Oct4-mediated inhibition of Lsd1 activity promotes the active and primed state of pluripotency enhancers

Lama AlAbdi[1], Debapriya Saha[1], Ming He[1], Mohd Saleem Dar[1], Sagar M. Utturkar[2], Putu Ayu Sudyanti[3], Stephen McCune[1], Brice, H. Spears[1], James A. Breedlove[1], Nadia A. Lanman[2,4], and Humaira Gowher[1,2*]

[1]Department of Biochemistry; [2]Purdue University Center for Cancer Research; [3]Department of Statistics; [4]Department of Comparative Pathobiology; Purdue University; West Lafayette, Indiana 47907 USA

*Corresponding Author and Lead Contact: hgowher@purdue.edu

Running title: Loss of Lsd1 regulation in carcinoma cells

Summary

An aberrant increase in pluripotency gene (PpG) expression due to enhancer reactivation could induce stemness and enhance tumorigenicity of cancer stem cells. Silencing of PpG enhancers (PpGe) during embryonic stem cell differentiation involves Lsd1–mediated H3K4me1 demethylation and DNA methylation. Here, we observed retention of H3K4me1 and DNA hypomethylation at PpGe associated with a partial repression of PpGs in F9 embryonal carcinoma cells (ECCs) post-differentiation. H3K4me1 demethylation in F9 ECCs could not be rescued by Lsd1 overexpression. Given our observation that H3K4me1 demethylation is accompanied by strong Oct4 repression in P19 ECCs, we tested if Oct4 interaction with Lsd1 affects its catalytic activity. Our data show a dose-dependent inhibition of Lsd1 activity by Oct4 and retention of H3K4me1 at PpGe in Oct4 overexpressing P19 ECCs. These data suggest that Lsd1-Oct4 interaction in cancer stem cells could establish a *primed* enhancer state that is susceptible to reactivation leading to aberrant PpG expression.

Keywords: Pluripotency, enhancers, Dnmt3a, DNA methylation, embryonal carcinoma cells, Lsd1, histone demethylation, Oct4, cancer stem cells, enhancer priming.

Introduction

Cell type specific gene expression is regulated by chromatin conformation which facilitates the interaction of distally placed enhancer elements with the specific gene promoter (Banerji et al., 1981, Bulger and Groudine, 2011, Ong and Corces, 2011, Plank and Dean, 2014). Enhancers house the majority of transcription factor binding sites and amplify basal transcription, thus playing a critical role in signal dependent transcriptional responses (summarized in (Heinz et al., 2015)). Epigenome profiling combined with the transcriptional activity in various cell types led to identification of potential enhancers, which are annotated as silent, primed, or active based on their epigenetic features. These epigenetic features include histone modifications, and DNA

45 methylation (Ernst and Kellis, 2010, Ernst et al., 2011, Calo and Wysocka, 2013). Whereas
46 histone H3K4me1 (monomethylation) and H3K4me2 (dimethylation) is present at both active and
47 primed enhancers, active enhancers invariantly are marked by histone H3K27Ac (acetylation)
48 and/or transcribed to produced enhancer RNA (eRNA) (Heintzman et al., 2007, Heinz et al., 2010,
49 Rada-Iglesias et al., 2011, Creyghton et al., 2010, Zentner et al., 2011, Zhu et al., 2013b).

50 During embryonic stem cell (ESC) differentiation, pluripotency gene (PpG) specific enhancers are
51 silenced via changes in histone modifications and a gain of DNA methylation (Whyte et al., 2012,
52 Mendenhall et al., 2013, Petell et al., 2016). In response to the differentiation signal, the
53 coactivator complex (Oct4, Sox2, Nanog and Mediator complex) dissociates from the enhancer
54 followed by the activation of pre-bound Lsd1-Mi2/NuRD enzymes. The histone demethylase Lsd1
55 demethylates H3K4me1, and the HDAC activity of the NuRD complex deacetylates H3K27Ac
56 (Whyte et al., 2012). Our previous studies have shown that the histone demethylation event is
57 critical for the activation of DNA methyltransferase Dnmt3a, which interacts with the demethylated
58 histone H3 tails through its chromatin-interacting ADD (ATRX-Dnmt3a-Dnmt3L) domain, allowing
59 site-specific methylation at pluripotency gene enhancers (PpGe) (Petell et al., 2016). These
60 findings were further supported by biochemical studies showing that Dnmt3a- ADD domain
61 interacts with histone H3 tail and this interaction is inhibited by H3K4 methylation (Guo et al.,
62 2015, Li et al., 2011a, Ooi et al., 2007, Otani et al., 2009), together suggesting that aberrant
63 inhibition of Lsd1 demethylase activity could cause a failure to gain DNA methylation, leading to
64 incomplete repression of PpGs.

65 Several studies have reported on potential mechanisms that control site specific targeting and
66 catalytic activity of Lsd1. Whereas Lsd1 interaction with CoREST activates the enzyme, BHC80
67 inhibits Lsd1 demethylation activity (Shi et al., 2005). The substrate specificity of Lsd1 is regulated
68 by its interaction with androgen receptor and estrogen related receptor α, or by alternative splicing
69 which adds four or eight amino acids to the Lsd1 enzyme (Carnesecchi et al., 2017, Metzger et

al., 2005, Laurent et al., 2015, Zibetti et al., 2010, Wang et al., 2015a). Lsd1 is targeted to various genomic regions through its interaction with SNAG domain containing TFs, such as Snail and GFI1B (McClellan et al., 2019, Vinyard et al., 2019). The SNAG domain binds to the active site of Lsd1 by mimicking the histone H3 tail and could potentially inhibit its activity (Baron et al., 2011). Interaction of the p53 C' terminal domain with the Lsd1 active site inhibits Lsd1 enzymatic activity (Speranzini et al., 2017). Lsd1 was also shown to be present in the Oct4 interaction network, and therefore could be targeted to Oct4-bound regulatory elements, which largely control pluripotency and stemness (van den Berg et al., 2010, Pardo et al., 2010, Ding et al., 2012).

Studies by the Cancer Genome Anatomy Project (CGAP) show that one out of three cancers express PpGs, suggesting their role in dysregulated proliferation during tumorigenesis (Zhang et al., 2013, Liu et al., 2013). Further, expression of PpGs, *Oct4*, *Sox2,* and *Nanog* potentiates self-renewal of putative cancer stem cells (CSCs) (Ben-Porath et al., 2008, Feske, 2007, Linn et al., 2010, Peng et al., 2010, Kumar et al., 2012, Wang et al., 2013, Mak et al., 2012, Wen et al., 2010, Jeter et al., 2011). CSCs proliferate as well as differentiate to give rise to cancer cells of various lineages (Iglesias et al., 2017). However, in order to retain the ability to proliferate, many cancer cells maintain expression of PpGs (Gwak et al., 2017, Yang et al., 2018). This has led to the development of terminal differentiation therapy, which aims to limit the proliferating cancer cell population (de The, 2018). Embryonal Carcinoma Cells (ECCs) have been used as a model cell line to study CSCs. ECCs were derived from developing mouse embryos at E6-7.5 and share regulatory characteristics with ESCs, including their ability to differentiate into various somatic lineages (Alonso et al., 1991, Han et al., 2017, Andrews et al., 2005, Zhu et al., 2013a). To understand the mechanism by which cancer cells retain PpG expression, we investigated the mechanism of enhancer-mediated regulation of PpG expression in ECCs. Our data showed that in differentiating F9 ECCs, the PpGs are only partially repressed. This was concomitant with H3K27 deacetylation, but with an absence of Lsd1-mediated H3K4me1 demethylation at PpGe.

