DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

Locus-specific editing of histone modifications at endogenous enhancers using programmable TALE-LSD1 fusions

The Harvard community has made this article openly available.
Please share how this access benefits you. Your story matters.

Citation	Mendenhall, Eric M., Kaylyn E. Williamson, Deepak Reyon, James Y. Zou, Oren Ram, J. Keith Joung, and Bradley E. Bernstein. 2013. "Locus-specific editing of histone modifications at endogenous enhancers using programmable TALE-LSD1 fusions." Nature biotechnology 31 (12): 10.1038/nbt.2701. doi:10.1038/nbt.2701. http://dx.doi.org/10.1038/nbt.2701.
Published Version	doi:10.1038/nbt.2701
Accessed	February 16, 2015 11:05:57 AM EST
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:12406682
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

(Article begins on next page)

Locus-specific editing of histone modifications at endogenous enhancers using programmable TALE-LSD1 fusions

Eric M. Mendenhall[1,2,3,4,*], **Kaylyn E. Williamson**[1,2,3,4], **Deepak Reyon**[2], **James Y. Zou**[1,2,3,5], **Oren Ram**[1,2,3,4], **J. Keith Joung**[2], and **Bradley E. Bernstein**[1,2,3,4]

[1]Howard Hughes Medical Institute, Chevy Chase, MD 20815

[2]Department of Pathology and Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114

[3]Broad Institute of Harvard and MIT, Cambridge MA 02142

[4]Center for Systems Biology, Massachusetts General Hospital, Boston, MA 02114

[5]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138

Abstract

Mammalian gene regulation is dependent on tissue-specific enhancers that can act across large distances to influence transcriptional activity[1-3]. Mapping experiments have identified hundreds of thousands of putative enhancers whose functionality is supported by cell type–specific chromatin signatures and striking enrichments for disease-associated sequence variants[4-11]. However, these studies did not address the *in vivo* functions of the putative elements or their chromatin states and could not determine which genes, if any, a given enhancer regulates. Here we present a strategy to investigate endogenous regulatory elements by selectively altering their chromatin state using programmable reagents. Transcription activator–like (TAL) effector repeat domains fused to the LSD1 histone demethylase efficiently remove enhancer-associated chromatin modifications from target loci, without affecting control regions. We find that inactivation of enhancer chromatin by these fusion proteins frequently causes down-regulation of proximal genes, revealing enhancer target genes. Our study demonstrates the potential of 'epigenome editing' tools to characterize an important class of functional genomic elements.

Here we sought to develop a strategy for testing the functions of genomic elements and associated chromatin states in their endogenous context. We focused on active enhancers, which are marked by histone H3 K4 mono- and di-methylation (H3K4me1 and H3K4me2) and K27 acetylation (H3K27ac)[4,6,9,12,13]. We hypothesized that a given enhancer could be inactivated by removal of these chromatin marks. To test this hypothesis, we engineered monomeric fusions between TAL effector repeat arrays and the lysine-specific demethylase 1 (LSD1)[14]. TAL effector repeats are modular DNA-binding domains that can be designed to bind essentially any genomic sequence of interest[15,16]. LSD1 catalyzes the removal of H3

K4 and H3 K9 methylation[1-3,14]. Although prior studies have used TAL effector nucleases to edit specific genomic regions to disrupt coding sequences[4-11,17,18], we reasoned that TAL effector-LSD1 fusions could provide a more versatile means for modulating the activity of noncoding elements and evaluating the significance of their chromatin states.

We initially focused on a candidate enhancer in the stem cell leukemia (SCL) locus that is enriched for H3K4me2 and H3K27ac in K562 erythroleukemia cells[4,6,9,12,13,19]. SCL encodes a developmental transcription factor with critical functions in hematopoiesis that is expressed in K562 cells. We designed a TAL effector array to bind an 18 base sequence in a segment of this enhancer predicted to be nucleosome-free based on DNase hypersensitivity (Fig. 1A, see **Methods**). As the binding specificity of monomeric TAL effectors has yet to be thoroughly characterized, we first created an expression construct encoding this TAL effector array fused to a 3X FLAG epitope. We transfected this construct into K562 cells, confirmed expression by Western blot, and mapped genome-wide binding by chromatin immunoprecipitation and sequencing (ChIP-seq). We found that the top ranked binding site corresponds precisely to the target sequence within the SCL locus (Figure 1B, Supplementary Fig. 1). We did not identify any other ChIP-seq peaks that were reproducibly detected in the two biological replicates. We also scanned the genome for sequence motifs with one or two mismatches from the TAL effector recognition motif, but did not detect any significant ChIP-seq enrichments at these sites either (Supplementary Fig. 1). These data support the specificity of TAL effector binding and are consistent with prior demonstrations of TAL effector activator domain fusions that selectively induce target genes[14,18,20].

