

1 **LSD1/KDM1A Maintains Genome-wide Homeostasis of Transcriptional Enhancers**

2 Saurabh Agarwal<sup>1,2</sup>, Patricia Marie Garay<sup>3</sup>, Robert Scott Porter<sup>2</sup>, Emily Brookes<sup>4</sup>, Yumie Murata-  
3 Nakamura<sup>2</sup>, Todd S. Macfarlan<sup>5</sup>, Bing Ren<sup>1\*</sup>, and Shigeki Iwase<sup>2\*</sup>

4 1. Ludwig Institute for Cancer Research, Department of Cellular and Molecular Medicine,  
5 University of California, San Diego School of Medicine, La Jolla, California 92093-0653, USA

6 2. Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48109, USA

7 3. Neuroscience Graduate Program, The University of Michigan Medical School, Ann Arbor, MI  
8 48109, USA

9 4. Division of Newborn Medicine, Boston Children's Hospital and Department of Cell Biology,  
10 Harvard Medical School, Boston, MA 02115, USA

11 5. Eunice Kennedy Shriver National Institute of Child Health and Human Development,  
12 National Institutes of Health, Bethesda, Maryland 20892, USA

13

14 \*Correspondence: [siwase@umich.edu](mailto:siwase@umich.edu) (S.I.), [biren@ucsd.edu](mailto:biren@ucsd.edu) (B.R.)

15

16 **Abstract**

17 Transcriptional enhancers enable exquisite spatiotemporal control of gene expression in metazoans.  
18 Enrichment of mono-methylation of histone H3 lysine 4 (H3K4me1) is a major chromatin signature  
19 that distinguishes enhancers from gene promoters. Lysine Specific Demethylase 1 (LSD1, aka  
20 KDM1A), an enzyme specific for demethylating H3K4me2/me1, has been shown to “decommission”  
21 stem cell enhancers during the differentiation of mouse embryonic stem cells (mESC). However, the  
22 roles of LSD1 in undifferentiated mESC remain obscure. Here, we show that LSD1 occupies a large  
23 fraction of enhancers (63%) that are primed with binding of transcription factors (TFs) and H3K4me1  
24 in mESC. In contrast, LSD1 is largely absent at latent enhancers, which are not yet primed by TF  
25 binding. Unexpectedly, LSD1 levels at enhancers exhibited a clear positive correlation with its  
26 substrate, H3K4me2 and enhancer activity. These enhancers gain additional H3K4 methylation upon  
27 the loss of LSD1 in mESC. The aberrant increase in H3K4me at enhancers was accompanied with  
28 increases in enhancer H3K27 acetylation and expression of enhancer RNAs (eRNAs) and their  
29 target genes. In post-mitotic neurons, loss of LSD1 resulted in premature activation of enhancers  
30 and genes that are normally induced after neuronal activation. These results demonstrate that LSD1  
31 is a versatile suppressor of primed enhancers, and is involved in homeostasis of enhancer activity.

32

33

34 **Introduction**

35 Transcriptional enhancers were discovered as potent gene regulatory elements that act  
36 independently of the distance and orientation to the target promoters<sup>1,2</sup>. A broad range of  
37 physiological and developmental processes rely on coordinated actions of transcriptional enhancers  
38 to achieve cell type-specific and temporally-controlled gene expression<sup>3,4</sup>. Numerous non-coding  
39 variants associated with a variety of human traits have been observed at enhancers, implicating their  
40 importance in normal physiology and disease pathogenesis<sup>5-7</sup>.

41 Recent studies have begun to reveal how the life cycle of enhancers progresses to induce  
42 gene expression changes during development<sup>8</sup>. Genome-wide discovery of thousands of potential

43 enhancer elements has been facilitated by profiling 1) binding of pioneer transcription factors (TFs),  
44 2) chromatin accessibility as measured by hypersensitivity to DNase I, and 3) patterns of histone  
45 modifications (reviewed in <sup>9</sup>). Monomethylation of histone H3 lysine 4 (H3K4me1) can distinguish  
46 enhancers from promoters, which in contrast are modified with H3K4me3 <sup>10, 11</sup>. In response to  
47 various environmental and developmental cues, TFs bind to specific DNA elements <sup>12, 13</sup>, and  
48 subsequent recruitment of methyltransferases KMT2C and KMT2D (aka MLL3 and MLL4,  
49 respectively) leads to H3K4me1 at enhancers <sup>14-16</sup>.

50         Once installation of H3K4me1 “primes” enhancers, they can become either “active” or  
51 “poised”, depending on the acetylation or tri-methylation of H3 lysine 27 (H3K27ac or H3K27me3) <sup>17</sup>,  
52 <sup>18</sup>, respectively. A recent report showed that active enhancers can be negatively regulated by  
53 RACK7-mediated recruitment of KDM5C, an H3K4me3/2 demethylase <sup>19</sup>. When genes need to be  
54 turned off, e.g. pluripotency genes during differentiation of stem cell, their enhancers undergo  
55 “decommissioning” by LSD1-mediated removal of the priming mark, H3K4me1 <sup>20</sup>. Notably, “latent”  
56 enhancers, i.e. DNA elements that lack TF binding or H3K4me1, can gain enhancer-like marks in  
57 response to extra-cellular stimuli and promote gene expression in fully-differentiated macrophages <sup>21</sup>.  
58 These findings highlight the dynamic H3K4 methylation of an enhancer during its life cycle.

59         Two H3K4 demethylases, LSD1 and KDM5C, have been shown to play important roles in  
60 regulation of enhancers <sup>19, 20, 22</sup>. While KDM5C reverses H3K4me3/2 leaving H3K4me1 intact <sup>23, 24</sup>,  
61 LSD1 can demethylate only H3K4me2/1 <sup>25</sup>. The distinct substrate specificities raise a possibility that  
62 these two H3K4 demethylases may cooperate to generate and/or maintain the balance of H3K4me  
63 landscape at different classes of enhancers. For example, absence of H3K4 methylation at latent  
64 enhancers could potentially be attributed to LSD1-mediated demethylation of H3K4me1. Besides the  
65 decommissioning of stem cell genes and enhancers during differentiation, LSD1 has also been  
66 shown to repress developmental genes <sup>26, 27</sup> and retrotransposons <sup>28</sup> in ES cells. However, it  
67 remains unclear whether LSD1 and KDM5C play any role at other classes of enhancers.

68         In the present study, we demonstrate that in addition to active enhancers, LSD1 also  
69 occupies poised enhancers, some of which are quickly activated (inducible enhancers) upon

70 differentiation of mESC or depolarization of post-mitotic neurons. Interestingly, LSD1 does not bind  
71 to latent enhancers, e.g. neuron-specific enhancers that are unprimed in mESC, and LSD1  
72 occupancy shows a clear positive correlation with its substrate, H3K4me2. Loss of LSD1, but not  
73 KDM5C, leads to a global upregulation of enhancer RNAs accompanied with increased H3K4  
74 methylation and H3K27ac at active, poised, and inducible enhancers, and their target genes. These  
75 results indicate that LSD1 is a pervasive suppressor of primed enhancers, involved in negative-  
76 feedback mechanisms to maintain homeostasis of histone-modification landscapes at enhancers.

77

## 78 **Results**

### 79 **LSD1 occupies a large fraction of primed enhancers**

80 To study the role of LSD1 in regulation of enhancers and gene expression, we first examined  
81 the genome-wide distribution of LSD1 at various regulatory elements. We analyzed the previously  
82 published CHIP-Seq datasets of LSD1<sup>20</sup>, p300<sup>29,30</sup>, CTCF<sup>29,31</sup>, DNase-Hypersensitivity (DHS)<sup>32</sup>  
83 and other histone modifications (see Supplementary Table 1). p300, a histone acetyltransferase and  
84 a transcriptional coactivator, has been shown to occupy both promoters and enhancers<sup>10,33</sup>,  
85 whereas CTCF binding sites anchor chromatin loops<sup>34</sup> and insulating domains<sup>35-37</sup>. By examining  
86 the overlaps of binding sites of LSD1 (109,541,  $q < 0.05$ ), p300 (86,426,  $q < 0.01$ ), CTCF (58,899,  $p$   
87  $< 10^{-12}$ ) and DHS sites (299,799,  $q < 0.05$ ), we found that 1) a majority of p300 binding sites (70.5%,  
88 Figure 1a) were co-occupied by LSD1, 2) in contrast, only 14.7% of non-p300 CTCF-binding sites  
89 were occupied by LSD1, and 3) most of the LSD1 binding sites (86%) showed an overlap ( $\pm 250$   
90 bases) with DHS sites. The higher degree of overlap of LSD1 with p300 compared to CTCF-only  
91 sites was observed at promoter, genic, and intergenic regions (Supplementary Figure 1). These  
92 observations indicate that LSD1 occupies a large fraction of primed enhancers.

93 Next, we sought to identify regulatory elements that could potentially act as enhancers.  
94 Previous studies have utilized a high H3K4me1:me3 ratio, either alone<sup>10,17</sup> or in conjunction with TF  
95 binding<sup>38,39</sup>, DHS or binding by CBP/p300<sup>40,41</sup>, to distinguish enhancers from promoters in a given  
96 cell-type. H3K4me2 is observed at both promoters and enhancers and has been shown to be a

97 signature to predict enhancers<sup>38</sup>. We therefore included H3K4me2 data to increase the sensitivity  
98 and precision of enhancer mapping. Enhancers also differ from promoters in that promoters are  
99 associated with stable transcripts<sup>42</sup> while active enhancers are associated with expression of  
100 enhancer RNA transcripts (eRNAs)<sup>42-44</sup>, which are short-lived due to exosome-mediated  
101 degradation<sup>44-46</sup>. This degradation of nascent transcripts from enhancers results in a very low, albeit,  
102 detectable levels of eRNAs in RNA-Seq (Figure 1b, right panel). Global Run-On, an *in vitro* assay,  
103 followed by high-throughput sequencing (GRO-Seq)<sup>47</sup> enables a sensitive and quantitative  
104 evaluation of transcriptionally-engaged RNA polymerase molecules. GRO-Seq, thus, serves as an  
105 indirect measure of nascent transcription at promoters and enhancers, irrespective of the  
106 subsequent stability of the transcripts<sup>42</sup>. Therefore, we employed a high ratio of GRO-Seq:RNA-Seq  
107 signals to further refine the prediction of enhancers in mESC. We focused on only intergenic  
108 enhancers as we found it difficult to differentiate the eRNAs from gene-coding and promoter-  
109 upstream<sup>46</sup> transcripts. In summary, intergenic enhancers were defined as  $\pm 500$  base regions  
110 around p300/DHS summits with i) H3K4me1 enrichment ( $rpk_m \geq 1$  and ChIP:Input  $> 1.5$ ), ii)  
111 H3K4me3 lower than either H3K4me1 or H3K4me2, iii) a low rate of transcription (RNA-Seq  $f_{pk_m} <$   
112 0.5), iv) a GRO-Seq:RNA-Seq ratio  $> 5$ , and v) a high average mappability to exclude repetitive  
113 regions. This pipeline predicted a total of 22,047 intergenic enhancers in mESC (Supplementary  
114 Table 2).

115 LSD1 has been shown to occupy enhancers in various cell types<sup>20, 48, 49</sup>. However, the  
116 genome-wide relationship between LSD1 binding and chromatin states at enhancers, such as  
117 histone-modification landscapes and eRNA transcription, remains unclear. To address this issue, we  
118 first subdivided the 22,047 predicted intergenic enhancers into quartiles (Q1-Q4, Figure 1b) based  
119 on the enrichment of H3K4me2 relative to H3K4me1. Similar to a previous observation in K562 cells  
120<sup>42</sup>, we noted a positive correlation between eRNA levels, measured by GRO-Seq, Nuclear RNA-Seq  
121 or RNA-Seq, and H3K4me2 levels (Figure 1b). Acetylations of H3K9 and H3K27 and eRNA  
122 expression have been established as signatures of active enhancers<sup>9, 17, 18</sup>. Consistently, we also

123 observed that enhancers with higher transcription levels displayed higher acetylation levels of H3K9  
124 and H3K27 relative to trimethylation (Supplementary Figure 2).

125 We then determined the extent of LSD1 binding across various enhancer classes and found  
126 that a large fraction (63.2%) of the predicted 22,047 enhancers is bound by LSD1 (Figure 1c).  
127 Surprisingly, we found that LSD1 occupancy at enhancers increased with increasing levels of  
128 H3K4me2 or increasing levels of GRO-Seq signals (Supplementary Figures 3 and 4a). We then  
129 calculated the correlation coefficients between LSD1 levels and levels of various histone  
130 modifications at promoter distal regions, i.e. excluding TSS  $\pm$  1500 bases. Compared to H3K4me1 ( $r$   
131 = 0.5745) and H3K4me3 ( $r$  = 0.486), we found that LSD1 levels showed the highest correlations with  
132 its primary substrate, H3K4me2 ( $r$  = 0.627, Supplementary Figure 4b), and H3K27ac ( $r$  = 0.604), a  
133 marker for enhancer activity. In contrast, H3K27me3 and H3K9me3 were inversely correlated with  
134 LSD1 levels (Supplementary Figure 4b). We also classified 22,047 intergenic enhancers into poised,  
135 active, and intermediate enhancers based on their H3K27me3 and H3K27ac levels, according to the  
136 previous report<sup>18</sup>. We found that LSD1 occupies substantial fractions of each of the three enhancer  
137 classes with increased occupancy of active enhancers compared to other classes (Figure 1d).  
138 Similar patterns were observed when we used the levels of H3K9me3 and H3K9ac to classify  
139 enhancers as described previously<sup>41</sup>(Figure 1e). These results indicate that LSD1 binds to multiple  
140 enhancer classes with positive correlations to H3K4me2, H3K27/K9 acetylation and eRNA levels.

141

#### 142 **LSD1 rarely binds to cell-type specific “latent” enhancers**

143 Higher LSD1 occupancy at more active enhancers may contradict with the LSD1’s classical  
144 role as a transcriptional repressor of neuron-specific genes in non-neuronal tissues<sup>25, 50</sup>. RE1-  
145 Silencing Transcription factor (REST) is known to be expressed in non-neuronal cells with the role of  
146 repression of neuronal genes in these cell-types through the Corepressor of REST(CoREST)  
147 complex<sup>50</sup>. We analyzed the previously published REST ChIP-Seq dataset<sup>20</sup> and found that only  
148 1.31% of predicted enhancers (288 out of 22,047) were bound by REST in mESC. However, a  
149 majority (91%) of these REST-positive enhancers were bound by LSD1. These data suggest that

150 REST/LSD1-mediated suppression is not the primary mechanism of enhancer regulation in mESC.  
151 Enhancers that are inert, free of TFs, and thus, insensitive to DNase I, are referred to as “latent”  
152 enhancers in a given cell type <sup>21</sup>. The low occupancy of primed enhancers by REST/LSD1 prompted  
153 us to test if LSD1 contributes to regulation of latent enhancers. We first looked at the *mir290* cluster,  
154 which is specifically expressed in mESC and early developmental stages <sup>51</sup>. In mESC, several  
155 enhancers upstream of its promoter, show high DHS, p300 binding and high GRO-Seq signals  
156 accompanied with strong LSD1 binding events (Figure 2a). These *mir290* enhancers lack brain-  
157 derived DHS, GRO-Seq signal and H3K4 methylation in cortical neurons (CN), demonstrating that  
158 these enhancers are primed in mESC but are latent in CN. LSD1 ChIP-Seq data from neural stem  
159 cells (NSC) <sup>52</sup>, however, showed a lack of LSD1 binding at these latent enhancers. Conversely, two  
160 enhancers upstream of *Npas4* (marked with asterisks, Figure 2b), a gene predominantly expressed  
161 in the brain, showed brain-specific DHS and LSD1 occupancy, GRO-Seq signals and high H3K4me1,  
162 specifically in the neuronal cell types (NSC or CN). In mESC, LSD1 is absent at these brain-specific  
163 DHS sites upstream of the *Npas4* promoter (Figure 2b). These two examples suggest that LSD1  
164 could primarily be recruited to primed enhancers in a given tissue in a TF-binding dependent manner.