95 The presence of H3K4me1 prevented Dnmt3a from methylating the DNA at these sites, potentially
96 abrogating PpGe silencing. Drug-mediated inhibition as well as overexpression of Lsd1 had little
97 or no effect on enhancer silencing and PpG repression, confirming an absence of Lsd1
98 dependence in differentiating ECCs. Given that Oct4 was expressed at substantial levels in F9
99 ECCs post-differentiation, we investigated the effect of Lsd1-Oct4 interaction on Lsd1 catalytic
100 activity. Using *in vitro* histone demethylation assays, we discovered that Lsd1-Oct4 interaction
101 inhibits Lsd1 activity, which could potentially result in the retention of H3K4me1 at PpGe in F9
102 ECCs. We tested this prediction in P19 ECCs in which Oct4 expression is strongly reduced post-
103 differentiation and H3K4me1 is demethylated at PpGe. The observation that overexpression of
104 Oct4 in differentiating P19 ECCs led to retention of H3K4me1 at PpGe, confirmed the role of Oct4-
105 mediated Lsd1 inhibition at these sites. Taken together our data show that inhibition of Lsd1
106 activity and Dnmt3a leads to the establishment of a "primed" enhancer state, which is open for
107 coactivator binding and prone to reactivation. We speculate that aberrant expression of Oct4 in
108 CSCs facilitates the establishment of "primed" enhancers, reactivation of which support
109 tumorigenicity.

110 **Results:**

111 **PpGs are partially repressed in differentiating F9 ECCs**

112 ECCs share many characteristics with ESCs, including mechanisms governing regulation of
113 gene expression and differentiation (Alonso et al., 1991). Based on the observation that
114 aberrant PpG expression is commonly found in cancers (Zhang et al., 2013, Liu et al., 2013), we
115 compared the magnitude of PpG repression in F9 ECCs with that in ESCs pre- and post-
116 differentiation. ESCs and F9 ECCs were induced to differentiate with retinoic acid (RA) and
117 expression of a subset of PpGs at 4 days (D4) post-induction was measured by RT-qPCR. In
118 ESCs, the expression of most PpGs was reduced by more than 80% post-differentiation. The
119 expression of *Sox2* and *Trim28* was maintained as an anticipated response to RA signaling

120 guiding ESCs towards neural lineage (Figures 1A). However, in differentiating F9 ECCs, several
121 PpGs were incompletely repressed, of which genes encoding the pioneer factors Oct4 showed
122 75% loss of expression whilst Nanog remain unchanged. A substantial increase in the
123 expression of the genes *Lefty1* and *Lefty2* in F9 ECCs suggests potential transient activation of
124 germ cell and testis developmental programs (Zhu et al., 2013a) (Figure 1B). To ensure that
125 incomplete repression of PpGs was not the consequence of delayed response to RA signaling,
126 we cultured cells for 8 days (D8). A variable decrease in PpGs was observed in the range of
127 25% - 75% with highest repression in *Lefty 2* (Figure 1B). Continued expression of PpGs in F9
128 ECCs was also evident by positive alkaline phosphatase staining and SSEA-1
129 immunofluorescence in differentiating F9 ECCs that is completely lost in ESCs post-
130 differentiation (Figures 1C and 1D). We asked if failure to exit pluripotency in F9 ECCs was
131 caused by an inability to activate lineage-specific genes. Our data showed a 5 to 60-fold
132 increase in the expression of the lineage specific genes *Gata4, Foxa2, Olig2, Gata6, Cxcr4, and*
133 *Fgf5* (Figure 1E), reflecting a standard response to differentiation signal. Given that previous
134 studies in ESCs have established a critical role of enhancer silencing incomplete PpG
135 repression, we next investigated if PpGe were fully decommissioned in F9 ECCs post-
136 differentiation (Petell et al., 2016).

137 **DNA methylation is not established at PpGe during F9 ECC differentiation**

138 In differentiating ESCs, PpGe silencing involves gain of DNA methylation, which is required for
139 complete PpG repression (Petell et al., 2016). We used bisulfite sequencing (Bis-Seq) to
140 compare DNA methylation changes at a subset of PpGe in F9 ECCs to that in ESCs post-
141 differentiation. Whereas DNA methylation was significantly increased at most PpGe in ESCs 4
142 days post-differentiation, these sites remained hypomethylated (<10% methylation, (Tierling et
143 al., 2018)) in F9 ECCs (Figures 2A and 2B). A similar hypomethylated state persisted at PpG
144 promoters in F9 ECCs except the highly methylated *Lefty2* promoter, where DNA methylation

145 was partially lost post-differentiation (Figure S1A). This result is consistent with the observed
146 partial repression of most PpGs and an induction of *Lefty2* expression in these cells (Figure 1B).
147 Furthermore, in ESCs, gain of DNA methylation at *Trim28* and *Sox2* enhancers suggests
148 enhancer-switching, which involves a potential use of neural lineage specific enhancers to
149 maintain a high expression of these genes post-differentiation (Figure 1A and 2A).

150 We confirmed that absence of DNA methylation at PpGe was not due to low expression of
151 Dnmt3a in F9 ECCs post-differentiation (Figures S1B and S1C). Based on previous
152 observations in cancers that overexpression of DNA methyltransferases leads to DNA
153 hypermethylation (Yu et al., 2015b, Gao et al., 2015, Ma et al., 2018, Jones et al., 2016,
154 Schübeler, 2015), we tested if overexpression of Dnmt3a could rescue DNA methylation at
155 PpGe. F9 ECCs were transfected with Myc-Dnmt3a and differentiated at 24 hr post-transfection
156 to ensure expression of recombinant Dnmt3a during early differentiation (Figure S1D). We
157 anticipate the DNA methylation established by transiently over expressing Dnmt3a will be
158 maintained by Dnmt1 during multiple cell divisions (Lyko, 2018). To ensure the detection of
159 methylation established by recombinant Dnmt3a, we differentiated the cells for 8 days.
160 However, DNA methylation levels at PpGe except Oct4 (15%) were well below <10% which is
161 within the range of detection error by this method (Tierling et al., 2018). A small gain within this
162 range was observed suggesting a spurious low methylation during cell divisions (Figure 2C). 5
163 out of 8 PpGs showed no additional decrease in expression (p value >0.1) when compared to
164 untransfected differentiated cells (Figure 2D) indicating that overexpression of Dnmt3a is unable
165 to rescue the differentiation defects observed in F9 ECCs. The absence of complete PpG
166 repression concomitant with little or no significant gain in DNA methylation at PpGe predicts a
167 potential disruption in the mechanism that mediates PpGe decommissioning.

168 To determine the extent of DNA methylation defect, we perfromed MethylRAD sequencing to
169 analyze changes in DNA methylation at all PpGe in F9 ECCs pre- and post-differentiation

170 (Wang et al., 2015b). This method uses *Fsp*EI restriction enzyme, which cuts DNA
171 bidirectionally from mC to create 31-32 bp fragments (Cohen-Karni et al., 2011, Zheng et al.,
172 2010). The restriction fragments were isolated for library preparation and high throughput
173 sequencing. Using this method, we captured DNA methylation at 1,370,254 cytosines genome-
174 wide. The reads were distributed among all chromosomes representing all annotated genomic
175 elements (Figure S2A). DNA methylation levels at enhancers (low-intermediate-high) were
176 calculated based on highest (75th percentile) and lowest (25th percentile) number of reads at all
177 annotated enhancers in the genome, which were obtained from the EnhancerAtlas 2.0. Previous
178 studies reported that PpGe are bound by Lsd1 in ESCs and post-differentiation, the PpGe
179 decommissioning requires H3K4me1 demethylation by Lsd1 (Whyte et al., 2012). In succession,
180 our studies showed that H3K4me1 demethylation was critical for Dnmt3a catalyzed DNA
181 methylation at these sites (Petell et al., 2016). Therefore, we filtered the data to focus our
182 analysis on DNA methylation changes at 3840 PpGe previously annotated in ESCs as Lsd1
183 bound regions (Whyte et al., 2012). Our method identified 1,865 PpGe in F9 ECCs. Compared
184 to methylation levels at all other known enhancers, the PpGe cluster into the low/intermediate
185 methylation group (Figure S2B). The difference in methylation for each PpGe region was
186 computed by subtracting the DNA methylation level in D4 differentiated from that in
187 undifferentiated F9 ECCs. The data show 1488 (82%) regions fail to gain DNA methylation post-
188 differentiation of F9 ECCs (No Change in DNA Methylation; NCDM) (Figure 2E). To determine
189 the function of genes associated with NCDM PpGe, we performed *Ingenuity* pathway analysis
190 (IPA) (www.qiagen.com/ingenuity), which showed a significant enrichment of Oct4-regulated
191 mammalian embryonic stem cell and molecular mechanisms of cancer pathways (Figure 2F).