To modulate chromatin state at the SCL enhancer, we combined the corresponding TAL effector with the LSD1 demethylase. We transfected K562 cells with a construct encoding this TAL effector-LSD1 (TALE-LSD1) fusion or a control mCherry vector, cultured the cells for three days and measured histone modification levels by ChIP-qPCR. We found that the fusion reduced H3K4me2 signals at the target locus by ~3-fold relative to control, but had no effect at several non-target control enhancers (Fig. 1C, Supplementary Fig. 2). In addition to its enzymatic activity, LSD1 physically interacts with other chromatin modifying enzymes, including histone deacetylases[21]. We therefore also tested for changes in H3K27ac, another characteristic enhancer mark. We found that the fusion reduced H3K27ac levels by >4-fold, suggesting that LSD1 recruitment leads to generalized chromatin inactivation at the target enhancer.

To eliminate the possibility that the chromatin changes reflect displacement of other transcription factors by the TAL effector, we tested a construct encoding the TAL effector without LSD1. We also examined a TALE-LSD1 fusion with a scrambled target sequence not present in the human genome to control for non-specific effects of LSD1 overexpression. Neither construct altered H3K4me2 or H3K27ac levels at the SCL locus (Fig. 1C, Supplementary Fig. 3). Lastly, to evaluate the specificity of the fusion comprehensively, we used ChIP-seq to map H3K4me2 and H3K27ac genome-wide in TALE-LSD1 and control transfected K562 cells. These data confirmed loss of H3K4me2 and H3K27ac across a 2 kb region surrounding the target sequence within the SCL locus (Fig. 1D).

These results indicate that directed LSD1 recruitment results in locus-specific reduction of H3K4me2 and H3K27ac. The generalized effect on chromatin state may be a direct consequence of H3K4 demethylation or, alternatively, may depend on partner proteins that associate with LSD1[15,16,22,23]. Regardless, prior studies indicate that sequence elements enriched for H3K4me2 and H3K27ac exhibit enhancer activity in corresponding cell types, whereas elements lacking these marks are rarely active[4,6,12]. Hence, our results suggest that this TALE-LSD1 fusion efficiently and selectively inactivates its target enhancer.

We therefore expanded our study to investigate a larger set of candidate enhancers with active chromatin in K562 cells. These include nine elements in developmental loci, sixteen additional highly cell type-specific elements, and fifteen intergenic elements. We designed and produced TALE repeat arrays for sequences in these 40 enhancers using the Fast Ligation-based Automatable Solid-phase High-throughput (FLASH) assembly method[24](Supplementary Table 1). We then cloned LSD1 fusion constructs for each TALE and transfected them individually into K562 cells, alongside mCherry control plasmid transfected separately into cells. At three days post transfection, we measured H3K4me2 and H3K27ac by ChIP-qPCR using two primer sets per target enhancer. We found that 26 of the 40 TALE-LSD1 constructs (65%) substantially reduced levels of these modifications at their target loci, relative to control transfected cells (Fig. 2; see Methods). An additional 8 constructs caused more modest reductions at their targets, suggesting that the strategy can be effective at most enhancers (Fig. 2). ChIP-qPCR measurements of H3K4me1 and H3K4me3 confirm that the reagents also reduce these alternative H3K4 methylation states (Supplementary Fig. 4). The induced changes were specific to the target loci, as analogous measurements at non-target enhancers did not reveal substantial changes (Supplementary Fig. 5). Furthermore, genome-wide ChIP-seq analysis of two TALE-LSD1 fusions that were positive by ChIP-qPCR confirmed the robustness and specificity with which they reduce chromatin signals at target loci (Supplementary Fig. 6).

We next considered whether reduced chromatin activity at specific enhancers affects the transcriptional output of nearby genes. We initially focused on 9 TALE-LSD1 fusions that robustly alter chromatin state (Fig. 2), and systematically screened for regulated genes using a modified RNA-seq procedure termed 3′ Digital Gene Expression (3′DGE). By only sequencing the 3′ ends of mRNAs, this procedure enables quantitative analysis of transcript levels at modest sequencing depths[25] (Garber M., manuscript in review). We transfected the 9 TALE-LSD1 constructs individually into K562 cells, alongside with control mCherry plasmids and measured mRNA levels in biological replicate. We normalized each 3′DGE dataset based on a negative binomial distribution and excluded any libraries that did not satisfy quality controls (see **Methods**)[26]. We then examined whether any of the TALE-LSD1 reagents substantially altered the expression of genes in the vicinity of its target enhancer. Four of the nine tested fusions (44%) caused a nearby gene to be down-regulated by at least 1.5-fold, with both biological replicates for the tested fusion exhibiting larger expression change than any of the other effectors or controls (**see Methods**, Fig. 3A, Supplementary Fig. 7).

