3K27ac, and H3K4me1 ChIP-seq in cells treated with either DMSO or E₂ (100 nM for 45 min). We found that ERα was recruited to three putative enhancer elements (E1, E2, and E3) upstream of the TET2 gene in an E₂-dependent manner (Fig. 3D).

To further test whether TET2 is a direct target of ERα, we treated MCF7 cells with the ER antagonist tamoxifen for 72 hours. Tamoxifen treatment led to significant decreases of TET2 at both the mRNA and protein levels (Fig. 3, E and F). We next tested whether tamoxifen treatment affected chromatin occupancy of TET2. We found that treatment of tamoxifen broadly led to reduced TET2 chromatin occupancy (Fig. 3, G and H). Together, these results demonstrate that TET2 is a direct target of E₂, and that the TET2-ERα positive feedback loop is disrupted by tamoxifen treatment, raising the possibility that disruption of the TET2-ERα axis may be involved in the development of resistance to endocrine therapy.

**Loss of MLL3, but not MLL4, COMPASS disrupts the TET2-ERα axis by reducing TET2 expression**

Because MLL3 and MLL4 members of the COMPASS family are major regulators of enhancer activity in mammals (26, 27) (Fig. 4A), we hypothesized that MLL3 and MLL4 COMPASS may be involved in the E₂-regulated TET2 expression at enhancers. To test our model, we knocked down MLL3 and MLL4 levels with two different shRNAs and performed real-time PCR and Western blotting to determine the TET2 expression in MCF7 cells. Depleting cells of MLL3, but not MLL4, led to decreased levels of TET2 RNA and protein (Fig. 4, B and C). To determine whether MLL3 COMPASS chromatin occupancy is altered upon E₂ treatment, we performed MLL3, H3K4me1, and H3K27ac ChIP-seq in MCF7 cells treated with DMSO or E₂. E₂ treatment led to increased recruitment of MLL3, H3K4me1, and H3K27ac levels at the cluster 1 active enhancers (Fig. 4D). When MCF7 cells were cultured in regular serum-containing medium, MLL3 COMPASS was found to be highly enriched at Enhancer 2 of the TET2 gene (Fig. 4E). In addition, induction of the estrogen response with E₂ led to increased occupancy of MLL3 to Enhancers 1 and 2 of the TET2 gene (Fig. 4F). Similar to TET2 depletion, depletion of MLL3 led to a global decrease in the E₂ response in MCF7 cells (fig. S3, A to D). Consistently, induction of TET2 expression, but not TET1 or TET3 expression, by E₂ was attenuated in MLL3-depleted cells (Fig. 4G and fig. S3B). In contrast, depletion of MLL3 does not affect TET2 gene expression in ER− cells (fig. S4A). Consistent with the RNA-seq data, MLL3 is not enriched at the TET2 enhancers in ER− cells, as seen by ChIP-seq (fig. S4B). Moreover, consistent with our cell line–based studies implicating MLL3 in the TET2-ERα axis, examination of data from human breast cancer patient samples finds a greater correlation between MLL3 and TET2 in ERα− than in ERα+, tumors (fig. S4C).
Fig. 2. Enhancer TET2 mediates proper recruitment of ERα. (A) TET2-WT and TET2-KO MCF7 cells were maintained in phenol red–free medium containing 5% charcoal-stripped FBS for 4 days, followed by treatment with either DMSO or E2 at 100 nM for 45 min. Heat maps generated from ChIP-seq data showing the occupancy of ERα in DMSO- and E2-treated cells. All rows are centered on ERα peaks and further divided into TSS and non-TSS regions. TSS and non-TSS regions were further divided into three clusters each by k-means (n = 2). (B) Left: Log fold change heat map shows the comparison of ERα occupancy between TET2-WT and TET2-KO cells treated with DMSO versus E2. Right: Log (fold change) of nearby gene expression in Tet2-WT or Tet2-KO cells treated with either DMSO or E2 (n = 2). (C) Heat maps generated from ChIP-seq data showing the occupancy of TET2, which is centered on the ERα peaks, and rows ordered as in (A) (n = 2). (D) Representative genome browser tracks of TET2 and ERα occupancy at enhancers. (E) Venn diagram showing common peaks between total TET2 peaks and ERα cluster 1 peaks. (F) Box plot quantifying changes on H3K4me3, H3K4me1, H3K27me3, and H3K27Ac occupancy at ERα alone and ERα/TET2 co-occupied cluster 1 peaks. (G to J) Well-observed CpG methylation around ERα-binding sites overlapping TSS sites (G) and non-TSS regions (H to J) separated by clusters identified in (A). Average CpG methylation values ± SEM for two biological replicates per cell type are plotted for the center of ERα-binding sites ± 2500 base pairs (bp). NTD, N-terminal domain.
DISCUSSION