192 Given that DNA methylation by Dnmt3a at PpGe requires H3K27 deacetylation and H3K4
193 demethylation by Lsd1/Mi2/NurD complex (Whyte et al., 2012), we anticipate a potential
194 impediment in this process causing a widespread failure to acquire DNA methylation at PpGe.

195 **A "primed" PpGe state is established during F9 ECC differentiation**

196 We asked if the chromatin state at PpGe in F9 ECCs is refractory to DNA methylation. To
197 examine histone H3K27 deacetylation, chromatin immunoprecipitation followed by qPCR (ChIP-
198 qPCR) was performed that showed a decrease in H3K27Ac at PpGe in F9 ECCs post-
199 differentiation (Figure 3A). This result suggests that, similar to our observations in ESCs (Figure
200 S3A), PpGe are active in undifferentiated F9 ECCs and initiate the decommissioning process
201 post-differentiation. Furthermore, deacetylation at *Lefty1* and *Lefty2* enhancers suggests
202 enhancer-switching involving the potential use of germline-specific enhancers post-
203 differentiation, leads to an observed increase in *Lefty1* and *Lefty2* expression (Figure 1A).
204 Numerous studies have proposed that deacetylation of H3K27Ac followed by H3K27
205 methylation by the PRC2 enzyme complex establishes a silenced state (Lindroth et al., 2008,
206 Barski et al., 2007, Wang et al., 2008). Our data showed no increase of H3K27me3 at the PpGe
207 in both ESCs as well as F9 ECCs post-differentiation, suggesting that PRC2 activity is
208 nonessential for PpGe silencing (Figures S3B and S3C). H3 occupancy at the PpGe in F9
209 ECCs between pre- and post- differentiation was similar, supporting these conclusions (Figure
210 S3D).

211 Using ChIP-qPCR, we next monitored H3K4me1 demethylation at PpGe during F9 ECC
212 differentiation. Surprisingly, we observed a retention of H3K4me1 modification at most
213 enhancers and a significant increase at 1 out of 7 tested PpGe post-differentiation (Figure 3B),
214 suggesting a potential disruption of Lsd1 activity. We verified similar expression levels of Lsd1 in
215 F9 ECCs compared to ESCs (Figures S1B and S1C). Previous studies have shown Lsd1 to
216 interact with the components of NuRD complex in ESCs and cancer cells to facilitate
217 deacetylation of H3K27 followed by demethylation of H3K4me1 (Whyte et al., 2012, Petell et al.,
218 2016, Li et al., 2011b, Patel et al., 2018, Wang et al., 2009b). To examine if Lsd1 interacts with
219 the Mi2/NuRD complex in F9 ECCs, we performed co-immunoprecipitation (Co-IP) experiments

220 using whole cell extracts. We used the whole cell extract from ESCs as a positive control.
221 Antibodies against Lsd1 and HDAC1 were used for reciprocal Co-IP. A strong signal for
222 HDAC1 and Lsd1 and a weak signal for CHD4 in both F9 ECCs and ESCs was observed
223 (Figure S3E). An absence of signal for the acetyltransferase HBO1 in Lsd1 Co-IP and presence
224 of CHD4 supports the specificity of the interaction between Lsd1 and NuRD complex in F9
225 ECCs (Figure S3F). To test whether Lsd1 is recruited to PpGe, ChIP-qPCR showed similar
226 enrichment of Lsd1 in F9 ECCs and ESCs at PpGe (Figure 3C). These data suggest that
227 retention of H3K4me1 at PpGe is not caused by the absence of Lsd1 but rather due to lack of its
228 activity at these sites in F9 ECCs post-differentiation.

229 To confirm the above conclusion, we next tested the effect of overexpression or inhibition of Lsd1
230 on PpGe silencing and PpG repression in differentiating F9 ECCs. F9 ECCs were transfected
231 with a FLAG-Lsd1 overexpressing plasmid and differentiated at 24 hr post-transfection (Figure
232 S4A). The recombinant Lsd1 was not able to rescue H3K4me1 demethylation at PpGe shown by
233 no decrease in H3K4me1 at most PpGe (Figure 3D). However, a significant increase was
234 observed at 3 out of 7 PpGe post-differentiation. This could result from incomplete demethylation
235 of H3K4me2 to H3K4me1 at these sites by recombinant Lsd1, confirmed by a decrease in
236 H3K4me2 signal at most of the PpGe in Lsd1 overexpressing F9 ECCs post-differentiation (Figure
237 3E). These data suggest that an inhibitory mechanism affects Lsd1 demethylation activity
238 irrespective of its origin of expression. Additionally, Lsd1 overexpression had no significant effect
239 on PpG repression or DNA methylation at PpGe post-differentiation (Figures S4B and S4C).

240 Previous studies in ESCs have shown that treatment with the Lsd1 inhibitor pargyline at the
241 onset of differentiation results in H3K4me1 retention at PpGe and incomplete repression of
242 PpGs (Whyte et al., 2012). F9 ECCs (WT and Lsd1 overexpressing) were treated with pargyline
243 6 hr prior to induction of differentiation. In contrast to 70-80% cell death caused by the Lsd1
244 inhibitor in ESCs, WT F9 ECCs remained largely viable upon treatment (Figure S4D) and a

245 slight relief of repression was observed for some PpGs in these cells (Figure S4E). Whereas no
246 decrease in H3K4me1 was observed at most PpGe in WT F9 ECCs, pargyline treatment
247 affected the H3K4me1 enrichment at 4 out of 7 PpGe in Lsd1 overexpressing cells post-
248 differentiation. This confirms the activity of Lsd1 at these sites by which it converts H3K4me2 to
249 H3K4me1 (Figure 3F). To test if Lsd1 activity contributed to the generation of H3K4me1 at
250 PpGe in the undifferentiated state, we transiently overexpressed Lsd1 in F9 ECCs and allowed
251 cells to grow for 72 hr. ChIP-qPCR analysis showed no increase in H3K4me1 levels in Lsd1
252 overexpressing cells (Figure S4F). This result suggests that, similar to what has been previously
253 reported in ESCs, the deposition of H3K4me1 in the undifferentiated F9 ECCs is largely
254 accomplished by MLL3/4 histone methyltransferases, and Lsd1 activity at these sites is initiated
255 only in response to a differentiation signal (Wang et al., 2017, Wang et al., 2016, Cao et al.,
256 2018, Whyte et al., 2012).

257 Taken together, these observations suggest that the restricted activity of Lsd1 at PpGe leads to
258 retention of H3K4me1 post-differentiation. Moreover, following the deacetylation of H3K27, the
259 absence of DNA methylation and the presence of H3K4me1 switches these enhancers to a
260 "primed" state, prone to reactivation.

261 **High-throughput analysis of changes in H3K4me1 at PpGe**

262 To identify all PpGe with aberrant retention of H3K4me1 post-differentiation, we performed
263 ChIP-Seq analysis of H3K4me1 genome-wide in undifferentiated and D4 differentiated F9
264 ECCs. Peak calling was performed using Epic2 for each input-ChIP pair and the list was filtered
265 to calculate the number of peaks based on defined cut offs (FDR<=0.05 and Log2FC>=2)
266 (Figure S5A). We analyzed the distribution of H3K4me1 peaks at the regulatory elements
267 across the genome (Figure S5B). For further analysis, we identified 1425 H3K4me1 peaks
268 found within 1 kbp of PpGe previously annotated in ESCs (Whyte et al., 2012). The difference
269 in peak enrichment (i.e. log2FoldChange) between differentiated and undifferentiated was

270 calculated to score for changes in H3K4me1 post-differentiation. The data showed no change,
271 increase, or decrease in 733, 510, and 182 PpGe, respectively. Therefore 87% of PpGe showed
272 no significant decrease in H3K4me1 (no decrease in histone methylation; NDHM PpGe) (Figure
273 4A). We next computed the correlation between the three PpGe sub-groups and the 1792 PpGe
274 that undergo H3K4me1 demethylation in differentiating ESCs (Whyte et al., 2012). Among the
275 NDHM PpGe, 74%PpGe with increased H3K4me1 and 69% PpGe with unchanged H3K4me1 in
276 F9 ECCs, overlap with PpGe that are H3K4me1 demethylated in ESCs (Figures 4B, 4C, and
277 S5C). These observations strongly support our previous conclusion that in F9 ECCs, Lsd1
278 activity is inhibited leading to retention of H3K4me1 at PpGe. IPA of NDHM PpGe–associated
279 genes showed highest enrichment for Oct4-regulated and stem cell pathways (Figure 4D).
280 Comparatively, genes associated with PpGe that undergo H3K4me1 demethylation, show
281 enrichment for signaling pathways (Figure S5D). A correlation between NCDM and NDHM
282 PpGe showed that 65% of the NDHM PpGe fail to acquire DNA methylation, underpinning the
283 role of histone demethylation in the regulation of DNA methylation at PpGe (Figure 4E) (Petell et
284 al., 2016).