165 To ascertain this specificity of LSD1 recruitment to primed enhancers on a genome-wide  
166 scale, we sought to identify genomic elements that are latent in mESC but are primed in other cell  
167 types. We identified DHS sites from mESC (398,675,  $q < 0.01$ ) and four additional mouse tissues,  
168 including adult brain (415,400), heart (320,416), liver (207,046), and lung (358,575), using Hotspot  
169 (v4.1) <sup>53</sup>. Similar to our earlier observation (Supplementary Figure 1), we found that most of mESC  
170 LSD1-binding sites (86%) overlapped with mESC hotspots (Figure 2c). Next, we performed an  
171 intersection of hotspots from the five tissues. This resulted in tens of thousands of hotspots, which  
172 could potentially act as tissue-specific enhancers in a given tissue and latent in others (Figure 2d, e).  
173 Motif analysis on promoter-distal hotspots revealed that these tissue-specific hotspots are indeed  
174 enriched with binding sites for lineage-specific TFs (Supplementary Figure 5). In agreement with the  
175 *mir290* and *Npas4* loci, mESC LSD1 binding sites showed negligible overlaps with tissue-specific  
176 hotspots (0.40-0.69%), whereas 14.31% of mESC-specific hotspots were bound by LSD1 (Figure 2d,

177 e). Based on these data, we concluded that LSD1 is predominantly recruited to primed enhancers  
178 and is not actively involved in maintaining inactivity of latent enhancers in mESC.

179

### 180 **Loss of LSD1 results in a genome-wide increase in enhancer H3K4 methylation and H3K27** 181 **acetylation**

182 The unexpected positive correlation between LSD1 binding and enhancer H3K4me2 levels  
183 raises a possibility that LSD1 may not be demethylating H3K4me2 at enhancers. LSD1 has been  
184 implicated in demethylation of H3K9 instead of H3K4 when it binds androgen receptors<sup>54, 55</sup>, though  
185 a recent study reported otherwise<sup>56</sup>. Phosphorylation of H3T6 appears to interfere with LSD1-  
186 mediated H3K4 demethylation<sup>57</sup>. Alternatively, the positive correlation may reflect a negative  
187 feedback mechanism, in which LSD1 searches for and binds to genomic regions with high H3K4me2  
188 levels and reverses this modification to regulate optimal enhancer activity.

189 To test whether LSD1 is involved in maintaining precise levels of H3K4 methylation at  
190 enhancers, we investigated the previously generated mESC line that lacks LSD1 due to the insertion  
191 of a gene-trap cassette (*Lsd1*-GT)<sup>28</sup>. Western blot analysis of mESC carrying either wild-type (WT)  
192 *Lsd1* or *Lsd1*-GT did not show any detectable differences in total H3K4me1, H3K4me2, H3K4me3 or  
193 H3K27ac levels (Supplementary Figure 6). We then performed ChIP-Seq to measure H3K4me  
194 levels across the genomes of these two mESC lines. Since LSD1 is known to associate with multiple  
195 HDAC-containing co-repressor complexes, including the CoREST<sup>50, 58</sup> and NuRD<sup>59</sup> complexes, we  
196 also included H3K27ac and HDAC1 in our ChIP-Seq analysis.

197 Genome-wide localization analysis (ChIP-Seq) profiles, which reflect the spatial distribution of  
198 these marks, looked highly similar between the two genotypes at most of the loci. Upon the loss of  
199 LSD1, however, H3K4 methylations displayed statistically-significant increases at active, poised, and  
200 intermediate LSD1-target enhancers, which were accompanied with conspicuous increases in  
201 H3K27ac (Figure 3a, Supplementary Figure 7a). Similar changes were also observed at enhancers  
202 that showed a significant increase in eRNA expression (see next section) in the *Lsd1*-GT mESC  
203 (Supplementary Figure 7). Interestingly, HDAC1 levels did not change significantly at poised



204 enhancers, while active or poised enhancers showed a small but significant increase in HDAC1  
205 binding (Figure 3a), which could be attributed to either experimental variations or unknown  
206 mechanisms to compensate for the loss of LSD1. The inability of HDAC1 to remove H3K27ac in  
207 *Lsd1*-GT cells is consistent with the previous observations that HDAC activity is negatively  
208 influenced by the presence of H3K4me<sup>27, 60</sup>. Representative genes *Pou5f1* (Figure 3b) and *Cbln4*  
209 (Figure 3c), which are normally active or poised in undifferentiated mESC, respectively, showed  
210 relatively higher H3K4me and H3K27ac at both promoters and enhancers in *Lsd1*-GT mESC. We  
211 also observed a concomitant increase in both promoter- and enhancer- associated GRO-Seq signals  
212 at these loci in *Lsd1*-GT mESC (Figure 3b, c). The increases in H3K27ac and nascent transcription  
213 at these poised and intermediate enhancers upon the loss of LSD1 suggest a shift in their identity  
214 towards active enhancers. These results indicate that LSD1 functions as an H3K4 demethylase at  
215 enhancers, and is required for maintenance of optimal H3K4me and H3K27ac levels in mES cells.

216

### 217 **Loss of LSD1 but not KDM5C results in aberrant activation of transcriptional enhancers**

218 In the above CHIP-Seq study, we found a genome-wide elevation of all three H3K4me  
219 statuses, including H3K4me<sub>3</sub>, at LSD1-target enhancers in *Lsd1*-GT mESC (Figure 3a,  
220 Supplementary Figure 7a). This increase of H3K4me<sub>3</sub> at enhancers cannot be explained directly by  
221 the loss of LSD1, as LSD1 is incapable of demethylating H3K4me<sub>3</sub><sup>25</sup>; therefore, one or more  
222 H3K4me<sub>3</sub> demethylases might be involved in maintaining low levels of H3K4me<sub>3</sub> at enhancers.  
223 LSD1 and KDM5C, an H3K4me<sub>3</sub>/me<sub>2</sub> demethylase<sup>23</sup>, have been previously shown to be in the  
224 same complex<sup>19</sup> and that KDM5C suppresses over-activation of active enhancers in breast cancer  
225 cells<sup>19</sup>. To elucidate the interrelationship between LSD1 and KDM5C in suppression of enhancer  
226 activity, we first performed KDM5C ChIP-Seq in mESC and identified 113,166 KDM5C binding sites  
227 (MACS2,  $q < 0.05$ ). Most of the 22,047 predicted intergenic enhancers (78.3%) were bound by either  
228 LSD1 or KDM5C and 52.1% of total were bound by both (Figure 4a).

229 We generated *Kdm5c*-knockout (KO) mESC by transfecting a *Cre*-expression plasmid into  
230 the mESC harboring the floxed exons 11 and 12, which encode the catalytic JmjC domain<sup>61, 62</sup>, and

231 confirmed the loss of KDM5C (Supplementary Figure 8). We then asked if the loss of either LSD1 or  
232 KDM5C leads to aberrant enhancer activity by quantifying changes in GRO-Seq signals at  
233 enhancers. To identify misregulated enhancers, we calculated the number of GRO-Seq reads  
234 mapping within  $\pm 500$  bases of the center of the predicted enhancers and normalized them against  
235 199,209 p300/DHS sites across the whole genome using DESeq<sup>63</sup>. Upon the loss of LSD1, a large  
236 fraction (24.8%, 5,471) of total intergenic enhancers showed a significant elevation in associated  
237 GRO-Seq transcripts, while a small number 674 (3.06%) displayed a reduced activity with a stringent  
238 cutoff of  $q < 0.05$  (Figure 4b). Next, we tested if this elevation of GRO-Seq signals is specific to  
239 poised, intermediate or active enhancers. We found that all three enhancer classes showed a  
240 significant increase in associated nascent transcripts ( $p < 2.2e-16$ , Wilcoxon signed-rank test, Figure  
241 4c), indicating that LSD1 is required for genome-wide suppression of aberrant enhancer activities.

242 Using the same DESeq cutoff, however, we were not able to identify any misregulated  
243 enhancers in *Kdm5c*-KO mESC. After relaxing the cutoff to  $p < 0.05$ , we could identify only 63  
244 upregulated and 102 downregulated enhancers upon the loss of KDM5C (Figure 4d). To confirm that  
245 these observations were not dependent on differences in sequencing depths or inter-replicate  
246 variability, equal number of reads were randomly selected from each GRO-Seq sample and pairwise  
247 DESeq comparisons between individual replicates of either genotypes were repeated. Thus, in  
248 contrast to the crucial role of LSD1, KDM5C is largely dispensable for enhancer suppression in  
249 mESC.

250 Since GRO-Seq is an *in vitro* transcription assay, we sought to validate this global  
251 upregulation of enhancers upon LSD1 depletion in mESC under physiological conditions by  
252 sequencing total cellular RNAs (RNA-Seq) and nuclear RNAs (Nuclear RNA-Seq). Either RNA-Seq  
253 or Nuclear RNA-Seq could not provide sufficiently high eRNA signals to call differentially-expressed  
254 enhancers likely due to the aforementioned exosome-mediated degradation of eRNAs. However,  
255 when we evaluated eRNA levels at all the intergenic enhancers as a group, RNA-Seq and Nuclear  
256 RNA-Seq corroborated our GRO-Seq results (Supplementary Figure 9). These data demonstrate  
257 that LSD1, but not KDM5C, is required for suppression of aberrant enhancer activities in mESC.

258

259 **Aberrant changes in enhancer activity are associated with misregulation of physically-**  
260 **interacting genes**

261 The standard approach to gauge the influence of enhancer misregulation on gene  
262 expression has been to quantify changes in expression of genes that are located in proximity to the  
263 enhancers of interest. However, recent advances in genome-wide profiling of chromatin interactions  
264 <sup>64-66</sup> have paved the way for a more precise determination of enhancer-promoter interactions. To  
265 identify genes that physically interact with our set of predicted enhancers, we utilized the recently-  
266 published “HiCap” data set, which is a high-resolution map of promoter-anchored chromatin  
267 interactions in mESC <sup>67</sup>. For instance, our enhancer prediction identified a ~ 6 kb-wide enhancer  
268 cluster downstream of *Dusp5* and the analysis of HiCap data revealed that one of the three  
269 individual enhancers within the cluster appears to interact with the *Dusp5* promoter (Figure 5a). This  
270 enhancer cluster was significantly upregulated in *Lsd1*-GT mESC and showed a pronounced  
271 increase in H3K4me2 and H3K27ac levels, and *Dusp5* transcription (Figure 5a). A concomitant  
272 misregulation of enhancers and the interacting gene was also observed for the aforementioned  
273 *mir290* cluster (Supplementary Figure 10a).

274 To evaluate the genome-wide impact of enhancer misregulation on gene expression, we first  
275 categorized enhancers based on the statistical significance of their differential expression in *Lsd1*-  
276 GT mESC: significantly misregulated ( $q < 0.05$ ) and moderately misregulated ( $0.05 \leq q < 0.25$ )  
277 enhancers from our GRO-Seq analysis. We then retrieved the physically-interacting promoters from  
278 the HiCap data, and plotted the changes in mRNA levels (RNA-Seq, Figure 5b) or rates of nascent  
279 transcription (GRO-Seq, Supplementary Figure 10b). We observed a general trend that genes  
280 associated with upregulated enhancers showed an increased expression and *vice versa*, and the  
281 genes that are anchored to unaffected enhancers did not exhibit any significant changes in  
282 expression in *Lsd1*-GT mESC (Figure 5b, Supplementary Figure 10b, c). Importantly, for each  
283 category, the magnitude and the statistical significance of median change in gene expression  
284 correlated positively with those of changes in enhancer activity (Supplementary Figure 10c).

285 Additionally, when interacting genes were called on the basis of genomic proximity to the enhancers,  
286 we observed a similar trend (Supplementary Figure 11). These results indicate that LSD1's role at  
287 enhancers is important for a precise transcription of their cognate genes.

288 To further corroborate if LSD1 and its catalytic activity are required for suppression of  
289 enhancer activity and associated genes, we utilized luciferase reporter assays in *Lsd1*-GT mESC.  
290 We selected 11 enhancers with two latent enhancers and at least one enhancer from each of poised,  
291 intermediate, and active enhancers that showed significant upregulation in *Lsd1*-GT mESC  
292 compared to WT-mESC and also showed an upregulation of the associated gene. 1.0 ~ 1.2 kb of the  
293 enhancer-containing regions were cloned downstream of the HSV-Thymidine Kinase promoter-  
294 driven firefly luciferase gene. *Lsd1*-GT mESC were transfected with a control plasmid or plasmids  
295 expressing either human LSD1 or the catalytically inactive LSD1-K661A mutant<sup>68</sup> along with the  
296 reporter plasmids. We found that LSD1's catalytic activity is indeed required for suppression of all  
297 the active enhancers tested ( $p < 0.1$ , Student's *t*-tests, Supplementary Figure 12), consistent with  
298 high levels of LSD1 at these enhancers in mESC. In contrast, one of the *Nanog* enhancers, which  
299 had not displayed a change upon the loss of LSD1 in mESC, was unaffected by LSD1 expression.  
300 We observed lower enhancer activities of the latent, poised, and intermediate enhancers compared  
301 to the active enhancers, indicating that our enhancer classification could accurately predict enhancer  
302 activity. However, we found it difficult to interpret the effect of LSD1 or its catalytic activity at these  
303 weak enhancers as they failed to enhance the activity of the promoter.

304

### 305 **Both mESC-specific and differentiation genes are upregulated upon the loss of LSD1 in** 306 **undifferentiated mES cells**

307 A previous study has implicated LSD1 in decommissioning of enhancers of pluripotency  
308 genes during differentiation of mES cells<sup>20</sup>. However, the roles of LSD1 in undifferentiated mES  
309 cells remain elusive. The global upregulation of enhancers (Figure 4b), prompted us to investigate if  
310 the loss of LSD1 in undifferentiated mES cells affected the expression of pluripotency and/or  
311 differentiation genes. To this end, we first analyzed a published RNA-Seq datasets for mESC and

312 epiblast stem cells (EpiSC), which were derived from mESC by treatment with Activin A and FGF2  
313 for 4 days<sup>69</sup>. We first selected the genes that showed a significant ( $q < 0.01$ , DESeq) and at least 5-  
314 fold change in expression during differentiation. This analysis yielded 710 induced and 745  
315 downregulated (decommissioned) genes during the differentiation of mESC to EpiSC (Figure 6a).  
316 mESC and EpiSC represent two consecutive stages of embryonic development, namely pre-  
317 implantation and post-implantation, respectively. Therefore, these sets of upregulated and  
318 downregulated genes may represent the earliest transcriptional response of mES cells to the  
319 differentiation cue.