Here, through the generation of TET2-specific antibodies, we have defined a role for TET2 as a transcriptional coactivator functioning through epigenetic regulation of gene expression by maintaining unmethylated DNA at enhancers. We have further identified a transcriptional feedback loop between TET2 and estrogen signaling, whereby the TET2 gene is a direct transcriptional target of ERα, and the TET2 protein serves as a coactivator for ERα. TET2 enhances ERα occupancy at enhancers by maintaining low levels of CpG methylation. We further identified putative cell type–specific enhancers for TET2 that are occupied by ERα, and these and other TET2-ERα enhancers are regulated by the MLL3, but not the MLL4, branch of the COMPASS family.
Although there are several studies that provided evidence of TET2 chromatin binding (28–30), the genome-wide binding pattern of TET2 remains under debate. Although it had been reported that TET2 mainly binds to, and functions at, promoter regions, it was demonstrated by multiple groups that loss of TET2 primarily affects 5hmC level at active enhancers in the same cell lines (9). These contradictory results led us to generate our own TET2-specific antibodies, which we validated by ChIP-seq in TET2 knockdown

Fig. 4. Loss of MLL3, but not MLL4, COMPASS disrupts the TET2-ERα axis. (A) Cartoon of MLL3 and MLL4 COMPASS regulation of gene expression from enhancers. (B) Level of expression of MLL3, MLL4, and TET2, as assessed by real-time PCR in MCF7 cells treated with shNONT, shMLL3, and shMLL4. Data are means ± SD; n = 3 independent experiments. **P < 0.01, two-tailed unpaired Student’s t test. (C) Protein levels of MLL3, MLL4, and TET2 as determined by Western blotting in MCF7 cells treated with shNONT, shMLL3, and shMLL4 (n = 3). (D) Heat maps of MLL3, H3K4me1, and H3K27ac log2 fold changes in response to E2 treatment. Rows are centered on the non-TSS ERα peaks and ordered as in Fig. 2A (n = 2). (E) Representative genome browser tracks showing the occupancy of MLL3 at TET2 enhancers (n = 2). (F) Representative tracks showing the recruitment of MLL3 and increased occupancy of H3K4me1 at TET2 enhancers induced by E2 treatment (n = 2). (G) Representative RNA-seq tracks showing the expression of TET2 induced by E2 in shNONT and shMLL3 cells (n = 2).
and KO cells. In our current studies, although we also observed that a portion of TET2 peaks are localized to TSS region, we did not see an obvious change of DNA methylation pattern at TSS regions after TET2-KO. In agreement with previous studies (8, 9), we found that loss of TET2 altered the methylation level of DNA at enhancers, which we found were directly bound by TET2. Consistent with our findings, a most recent study shows that TET2 binds to enhancers and facilitates transcription factor recruitment in hematopoietic cells (31). We cannot rule out at this point whether TET2 may function redundantly at promoter regions with other TET enzymes or may have catalytic-independent activities at these regions (32).

The function of ER in breast cancer has been well studied; however, the regulation of ERα recruitment remains unclear. Recently, emerging studies revealed that DNA methylation at enhancers may affect ER binding (33) to these regions and participate in ERα breast cancer resistance to anti-estrogen treatment (34). These studies strongly implied that there might be unknown enhancer binding factors that maintain enhancer DNA methylation status in breast cancer pathogenesis (35). In our current study, we demonstrated that loss of TET2 results in increased enhancer DNA methylation, which is accompanied by decreased ERα recruitment. Notably, we found that cells lacking TET2 have limited E2-dependent growth, indicating that loss of TET2 may be involved in the development of resistance to endocrine therapy in breast cancer.

MLL3 and MLL4 COMPASS are enhancer binding factors that are responsible for enhancer activation and downstream gene expression (17, 23). MLL3 and MLL4 are among the most mutated histone modifiers in multiple human cancers (22). Our genome-wide studies demonstrated that the loss of MLL3, but not MLL4, attenuates TET2 induction by E2. A recent study found that loss of MLL3 was found to directly promote hormone-independent outgrowth (36), which strongly supports our model because loss of MLL3 also leads to decreased TET2 expression, thereby disrupting the TET2-ERα epigenetic axis.

In human prostate cancer, TET2 was found to be repressed by androgen signaling. The androgen receptor (AR) induces the expression of mIR-29, which directly binds to the 3′ untranslated region of the TET2 gene, and further regulates the stability of TET2 mRNA. Knocking down TET2 in prostate cancer stimulates the gene expression response by androgen, suggesting that there is a negative feedback between TET2 and AR signaling (29). In contrast, our study reveals a positive feedback between TET2 and ERα signaling, which is further facilitated by MLL3 COMPASS (fig. S4D). Mutations or dysregulations of MLL3 or TET2, which is a common feature in numerous human cancers, may disrupt the transcriptional axis, leading to malignant progression of cancer or resistance to endocrine therapy. Our study also suggested distinct regulation of TET2 in male and female cancers, which could provide new transcriptional targets for endocrine therapy.

In conclusion, our study shows that TET2 specifically functions at enhancers by demethylating these loci and preparing them for recruitment of transcription factors. Furthermore, we have found that MLL3 COMPASS is required for this process through direct regulation of TET2 expression in an ERα-dependent manner. These findings reveal the existence of an epigenetic axis coordinating a transcriptional program through enhancer activity requiring DNA demethylation.