285 **Lsd1 activity at PpGe is inhibited by its interaction with Oct4**

286 Next, we sought to determine the mechanism that inhibits Lsd1 activity. Due to its continued
287 expression in F9 ECCs post-differentiation (Figure S6A), we assumed that Oct4 remains
288 associated with PpGe and prevents demethylation of H3K4me1 at these sites. We tested this
289 hypothesis in P19 ECCs in which Oct4 expression was reported to be strongly repressed post-
290 differentiation (Li et al., 2007, Wei et al., 2007, Li et al., 2013, Palmieri et al., 1994, Fuhrmann et
291 al., 2001, Liu et al., 2011, Marikawa et al., 2011), an observation distinct from what we report in
292 F9 ECCs. After confirming a 90% reduction in Oct4 expression (Figure 5A), we probed
293 H3K4me1 demethylation at PpGe during P19 ECC differentiation. Indeed, our data show
294 decreased enrichment of H3K4me1 at the 5 tested PpGe, 4 days post-differentiation (Figure

295 5B). The change in chromatin state included up to 40% gain of DNA methylation at 3 out of 5
296 enhancers (Figure 5C). Similar to ESCs, we also observed a massive cell death when P19
297 ECCs were exposed to the Lsd1 inhibitor during differentiation (Figure S6B).

298 Together with the pathway analysis showing enrichment of Oct4-regulated genes in NDHM and
299 NCDM enhancers, the above data suggested that Oct4 might regulate the demethylase activity
300 of Lsd1 at PpGe. Previous observations showing an interaction of Oct4 with Lsd1 and the
301 Mi2/NuRD complex support this assumption (van den Berg et al., 2010, Pardo et al., 2010, Ding
302 et al., 2012). Interaction of Lsd1 and Oct4 was confirmed by co-immunoprecipitation of Lsd1 with
303 anti-Oct4 antibody using nuclear extract from both undifferentiated and D4 differentiated F9 ECCs
304 (Figure 6A). Co-precipitation experiments using recombinant proteins, GST-Lsd1 and Oct4,
305 reveal a direct interaction between the two proteins (Figure 6B). To test the effect of Oct4
306 interaction on Lsd1 catalytic activity, we performed *in vitro* Lsd1 demethylation assays using
307 H3K4me2 peptide as a substrate. In the presence of Oct4, Lsd1 activity was reduced by 60-70%
308 in a dose dependent manner. The specificity of Oct4-mediated inhibition was demonstrated by no
309 effect of recombinant Dnmt3a protein on Lsd1 activity. Complete loss of Lsd1 activity in the
310 presence of its inhibitor TCP (tranylcypromine) confirmed the specificity of the signal measured
311 in this assay (Figures 6C and 6D).

312 We also performed Lsd1 demethylation assays using purified histones as a substrate and
313 detected H3K4me2 demethylation on a Western blot. Activity of Lsd1 was assessed by reduced
314 H3K4me2 signal, which was rescued in the presence of 0.1 mM TCP. An accumulation of
315 H3K4me2 signal with an increase in Oct4 concentration in the reaction mix clearly showed
316 increased inhibition of Lsd1 activity by Oct4 (Figure 6E). These data suggest that Oct4 could
317 inhibit Lsd1 activity at PpGe in ECCs post-differentiation. We tested this by stably expressing
318 recombinant Oct4 in P19 ECCs (Figure S6C). Upon differentiation, the retention of H3K4me1 at
319 PpGe indicates the inhibition of Lsd1 by the recombinant Oct4 (Figure 6F). Moreover, several

320 PpGs were incompletely repressed with reduced gain in DNA methylation at their respective
321 enhancers (Figures 6G and 6H). These data suggest that in F9 ECCs, due to its continued
322 expression post-differentiation, Oct4 remains bound at PpGe and inhibits Lsd1 activity.

323 Taken together, we propose the following model to explain the regulation of Lsd1 activity at
324 PpGe. At the active PpGe, Lsd1 is inhibited by its interaction with bound Oct4. Post-
325 differentiation, this inhibition is relieved by dissociation of Oct4 from its binding sites. However,
326 in cancer cells where Oct4 expression is maintained, Lsd1 is held in its inhibited state, leading
327 to incomplete histone demethylation. Retention of H3K4me1 in turn blocks the activation of
328 Dnmt3a from its autoinhibited state, resulting in an absence of DNA methylation at PpGe (Petell
329 et al., 2016, Guo et al., 2015, Li et al., 2011a, Ooi et al., 2007, Otani et al., 2009). The absence
330 of DNA methylation and the presence of H3K4me1 switches these enhancers to a "primed"
331 state, prone to activation in the presence of the coactivator (Figure 7). Based on previous
332 reports that Oct4 as well as Lsd1 are aberrantly expressed in several cancers (Kim et al., 2015,
333 Wang et al., 2010, Schoenhals et al., 2009, Wang et al., 2009b, Kashyap et al., 2013, Lv et al.,
334 2012, Hosseini and Minucci, 2017), we suspect that Oct4-Lsd1 interaction at Oct4-bound
335 regions will disrupt Lsd1 activity leading to aberrant gene expression.

336 **Discussion**

337 The epigenetic state of enhancers is preserved in various cell types suggesting that aberrant
338 changes could promote tumorigenesis. This phenomenon is supported by recent studies
339 revealing a crucial role for enhancer-mediated activation of oncogenes (Hnisz et al., 2015,
340 Mansour et al., 2014, Chapuy et al., 2013, Groschel et al., 2014, Loven et al., 2013). Changes in
341 H3K4me1 levels and DNA accessibility at various enhancers have been reported in many cancers
342 (Akhtar-Zaidi et al., 2012). Some studies also propose that changes in enhancer states play a
343 role in the development of therapy resistant cancer cells. These studies showed loss or gain of
344 H3K4me1/2 at the enhancers in resistant breast cancer cells and loss of H3K27 acetylation at

345 enhancers in T cell acute lymphoblastic leukemia (T-ALL) (Magnani et al., 2013, Knoechel et al.,
346 2014). In addition, DNA hypermethylation, concomitant with overexpression of DNA
347 methyltransferases, is a hallmark of many cancers (Jones et al., 2016, Schübeler, 2015).
348 Changes in DNA methylation at enhancers occur in breast, lung, prostate, and cervical cancers
349 (Taberlay et al., 2014, Aran and Hellman, 2013, Aran et al., 2013, Yegnasubramanian et al.,
350 2011). These studies suggest that the chromatin state of tissue-specific enhancers can be used
351 as a diagnostic to predict aberrant expression of their respective genes in cancer. This prediction
352 is supported by our data here and previous studies showing that PpGe gain DNA methylation and
353 lose H3K4me1 in differentiating ESCs, whereas in ECCs, absence of DNA methylation is
354 accompanied by retention of H3K4me1 at these sites. The tissue specificity of the enhancer state
355 is further highlighted by our data showing that during ESC differentiation, all tested PpGe undergo
356 repressive chromatin changes irrespective of the transcriptional status of the associated gene.
357 This observation is exemplified by H3K27Ac deacetylation and gain of DNA methylation at *Sox2*
358 and *Trim28* enhancers despite the maintained expression of these genes.