320 Compared to WT mESC, *Lsd1*-GT mESC displayed roughly equal number of genes being  
321 significant-upregulated or -downregulated (55% vs. 45%, 1493 upregulated and 1203 downregulated,  
322 respectively,  $q < 0.05$ ). Quantitation of gene expression changes of the induced and the  
323 decommissioned genes, using our RNA-Seq and GRO-Seq datasets revealed that many of these  
324 genes are upregulated in undifferentiated *Lsd1*-GT mESC and both of these gene sets showed  
325 statistically-significant upregulation as group ( $p < 2.2e-16$ , Wilcoxon signed-rank test, Figure 6b, c).  
326 For example, *Hmga2* is a gene which is induced upon differentiation of mES cells and is required for  
327 the exit from naïve pluripotency<sup>70</sup>. As shown in Figure 6d, LSD1 loss led to a marked increase in  
328 *Hmga2* expression which was also associated with increased H3K4 methylations, H3K27ac, and  
329 GRO-Seq signals at nearby enhancers. We have shown earlier that some pluripotency genes  
330 including *Pou5f1* (Figure 3b) and *mir290* (Supplementary Figure 10a) and their nearby enhancers  
331 were upregulated in *Lsd1*-GT mESC. Upregulation of both pluripotency and differentiation genes  
332 upon the loss of LSD1 in mES cells suggests that LSD1 does not instruct the fate of mES cells to a  
333 particular direction.

334

### 335 **LSD1 is required for suppression of inducible enhancers in terminally-differentiated neurons**

336 We next sought to test if LSD1-mediated regulation of enhancers plays a role in gene  
337 expression program of terminally-differentiated cells using cortical neuron (CN) culture as a model.  
338 Using lentiviral delivery of two independent short hairpin RNAs (shRNAs) at 7 days *in vitro* (DIV), we

339 knocked down (KD) LSD1 in primary cultures of mouse CN (Figure 7a). We employed BrU-Seq, a  
340 recently-developed nascent-RNA sequencing technique<sup>71</sup>, that allows an accurate evaluation of any  
341 changes in active transcription of both mRNAs and eRNAs. At DIV 11, i.e. after 4 days of control or  
342 *Lsd1* shRNA delivery, neuronal cultures were treated with 5-bromouridine for 32 min, followed by  
343 enrichment of BrU-containing nascent transcripts using anti-BrdU beads and high-throughput  
344 sequencing. DESeq analysis indicated a significant misregulation of 1,500 genes ( $q < 0.05$ , 778  
345 downregulated and 722 upregulated). Interestingly, many well-characterized activity-regulated genes  
346 (ARGs)<sup>43</sup>, including *Arc*, *Fos*, *Fosb*, *Npas4*, *Egr1-4*, and *Nr4a1-3*, were among the most  
347 significantly-upregulated genes upon LSD1-KD in unstimulated neurons (Figure 7b). ARGs are  
348 expressed at low levels in resting neurons and are rapidly induced by depolarization of neurons via  
349 sensory inputs, thereby representing a stimulus-responsive gene regulatory program. Since products  
350 of ARGs play important roles in synaptic plasticity underlying cognitive development, learning and  
351 adaptive processes<sup>72-74</sup>, we narrowed our focus on these inducible genes.

352 To evaluate this ARG upregulation on a genome-wide scale, we analyzed our previously  
353 published RNA-Seq data set<sup>75</sup> and identified 140 ARGs that were induced by KCl-mediated  
354 depolarization of CN in culture. Both of the two independent *Lsd1* shRNAs led to a spurious  
355 induction of many ARGs in the resting neurons (Figure 7b, Supplementary Figure 13), indicating that  
356 LSD1 suppresses premature induction of ARGs in CN. We next examined the *Npas4* locus to check  
357 if enhancer misregulation upon the loss of LSD1 could be involved in the premature induction of  
358 ARGs. We identified three putative enhancers upstream of the *Npas4* promoter based on DHS and  
359 H3K4me1 enrichment (Figure 7c). These three enhancers appear to respond to membrane  
360 depolarization, as they show activity-dependent increases in NPAS4 binding<sup>43</sup> and H3K27ac levels  
361<sup>76</sup>. These activity-regulated *Npas4* enhancers are bound by LSD1 in NSC<sup>52</sup>. An increase in nascent  
362 transcription across these enhancers, concomitant with an increased *Npas4* expression, indicates  
363 that these enhancers are deregulated upon the loss of LSD1 (Figure 7c). A previous study had  
364 reported more than ten thousand putative activity-regulated enhancers based on increased CBP  
365 binding in response to membrane depolarization<sup>43</sup>. Subsequent work categorized these candidate

366 enhancers into four groups on the basis of activity-dependent changes in H3K27ac<sup>76</sup> (Figure 7d).  
367 The study found that only the enhancers that displayed activity dependent changes in H3K27ac  
368 were involved in promoting ARG transcription<sup>76</sup>. Next, we investigated if the premature upregulation  
369 of ARGs in *Lsd1*-KD neurons was accompanied with misregulation of any of these activity-regulated  
370 enhancer groups. Analysis of BrU-Seq data revealed that loss of LSD1 did not have a significant  
371 impact on enhancers that do not display any activity-dependent changes in H3K27ac (Wilcoxon  
372 signed-ranked test, Figure 7d). However, *Lsd1*-KD led to a significant upregulation of eRNA levels at  
373 the enhancers that gain or lose H3K27ac upon KCl treatment (Figure 7d, Supplementary Figure 14a).  
374 Interestingly, the group of enhancers with no H3K27ac either before or after depolarization, and are  
375 presumably poised neuronal enhancers, also showed an upregulation upon the loss of LSD1 (Figure  
376 7d, Supplementary Figure 14a). Similar to our earlier observations with RNA-Seq and Nuclear RNA-  
377 Seq in mESC, the eRNA signals with BrU-Seq were considerably lower than those with mESC GRO-  
378 Seq to obtain sufficiently-high statistical power for comparison; therefore, we aggregated eRNA  
379 signals from the two control and the two *Lsd1*-KD experimental groups for this analysis (Figure 7d,  
380 Supplementary Figure 14a). Similar trends were observed in analysis without grouping the samples  
381 (Supplementary Figure 14b). These data indicate that LSD1 is required for genome-wide  
382 suppression of premature enhancer activation in resting neurons.

383         Activation of ARGs upon *Lsd1*-KD could also be a result of extraneous activation of signaling  
384 pathways upstream of ARG induction. Extracellular signal-regulated kinases 1 and 2 (ERK1/2) are  
385 rapidly phosphorylated in response to a variety of extracellular stimuli, including membrane  
386 depolarization, and play critical roles to mediate the transcriptional response<sup>72,77</sup>. A lack of  
387 noticeable differences in phosphorylation levels of ERK1/2 upon *Lsd1*-KD (Supplementary Figure  
388 14c), further support a direct role of LSD1 in suppression of activity-regulated enhancers and genes.

389

## 390 **Discussion**

391         Early embryonic lethality of homozygous *Lsd1*-KO mice indicates an essential role of LSD1  
392 in development<sup>78</sup>. However, the roles of LSD1 in early embryogenesis have not been fully

393 elucidated. A previous study has shown that the silencing of pluripotency genes during differentiation  
394 of mES cells is mediated by the decommissioning of pluripotency enhancers by LSD1<sup>20</sup>. This study  
395 employed Tranylcypromine (TCP), a pharmacological agent to block LSD1's enzymatic activity<sup>20</sup>.  
396 However, TCP also inhibits the H3K4 demethylase activity of LSD2 (aka KDM1B)<sup>79</sup>, the paralog of  
397 LSD1, which is involved in regulation of transcriptional elongation<sup>80</sup>. Thus, it remains unclear  
398 whether the observed impact of TCP treatment on enhancer dysregulation in mES cells was  
399 mediated by inhibition of either LSD1 or LSD2 or both. By employing genetic ablation of *Lsd1* in  
400 mES cells, we demonstrate that LSD1 suppresses the activity of a large fraction of primed  
401 enhancers, including the pluripotency enhancers and poised enhancers of differentiation genes.  
402 Notably, several of these key pluripotency enhancers and genes are already upregulated in  
403 undifferentiated *Lsd1*-deficient mESC (Figures 3 and 6). Similar to our observations, another group  
404 had found that *Lsd1*-KD by siRNA led to an upregulation of several stem cell genes in  
405 undifferentiated mES cells<sup>81</sup>. Loss of LSD1 in stem cells has been implicated in multiple  
406 differentiation defects, including de-differentiation of the pluripotent mESC state towards the  
407 totipotent 2-cell state<sup>82</sup>, or the premature differentiation of human ES cells to endodermal and  
408 mesodermal lineages<sup>26</sup>. These observations could be reconciled by our findings that LSD1 is  
409 required for suppressing both pluripotency genes and differentiation genes in mES cells, possibly  
410 through maintenance of proper enhancer activity.

411 We provide several lines of evidence that LSD1 plays an essential role in genome-wide  
412 homeostasis of primed enhancers. We show that recruitment of LSD1 correlates positively with  
413 levels of enhancer H3K4me2, H3K27ac and eRNA transcription (Figure 1) and this recruitment is  
414 specific to primed enhancers (Figure 2). Loss of LSD1 led to an upregulation of a large number of  
415 enhancers, as demonstrated by increased H3K4 methylation, H3K27ac (Figure 3), and eRNA  
416 transcription (Figure 4), concomitant with an upregulation of the associated genes (Figure 5). These  
417 results support the following model of LSD1-mediated homeostasis of the histone modification  
418 landscape during the life cycle of an enhancer. Binding of TF and subsequent recruitment of MLL3/4  
419<sup>14-16</sup> prime the enhancers with H3K4me1/me2, which attract LSD1 irrespective of whether the



420 enhancers are destined to be either “active” or “poised” (Supplementary Figure 15). LSD1 then  
421 counteracts with MLL3/4 to maintain an optimal H3K4me levels. Enhancers with a relatively low  
422 H3K4me2 may represent early stages of priming by TFs and MLLs. It is possible that enhancers with  
423 low levels of H3K4me2 recruit little LSD1, which is not detectable by ChIP-Seq (Q1, Figure 1b).  
424 When gene expression needs to be increased, recruitment of additional factors and/or MLL3/4 may  
425 convert these less active enhancers to more active enhancers with higher H3K4 methylation and  
426 H3K27ac, which would then require higher levels of LSD1. LSD1’s recruitment might also serve as a  
427 surveillance mechanism to suppress ectopic installation of H3K4 methylation and spurious activation  
428 of enhancers.

429         During differentiation of mES cells, the pluripotency enhancers may be unprimed by the loss  
430 of ES-specific TFs followed by the loss of MLL3/4 and H3K4 methylation. Our model is not mutually  
431 exclusive to LSD1-mediated decommissioning of enhancers, as LSD1 could remove remnant  
432 H3K4me2/me1 to completely disengage the enhancer from active regulation. To further elucidate the  
433 mechanisms of decommissioning of pluripotency enhancers, it will be important to determine how  
434 early differentiation cues shift the balance of MLL3/4-mediated H3K4 methylation and LSD1-  
435 mediated demethylation.

436         Enrichment of H3K4me1 and depletion of H3K4me3 was the first combination of chromatin  
437 signatures to predict a large number of transcriptional enhancers in a mammalian genome <sup>10, 11</sup>.  
438 More recent studies have shown H3K4me3 to be present at a subset of active enhancers <sup>83</sup> with a  
439 positive correlation between the H3K4me3/me1 ratio and enhancer transcription levels <sup>42</sup>. We found  
440 that KDM5C and LSD1 can co-occupy enhancers in mES cells (Figure 4a). However, only the loss of  
441 LSD1, but not KDM5C, displayed significant changes in enhancer activity and gene expression,  
442 highlighting an essential and non-redundant role of LSD1 in mES cells. KDM5C has been implicated  
443 in both promotion of enhancer activity by generating H3K4me1 in mES cells <sup>22</sup>, and suppression of  
444 over-activation of enhancers in breast cancer cells <sup>19</sup>. Consistent with the former study, our analysis  
445 found a small reduction in eRNA levels in *Kdm5c*-KO mESC (Supplementary Figure 9). In addition to  
446 *Kdm5c*, other KDM5 family members, *Kdm5a* and *Kdm5b*, are also expressed in mES cells at similar

447 levels and could possibly compensate for the loss of KDM5C. These observations suggest  
448 differential requirement of KDM5C in context of either different enhancers or different cell types.

449 Repeated “write-and-read” of histone modifications can form a feed-forward loop to allow TF-  
450 independent maintenance and propagation of chromatin status. Such models have been well  
451 established for the propagation of H3K27me3<sup>84</sup> and H3K9me3<sup>85</sup>. Maintenance and propagation of  
452 H3K4me epigenetic memory across generations by LSD1 have been observed in worms<sup>86</sup> and mice  
453<sup>87</sup>. PHF21A (aka BHC80), another member of the LSD1 complex<sup>50,58</sup>, was the first known reader  
454 protein to recognize unmethylated H3K4<sup>88</sup>. This unique combination of an H3K4 demethylase and a  
455 reader of unmethylated H3K4, makes the LSD1-PHF21A complex an ideal candidate to exert a self-  
456 perpetuating “erase-and-read” mechanism. However, a positive correlation between LSD1 and  
457 H3K4me2 and LSD1’s absence at latent enhancers suggest that the role of LSD1 in maintaining the  
458 epigenetic memory could be limited to other genomic elements and warrants further investigation.

459 LSD1’s role in suppression of primed enhancers does not appear to be restricted to mES cells.  
460 Similar to our observations in mESC, our BrU-Seq analyses in post-mitotic neurons revealed that  
461 LSD1 suppresses premature activation of neuronal activity-regulated genes and enhancers (Figure  
462 7). LSD1 has a neuron-specific isoform (neuroLSD1 or LSD1n) with four extra amino acids in the  
463 catalytic domain<sup>89</sup> and an altered substrate specificity which remains ambiguous<sup>49,90</sup>. Since our  
464 RNAi approach depleted both neuroLSD1 and canonical LSD1 in CN, it remains unclear if one or  
465 both of LSD1 isoforms mediate the suppression of activity-regulated enhancers and ARGs. Given  
466 that the genetic ablation of neuroLSD1 led to a downregulation of ARG expression<sup>49,91</sup>, it is more  
467 likely that the canonical LSD1, and not neuroLSD1, is involved in the suppression of activity-  
468 regulated enhancers. Loss-of-function *LSD1/KDM1A* mutations have been genetically associated  
469 with several neurodevelopmental conditions<sup>92-94</sup>. These disorders could possibly be attributed to  
470 uncontrolled activation of activity-regulated genes and enhancers upon the loss of LSD1 and/or  
471 neuroLSD1. LSD1-mediated homeostasis of transcriptional enhancers, therefore, underlies various  
472 physiological processes including embryonic development and human cognitive function.