The significance of these transcriptional changes is supported by a simulated analysis of a random sampling of 1000 genomic locations that did not yield any false-positives in which an adjacent gene scored as regulated (FDR<0.1%). The expression changes were also confirmed by quantitative RT-PCR (Supplementary Fig. 8). Two of the enhancers that significantly regulated genes are intergenic, wheras a third coincides with the 3′ end of a gene, but affects the activity of the next downstream gene. The fourth scoring enhancer resides in the first intron of ZFPM2. We confirmed that ZFPM2 down-regulation requires LSD1 recruitment, as a TALE lacking the demethylase did not affect its expression (Supplementary Fig. 8). We cannot distinguish whether the other five putative enhancers have weak transcriptional effects below our detection threshold or, alternatively, do not regulate any genes in K562 cells. Regardless, our results indicate that TALE-LSD1 fusions can alter enhancer activity in a targeted, loss-of-function manner, and thereby enable identification and modulation of their target genes.

The high prevalence of putative enhancers in the genome suggests that many act redundantly or function only in specific contexts, which could explain our inability to assign target genes to roughly half of the tested elements. To address the former, we examined three putative

enhancers within the developmental locus encoding ZFPM2 (Fig. 3B). In addition to the TALE-LSD1 fusion targeted to the intronic enhancer described above (Fig. 3A, 3B; enhancer +10), we designed and validated TALE-LSD1 fusions that reduced modification levels at two additional intronic ZFPM2 enhancers (enhancers +16, +45) (Fig. 2, 3B). First, we transfected each TALE-LSD1 fusion individually and tested their effects on ZFPM2 expression by qPCR. Whereas the fusion targeting the original +10 enhancer reduced ZFPM2 expression by ~2-fold, the fusions targeting the +16 and +45 enhancers showed only modest reductions of ~13% and ~22%, respectively, that did not reach statistical significance (Fig. 3C). To determine if these enhancers act additively or synergistically, we transfected the fusions in pairwise combinations. Although targeting pairs of enhancers tended to reduce gene expression more than hitting a single enhancer, the cumulative effects were substantially less than the sum of the two individual effects. This suggests that the multiple enhancers in this locus function redundantly to maintain ZFPM2 expression in K562 cells. These results indicate the potential of programmable TALE-LSD1 fusions to shed light on complex regulatory interactions among multiple enhancers and genes in a locus.

In conclusion, our study presents epigenome editing tools to modulate the activity of a poorly characterized class of functional genomic elements in their native contexts. The approach should also be useful for directing alterations of other epigenomic features, including repressive chromatin states and potentially with temporal control[27]. We demonstrate that programmable TALE-LSD1 fusions can be used to modulate the chromatin state and regulatory activity of individual enhancers with high specificity. These reagents should be generally useful for evaluating candidate enhancers identified in genomic mapping studies with higher throughput than direct genetic manipulations, particularly when combined with high-throughput methods for engineered TAL effector-based proteins[24]. Moreover, the approach may allow researchers to modulate developmental or disease-associated genes in specific contexts by inactivating their tissue-specific enhancers, and thus ultimately yield new therapeutic strategies.

Online Methods

Construction of TAL effector fusions

The open reading frame for LSD1 was amplified from a cDNA library from K562 cells using the following primers (F:gttcaagatctttatctgggaagaaggcgg, R:gaccttaattaaatgggcctcttcccttagaa). The PCR product was cloned into a TAL effector compatible expression vector[28] using PacI and BamHI/BglII such that LSD1 is fused to the C-terminal end of the TAL effector. TAL effector repeat array monomers were designed and assembled using FLASH as described[24]. These assembled DNA fragments were cloned into the expression vector using BsmBI sites and verified by restriction enzyme digestion and sequencing. The mCherry control vector was created by incorporating an mCherry open reading frame in place of the TAL effector array using NotI and PacI. Control TAL effector vectors lacking LSD1 were constructed using BamHI and PacI to remove LSD1, followed by blunt end ligation. The 3X Flag Tagged TAL effector vector was created by designing a gBlock (IDT) encoding a 29 amino acid Glycine:Serine linker followed by the 3X Flag sequence and cloning into the BamHI and PacI sites at the C-terminal end of the TAL effector repeat. Plasmids for construction of LSD1 and 3X Flag fusions will be available from Addgene.

Cell culture and transfection

The human erythroleukemia cell line, K562 (ATCC, CLL-243), was cultured in RPMI with 10% FBS, 1% Pen/Strep (Life Technologies). For transfections, 5×10^6 cells per

transfection were washed once with PBS. Cells were then transfected with 20 ug of TAL effector plasmid DNA or control mCherry plasmid by nucleofection with Lonza Kit V, as described by the manufacturer (Program T-016). Cells were immediately resuspended in K562 media at a density of 0.25×10^6 cells/ml. Cells were harvested at 72 hours for ChIP or RNA extraction. For ZFPM2 gene expression analysis, we standardized the total amount of DNA per transfection by co-transfecting either 10ug of a single TALE-LSD1 plasmid plus 10 ug of a scrambled TALE-LSD1 plasmid, or 10ug each of two TALE-LSD1 plasmids. Transfection efficiency, determined by flow cytometry analysis of mCherry control transfected cells, ranged from 89-94% across multiple biological replicates.