**Materials and Methods**

**Antibodies**

ERα (sc-543) and HSP90 (sc-7947) were purchased from Santa Cruz Biotechnology, and TET2 [A304-247A, for Western blotting and immunoprecipitation (IP)] was purchased from Bethyl Laboratories. TET2 (#18950, for ChIP-seq) and H3K27ac (#8173) were purchased from Cell Signaling Technology. MLL3, MLL4, H3K4me1, and H3K4me3 antibodies were made in-house, as described before (24). Rabbit anti-TET2 (for ChIP-seq and Western blotting) was generated against the N-terminal peptide of TET2 at Pocono Rabbit Farm and Laboratory.

**Cell lines and RNA interference**

MCF7 cells were obtained from the American Type Culture Collection, and CAL51 cells were obtained from Leibniz Institute DSMZ (German Collection of Microorganisms and Cell Cultures). All these cells were maintained with Dulbecco’s modified Eagle’s medium (DMEM) (Gibco, Gaithersburg, MD) containing 10% FBS (Sigma). For E2 induction, the MCF7 cells were maintained in phenol red–free DMEM (Gibco, Gaithersburg, MD) containing 5% charcoal-stripped FBS (Sigma). For shTET2, shMLL3, and shMLL4 infection, cells were infected with lentivirus containing shRNAs in the presence of polybrene (4 μg/ml; Sigma) for 24 hours in DMEM supplemented with 10% FBS. The infected cells were selected with puromycin (2 μg/ml) for an extra 48 hours before harvest.

**CRISPR-mediated KOs**

Single-guide RNAs were designed with CRISPRtool (http://crispr.mit.edu) and then cloned into lentCRISPR v2 (Addgene, 52961) vector. Targeting vector and single-stranded DNA donor were cotransfected in cells for 24 hours and followed by 2 days of puromycin selection. Targeted single-cell clones were screened by PCR.

**Next-generation sequencing sample preparation**

ChIP-seq libraries were prepared using the KAPA HTP Library Preparation Kit and multiplexed with NEXTflex DNA Barcodes from Bioo Scientific. DNA (10 ng) was used as starting material for input and IP samples. Libraries were amplified using 13 cycles on the thermocycler. Post-amplification libraries were size selected at 250 to 450 bp in length using Agencourt AMPure XP beads from Beckman Coulter. Libraries were validated using the Agilent High Sensitivity DNA Kit. RNA-seq libraries were prepared using the Illumina TruSeq Stranded Total RNA Preparation Kit with Ribo-Depletion. Input RNA quality was validated using the Agilent RNA 6000 Nano Kit. Total RNA (1 μg) was used as starting material. Libraries were validated using the Agilent DNA 1000 Kit.

**RNA-seq analysis**

Gene counts were computed by HTSeq (37) and used as input for edgeR 3.0.8 (38). Genes with Benjamini–Hochberg–adjusted P values less than 0.01 were considered to be differentially expressed, unless otherwise specified. RNA-seq heat maps adjacent to ChIP-seq heat maps display log2 fold change values of genes corresponding to TSSs nearest to ChIP-seq peaks and were displayed using Java TreeView (39). Gene Ontology functional analysis was carried out using Metascape with default parameters (40).

**ChIP-seq analysis**

ChIP-seq was performed as previously described (41). For ChIP-seq analysis, TET2 peaks in both TET2-WT and TET2-KO or shNONT and shTET2 conditions were called with the MACS v1.4.2 software (42) using default parameters and corresponding input samples. Metaplots and heat maps were generated using ngseqplot (43). The
**REFERENCES AND NOTES**


Acknowledgments: We thank F. Zheng for the gifts of the P330 and lentiCRISPR v2 vectors. Funding: This work was supported by the Training Program in Signal Transduction and Cancer (grant T32 CA070085 to L.W.); NIH/NCI training grant T32CA070085, Alex’s Lemonade Stand Foundation, and Northwestern Mutual (to Z.Z.); NIH grant R50CA11428 (to E.R.S.); and NIH grant K08HL128867 and the Francis Family Foundation’s Parker B. Francis Research Opportunity Award (to B.D.S.). Studies in the Shilatifard laboratory related to COMPASS are supported by the NCI’s Outstanding Investigator Award R35CA197569 to A.S. Author contributions: L.W. and A.S. designed the study. L.W., Z.Z., C.R., and A.L.W. performed most of the experiments and part of the analyses. L.W., E.R.S., M.A.M., and A.S. wrote and edited the manuscript. S.M.A. and E.J.R. generated and sequenced the next-generation sequencing (NGS) libraries. P.A.D. performed all of the RNA-seq and ChiP-seq bioinformatic analyses. L.Z. performed the clinical data analysis. A.P., K.A.H., and B.D.S. performed the DNA methylation analysis. B.D.S. analyzed the DNA methylation data. L.W., E.R.S., and A.S. revised 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. Additional data related to this paper may be requested from the authors. Raw and processed NGS data will be deposited in the National Center for Biotechnology Information GEO database under accession number GSE120756. Submitted 6 July 2018 Accepted 11 October 2018 Published 7 November 2018 10.1126/sciadv.aau6986