473

474 **Methods**

475 See supplementary methods for details.

476 **Cell culture**

477 *Lsd1*-WT and *Lsd1*-GT mESC have been described previously<sup>28</sup>. *Kdm5c*-KO mESC were  
478 derived from the previously described mESC that carry the floxed *Kdm5c* allele<sup>75</sup> by Cre-mediated  
479 deletion of exons 11 and 12, which encode the enzymatic JmjC domain. mESC were grown on  
480 gelatin coated plates.

481 **Western blot analysis**

482 mESC or CN were lysed in Laemmli sample buffer, sonicated, and subjected to SDS-PAGE.  
483 Western blot analyses were carried out using standard protocols using anti-H3K4me1 (ab8895,  
484 Abcam), anti-H3K4me2 (ab7766, Abcam), anti-H3K4me3 (ab8580, Abcam), anti-H3K27ac (39135,  
485 Active Motif), anti-KDM5C<sup>75</sup>, anti-LSD1 (ab17721, Abcam)<sup>20</sup> and anti-phospho-ERK1/2 (4370, Cell  
486 Signaling Technology).

487 ***Lsd1* knockdown in mouse cortical neurons**

488 Primary cultures of cortical neurons were carried out as described previously<sup>75</sup>. *Lsd1*-KD in  
489 CN was achieved by lentiviral delivery of either scramble shRNA (SHC202, Sigma) or *Lsd1*-shRNAs  
490 (A: TRCN0000071375 and B: TRCN0000071376, Sigma)<sup>95</sup> on 7 days *in vitro* (DIV) and 5-  
491 Bromouridine incorporation was performed on DIV 11.

492 **ChIP-Seq**

493 Antibodies used for chromatin immunoprecipitation (ChIP) were anti-H3K4me1 (ab8895,  
494 Abcam and 07-436, EMD Millipore), anti-H3K4me2 (05-790, EMD Millipore), anti-H3K4me3 (04-745,  
495 EMD Millipore), anti-H3K27ac (39135, Active Motif), anti-HDAC1 (A300-713A, Bethyl Laboratories  
496 and sc-6298, Santa Cruz Biotechnology) and anti-KDM5C<sup>75</sup>. KDM5C ChIP-Seq experiments were  
497 performed as described previously<sup>75</sup>. Other ChIP experiments were performed as described  
498 previously<sup>96</sup> with minor modifications.

499 **RNA-Seq and Nuclear RNA-Seq**

500 RNA-Seq libraries have been described in detail previously<sup>97</sup>. For sequencing of nuclear  
501 RNA, nuclei were isolated as described previously<sup>98</sup> with minor modifications. Libraries from rRNA-  
502 depleted RNA were prepared using Direct Ligation of Adapters to First-strand cDNA (DLAF)<sup>97</sup>.

### 503 **Global Run-On**

504 GRO was modified from the method described previously<sup>47, 98</sup>. In addition to the presence of  
505 0.2% IGEPAL CA-630, GRO on *Kdm5c* mESC, *Lsd1* mESC and CN were done in presence of 0.5%,  
506 0.25% and 0.2% of *N*-Lauroylsarcosine, respectively for 8 min at 30°C.

### 507 **BrU-Seq**

508 Cortical neurons (DIV 11), after shRNA treatment for four days, were incubated with 2 mM 5-  
509 Bromouridine (850187, Sigma) for 32 min at 37°C. To reduce the number of steps for library  
510 preparation, we developed Direct Ligation of Adaptor to the 3' end of RNA (DLAR), a method  
511 suitable for preparation of libraries for BrU-Seq.

512 All sequencing experiments were conducted in biological duplicates concurrently with  
513 different genotypes to minimize technical variations.

### 514 **Sequencing and Alignment**

515 Multiplexed libraries were subjected to single-end sequencing on Illumina HiSeq 2000/2500  
516 instruments using standard oligonucleotides designed for multiplexed paired-end sequencing, except  
517 that BrU-Seq indices were sequenced with DLAR\_Index\_Read:5'-  
518 CATAGGAAGAGCACACGTCTGAACTCCAGTCAC-3'. ChIP-Seq reads were mapped to the mm9  
519 genome using Bowtie1 (v1.1.2)<sup>99</sup> allowing for up to two mismatches. PCR duplicates from ChIP-Seq  
520 reads were removed using samtools rmdup utility (v1.3)<sup>100</sup> and coverage along the genome was  
521 calculated using BEDTools (v2.25.0)<sup>101</sup> after extending the ChIP-Seq reads to a total length of 180  
522 bases. RNA-Seq libraries were mapped to the *mm9* genome and transcriptome using TopHat2  
523 (v2.1.0)<sup>102</sup> with Bowtie2 (v2.2.6)<sup>103</sup>. For GRO-Seq and BrU-Seq, full length reads were first aligned  
524 using Bowtie1 or Tophat2, respectively. Adaptor sequences were trimmed from the unmapped reads  
525 using BBDuk utility<sup>104</sup> and reads were remapped and merged to the reads from the initial alignment.

526 Only uniquely mapping reads were retained for further analysis and libraries were normalized to total  
527 number of non-mitochondrial and non-ribosomal reads.

## 528 **Analysis**

529 MACS2 (v 2.1.0)<sup>105</sup> was used to call DHS or ChIP-Seq peaks. For selection of candidate  
530 p300/DHS sites for enhancer prediction, we first scanned the genome for the strongest (with highest  
531 MACS2 signal) p300 or DHS site in a 1,250 base sliding window. When both p300 and DHS sites  
532 were present in the same window, p300 binding site was given higher precedence over any DHS  
533 sites. Intergenic p300/DHS sites were defined as sites that were outside of 1.25 kbp upstream to 3  
534 kbp downstream of the genes. LSD1 has been shown to be involved in silencing of repetitive  
535 elements including endogenous retroviral elements (ERVs)<sup>28</sup>. Therefore, to focus on prototypical  
536 enhancers in this study, we excluded p300/DHS sites with a low mappability ( $M_1 < 0.75$  and  $M_2 <$   
537  $0.75$ ), where  $M_1$  and  $M_2$  indicate the fraction of uniquely mapping bases<sup>106</sup> within  $\pm 500$  and  $\pm 100$   
538 bases, respectively, of the p300/DHS site. p300/DHS sites within the ENCODE blacklisted regions<sup>5</sup>  
539 were also excluded.

540 FeatureCounts<sup>107</sup> was used for calculating the number of reads overlapping various  
541 genomic features. Intersection analyses were done using BEDTools. DESeq (v1.22.1)<sup>63</sup> was used  
542 for normalization and differential gene expression analysis. 10-20 genes with exceptionally high  
543 expression and miRNAs and were excluded from further analysis. ChIP-Seq enrichment for  
544 H3K4me2, H3K4me3 and H3K27ac were normalized using MAnorm<sup>108</sup> with the fraction of reads  
545 aligning within the common peaks of WT and *Lsd1*-GT mESC samples from each replicate. ChIP-  
546 Seq coverage profiles from only one replicate, which did not require MAnorm normalization, were  
547 used in the browser snapshots in the figures. Prioritization of enhancer assignment is detailed in the  
548 supplemental information. Activity-regulated genes were identified as genes showing significant  
549 upregulation ( $p < 0.05$ , DESeq) in each of the two independent replicates of previously published  
550 RNA-Seq datasets from untreated and KCl-treated CN<sup>75</sup>. Wilcoxon signed-rank tests were  
551 performed after log transformation of changes in expression or ChIP enrichment. The Perl scripts  
552 used for analyses are available upon request.

553 **Accession Numbers**

554 Raw and processed sequence data files are available on the Gene Expression Omnibus  
555 (GEO) under accession GSE93952. The data can be accessed by the reviewers at:

556 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=ulchsiocxvoitaj&acc=GSE93952>

557

558 **Acknowledgements**

559 We thank Dr. Fulai Jin and Dr. Gary Hon at LICR, San Diego, for helpful discussions and  
560 guidance with bioinformatics. We are also grateful to Dr. Nima Mosammaparast at Washington  
561 University School of Medicine in St. Louis for the kind gift of WT and catalytically inactive *LSD1*  
562 expression plasmids. The work was supported by grants from the LICR (to BR), the California  
563 Institute for Regenerative Medicine (RN2-00905, to BR), the University of Michigan Medical School  
564 (to SI), Cooley's Anemia Foundation Fellowship (to SI), NIH (NS089896, to SI), the Farrehi research  
565 fund (to SI), NSF Graduate Research Fellowship Program (DGE #1256260, to PMG), University of  
566 Michigan Career Training in Reproductive Biology (T32HD079342, to RSP), and the Eunice  
567 Kennedy Shriver National Institute of Child Health and Human Development DIR (HD008933, to  
568 TSM).

569 **Author contributions**

570 SA and BR conceived the project and designed the experiments. SI and SA performed the  
571 RNA-Seq, Nuclear RNA-Seq, and GRO-Seq experiments. SA, EB, and TSM performed the CHIP-  
572 Seq experiments. SI, PMG, and SA designed and performed the BrU-Seq experiment. SA and RSP  
573 performed the luciferase reporter assays. YMN, PMG, and SA performed the western blot analyses.  
574 SA established the protocols and performed the computational analysis. SA and SI designed the  
575 analysis and wrote the manuscript. All authors approved and edited the manuscript.

576

577 **Competing financial interests**

578 The authors declare no competing financial interests.

579

580 **Figure legends**

581 **Figure 1. LSD1 occupies a large fraction of primed enhancers in mES cells.** (a) Overlap of  
582 binding sites of p300, CTCF, and LSD1 in mESC. (b) Intergenic enhancers were divided into  
583 quartiles (Q1-Q4) based on the enrichment of H3K4me2 relative to H3K4me1 (left panel). Boxplots  
584 of enrichment of indicated histone modifications, LSD1, and p300, as measured by ChIP-Seq, and  
585 eRNA levels (GRO-Seq, Nuclear RNA-Seq, and RNA-Seq) at each quartile of intergenic enhancers.  
586 Levels of LSD1 show positive correlations with increases in H3K4me2 and eRNA expression from  
587 Q1 to Q4. (c) The percentage of intergenic enhancers with LSD1 peaks. (d, e) LSD1 occupancy at  
588 active, poised, and intermediate enhancers classified based on enrichment of either trimethylation or  
589 acetylation of H3K27 (d) or H3K9 (e). LSD1 occupancy at enhancers increases with higher activity.  
590 In all figures, the bottom and top boxes signify the second and third quartiles, respectively, and the  
591 middle band represents the median of the population. Whiskers represent 1.5 times the inter-quartile  
592 range (IQR) and the notch represents the 95% confidence interval of the median.

593  
594 **Figure 2. LSD1 rarely binds to cell-type specific “latent” enhancers.** (a) UCSC genome browser  
595 snapshot of the *mir290* locus. LSD1 ChIP-Seq peaks in mESC coincide with enhancer signatures,  
596 including DHS, H3K4me1, and divergent GRO-Seq signals, upstream of the *mir290* promoter in  
597 mESC but not in neuronal cells. (b) LSD1 binding at the *Npas4* locus in mESC and NSC. LSD1 is  
598 present at the two brain-specific DHS sites (marked by asterisks) in neuronal cells but not in mESC.  
599 Note that the DHS sites common in adult brain and mESC are occupied by LSD1 in both mESC and  
600 NSC. NSC: Neural stem cells, CN: Cortical neurons. CTCF, DHS and LSD1 tracks were generated  
601 from previously published datasets. (c) Fraction of LSD1 peaks overlapping with mESC hotspots. (d)  
602 Fraction of mESC-specific hotspots overlapping with mESC-specific LSD1 peaks. (e) Fractions of  
603 tissue-specific hotspots overlapping with LSD1 peaks with no mESC-derived DHS.

604  
605 **Figure 3. Loss of LSD1 results in increases in H3K4 methylation and H3K27 acetylation at**  
606 **enhancers.** (a) H3K4me1, H3K4me2, H3K4me3, H3K27ac and HDAC1 levels on LSD1-bound

607 enhancers in WT (gray boxes) and *Lsd1*-GT mESC (red boxes). Enhancers were classified into  
608 poised (P), intermediate (I), and active (A) enhancers based on the enrichment of either H3K27ac or  
609 H3K27me3. Geometric mean of ChIP:Input ratios from the two independent ChIP-Seq replicates are  
610 shown. *P*-values (*p*) from Wilcoxon signed-rank tests on differences,  $\log_2(Lsd1\text{-GT}/WT)$ , are  
611 denoted in blue beneath each panel. *n* indicates the number of enhancers in each category. (b)  
612 Dysregulation of active enhancers at the *Pou5f1* (aka *Oct4*) locus. A cluster of enhancers is co-  
613 occupied by p300 and LSD1. Some of the individual enhancers show an increase in H3K4me2,  
614 H3K27ac, and GRO-Seq signals in *Lsd1*-GT mESC (red) compared to WT mESC (gray). (c)  
615 Misregulation of a poised enhancer (red bar) at the *Cbln4* locus. This locus is decorated with a broad  
616 H3K27me3 domain, and shows elevation in H3K4me1, H3K4me2, and GRO-Seq signals upon the  
617 loss of LSD1. Gray bar: Predicted enhancer. Blue bar: significantly-upregulated enhancer in *Lsd1*-  
618 GT mESC compared to WT mESC based on changes in GRO-Seq signal (See Figure 4).

619

620 **Figure 4. Loss of LSD1 but not KDM5C results in aberrant activation of enhancers. (a)**

621 Fractions of intergenic enhancers bound by LSD1 and/or KDM5C in mESC. (b, d) Volcano plots of  
622 GRO-Seq signals at all intergenic enhancers from DESeq analysis. While the loss of LSD1 resulted  
623 in a large-scale increase in GRO-Seq signals at enhancers, deletion of KDM5C had a minimal  
624 impact. X-axis and Y-axis indicate the  $\log_2$  fold-change and significance, respectively of differential  
625 expression in WT and mutant mES cell lines. (c) Scatter plots of GRO-Seq levels at poised,  
626 intermediate, and active enhancers classified of the basis of enrichment of either H3K27me3 or  
627 H3K27ac. Significantly-upregulated and -downregulated enhancers ( $q < 0.05$ , DESeq) are shown in  
628 blue and orange, respectively. Red curve indicates the LOWESS curve for each class of enhancers.  
629 Total number (*n*) of all, significantly-upregulated, and -downregulated enhancers in each group are  
630 indicated in black, blue, and orange, respectively. Each class of enhancers shows a significant  
631 upregulation ( $p < 2.2e-16$ , Wilcoxon signed-rank test) in *Lsd1*-GT mESC compared to WT mESC.

632



633 **Figure 5. Aberrant changes in enhancer activity are associated with misregulation of**  
634 **physically-interacting genes.** (a) An example of long-range promoter-enhancer interactions (top  
635 track) obtained from the mESC HiCap data set<sup>67</sup> at the *Dusp5* locus. One of the three significantly-  
636 upregulated enhancers (blue bars) interacts with the *Dusp5* promoter. Upon the loss of LSD1, the  
637 gene and the enhancers show upregulation of H3K4me2, H3K27ac and GRO-Seq signals in *Lsd1*-  
638 GT mESC (red) compared to WT mESC (gray). (b) Volcano plots of changes in mRNA levels (RNA-  
639 Seq) of genes that physically interact with misregulated enhancers. Based on changes in enhancer-  
640 associated GRO-Seq signals upon the loss of LSD1, enhancers were subdivided as significantly up  
641 ( $q < 0.05$ , DESeq), significantly down, moderately up ( $0.05 \leq q < 0.25$ ), moderately down, and the  
642 rest. When multiple enhancers showed interactions with a single promoter, assignment of the genes  
643 to each enhancer subgroup was prioritized in the aforementioned order. Total number of associated  
644 genes ( $n$ ) and  $p$ -values ( $p$ ) from Wilcoxon signed-rank tests on differences between mRNA levels in  
645 *Lsd1*-GT and WT mESC are shown beneath each panel. Note that more genes anchored to  
646 upregulated enhancers are upregulated compared to genes that interact with downregulated  
647 enhancers.