Flag tagged ChIP

TALE-3X Flag transfected K562 cells were crosslinked with 0.5% formaldehyde for 5 minutes at room temperature. Nuclei were isolated and lysed as described[29]. After sonication, solubilized chromatin was incubated with protein G Dynabeads (Invitrogen) and 0.5 ug anti-FLAG M2 antibody (Sigma) at 4C overnight. Samples were washed with TBS-T, low salt (150 mM NaCl, 2mM Tris-HCl, 1% Triton-X), LiCl (250mM LiCl, 1mM Tris-HCl, 1% Triton-X), and high salt (750mM NaCl, 2 mM Tris-HCl, 1% Triton-X) buffers at room temperature. Enriched chromatin was eluted (1% SDS, 5mM DTT) at 65 C for 20 minutes, purified and used directly for Illumina library prep. A control library was made from input DNA diluted to 50 picograms. Reads were aligned using Bowtie, and peak analysis was done using MACS with input controls, and masking genomic regions repetitive in Hg19 or K562[30].

Native ChIP

Quantitative measurements of histone modification levels were preformed in parallel using native ChIP. 0.01 U of MNase (ThermoScientific) was added to 1 ml lysis buffer (50mM Tris-HCl, 150 mM NaCl, 1% Triton X-100, 0.1% sodium deoxycholate, 1mM CaCl2) with EDTA free proteinase inhibitor. For each transfected sample, 260 ul of MNase:Lysis buffer was added and incubated for 15 minutes at 25 C, and 20 minutes at 37C. MNase was inactivated by adding 20 mM EGTA. The lysed sample was split into 96 well plate format for ChIP with H3K4me2 (abcam ab32356), H3K27ac (Active Motif 39133), H3K4me3 (Abcam ab8580), or H3K4me1 (Abcam ab8895). Antibody binding, bead washing, DNA elution and sample clean-up were performed as described[31]. ChIP DNA was analyzed by real-time PCR using FastStart Universal SYBR Green Master (Applied Biosystems), and enrichment ratios were calculated relative to equal amount of input DNA. Enrichment was normalized across ChIP samples to two standard off-target control enhancers (Supplementary Table 2), and fold-ratios were calculated relative to mCherry plasmid transfected cells assayed in parallel. Each TAL effector ChIP experiment was performed in a minimum of 3 biological replicates. TAL effector-LSD1 reagents were scored based on the fold-changes of H3K4me2 and H3K27ac for two primers flanking the target sequence. A given reagent was scored 'positive' if it induced a 2-fold or greater reduction in modification signal for at least 2 of these 4 values, with a p-value<0.05 using a one-tailed t-test. For ChIP-seq maps, 5 ng of ChIP DNA was used for library preparation as described[31].

Gene expression analysis

Genome-wide RNA expression analysis was performed using 3′DGE RNA-seq. Total RNA from 1 million TALE-LSD1 transfected or control (K562 alone or mCherry plasmid transfected) cells in biological replicate using RNeasy Mini kit (Qiagen). 2 ug of total RNA was fragmented and the 3′ ends of polyA mRNAs were isolated using Dynabeads (Invitrogen), and used to generate Illumina sequencing libraries, as described[25]. To precisely quantify the gene expression, we used a 3′ DGE analysis pipeline (Garber M., in preparation, http://garberlab.umassmed.edu/software/esat/). The pipeline estimates gene

expression based on the maximum number of reads in any 500 basepair window within 10 kb of the annotated 3′ gene end. This approach compensates for the fact that annotated ends for some genes are imprecise and may be cell type dependent and yields accurate quantifications. We then normalized the gene expression levels, scaling samples by the median gene inter-sample variation, as described in[26]. This approach controls for differences in sequencing depth between libraries and in the overall transcript abundance distribution. We excluded libraries with extreme normalization coefficients below 0.7 or above 1.5. To identify candidate regulated genes, we examined the three closest upstream and three closest downstream genes. We scored a gene as regulated if (i) it was detected in control K562 cells with a normalized RNA-seq value >10, i.e. the top 50th percentile of expression; (ii) its mean expression value was at least 1.5-fold lower in the corresponding on-target TALE-LSD1 libraries compared to all other libraries, $p < 0.05$ calculated using DESeq[26] and (iii) its normalized 3′DGE values in the on-target TALE-LSD1 libraries were the two lowest over all 22 datasets. To simulate the 1000 random binding sites, we sample genomic positions uniformly at random and use rejection sampling to ensure that the random set has a similar distribution relative to genomic annotations (intergenic, promoter, gene body, UTR) to the actual TAL effector binding sites. We then used significance testing criteria identical to that applied to the actual TAL effector experiments.

For RT-PCR based expression analysis, total RNA was extracted and reverse transcribed into cDNA using Superscript III First-Strand Synthesis system for RT-PCR (Invitrogen). Quantitative PCR was performed with FastStart Universal SYBR Green Master (Invitrogen) with primer sequences listed in Table S2 on an ABI 7500 machine. Gene expression values are presented as log2 Ct ratios relative to 2 housekeeping control genes (TBP and SDHA), and represents an average of four independent biological replicates each assayed in two technical replicates.