359 Histone demethylase, Lsd1, catalyzes demethylation of dimethyl (me2) and monomethyl (me1)
360 at K4 of histone H3 (Shi et al., 2004). The catalytic activity of Lsd1 was shown to be regulated
361 by post-translational modification, alternative splicing and interaction with several factors
362 (Carnesecchi et al., 2017, Laurent et al., 2015, Metzger et al., 2005, Shi et al., 2005, Speranzini
363 et al., 2017, Wang et al., 2015a, Zibetti et al., 2010, McClellan et al., 2019, Vinyard et al., 2019).
364 Genetic knockout of the histone demethylase, Lsd1, was also shown to cause genome-wide
365 loss of DNA methylation in late passage ESCs due to degradation of DNMT1 enzyme (Wang et
366 al., 2009a). Our previous studies demonstrated the critical role of H3K4me1 demethylation by
367 Lsd1 in guiding DNMT3A-mediated DNA methylation at PpGe causing enhancer silencing in
368 ESCs (Petell et al., 2016). In contrary to ESCs, our data in differentiating F9 ECCs shows that
369 PpGs are partially repressed and their respective enhancers retain H3K4me1 and DNA

370 hypomethylation. The transgenic overexpression of Dnmt3A or Lsd1 in F9 ECCs was unable to
371 rescue H3K4me1 demethylation and gain of DNA methylation at these PpGe. A small gain of
372 DNA methylation (15%) was observed at the Oct4 enhancers in Dnmt3A –overexpressing cells,
373 however this was insufficient to further downregulate the expression of Oct4. This could be
374 explained by retention of histone H3K4me1 at the Oct4 enhancer that keeps the enhancer in a
375 partially active state. Moreover we show that when Oct4 gene is repressed during P19 ECCs
376 differentiation, Oct4 enhancer accumulates upto 40% DNA methylation accompanied by
377 H3K4me1 is also demethylated (Figure 5). This suggest that higher DNA methylation levels are
378 required for complete silencing of Oct4 enhancer. The data above is in agreement with previous
379 studies showing a similar accumulation DNA methylation at the Oct4 enhancer (40%-50%)
380 when Oct4 gene is repressed during ESCs post differentiation (Athanasiadou et al., 2010, Petell
381 et al., 2016). Taken together our data that compromised activity of Lsd1 results in the retention
382 of H3K4me1 and the absence of DNA methylation at the PpGe, leading to a primed state of
383 enhancers in cancer cells. Thus we reveal a mechanism by which developmental enhancers
384 could acquire aberrant histone modification and DNA methylation states that affect gene
385 expression. Unlike the silenced state, the “primed” enhancer state grants accessibility for
386 coactivator binding which renders cells highly vulnerable to a small increase in the expression of
387 oncogenic coactivators or master transcription factors (Zaret and Carroll, 2011, Calo and
388 Wysocka, 2013).

389 We also discovered that Lsd1 activity is inhibited by its interaction with the pioneer transcription
390 factor Oct4, which is expressed at a substantial level in F9 ECCs post-differentiation. Recently,
391 a similar observation was reported from flow cytometric analysis showing that compared to
392 ESCs, a significantly higher number of F9 ECCs have persistent *Oct4* expression post-
393 differentiation (Gordeeva and Khaydukov, 2017). Aberrant expression of *Oct4*, *Sox2*, and
394 *Nanog* is associated with tumor transformation, metastasis, and drug resistance (Sampieri and

395 Fodde, 2012, Ben-Porath et al., 2008). We speculate that during differentiation of cancer stem
396 cells, inhibition of Lsd1 by Oct4 leads to PpGe priming/reactivation leading to enhanced PpG
397 expression. Our studies further highlight the versatile regulation of Lsd1 binding and activity,
398 which can be fine-tuned by its interaction with numerous factors, allowing the enzyme to
399 function in various cellular processes including differentiation and disease.

400 **Acknowledgements**

401 We are thankful to Gowher lab members for discussions. This work was supported by
402 NIHR01GM118654-01 and graduate fellowship for LA from King Saud University. The authors
403 gratefully acknowledge the Walter Cancer Foundation, DNA Sequencing Facility, and support
404 from the Purdue University Center for Cancer Research, P30CA023168. We thank Dr. Phillip
405 SanMiguel from Genomic Core, Purdue for analysis of Bis-Seq data and Dr. Taiping Chen for
406 providing FLAG-Lsd1 and Myc-Dnmt3a2 expression plasmids.

407 **Authors Contributions**

408 L.A., S.M., B.H., J.B., M.H., D.S., and M.S.D performed the experiments. N.A., S.U., and P.A.S.
409 analyzed the genome-wide data. L.A. and H.G. wrote the manuscript

410 **Declaration of Interests**

411 None declared.

412 **Figure Legends**

413 **Figure 1. Pluripotency genes are partially repressed in embryonal carcinoma cells.**

414 UD: undifferentiated; D4, D8: Days post-induction of differentiation; ESCs: embryonic stem cells
415 and F9 ECCs: F9 embryonal carcinoma cells, PpGs: pluripotency genes.

416 (A, B, and E) Gene expression analysis by RT-qPCR of PpGs in (A) F9 ESCs, (B) ECCs (E)
417 lineage specific genes in F9 ECCs. The C_t values for each gene were normalized to *Gapdh* and
418 expression is shown relative to that in undifferentiated cells (dotted line). In F9 ECCs, the
419 lineage specific genes show a 5 to 60-fold induction of gene expression (E) whereas the
420 expression of PpGs is on average reduced to about 50% post-differentiation (B). Average and
421 SEM of two biological replicates are shown for each gene.

422 (C) Alkaline phosphatase staining and (D) SSEA-1 immunofluorescence of ESCs and F9 ECCs
423 pre- and post-differentiation. Positive signal indicates pluripotency that is lost post-differentiation
424 in ESCs. The scale bar is a 100 μm.

425 **Figure 2. Pluripotency gene enhancers do not gain DNA methylation in embryonal**
426 **carcinoma cells.**

427 UD: undifferentiated; D4, D8: Days post-induction of differentiation; D8 UT: untransfected F9
428 ECCs differentiated for 8 days; D8+Myc-Dnmt3a: F9 ECCs overexpressing Myc-Dnmt3a and
429 differentiated for 8 days; ESCs: embryonic stem cells and F9 ECCs: F9 embryonal carcinoma
430 cells; PpGe: pluripotency gene enhancers.

431 (A, B, and C) DNA methylation analysis using Bis-Seq. Genomic DNA was treated with bisulfite
432 and PpGe regions were amplified by PCR. The amplicons were sequenced on a high
433 throughput sequencing platform (Wide-Seq) and the data were analyzed using Bismark
434 software. DNA methylation of PpGe in (A) ESCs (B) F9 ECCs pre- and post-differentiation. Less

than 10% DNA methylation was recorded in F9 ECCs whereas the H19 imprinted region, used as a control, showed DNA methylation at 80%. At the same regions, DNA methylation increased up to 30% in ESCs. See also Figure S1A.

(C) DNA methylation of PpGe in F9 ECCs overexpressing Myc-Dnmt3a. (D) Gene expression analysis by RT-qPCR PpGs in F9 ECCs overexpressing Myc-Dnmt3a pre- and post-differentiation. (C) show low levels (less than 10%) in DNA methylation at most PpGe (D) no significant decrease in expression in 5 out of 8 tested PpGs (*P*-value >0.1). The C_t values for each gene were normalized to *Gapdh* and expression is shown relative to that in undifferentiated cells (dotted line). Data for A, B, C and D are the average and SEM of two biological replicates.

(E) Genome-wide DNA methylation analysis by MethylRAD sequencing. Genomic DNA was digested with the restriction enzyme *Fsp*EI that cuts methylated DNA into 31-32 bp fragments. The fragments were sequenced and mapped on mm10 mouse genome. The number of reads per region were used as a measure for extent of DNA methylation and compared between undifferentiated and D4 differentiated F9 ECCs. The waterfall plot shows DNA methylation changes at PpGe computed by subtracting normalized counts in D4 samples from normalized counts in undifferentiated samples. Upper and lower quartiles were used in thresholding regions as gaining or losing methylation. The pie chart shows fractions of PpGe with increase, decrease or no change in DNA methylation (NCDM). See also Figure S2B.

(F) Top ten statistically significant enriched canonical pathways amongst the genes associated with the NCDM enhancers, which showed no change. The x-axis shows the $\log_{10}$ (adjusted p-value), with the p-value adjusted for multiple testing using the Benjamini-Hochberg method.