648

649 **Figure 6. Both mESC-specific and differentiation genes are upregulated upon the loss of**  
650 **LSD1 in undifferentiated mES cells.** (a) Schematic showing the number of significantly “induced”  
651 and “decommissioned” genes upon differentiation of mESC to epiblast stem cells with Activin A and  
652 FGF2<sup>69</sup>. (b, c) Scatter plots of mRNA levels (b) and levels of nascent transcription (c), as measured  
653 by RNA-Seq and GRO-Seq, respectively in WT and *Lsd1*-GT mESC. Number ( $n$ ) of significantly-  
654 upregulated ( $q < 0.05$ ) and -downregulated genes in each category are shown in blue and orange,  
655 respectively. Upon the loss of LSD1 in mESC, both groups of “induced” and “decommissioned”  
656 genes show a significant increase ( $p < 2.2e-16$ , Wilcoxon signed-rank test) in mRNA levels and  
657 nascent transcription. (d) Elevated transcription of *Hmga2*, a differentiation gene that plays an  
658 important role in exit of mES cells from the ground pluripotency state<sup>70</sup>, and its nearby enhancers in

659 *Lsd1*-GT mESC. Gray bar: Predicted enhancer. Blue bar: significantly-upregulated enhancer in  
660 *Lsd1*-GT mESC.

661

662 **Figure 7. LSD1 is required for suppression of inducible enhancers in terminally-differentiated**

663 **neurons.** (a) Western blot analysis to confirm the knockdown (KD) of LSD1 in mouse cortical  
664 neurons (CN) at DIV 11, after 4 days of lentiviral-mediated delivery of either scrambled shRNA  
665 (control) or two independent *Lsd1* shRNAs (A and B). (b) Upregulation of activity-regulated genes  
666 (ARGs) in *Lsd1*-KD CN. Scatter plots of transcription levels of ARGs ( $n=140$ ), from BrU-Seq analysis,  
667 in CN treated with either *Lsd1* shRNAs (Y-axis) or control shRNA (X-axis). Significantly-upregulated  
668 ( $q < 0.05$ , DESeq) and -downregulated ARGs are shown in blue and orange, respectively, and ARGs  
669 displaying greater than a 2-fold difference upon the loss of LSD1 are labeled with gene symbols. *P*-  
670 values ( $p$ ) from Wilcoxon signed-rank tests are denoted in blue. (c) Aberrant induction of *Npas4*, an  
671 ARG, upon *Lsd1*-KD in resting CN. Boxed P: *Npas4* Promoter. Boxed E: Putative activity-regulated  
672 enhancers as evident from presence of DHS<sup>32</sup>, high H3K4me1, low H3K4me3<sup>75</sup>, activity-dependent  
673 binding of NPAS4, and an increase in H3K27ac after KCl treatment<sup>76</sup>. *Npas4* mRNA and eRNA are  
674 upregulated specifically in the *Lsd1*-KD neurons (red) (d) Increased eRNA levels at activity-regulated  
675 enhancers in *Lsd1*-KD CN. These enhancers have been previously divided into four groups based  
676 on the activity-induced changes in H3K27ac<sup>76</sup>. Three groups of enhancers showed a significant  
677 increase in eRNA levels upon *Lsd1*-KD (red boxes) compared to control conditions (untreated CN or  
678 control shRNA-treated CN, gray boxes). A+B: Geometric mean of eRNA levels in CN treated with  
679 either *Lsd1* shRNAs A or B. U+C: Geometric mean of eRNA levels in control neurons. *P*-values ( $p$ )  
680 from Wilcoxon signed-rank tests are denoted in blue beneath each panel.

681

682 **References**

- 683 1. Banerji, J., Rusconi, S. & Schaffner, W. Expression of a beta-globin gene is enhanced by  
684 remote SV40 DNA sequences. *Cell* **27**, 299-308 (1981).  
685 2. Moreau, P., *et al.* The SV40 72 base repair repeat has a striking effect on gene expression  
686 both in SV40 and other chimeric recombinants. *Nucleic Acids Res* **9**, 6047-6068 (1981).

- 687 3. Bulger, M. & Groudine, M. Functional and mechanistic diversity of distal transcription  
688 enhancers. *Cell* **144**, 327-339 (2011).
- 689 4. Levine, M., Cattoglio, C. & Tjian, R. Looping back to leap forward: transcription enters a new  
690 era. *Cell* **157**, 13-25 (2014).
- 691 5. Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature*  
692 **489**, 57-74 (2012).
- 693 6. Maurano, M. T., *et al.* Systematic localization of common disease-associated variation in  
694 regulatory DNA. *Science* **337**, 1190-1195 (2012).
- 695 7. Smith, E. & Shilatifard, A. Enhancer biology and enhanceropathies. *Nat Struct Mol Biol* **21**,  
696 210-219 (2014).
- 697 8. Long, H. K., Prescott, S. L. & Wysocka, J. Ever-Changing Landscapes: Transcriptional  
698 Enhancers in Development and Evolution. *Cell* **167**, 1170-1187 (2016).
- 699 9. Ren, B. & Yue, F. Transcriptional Enhancers: Bridging the Genome and Phenome. *Cold Spring*  
700 *Harb Symp Quant Biol* **80**, 17-26 (2015).
- 701 10. Heintzman, N. D., *et al.* Distinct and predictive chromatin signatures of transcriptional  
702 promoters and enhancers in the human genome. *Nat Genet* **39**, 311-318 (2007).
- 703 11. Heintzman, N. D., *et al.* Histone modifications at human enhancers reflect global cell-type-  
704 specific gene expression. *Nature* **459**, 108-112 (2009).
- 705 12. Zaret, K. S. & Carroll, J. S. Pioneer transcription factors: establishing competence for gene  
706 expression. *Genes Dev* **25**, 2227-2241 (2011).
- 707 13. Spitz, F. & Furlong, E. E. Transcription factors: from enhancer binding to developmental  
708 control. *Nat Rev Genet* **13**, 613-626 (2012).
- 709 14. Herz, H. M., *et al.* Enhancer-associated H3K4 monomethylation by Trithorax-related, the  
710 *Drosophila* homolog of mammalian Mll3/Mll4. *Genes Dev* **26**, 2604-2620 (2012).
- 711 15. Lee, J. E., *et al.* H3K4 mono- and di-methyltransferase MLL4 is required for enhancer  
712 activation during cell differentiation. *Elife* **2**, e01503 (2013).
- 713 16. Hu, D. Q., *et al.* The MLL3/MLL4 Branches of the COMPASS Family Function as Major  
714 Histone H3K4 Monomethylases at Enhancers. *Molecular and Cellular Biology* **33**, 4745-4754  
715 (2013).
- 716 17. Creighton, M. P., *et al.* Histone H3K27ac separates active from poised enhancers and  
717 predicts developmental state. *Proc Natl Acad Sci U S A* **107**, 21931-21936 (2010).
- 718 18. Rada-Iglesias, A., *et al.* A unique chromatin signature uncovers early developmental  
719 enhancers in humans. *Nature* **470**, 279-283 (2011).
- 720 19. Shen, H., *et al.* Suppression of Enhancer Overactivation by a RACK7-Histone Demethylase  
721 Complex. *Cell* **165**, 331-342 (2016).
- 722 20. Whyte, W. A., *et al.* Enhancer decommissioning by LSD1 during embryonic stem cell  
723 differentiation. *Nature* **482**, 221-225 (2012).
- 724 21. Ostuni, R., *et al.* Latent enhancers activated by stimulation in differentiated cells. *Cell* **152**,  
725 157-171 (2013).
- 726 22. Outchkourov, N. S., *et al.* Balancing of histone H3K4 methylation states by the Kdm5c/SMCX  
727 histone demethylase modulates promoter and enhancer function. *Cell Rep* **3**, 1071-1079  
728 (2013).
- 729 23. Iwase, S., *et al.* The X-linked mental retardation gene SMCX/JARID1C defines a family of  
730 histone H3 lysine 4 demethylases. *Cell* **128**, 1077-1088 (2007).
- 731 24. Tahiliani, M., *et al.* The histone H3K4 demethylase SMCX links REST target genes to X-linked  
732 mental retardation. *Nature* **447**, 601-605 (2007).
- 733 25. Shi, Y., *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1.  
734 *Cell* **119**, 941-953 (2004).
- 735 26. Adamo, A., *et al.* LSD1 regulates the balance between self-renewal and differentiation in  
736 human embryonic stem cells. *Nat Cell Biol* **13**, 652-659 (2011).
- 737 27. Yin, F., *et al.* LSD1 regulates pluripotency of embryonic stem/carcinoma cells through histone  
738 deacetylase 1-mediated deacetylation of histone H4 at lysine 16. *Mol Cell Biol* **34**, 158-179  
739 (2014).

- 740 28. Macfarlan, T. S., *et al.* Endogenous retroviruses and neighboring genes are coordinately  
741 repressed by LSD1/KDM1A. *Genes Dev* **25**, 594-607 (2011).
- 742 29. Shen, Y., *et al.* A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116-  
743 120 (2012).
- 744 30. Buecker, C., *et al.* Reorganization of enhancer patterns in transition from naive to primed  
745 pluripotency. *Cell Stem Cell* **14**, 838-853 (2014).
- 746 31. Handoko, L., *et al.* CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat*  
747 *Genet* **43**, 630-638 (2011).
- 748 32. Neph, S., *et al.* An expansive human regulatory lexicon encoded in transcription factor  
749 footprints. *Nature* **489**, 83-90 (2012).
- 750 33. Wang, Q., Carroll, J. S. & Brown, M. Spatial and temporal recruitment of androgen receptor  
751 and its coactivators involves chromosomal looping and polymerase tracking. *Mol Cell* **19**, 631-  
752 642 (2005).
- 753 34. Ghirlando, R. & Felsenfeld, G. CTCF: making the right connections. *Genes Dev* **30**, 881-891  
754 (2016).
- 755 35. Kim, T. H., *et al.* Analysis of the vertebrate insulator protein CTCF-binding sites in the human  
756 genome. *Cell* **128**, 1231-1245 (2007).
- 757 36. Dixon, J. R., *et al.* Topological domains in mammalian genomes identified by analysis of  
758 chromatin interactions. *Nature* **485**, 376-380 (2012).
- 759 37. Hnisz, D., Day, D. S. & Young, R. A. Insulated Neighborhoods: Structural and Functional Units  
760 of Mammalian Gene Control. *Cell* **167**, 1188-1200 (2016).
- 761 38. He, H. H., *et al.* Nucleosome dynamics define transcriptional enhancers. *Nat Genet* **42**, 343-  
762 347 (2010).
- 763 39. Heinz, S., *et al.* Simple combinations of lineage-determining transcription factors prime cis-  
764 regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589 (2010).
- 765 40. Visel, A., *et al.* ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**,  
766 854-858 (2009).
- 767 41. Zentner, G. E., Tesar, P. J. & Scacheri, P. C. Epigenetic signatures distinguish multiple  
768 classes of enhancers with distinct cellular functions. *Genome Res* **21**, 1273-1283 (2011).
- 769 42. Core, L. J., *et al.* Analysis of nascent RNA identifies a unified architecture of initiation regions  
770 at mammalian promoters and enhancers. *Nat Genet* **46**, 1311-1320 (2014).
- 771 43. Kim, T. K., *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature*  
772 **465**, 182-187 (2010).
- 773 44. De Santa, F., *et al.* A large fraction of extragenic RNA pol II transcription sites overlap  
774 enhancers. *PLoS Biol* **8**, e1000384 (2010).
- 775 45. Wyers, F., *et al.* Cryptic pol II transcripts are degraded by a nuclear quality control pathway  
776 involving a new poly(A) polymerase. *Cell* **121**, 725-737 (2005).
- 777 46. Preker, P., *et al.* RNA exosome depletion reveals transcription upstream of active human  
778 promoters. *Science* **322**, 1851-1854 (2008).
- 779 47. Core, L. J., Waterfall, J. J. & Lis, J. T. Nascent RNA sequencing reveals widespread pausing  
780 and divergent initiation at human promoters. *Science* **322**, 1845-1848 (2008).
- 781 48. Kerenyi, M. A., *et al.* Histone demethylase Lsd1 represses hematopoietic stem and progenitor  
782 cell signatures during blood cell maturation. *Elife* **2**, e00633 (2013).
- 783 49. Wang, J., *et al.* LSD1n is an H4K20 demethylase regulating memory formation via  
784 transcriptional elongation control. *Nat Neurosci* **18**, 1256-1264 (2015).
- 785 50. Hakimi, M. A., *et al.* A core-BRAF35 complex containing histone deacetylase mediates  
786 repression of neuronal-specific genes. *Proc Natl Acad Sci U S A* **99**, 7420-7425 (2002).
- 787 51. Jouneau, A., *et al.* Naive and primed murine pluripotent stem cells have distinct miRNA  
788 expression profiles. *RNA* **18**, 253-264 (2012).
- 789 52. Wang, Y., *et al.* LSD1 co-repressor Rcor2 orchestrates neurogenesis in the developing mouse  
790 brain. *Nat Commun* **7**, 10481 (2016).
- 791 53. John, S., *et al.* Chromatin accessibility pre-determines glucocorticoid receptor binding patterns.  
792 *Nat Genet* **43**, 264-268 (2011).