Supplementary Material

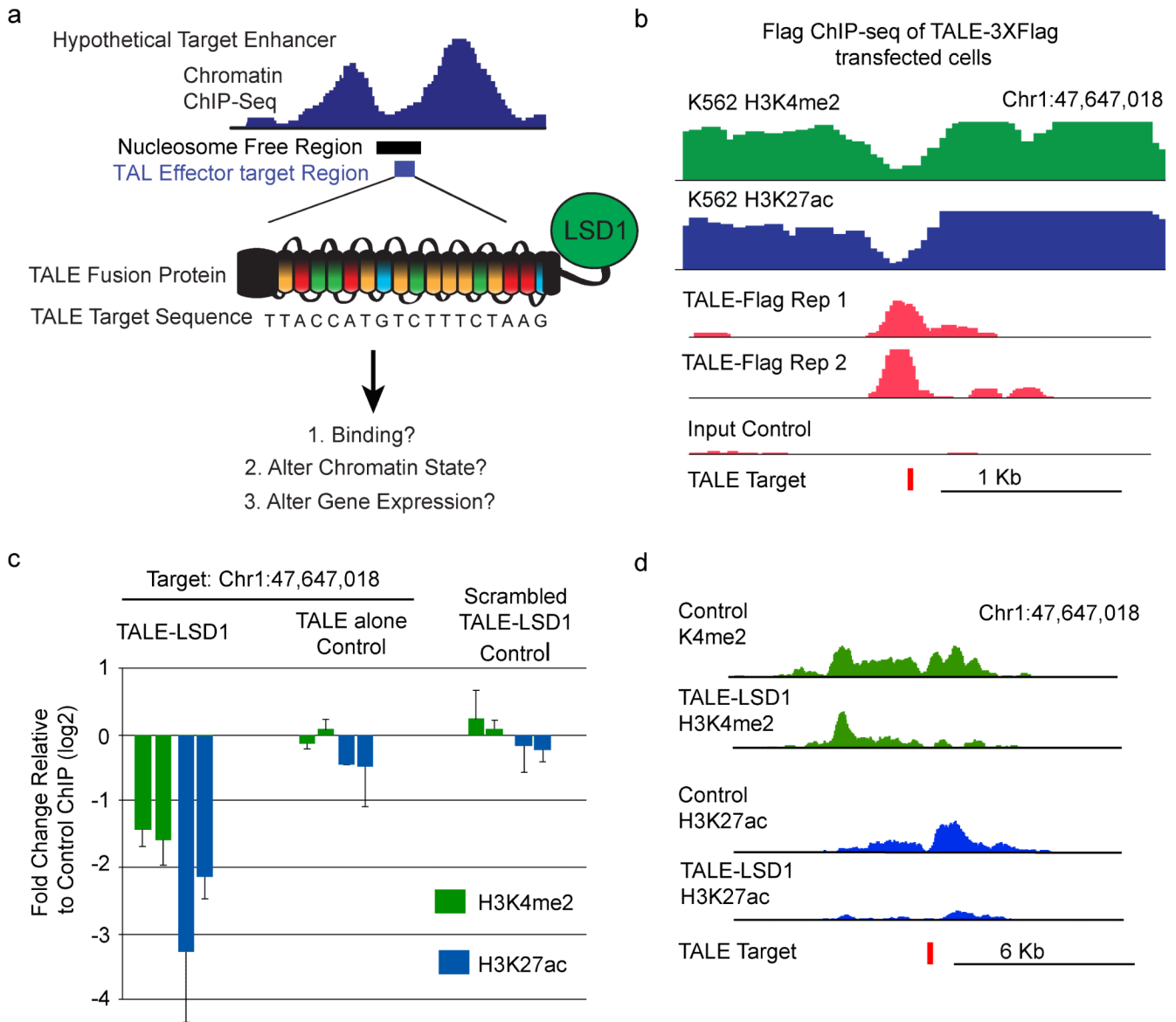Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Bulger M, Groudine M. Functional and mechanistic diversity of distal transcription enhancers. Cell. 2011; 144:327–339. [PubMed: 21295696]

2. Visel A, Rubin EM, Pennacchio LA. Genomic views of distant-acting enhancers. Nature. 2009; 461:199–205. [PubMed: 19741700]

3. Noonan JP, McCallion AS. Genomics of long-range regulatory elements. Annu Rev Genomics Hum Genet. 2010; 11:1–23. [PubMed: 20438361]

4. Heintzman ND, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature. 2009; 459:108–112. [PubMed: 19295514]

5. Boyle AP, et al. High-resolution mapping and characterization of open chromatin across the genome. Cell. 2008; 132:311–322. [PubMed: 18243105]