457 **Figure 3. A "primed" chromatin state is established at pluripotency gene enhancers in**
458 **embryonal carcinoma cells**

459 UD: undifferentiated; D4: Days post-induction of differentiation; D4+FLAG-Lsd1: F9 ECCs
460 overexpressing FLAG-Lsd1 and differentiated for 4 days; Prg: Pargyline; ESCs: embryonic stem
461 cells and F9 ECCs: F9 embryonal carcinoma cells; PpGe: pluripotency gene enhancers.

462 (A-F) Chromatin immunoprecipitation (ChIP)-qPCR assays. Histone modifications at PpGe (A)
463 H3K27Ac and (B) H3K4me1 in F9 ECCs pre- and D4 post-differentiation. Whereas
464 deacetylation of PpGe is observed as a decrease in the H3K27Ac signal, histone H3K4me1 is
465 retained post-differentiation.

466 (C) Lsd1 occupancy in undifferentiated ESCs and F9 ECCs.

467 (D, E) Enrichment of (D) H3K4me1 and (E) H3K4me2 in F9 ECCs expressing recombinant
468 FLAG-Lsd1 compared to undifferentiated transfected cells. The ECCs were differentiated 24hrs
469 post transfection with Lsd1 expressing plasmid. Whereas there is an increase in H3K4me1 at
470 some PpGe (D), there is a concomitant decrease in H3K4me2 enrichment (E).

471 (F) Fold change in enrichment of H3K4me1 at PpGe in pargyline treated and untreated, WT,
472 and FLAG-Lsd1 overexpressing F9 ECCs, at D4 post-differentiation. Fold change is
473 represented as relative to enrichment in the undifferentiated state (dotted line).

474 % Enrichment = Fold enrichment over input X100.*P*-values derived from Student's *t*-test where
475 *, **, and *** denote (*P*-value <0.05), (*P*-value <0.01), and (*P*-value <0.005), respectively.

476 **Figure 4. Global retention of H3K4me1 at pluripotency gene enhancers in embryonal**
477 **carcinoma cells.**

478 ESCs: embryonic stem cells and F9 ECCs: F9 embryonal carcinoma cells; PpGe: pluripotency
479 gene enhancers; D4: Days post-induction of differentiation

480 Genome-wide H3K4me1 levels in F9 ECCs pre- and post-differentiation were measured by
481 ChIP-Seq. Peak calling was performed using Epic2 for each input-ChIP pair. A total of 1425
482 H3K4me1 peaks were identified in F9 ECCs within 1 kbp of previously annotated PpGe in ESCs
483 (Whyte et al., 2012). See also Figure S5A. Histone demethylation activity of Lsd1 was surmised
484 by calculating the change in H3K4me1 peak enrichment at PpGe between D4 differentiated and
485 undifferentiated samples.

486 (A) Waterfall plot represents changes in H3K4me1, which were calculated as the difference
487 between log2FC of D4 and undifferentiated samples, and transformed to Z-score. Z-score
488 thresholds of +1 and -1 were used to define the fractions showing increase, no change, or
489 decrease in H3K4me1 shown in the pie chart. Taken together, 87% PpGe show an increase or
490 no change (NDCM) in H3K4me1 enrichment.

491 (B, C) Venn diagram showing an overlap between PpGe that show (B) an increase, or (C) no
492 change in peak enrichment in F9 ECCs but undergo histone H3K4me1 demethylation in ESCs
493 post-differentiation (Whyte et al., 2012).

494 (D) Top ten statistically significant enriched canonical pathways amongst the genes associated
495 with increase and no change in F9 ECCs. The x-axis shows the $\log_{10}$ (adjusted p-value), with
496 the p-value adjusted for multiple testing using the Benjamini-Hochberg method.

497 (E) Overlap between the PpGe showing no change in DNA methylation (NDHM) and PpGe that
498 show no decrease in H3K4me1 (NCDM).

499 **Figure 5. Pluripotency gene enhancers are decommissioned in P19 embryonal carcinoma**
500 **cells**.

501 UD: undifferentiated; D4, D8: Days post-induction of differentiation; P19 ECCs: P19 embryonal
502 carcinoma cells, PpGs: pluripotency genes; PpGe: Pluripotency gene enhancers.

503 (A) Gene expression analysis by RT-qPCR of PpGs in P19 ECCs. The C_t values for each gene
504 were normalized to *Gapdh* and expression is shown relative to that in undifferentiated cells
505 (dotted line). Similar to the repression of PpGs in ESCs (Figure 1A), PpGs, especially Oct4 and
506 Nanog, show more than 90% reduction in expression.

507 (B) ChIP-qPCR showing H3K4me1 enrichment in UD and D4 differentiated P19 ECCs. A
508 decrease in H3K4me1 was observed at all PpGe post-differentiation demonstrating histone
509 demethylation activity. % Enrichment = Fold enrichment over input X100.

510 (C) DNA methylation analysis of PpGe using Bis-Seq in UD, D4 and D8 differentiated P19
511 ECCs. Up to 40% increase in DNA methylation level was observed at 3 out of 5 PpGe post-
512 differentiation. Average and SEM of two biological replicates are shown for each gene.

513 **Figure 6. Oct4 interacts with Lsd1 and inhibits its catalytic activity.**

514 TCP: tranylcypromine; ECCs: embryonal carcinoma cells; PpGe: pluripotency gene enhancers
515 D4: Days post-induction of differentiation.

516 (A) Nuclear extract from undifferentiated and D4 differentiated F9 ECCs was used to perform
517 Co-IP with anti-Oct4 antibody and control IgG. 10% of the input and eluate from Co-IP were
518 probed for Oct4 and Lsd1 on Western blot. The vertical space denotes the extra lane in the gel
519 that was digitally removed.

520 (B) GST-pull down experiment showing direct interaction between Lsd1 and Oct4. Recombinant
521 GST-Lsd1 was incubated with Oct4 at about 1:2 molar ratio and precipitated using GST-
522 Sepharose. The co-precipitated Oct4 is detected using Anti-Oct4 antibody. The vertical space
523 denotes the extra lane in the gel that was digitally removed

524 (C) Lsd1 demethylase assay was performed using 0.25 μM of Lsd1 and H3K4me2 peptide as
525 substrate. Lsd1 demethylation activity was completely inhibited by 0.1 mM TCP

526 (tranylcypromine) in the reaction. To test the effect of Oct4 on Lsd1 activity, demethylation
527 assays were performed in the presence of 0.5 μM Oct4 at 1:2 (Lsd1:Oct4) molar ratio. The
528 catalytic domain of Dnmt3a at the same molar ratio was used as a control.

529 (D) Dose dependent inhibition assays were performed using increasing concentrations of Oct4
530 in the following molar ratios of Lsd1:Oct4, (1:0.5), (1:1), (1:2), (1:3), (1:4). Data are an average
531 and SD of at least 5 experimental replicates.

532 (E) Lsd1 demethylation assays were performed using 0.25 μM of Lsd1 and 30 μg bulk histones
533 as substrate with increasing concentrations of Oct4 in the reaction. Upper Panel: Histone
534 demethylation was detected by using anti H3K4me2 on Western blot showing a retention of
535 signal in presence of increasing concentration of Oct4. Lower panels: Amount of Lsd1 enzyme
536 and increasing amounts of Oct4 in the histone demethylation reaction. Ponceau S stain of bulk
537 histones shows equal loading on the gel. The bar graph on the right shows quantification of
538 H3K4me2 signal using imageJ software.

539 (F) ChIP-qPCR showing percent enrichment of H3K4me1 at PpGe in P19 ECCs stably
540 expressing recombinant Myc-Oct4 pre- and post-differentiation. The data show retention of
541 H3K4me1 post-differentiation. % Enrichment = Fold enrichment over input X100

542 (G) Gene expression analysis by RT-qPCR of PpGs in P19 ECCs expressing recombinant Myc-
543 Oct4. The C_t values were normalized to *Gapdh* and expression is shown relative to that in
544 undifferentiated cells (dotted line).

545 (H) DNA methylation analysis of PpGe using Bis-Seq in UD and D4 differentiated P19 ECCs
546 WT and expressing recombinant Myc-Oct4. Oct4 expressing cells show failure to gain DNA
547 methylation at PpGe post-differentiation compared to untransfected WT. Average and SEM of
548 two biological replicates are shown.