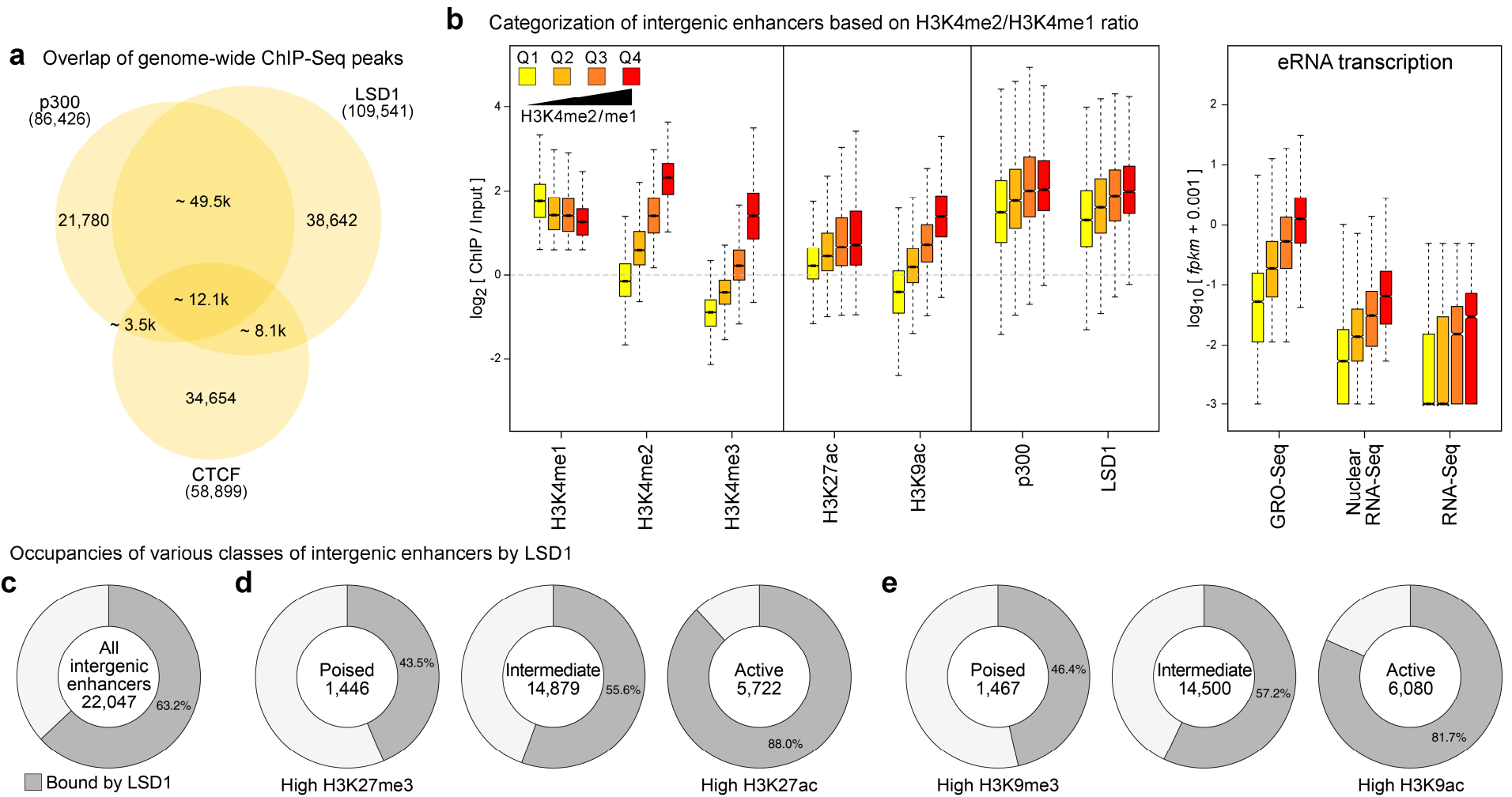
- 793 54. Metzger, E., *et al.* LSD1 demethylates repressive histone marks to promote androgen-  
794 receptor-dependent transcription. *Nature* **437**, 436-439 (2005).
- 795 55. Wissmann, M., *et al.* Cooperative demethylation by JMJD2C and LSD1 promotes androgen  
796 receptor-dependent gene expression. *Nat Cell Biol* **9**, 347-353 (2007).
- 797 56. Cai, C., *et al.* Lysine-specific demethylase 1 has dual functions as a major regulator of  
798 androgen receptor transcriptional activity. *Cell Rep* **9**, 1618-1627 (2014).
- 799 57. Metzger, E., *et al.* Phosphorylation of histone H3T6 by PKCbeta(I) controls demethylation at  
800 histone H3K4. *Nature* **464**, 792-796 (2010).
- 801 58. Shi, Y. J., *et al.* Regulation of LSD1 histone demethylase activity by its associated factors. *Mol*  
802 *Cell* **19**, 857-864 (2005).
- 803 59. Wang, Y., *et al.* LSD1 is a subunit of the NuRD complex and targets the metastasis programs  
804 in breast cancer. *Cell* **138**, 660-672 (2009).
- 805 60. Lee, M. G., *et al.* Functional interplay between histone demethylase and deacetylase enzymes.  
806 *Mol Cell Biol* **26**, 6395-6402 (2006).
- 807 61. Klose, R. J., Kallin, E. M. & Zhang, Y. JmjC-domain-containing proteins and histone  
808 demethylation. *Nat Rev Genet* **7**, 715-727 (2006).
- 809 62. Tsukada, Y., *et al.* Histone demethylation by a family of JmjC domain-containing proteins.  
810 *Nature* **439**, 811-816 (2006).
- 811 63. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol*  
812 **11**, R106 (2010).
- 813 64. Lieberman-Aiden, E., *et al.* Comprehensive mapping of long-range interactions reveals folding  
814 principles of the human genome. *Science* **326**, 289-293 (2009).
- 815 65. Fullwood, M. J. & Ruan, Y. ChIP-based methods for the identification of long-range chromatin  
816 interactions. *J Cell Biochem* **107**, 30-39 (2009).
- 817 66. Rao, S. S., *et al.* A 3D map of the human genome at kilobase resolution reveals principles of  
818 chromatin looping. *Cell* **159**, 1665-1680 (2014).
- 819 67. Sahlen, P., *et al.* Genome-wide mapping of promoter-anchored interactions with close to  
820 single-enhancer resolution. *Genome Biol* **16**, 156 (2015).
- 821 68. Lee, M. G., Wynder, C., Cooch, N. & Shiekhattar, R. An essential role for CoREST in  
822 nucleosomal histone 3 lysine 4 demethylation. *Nature* **437**, 432-435 (2005).
- 823 69. Acampora, D., *et al.* Loss of the Otx2-Binding Site in the Nanog Promoter Affects the Integrity  
824 of Embryonic Stem Cell Subtypes and Specification of Inner Cell Mass-Derived Epiblast. *Cell*  
825 *Rep* **15**, 2651-2664 (2016).
- 826 70. Navarra, A., *et al.* Hmga2 is necessary for Otx2-dependent exit of embryonic stem cells from  
827 the pluripotent ground state. *BMC Biol* **14**, 24 (2016).
- 828 71. Paulsen, M. T., *et al.* Coordinated regulation of synthesis and stability of RNA during the acute  
829 TNF-induced proinflammatory response. *Proc Natl Acad Sci U S A* **110**, 2240-2245 (2013).
- 830 72. Greer, P. L. & Greenberg, M. E. From synapse to nucleus: calcium-dependent gene  
831 transcription in the control of synapse development and function. *Neuron* **59**, 846-860 (2008).
- 832 73. Loebrich, S. & Nedivi, E. The function of activity-regulated genes in the nervous system.  
833 *Physiol Rev* **89**, 1079-1103 (2009).
- 834 74. Ebert, D. H. & Greenberg, M. E. Activity-dependent neuronal signalling and autism spectrum  
835 disorder. *Nature* **493**, 327-337 (2013).
- 836 75. Iwase, S., *et al.* A Mouse Model of X-linked Intellectual Disability Associated with Impaired  
837 Removal of Histone Methylation. *Cell Rep* **14**, 1000-1009 (2016).
- 838 76. Malik, A. N., *et al.* Genome-wide identification and characterization of functional neuronal  
839 activity-dependent enhancers. *Nat Neurosci* **17**, 1330-1339 (2014).
- 840 77. Wiegert, J. S. & Bading, H. Activity-dependent calcium signaling and ERK-MAP kinases in  
841 neurons: a link to structural plasticity of the nucleus and gene transcription regulation. *Cell*  
842 *Calcium* **49**, 296-305 (2011).
- 843 78. Wang, J., *et al.* Opposing LSD1 complexes function in developmental gene activation and  
844 repression programmes. *Nature* **446**, 882-887 (2007).

- 845 79. Karytinios, A., *et al.* A novel mammalian flavin-dependent histone demethylase. *J Biol Chem*  
846 **284**, 17775-17782 (2009).
- 847 80. Fang, R., *et al.* Human LSD2/KDM1b/AOF1 regulates gene transcription by modulating  
848 intragenic H3K4me2 methylation. *Mol Cell* **39**, 222-233 (2010).
- 849 81. Nair, V. D., *et al.* Involvement of histone demethylase LSD1 in short-time-scale gene  
850 expression changes during cell cycle progression in embryonic stem cells. *Mol Cell Biol* **32**,  
851 4861-4876 (2012).
- 852 82. Macfarlan, T. S., *et al.* Embryonic stem cell potency fluctuates with endogenous retrovirus  
853 activity. *Nature* **487**, 57-63 (2012).
- 854 83. Pekowska, A., *et al.* H3K4 tri-methylation provides an epigenetic signature of active enhancers.  
855 *EMBO J* **30**, 4198-4210 (2011).
- 856 84. Margueron, R., *et al.* Role of the polycomb protein EED in the propagation of repressive  
857 histone marks. *Nature* **461**, 762-767 (2009).
- 858 85. Rangunathan, K., Jih, G. & Moazed, D. Epigenetics. Epigenetic inheritance uncoupled from  
859 sequence-specific recruitment. *Science* **348**, 1258699 (2015).
- 860 86. Katz, D. J., Edwards, T. M., Reinke, V. & Kelly, W. G. A *C. elegans* LSD1 demethylase  
861 contributes to germline immortality by reprogramming epigenetic memory. *Cell* **137**, 308-320  
862 (2009).
- 863 87. Siklenka, K., *et al.* Disruption of histone methylation in developing sperm impairs offspring  
864 health transgenerationally. *Science* **350**, aab2006 (2015).
- 865 88. Lan, F., *et al.* Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated  
866 gene repression. *Nature* **448**, 718-722 (2007).
- 867 89. Zibetti, C., *et al.* Alternative splicing of the histone demethylase LSD1/KDM1 contributes to the  
868 modulation of neurite morphogenesis in the mammalian nervous system. *J Neurosci* **30**, 2521-  
869 2532 (2010).
- 870 90. Laurent, B., *et al.* A specific LSD1/KDM1A isoform regulates neuronal differentiation through  
871 H3K9 demethylation. *Mol Cell* **57**, 957-970 (2015).
- 872 91. Rusconi, F., *et al.* LSD1 modulates stress-evoked transcription of immediate early genes and  
873 emotional behavior. *Proc Natl Acad Sci U S A* **113**, 3651-3656 (2016).
- 874 92. Tunovic, S., Barkovich, J., Sherr, E. H. & Slavotinek, A. M. De novo ANKRD11 and KDM1A  
875 gene mutations in a male with features of KBG syndrome and Kabuki syndrome. *Am J Med*  
876 *Genet A* **164**, 1744-1749 (2014).
- 877 93. Chong, J. X., *et al.* Gene discovery for Mendelian conditions via social networking: de novo  
878 variants in KDM1A cause developmental delay and distinctive facial features. *Genet Med*,  
879 (2015).
- 880 94. Pilotto, S., *et al.* LSD1/KDM1A mutations associated to a newly described form of intellectual  
881 disability impair demethylase activity and binding to transcription factors. *Hum Mol Genet*,  
882 (2016).
- 883 95. Moffat, J., *et al.* A lentiviral RNAi library for human and mouse genes applied to an arrayed  
884 viral high-content screen. *Cell* **124**, 1283-1298 (2006).
- 885 96. Li, Z., *et al.* A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells. *Proc*  
886 *Natl Acad Sci U S A* **100**, 8164-8169 (2003).
- 887 97. Agarwal, S., Macfarlan, T. S., Sartor, M. A. & Iwase, S. Sequencing of first-strand cDNA library  
888 reveals full-length transcriptomes. *Nat Commun* **6**, 6002 (2015).
- 889 98. Wang, D., *et al.* Reprogramming transcription by distinct classes of enhancers functionally  
890 defined by eRNA. *Nature* **474**, 390-394 (2011).
- 891 99. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment  
892 of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).
- 893 100. Li, H., *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079  
894 (2009).
- 895 101. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic  
896 features. *Bioinformatics* **26**, 841-842 (2010).

- 897 102. Kim, D., *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions,  
898 deletions and gene fusions. *Genome Biol* **14**, R36 (2013).
- 899 103. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**,  
900 357-359 (2012).
- 901 104. Bushnell, B. BBMap short read aligner, and other bioinformatic tools.  
902 <https://sourceforge.net/projects/bbmap/>. (ed<sup>^</sup>(eds) (2014).
- 903 105. Zhang, Y., *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
- 904 106. Derrien, T., *et al.* Fast computation and applications of genome mappability. *PLoS One* **7**,  
905 e30377 (2012).
- 906 107. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for  
907 assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
- 908 108. Shao, Z., *et al.* MAnorm: a robust model for quantitative comparison of ChIP-Seq data sets.  
909 *Genome Biol* **13**, R16 (2012).

910

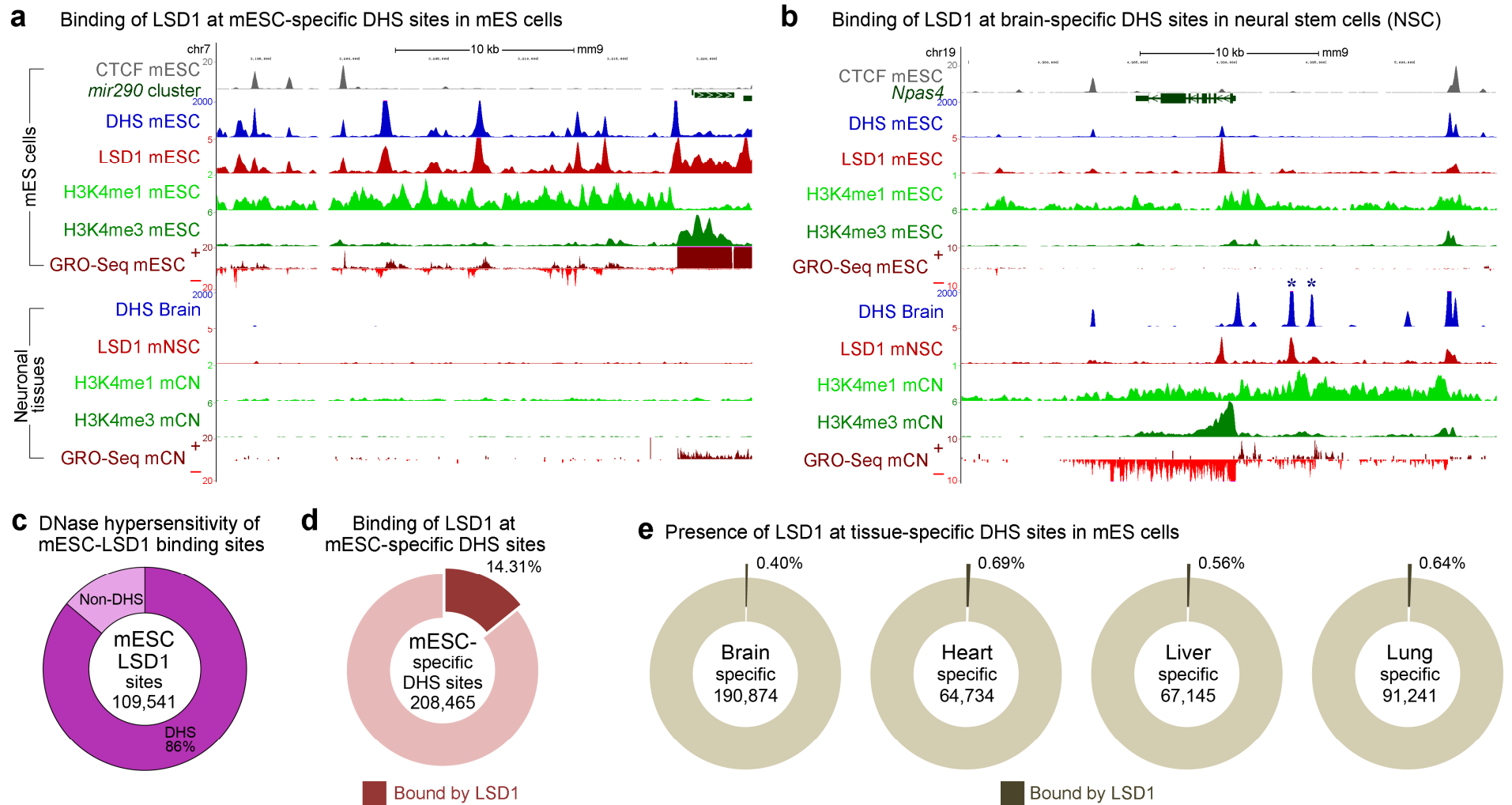
**Figure 1.**



**Figure 1. LSD1 occupies a large fraction of primed enhancers in mES cells.** (a) Overlap of binding sites of p300, CTCF, and LSD1 in mESC. (b) Intergenic enhancers were divided into quartiles (Q1-Q4) based on the enrichment of H3K4me2 relative to H3K4me1 (left panel). Boxplots of enrichment of indicated histone modifications, LSD1, and p300, as measured by ChIP-Seq, and eRNA levels (GRO-Seq, Nuclear RNA-Seq, and RNA-Seq) at each quartile of intergenic enhancers. Levels of LSD1 show positive correlations with increases in H3K4me2 and eRNA expression from Q1 to Q4. (c) The percentage of intergenic enhancers with LSD1 peaks. (d, e) LSD1 occupancy at active, poised, and intermediate enhancers classified based on enrichment of either trimethylation or acetylation of H3K27 (d) or H3K9 (e). LSD1 occupancy at enhancers increases with higher activity. In all figures, the bottom and top boxes signify the second and third quartiles, respectively, and the middle band represents the median of the population. Whiskers represent 1.5 times the inter-quartile range (IQR) and the notch represents the 95% confidence interval of the median.