6. Ernst J, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature. 2011; 473:43–49. [PubMed: 21441907]

7. Consortium TEP, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 488:57–74. [PubMed: 22832584]

8. Maurano MT, et al. Systematic localization of common disease-associated variation in regulatory DNA. Science. 2012; 337:1190–1195. [PubMed: 22955828]

9. Calo E, Wysocka J. Modification of Enhancer Chromatin: What, How, and Why? MOLCEL. 2013; 49:825–837.

10. Stadler MB, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature. 2011; 480:490–495. [PubMed: 22170606]

11. Ng J-H, et al. In vivo epigenomic profiling of germ cells reveals germ cell molecular signatures. Dev Cell. 2013; 24:324–333. [PubMed: 23352811]

12. Creyghton MP, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences. 2010; 107:21931–21936.

13. Rada-Iglesias A, et al. A unique chromatin signature uncovers early developmental enhancers in humans. Nature. 2011; 470:279–283. [PubMed: 21160473]

14. Shi Y, et al. Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. Cell. 2004; 119:941–953. [PubMed: 15620353]

15. Boch J, et al. Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. Science. 2009; 326:1509–1512. [PubMed: 19933107]

16. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009; 326:1501. [PubMed: 19933106]

17. Mussolino C, Cathomen T. TALE nucleases: tailored genome engineering made easy. Curr Opin Biotechnol. 2012; 23:644–650. [PubMed: 22342754]

18. Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. Nat. Rev. Mol. Cell Biol. 2013; 14:49–55. [PubMed: 23169466]

19. Dhami P, et al. Genomic Approaches Uncover Increasing Complexities in the Regulatory Landscape at the Human SCL (TAL1) Locus. PLoS ONE. 2010; 5:e9059. [PubMed: 20140202]

20. Zhang F, et al. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol. 2011; 29:149–153. [PubMed: 21248753]

21. Lee MG, et al. Functional interplay between histone demethylase and deacetylase enzymes. Molecular and Cellular Biology. 2006; 26:6395–6402. [PubMed: 16914725]

22. Lee MG, Wynder C, Cooch N, Shiekhattar R. An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature. 2005 doi:10.1038/nature04021.

23. Whyte WA, et al. Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. Nature. 2012:1–5. doi:10.1038/nature10805.

24. Reyon D, et al. FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol. 2012; 30:460–465. [PubMed: 22484455]

25. Yoon OK, Brem RB. Noncanonical transcript forms in yeast and their regulation during environmental stress. RNA. 2010; 16:1256–1267. [PubMed: 20421314]

26. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010; 11:R106. [PubMed: 20979621]

27. Konermann S, et al. NOT FINAL PROOF. Nature. 2013:1–17. doi:10.1038/nature12466.

28. Maeder ML, et al. Robust, synergistic regulation of human gene expression using TALE activators. Nat Meth. 2013; 10:243–245.

29. Ku M, et al. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet. 2008; 4:e1000242. [PubMed: 18974828]

30. Pickrell JK, Gaffney DJ, Gilad Y, Pritchard JK. False positive peaks in ChIP-seq and other sequencing-based functional assays caused by unannotated high copy number regions. Bioinformatics. 2011; 27:2144–2146. [PubMed: 21690102]

31. Ram O, et al. Combinatorial Patterning of Chromatin Regulators Uncovered by Genome-wide Location Analysis in Human Cells. Cell. 2011; 147:1628–1639. [PubMed: 22196736]

Figure 1.
Programmable TALE-LSD1 fusion modulates chromatin at an endogenous enhancer. **(a)** Schematic depicts workflow for identification of nucleosome-free target sequence (black stripe) within enhancer (blue peaks of histone modification) and design of corresponding TAL effector fusion. TAL effector arrays comprising ~18 repeats (colored ovals) that each bind a single DNA base are fused to the LSD1 histone H3K4 demethylase. Upon transient transfection, we assayed for binding to the target site, induced chromatin changes and altered gene expression. **(b)** ChIP-seq signal tracks show H3K4me2, H3K27ac and TALE binding in K562 cells across a targeted enhancer in the SCL locus. Control tracks show anti-FLAG ChIP-seq signals in mCherry transfected cells and input chromatin. The target sequence of the TALE is indicated below. **(c)** ChIP-qPCR data show fold-change of H3K4me2 and H3K27ac enrichment in cells transfected with constructs encoding TALE-LSD1, the same TALE but lacking LSD1, or a 'non-target' TALE-LSD1 whose cognate sequence is not present in the human genome. Data are presented as log2 ratios normalized to mCherry plasmid transfected control (error bars represent ±s.e.m. n=4 biological