549 **Figure 7. Model of epigenetic changes at pluripotency gene enhancers during stem cell**
550 **differentiation.**

551 In an undifferentiated state, the pluripotency gene enhancers (PpGe) are active, bound by the
552 coactivator complex and containing chromatin modifications including H3K4m2/1 and H3K27Ac.
553 In response to signal of differentiation, the dissociation of the coactivator complex including
554 Oct4 is followed by the activity of the Lsd1-Mi2/NuRD complex, which facilitates enhancer
555 silencing. The histone deacetylase (HDAC) removes H3K27Ac at PpGe, and Lsd1 demethylates
556 H3K4me1, followed by DNA methylation by Dnmt3a. However, in F9 ECCs, Lsd1 activity is
557 inhibited in the presence of Oct4, causing retention of H3K4me1. The ADD domain of Dnmt3a
558 cannot interact with H3K4 methylated histone tail and will potentially remain in the autoinhibited
559 state, thus preventing DNA methylation at these sites. Consequently, PpGe instead of being
560 silenced acquire a "primed" state. Black pins represent methylated CpGs.

561 **STAR Methods**

562 **LEAD CONTACT AND MATERIALS AVAILABILITY**

563 Further information and requests for resources and reagents should be directed to and will be
564 fulfilled by the Lead Contact, Humaira Gowher (hgowher@purdue.edu)

565 All unique/stable reagents generated in this study will be made available on request but we may
566 require a payment for processing and shipping and/or a completed Materials Transfer
567 Agreement if there is potential for commercial application.

568 **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

569 F9 embryonal carcinoma cells (F9 ECCs), P19 embryonal carcinoma cells (P19 ECCs), and
570 E14Tg2A Embryonic stem cells (ESCs) were cultured and maintained in gelatin-coated tissue
571 culture plates. All three above cell lines are male. Differentiation of ECCs was induced by
572 plating $20X10^6$ cells in low attachment 15 cm petri dishes and the addition of 1μM Retinoic acid

573 (RA). ESCs were differentiated by the same method with a concurrent withdrawal of LIF. The
574 medium was replenished every two days and samples were collected on Days 4 and 8 post-
575 differentiation.

576 Plasmids expressing Myc-Dnmt3a2 and FLAG-Lsd1 WT were transfected into F9 ECCs using
577 Lipofectamine 2000. One day post-transfection, a UD sample was collected (D0), and
578 transfected cells were induced to differentiate on gelatinized plates by the addition of RA. The
579 next day, differentiated cells were trypsinized and plated on low adherence petri dishes.
580 Samples were collected on Days 4 and 8 post-differentiation. Lsd1 inhibitor treatment was
581 performed as described 6 hr prior to induction of differentiation (Petell et al., 2016).

582 P19 ECCs were transfected with pCAG-Myc-Oct4 (Addgene 13460) using Lipofectamine 2000
583 per the manufacturer's instructions. Transfected cells were clonally propagated, and Myc-Oct4
584 expression was determined by Western blots with anti-cMyc antibody (Millipore, MABE282).

585 **METHOD DETAILS**

586 **DNA methylation analysis**

587 Bisulfite sequencing: Bisulfite conversion was performed using an EpiTect Fast Bisulfite
588 Conversion Kit (Qiagen, 59802) according to the manufacturer's protocol. PCR conditions for
589 outer and inner amplifications were performed (Petell et al., 2016). The pooled samples were
590 sequenced using NGS on a Wide-Seq platform. The reads were assembled and analyzed by
591 Bismark and Bowtie2. Methylated and unmethylated CpGs for each target were quantified,
592 averaged, and presented as percent CpG methylation. Number of CpGs for regions tested are
593 listed in Table S1. Total number of reads used to calculate percent CpG methylation are listed in
594 Table S2, S3. Primer sequences can be found in Table S4.

595 MethylRAD sequencing: Genomic DNA was isolated using a standard phenol:chloroform
596 extraction, followed by ethanol precipitation. DNA from various samples was digested with

597 *Fsp*EI for 4 hr at 37°C and subjected to electrophoresis through a 2% agarose gel. 30 base pair
598 fragments were cut out, purified, and adaptors were ligated at 4°C overnight (Wang et al.,
599 2015b). The ligated DNA was PCR amplified with index primers and sequenced using a
600 Novaseq 6000. The primers used for PCR amplification is in Table S4. The details of the
601 bioinformatics analysis of data are listed in Supplementary Methods.

602 **Chromatin Immunoprecipitation and ChIP-Seq**

603 ChIP was performed as described (Petell et al., 2016). Chromatin was sheared by sonication
604 using a Covaris E210 device, according to the manufacturer's protocol. A total of 8µg of sheared
605 crosslinked chromatin was incubated with 8µg of antibody pre-loaded on a 1:1 ratio of protein A
606 and protein G magnetic beads (Life Technologies, 10002D and 10004D, respectively). After
607 washing the beads, the samples were eluted in 1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH
608 8.0. Crosslinking was reversed by incubation at 65°C for 30 min with shaking. Samples were
609 treated with RNase (Roche, 11119915001) for 2 h at 37°C, and subsequently treated with
610 Proteinase K (Worthington, LS004222) for 2 h at 55°C. DNA was purified by phenol:chloroform
611 extraction followed by ethanol precipitation and quantified using PicoGreen (Life Technologies,
612 P11495) and NanoDrop 3300 fluorospectrometer. qPCR was then performed using equal
613 amounts of IN and IP samples. Fold enrichment was first calculated as: 2^(C_t(IN)-C_t(IP)).
614 Percent enrichment = Fold enrichment X 100. Significance of change was determined via *p*-
615 value, which was calculated by GraphPad Prism using Student's *t*-test.

616 Table S4 lists sequences of primers used.

617 **Gene expression analysis**

618 RNA was isolated using the TRIzol reagent (Invitrogen, 15596026) according to the
619 manufacturer's protocol. Samples were treated with DNAse (Roche, 04716728001) at 37°C,
620 and then purified using a Quick-RNA™ MiniPrep Plus Kit (ZymoReseach, R1057). Reverse-

621 transcription quantitative PCR was performed by using Verso One-Step RT-qPCR kits (Thermo
622 Scientific, AB-4104A) with 1µg of purified RNA. Gene expression was calculated as ΔC_t which is
623 C_t(Gene)- C_t(*Gapdh*). Change in gene expression is reported as fold change relative to that in
624 undifferentiated cells, which was set to 1. See Table S4 for primers used.

625 **Microscopy**

626 Bright field images of embryoid bodies (EBs) were obtained with Zeiss microscope using a 10X
627 objective. Alkaline phosphatase staining was performed using solutions supplied by an alkaline
628 phosphatase staining kit (Sigma, AB0300). Cells were cross-linked with 1% formaldehyde for 5
629 min, followed by quenching with a final concentration of 150 mM glycine. Cells were washed
630 twice with 1xPBS, then twice with combined staining solution (BCIP and NBT). The stain was
631 developed in the dark for 5 min, then washed three times with 1XPBS. SSEA-1
632 immunofluorescence was performed using the following antibodies: anti-SSEA-1 (Millipore,
633 MAB430) and AlexaFluor 555 nm (Life Technologies, A21422). SSEA-1 and Alkaline
634 phosphatase staining were imaged using 20X objectives under Nikon Ts and Zeiss
635 microscopes, respectively.

636 **Co-precipitation assays**

637 For co-immunoprecipitation (Co-IP), the nuclear extract was prepared according to
638 manufacturer's protocol (Active Motif, 40010) except that DNase was added for the release of
639 chromatin-associated proteins. The Co-IP was performed using 1:1 mix of Dynabeads™ Protein
640 A (Life Technologies, 1002D) and Dynabeads™ Protein G (Life Technologies, 1004D), and
641 conjugated with 5 µg antibody and with 50 µg of nuclear extract according to the manufacturer's
642 protocol. Antibodies used include: anti-Oct4 (Abcam, ab181557), anti-CHD4 (Abcam, ab72418),
643 anti-HDAC1 (Abcam, ab7028), anti-Lsd1 (Abcam, ab17721) and anti-HBO1 (Abcam,
644 ab124993).