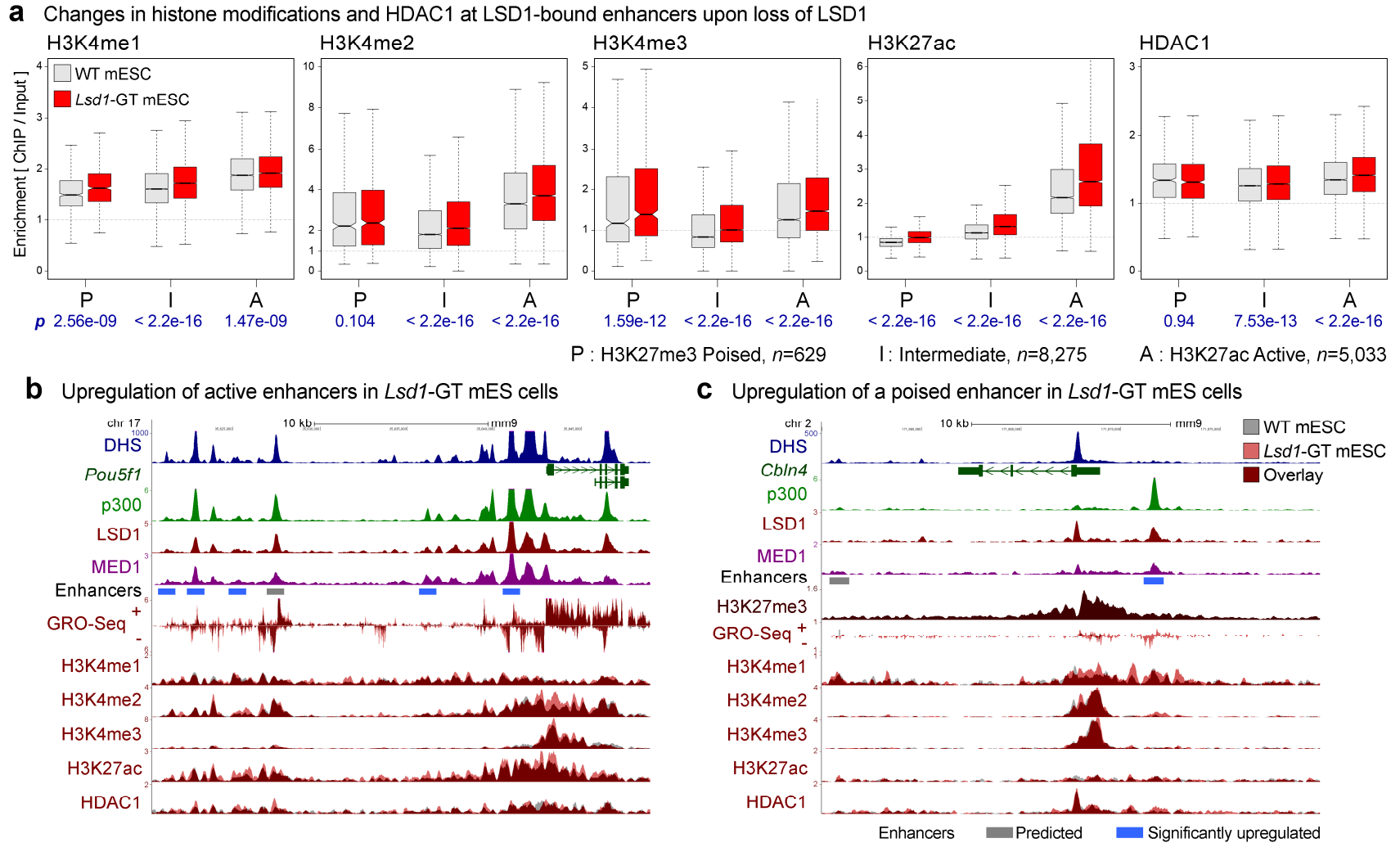


**Figure 2.**



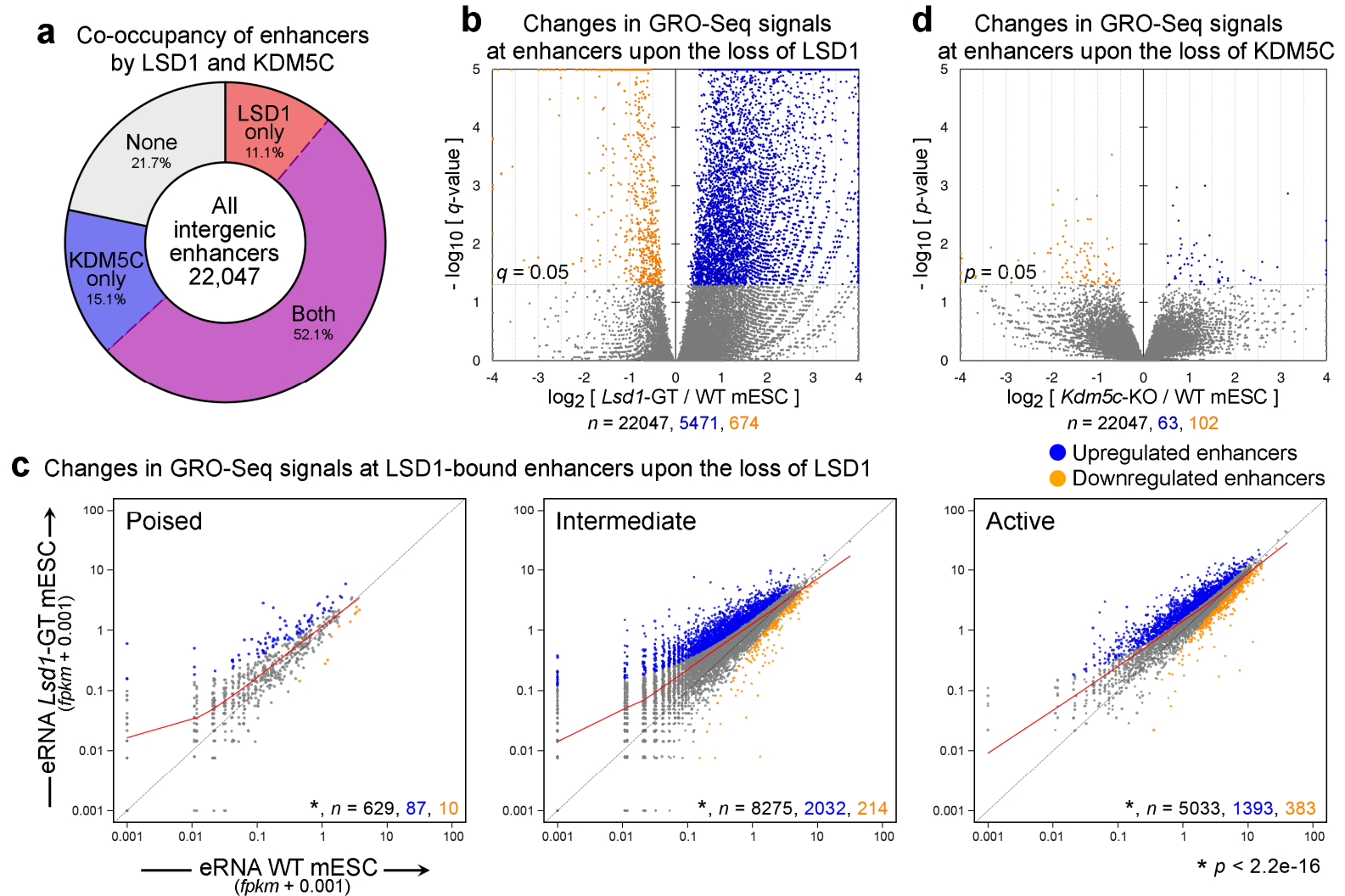
**Figure 2. LSD1 rarely binds to cell-type specific “latent” enhancers.** (a) UCSC genome browser snapshot of the *mir290* locus. LSD1 ChIP-Seq peaks in mESC coincide with enhancer signatures, including DHS, H3K4me1, and divergent GRO-Seq signals, upstream of the *mir290* promoter in mESC but not in neuronal cells. (b) LSD1 binding at the *Npas4* locus in mESC and NSC. LSD1 is present at the two brain-specific DHS sites (marked by asterisks) in neuronal cells but not in mESC. Note that the DHS sites common in adult brain and mESC are occupied by LSD1 in both mESC and NSC. NSC: Neural stem cells, CN: Cortical neurons. CTCF, DHS and LSD1 tracks were generated from previously published datasets. (c) Fraction of LSD1 peaks overlapping with mESC hotspots. (d) Fraction of mESC-specific hotspots overlapping with mESC-specific LSD1 peaks. (e) Fractions of tissue-specific hotspots overlapping with LSD1 peaks with no mESC-derived DHS.

**Figure 3.**



**Figure 3. Loss of LSD1 results in increases in H3K4 methylation and H3K27 acetylation at enhancers.** (a) H3K4me1, H3K4me2, H3K4me3, H3K27ac and HDAC1 levels on LSD1-bound enhancers in WT (gray boxes) and *Lsd1*-GT mESC (red boxes). Enhancers were classified into poised (P), intermediate (I), and active (A) enhancers based on the enrichment of either H3K27ac or H3K27me3. Geometric mean of ChIP:Input ratios from the two independent ChIP-Seq replicates are shown. *P*-values (*p*) from Wilcoxon signed-rank tests on differences,  $\log_2(Lsd1\text{-GT}/WT)$ , are denoted in blue beneath each panel. *n* indicates the number of enhancers in each category. (b) Dysregulation of active enhancers at the *Pou5f1* (aka *Oct4*) locus. A cluster of enhancers is co-occupied by p300 and LSD1. Some of the individual enhancers show an increase in H3K4me2, H3K27ac, and GRO-Seq signals in *Lsd1*-GT mESC (red) compared to WT mESC (gray). (c) Misregulation of a poised enhancer (red bar) at the *Cbln4* locus. This locus is decorated with a broad H3K27me3 domain, and shows elevation in H3K4me1, H3K4me2, and GRO-Seq signals upon the loss of LSD1. Gray bar: Predicted enhancer. Blue bar: significantly-upregulated enhancer in *Lsd1*-GT mESC compared to WT mESC based on changes in GRO-Seq signal (See Figure 4).

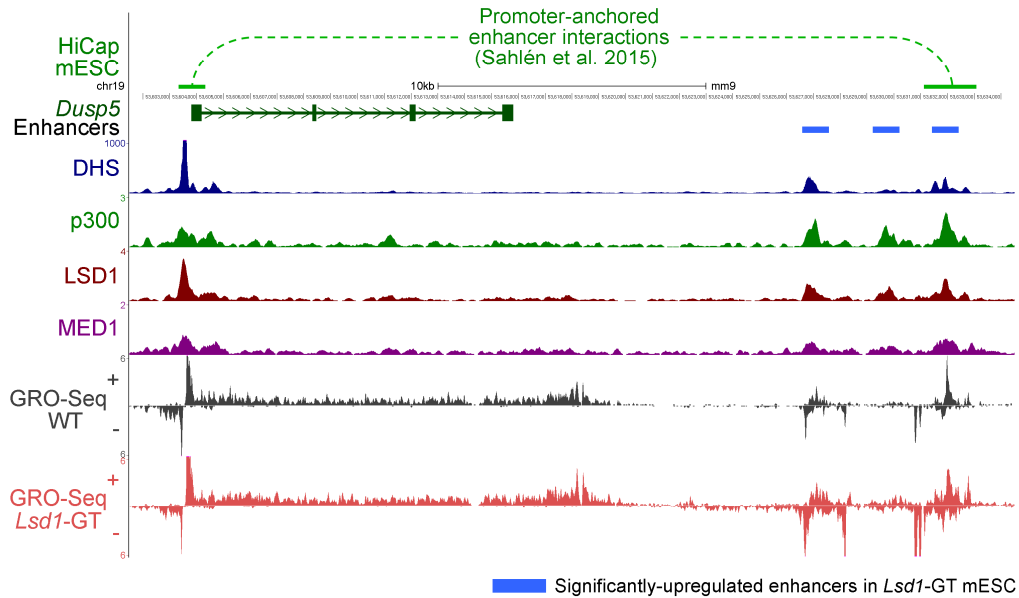
**Figure 4.**



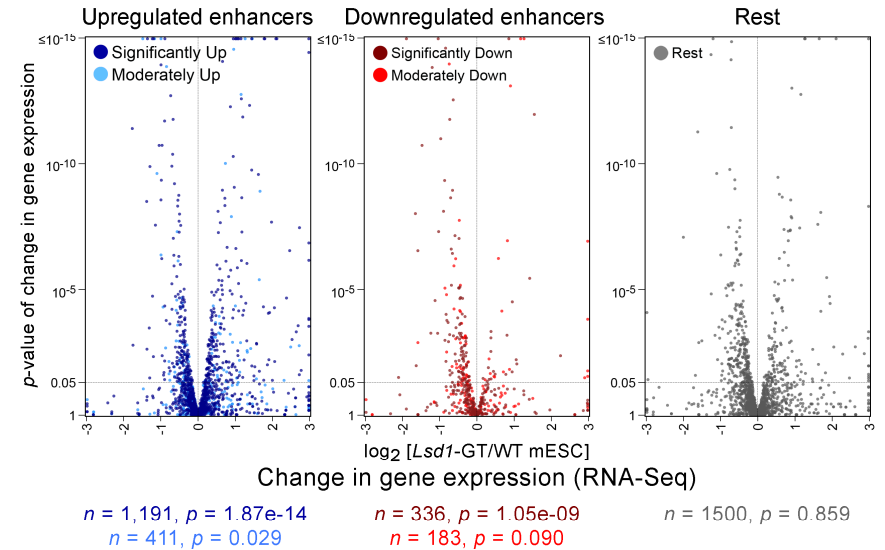
**Figure 4. Loss of LSD1 but not KDM5C results in aberrant activation of enhancers.** (a) Fractions of intergenic enhancers bound by LSD1 and/or KDM5C in mESC. (b, d) Volcano plots of GRO-Seq signals at all intergenic enhancers from DESeq analysis. While the loss of LSD1 resulted in a large-scale increase in GRO-Seq signals at enhancers, deletion of KDM5C had a minimal impact. X-axis and Y-axis indicate the log<sub>2</sub> fold-change and significance, respectively of differential expression in WT and mutant mES cell lines. (c) Scatter plots of GRO-Seq levels at poised, intermediate, and active enhancers classified on the basis of enrichment of either H3K27me<sub>3</sub> or H3K27ac. Significantly-upregulated and -downregulated enhancers ( $q < 0.05$ , DESeq) are shown in blue and orange, respectively. Red curve indicates the LOWESS curve for each class of enhancers. Total number ( $n$ ) of all, significantly-upregulated, and -downregulated enhancers in each group are indicated in black, blue, and orange, respectively. Each class of enhancers shows a significant upregulation ( $p < 2.2\text{e-}16$ , Wilcoxon signed-rank test) in *Lsd1*-GT mESC compared to WT mESC.