replicates). **(d)** ChIP-seq tracks show H3K4me2 and H3K27ac signals across the target SCL locus for K562 cells transfected with TAL effector-LSD1 or control mCherry plasmid.
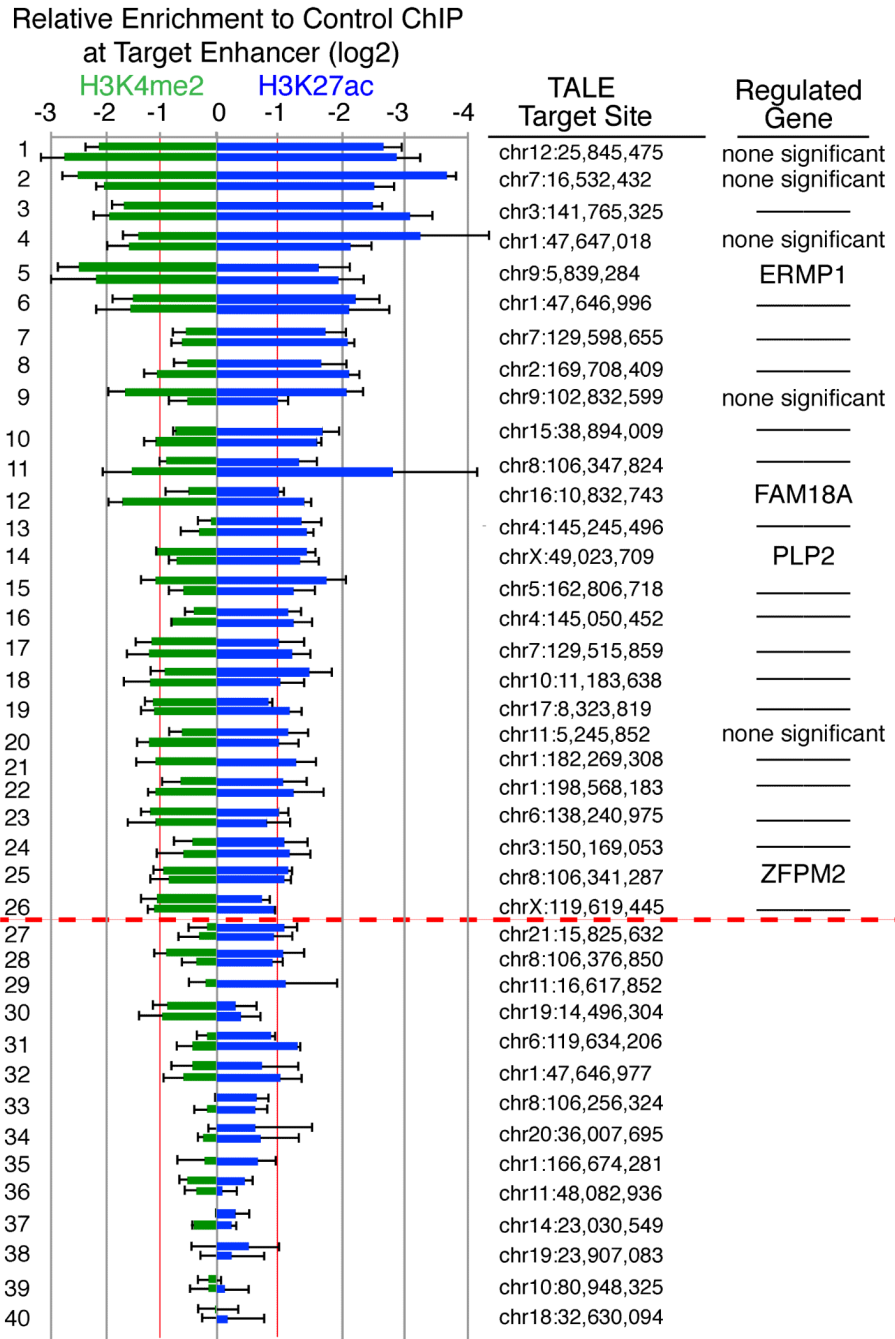
Figure 2.
TALE-LSD1 fusions targeting 40 candidate enhancers in K562 cells. The FLASH assembly method was used to engineer 40 TALE-LSD1 fusions that recognize 17 – 20 base sequences in nucleosome-free regions of candidate enhancers. These reagents were transfected into K562 cells and evaluated by ChIP-qPCR. Bi-directional plot shows fold change of H3K4me2 (green, left) and H3K27ac (blue, right) at the target locus for each of the 40 fusions, which are ordered by strength of effect and labeled by their target genomic site. Most target sites were evaluated using two qPCR primer sets. Data are presented as log2 ratios normalized to mCherry plasmid transfected control (error bars represent ±s.e.m., n=3 biological replicates). The solid red lines define a 2-fold difference (log2 = −1). The dashed

red line demarcates constructs that induce a 2-fold reduction in histone modification levels for two or more of the four values shown. Regulated genes for 9 tested fusions are shown at right (see text and Figure 3). The data indicate that TALE-LSD1 reagents provide a general means for modulating chromatin state at endogenous enhancers.
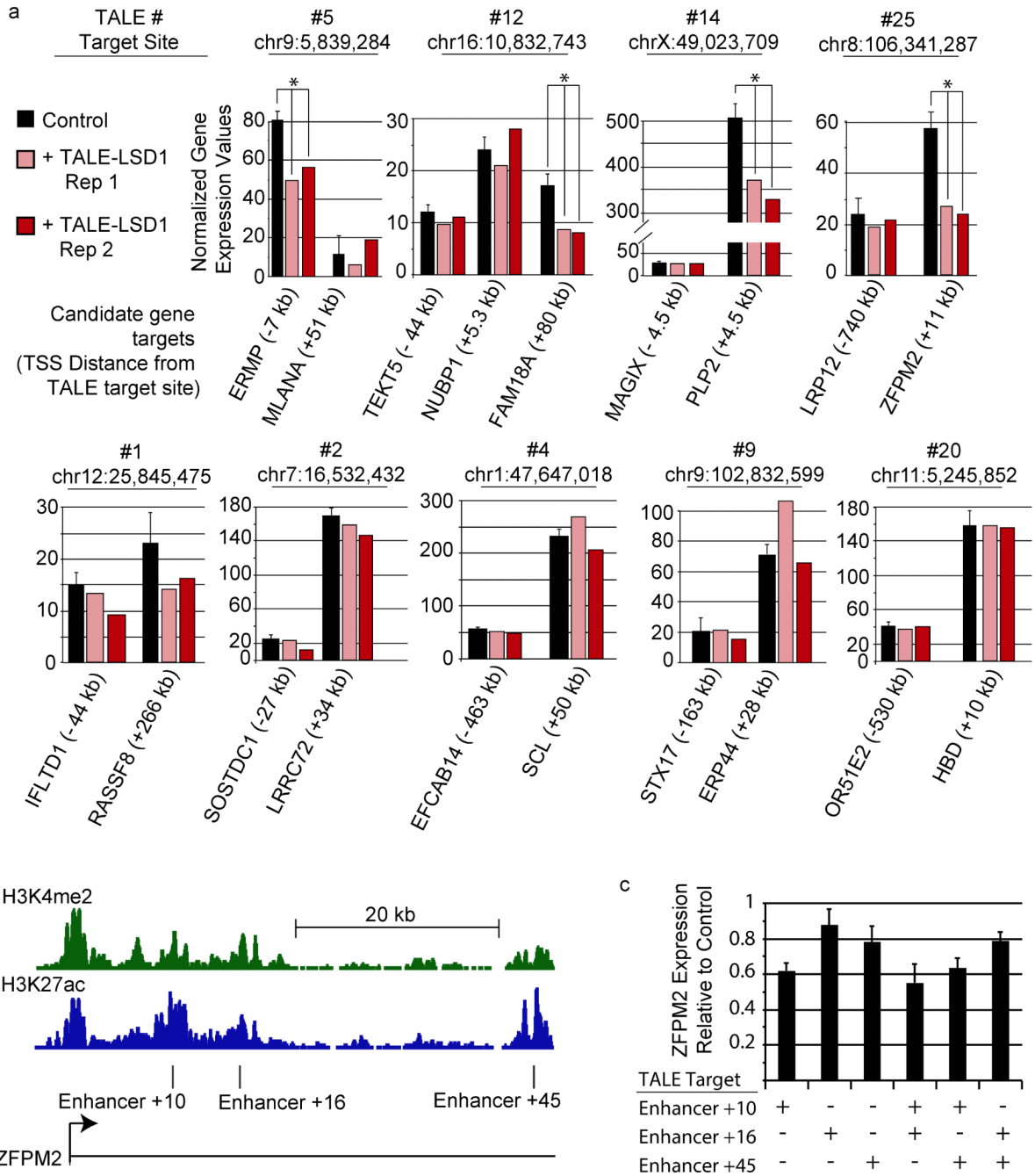
Figure 3.
TALE-LSD1 fusions to endogenous enhancers affect proximal gene expression. **(a)** Nine TALE-LSD1 fusions that robustly alter chromatin state (see Figure 2) were evaluated for their effects on gene expression by RNA-seq (see **Methods**). For each of the nine fusions, a bar graph shows normalized gene expression values for the closest expressed upstream and downstream genes. The red and pink bars indicate the gene expression value for two biological replicates in cells transfected with the corresponding 'on-target' TALE-LSD1 construct, and the black bars indicate the mean expression in cells transfected with control 'off-target' TALE constructs (error bars for the "Control" represent s.e.m, n=20 non-target libraries, see **Methods**, * indicates $p <0.05$ using an unpaired t-test). **(b)** ChIP-seq tracks

show H3K4me2 and H3K27ac signals across the Zfpm2 locus. TAL effector-LSD1 fusions were designed to target candidate enhancers (black bars) in the first intron. **(c)** Bar graph shows relative ZFPM2 expression in K562 cells transfected with the indicated combinations of TALE-LSD1 constructs. Error bars indicate ±s.e.m of 4 RT-qPCR measurements). The data suggest that these enhancers act redundantly in K562 cells to maintain ZFPM2 expression.