**Figure 5.**

**a** Concomitant changes in enhancer activity and gene expression at *Dusp5* locus

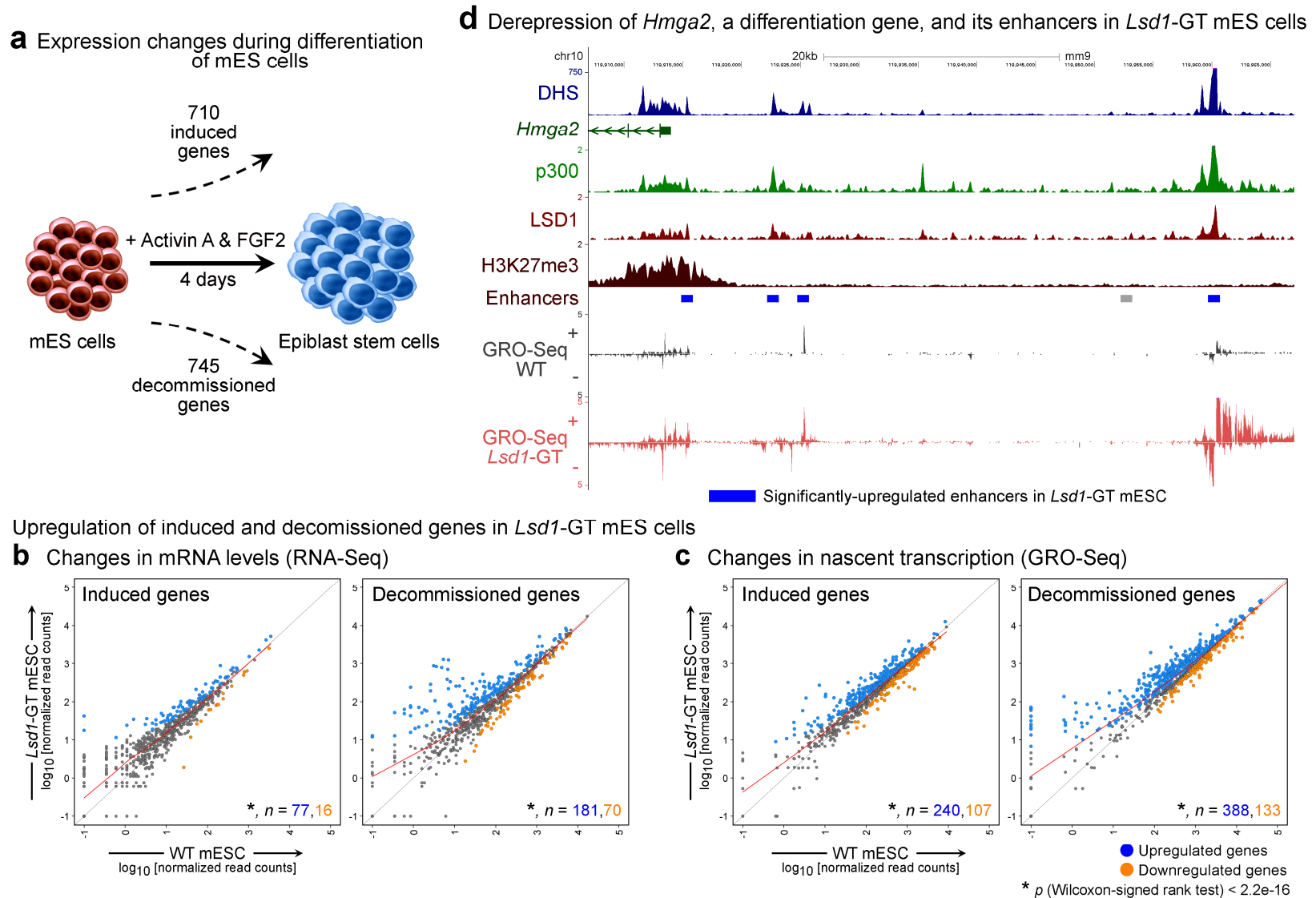


**b** Changes in mRNA levels of genes anchored to enhancers



**Figure 5. Aberrant changes in enhancer activity are associated with misregulation of physically-interacting genes.** (a) An example of long-range promoter-enhancer interactions (top track) obtained from the mESC HiCap data set<sup>67</sup> at the *Dusp5* locus. One of the three significantly-upregulated enhancers (blue bars) interacts with the *Dusp5* promoter. Upon the loss of LSD1, the gene and the enhancers show upregulation of H3K4me2, H3K27ac and GRO-Seq signals in *Lsd1-GT* mESC (red) compared to WT mESC (gray). (b) Volcano plots of changes in mRNA levels (RNA-Seq) of genes that physically interact with misregulated enhancers. Based on changes in enhancer-associated GRO-Seq signals upon the loss of LSD1, enhancers were subdivided as significantly up ( $q < 0.05$ , DESeq), significantly down, moderately up ( $0.05 \leq q < 0.25$ ), moderately down, and the rest. When multiple enhancers showed interactions with a single promoter, assignment of the genes to each enhancer subgroup was prioritized in the aforementioned order. Total number of associated genes ( $n$ ) and  $p$ -values ( $p$ ) from Wilcoxon signed-rank tests on differences between mRNA levels in *Lsd1-GT* and WT mESC are shown beneath each panel. Note that more genes anchored to upregulated enhancers are upregulated compared to genes that interact with downregulated enhancers.

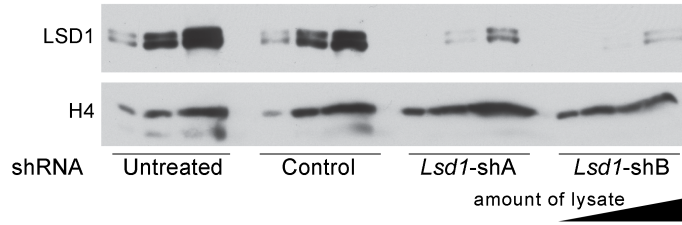
**Figure 6.**



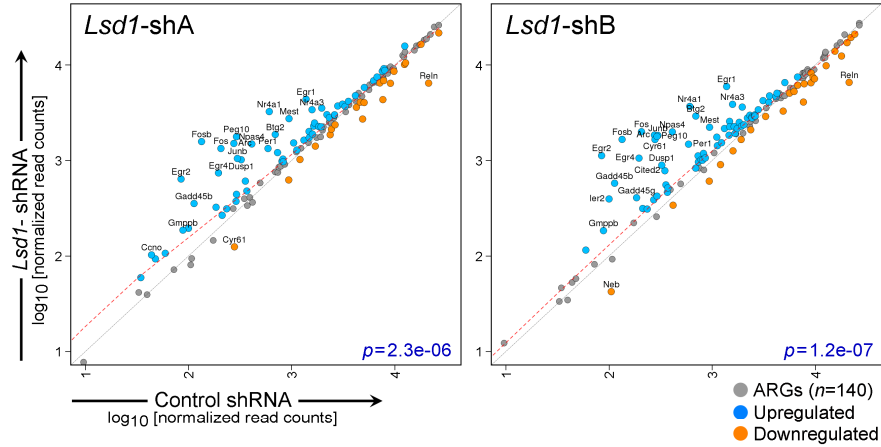
**Figure 6. Both mESC-specific and differentiation genes are upregulated upon the loss of LSD1 in undifferentiated mES cells.** (a) Schematic showing the number of significantly “induced” and “decommissioned” genes upon differentiation of mESC to epiblast stem cells with Activin A and FGF2<sup>69</sup>. (b, c) Scatter plots of mRNA levels (b) and levels of nascent transcription (c), as measured by RNA-Seq and GRO-Seq, respectively in WT and *Lsd1-GT* mESC. Number ( $n$ ) of significantly-upregulated ( $q < 0.05$ ) and -downregulated genes in each category are shown in blue and orange, respectively. Upon the loss of LSD1 in mESC, both groups of “induced” and “decommissioned” genes show a significant increase ( $p < 2.2e-16$ , Wilcoxon signed-rank test) in mRNA levels and nascent transcription. (d) Elevated transcription of *Hmga2*, a differentiation gene that plays an important role in exit of mES cells from the ground pluripotency state<sup>70</sup>, and its nearby enhancers in *Lsd1-GT* mESC. Grey bar: Predicted enhancer. Blue bar: significantly-upregulated enhancer in *Lsd1-GT* mESC.

**Figure 7.**

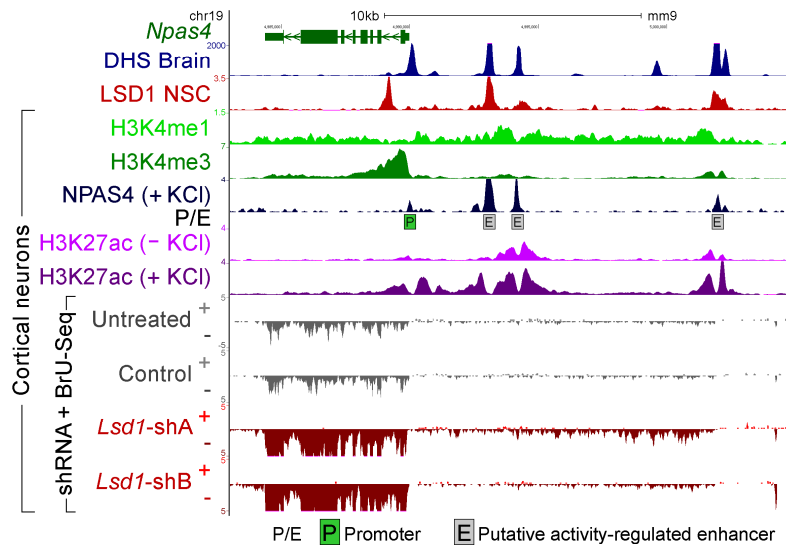
**a** Knock-down of LSD1 in mouse cortical neurons



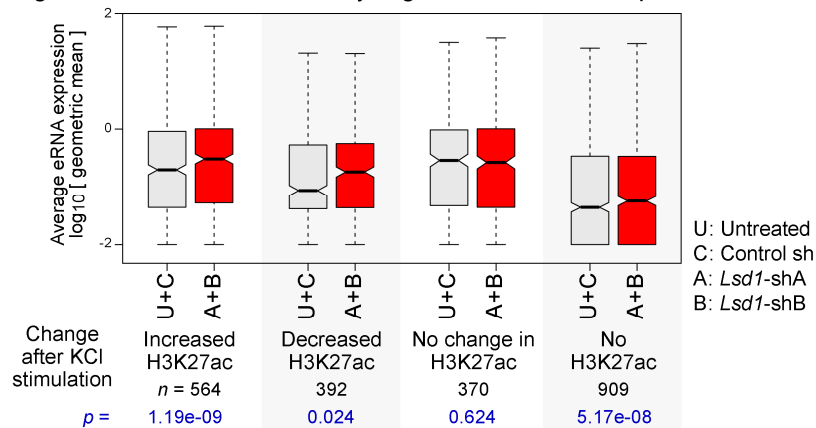
**b** Changes in expression levels of ARGs upon loss of LSD1



**c** Derepression of *Npas4* and nearby enhancers upon loss of LSD1



**d** Changes in eRNA levels of activity-regulated enhancers upon loss of LSD1



**Figure 7. LSD1 is required for suppression of inducible enhancers in terminally-differentiated neurons. (a)**

Western blot analysis to confirm the knockdown (KD) of LSD1 in mouse cortical neurons (CN) at DIV 11, after 4 days of lentiviral-mediated delivery of either scrambled shRNA (control) or two independent *Lsd1* shRNAs (A and B). **(b)** Upregulation of activity-regulated genes (ARGs) in *Lsd1*-KD CN. Scatter plots of transcription levels of ARGs ( $n=140$ ), from BrU-Seq analysis, in CN treated with either *Lsd1* shRNAs (Y-axis) or control shRNA (X-axis). Significantly-upregulated ( $q < 0.05$ , DESeq) and -downregulated ARGs are shown in blue and orange, respectively, and ARGs displaying greater than a 2-fold difference upon the loss of LSD1 are labeled with gene symbols. *P*-values ( $p$ ) from Wilcoxon signed-rank tests are denoted in blue. **(c)** Aberrant induction of *Npas4*, an ARG, upon *Lsd1*-KD in resting CN. Boxed P: *Npas4* Promoter. Boxed E: Putative activity-regulated enhancers as evident from presence of DHS<sup>32</sup>, high H3K4me1, low H3K4me3<sup>75</sup>, activity-dependent binding of NPAS4, and an increase in H3K27ac after KCl treatment<sup>76</sup>. *Npas4* mRNA and eRNA are upregulated specifically in the *Lsd1*-KD neurons (red) **(d)** Increased eRNA levels at activity-regulated enhancers in *Lsd1*-KD CN. These enhancers have been previously divided into four groups based on the activity-induced changes in H3K27ac<sup>76</sup>. Three groups of enhancers showed a significant increase in eRNA levels upon *Lsd1*-KD (red boxes) compared to control conditions (untreated CN or control shRNA-treated CN, gray boxes). A+B: Geometric mean of eRNA levels in CN treated with either *Lsd1* shRNAs A or B. U+C: Geometric mean of eRNA levels in control neurons. *P*-values ( $p$ ) from Wilcoxon signed-rank tests are denoted in blue beneath each panel.