645 Pull down assays were performed using 1 µg of GST-Lsd1 (Sigma, SRP0122) incubated with 1
646 µg of recombinant Oct4 (abcam, ab134876) and Glutathione sepharose 4B (GE healthcare, 17-
647 0756-01) resin in the binding buffer (50 mM Tris pH 8.5, 50 mM KCl, 5 mM MgCl, 0.5% BSA,
648 and 5% glycerol, complemented with a cocktail of protease inhibitors) overnight at 4°C with
649 gentle agitation. The resin was washed twice with binding buffer and proteins were eluted using
650 the elution buffer (50 mM Tris-HCl, 10 mM reduced glutathione, pH 8) according to the
651 manufacturer's instructions. Eluate and input were loaded onto a 10% SDS-PAGE gels and
652 blots were probed using anti-Lsd1 (Abcam, ab17721) and anti-Oct4 (Santa Cruz, sc-8628)
653 antibodies.

654 ***In vitro* Lsd1 demethylase activity assay**

655 An *in vitro* fluorometric assay was used to detect Lsd1 demethylase activity using an
656 Epigenase™ kit (Epigentek, P-0379) according to the manufacturer's protocol. 0.25 µM of Lsd1
657 (Sigma, SRP0122) was used for activity assays together with 0.5 µM (or as indicated) of Oct4
658 (abcam, ab134876 and ab169842) or BSA (Sigma, A3059) or the catalytic domain of Dnmt3a
659 (*purified in-house*) or 0.1 mM of the Lsd1 inhibitor Tranylcypromine (TCP). Signals were
660 measured using a CLARIOstar plate reader and analyzed using MARs software as described by
661 the manufacturer.

662 **Histone demethylation assay**

663 Lsd1 histone demethylation assays were performed as described (Shi et al., 2004). A total of 30
664 µg of bulk histones (Sigma, H9250) in a histone demethylation buffer (50 mM Tris pH 8.5, 50
665 mM KCl, 5 mM MgCl, 0.5% BSA, and 5% glycerol) were incubated with 0.25 µM of Lsd1
666 (Sigma, SRP0122) alone, or increasing concentrations (0.125 µM, 0.25 µM, 0.5 µM, 1 µM) of
667 Oct4 (Abcam, ab134876 and ab169842), or 0.1 µM TCP for 4 hr at 37°C. Lsd1 activity was

668 monitored by Western blot using anti-H3K4me2 antibody (Abcam, ab32356). The membrane
669 was stained by Ponceau S to determine equal loading of the reaction mix.

670 **Co-immunoprecipitation**

671 Co-immunoprecipitation experiments was also performed as described (Whyte et al., 2012).
672 Briefly, undifferentiated F9 ECCs and ESCs were washed and harvested in cold 1X PBS.
673 Cellular proteins were extracted using TNEN250 lysis buffer (50 mM Tris pH 7.5, 5 mM EDTA,
674 250 mM NaCl, 0.1% NP-40) complemented with protease inhibitors at 4°C with rotation for 30
675 min. Complexes were then immunoprecipitated overnight at 4°C with rotation by incubating the
676 supernatant solution supplemented with two volumes of TNENG (50 mM Tris pH 7.5, 5 mM
677 EDTA, 100 mM NaCl, 0.1% NP-40, 10% glycerol) with Dynabeads® M280 (Life Technologies,
678 11203D) bound to 5 ug of antibody. Beads were washed with TNEN125 (50 mM Tris pH 7.5, 5
679 mM EDTA, 125 mM NaCl, 0.1% NP-40) and samples were eluted by boiling for 10 min in
680 Laemmli's loading buffer containing 100 mM DTT. Western blots were performed with NuPAGE
681 4-12% Tris-Bis gels. Antibodies used included: Lsd1 (abcam, ab17721), HDAC1 (abcam,
682 ab7028), Mi-2b (abcam, ab72418).

683 **Western blot**

684 Western blot analysis was performed using the standard method and the following antibodies
685 and dilutions: anti-Dnmt3a, 1:1000 (Active Motif, 39206), Anti-Lsd1, 1:1000 (abcam, ab17721)
686 and anti-β Actin, 1:1000 (Santa Cruz, sc8628), and anti-Rabbit, 1:10,000 (Jackson
687 Immunoresearch, 111-035-003) or anti-Mouse, 1:10,000 (Jackson Immunoresearch, 115-035-
688 003). Chemiluminescence was performed according to the manufacturer's protocol (Thermo-
689 Fisher Scientific, 34580).

MethylRAD sequencing Analysis

Alignment and Quality Control: A total of 258,987,008 single end 1x150 sequencing was performed using a NovaSeq 6000 platform for undifferentiated and day 4 F9 samples. The program FastQC v. 0.11.7 (Andrews, 2010) was used to check data quality pre- and post-quality trimming/adapter removal. Adapters were removed from reads using Trimmomatic v. 0.36 (Bolger et al., 2014). Trimmomatic is a program that removes adapter sequences and trims short Illumina reads based on quality. Cutadapt version 2.2 (Martin, 2011) was used to trim reads further, removing the first two and last two bases of each read. Reads containing greater than 0 N's were discarded. After trimming, a total of 112,289,569 reads remained in the undifferentiated and day 4 samples. Reads which do not have *Fsp*E1 sites present anywhere in the read were removed using 'grep', leaving a total of 71,730,977 reads. Finally, Bowtie2 version 2.3.3 (Langmead et al., 2009, Langmead and Salzberg, 2012) was used to map reads to the ENSEMBL *Mus musculus* reference genome version GRCm38.93. A maximum of 1 mismatch was allowed in read mapping. The mapping rate of reads was 89%, with 64% of the reads mapped to the genome exactly 1 time and included in further analyses.

Data preprocessing: Methylated sites were catalogued by iterating through all read sequences to find a matched pattern of methylation (i.e. CCGG, CCAGG, and CCTGG) and recording its location in the genome. The number of reads mapping to each methylated site was recorded and adjusted for substitution, deletion, and insertion accordingly. Sites were also matched with the reference genome for verification. Sites that had less than 5 reads were removed from downstream analysis; counts from duplicate sites between patterns were summed as one site. Python and R scripts used in this analysis are included at www.github.com/natallah.

Annotation to LSD1 Enhancers: Sites in the LSD1 enhancers were modified to include 1 kb up- and downstream of the identified start site. Both the undifferentiated and day 4 differentiated F9 samples were annotated to the modified LSD1 regions using BEDTools intersect. The total

amount of methylation in a region was determined by combining the read counts of all sites in that region. Upper and lower quartiles were used in thresholding regions as gaining or losing methylation. Specifically, all regions with at least 22 counts more in undifferentiated than in day 4 differentiated samples were identified as losing DNA methylation as differentiation occurred. All regions with at least 30 counts more in day 4 samples than in undifferentiated samples were identified as gaining DNA methylation.

Quantifying methylated regions: The union of enhancers found in both the undifferentiated and day 4 differentiated F9 samples were determined. The difference in methylation for each combined enhancer regions were computed by subtracting the total methylation in day 4 differentiated F9 samples to that of the undifferentiated sample. Comparative figures were produced from these final data.

Determining methylation level of enhancers: DNA sequences of all known enhancers for the ESC_J1 strain of mouse were downloaded from EnhancerAtlas V2.0 (Gao et al., 2016). Overlaps between MethylRAD sites and enhancers were found using BEDTools intersect. Reads that overlapped the enhancer for the same gene were then summed together. Raw counts were normalized for length. The average length of enhancers in the EnhancerAtlast database was computed for each gene. If different studies in the database reported variable lengths for enhancers, the average length of the enhancer was computed. Counts for each sample were then divided by the length of the enhancer and subsequently multiplied by 1000 (to enhance readability). The 25th and 75th percentiles were computed for all enhancers separately for undifferentiated and day four samples and were used as cutoffs for low and intermediate methylation. Thus, for undifferentiated samples, enhancers annotated as having low methylation have normalized counts between (0, 24.57], intermediate methylation are between (24.57, 243.68], and high methylation have greater than 243.58 normalized counts. Enhancers in day 4 samples are annotated as highly methylated if normalized counts are observed to be between (0,

24.06], intermediate counts are between (24.06, 256.56], and high counts have greater than 256.56 normalized counts.

ChIP-Seq Analysis

***Quality Control and Mapping*:** Sequencing was performed using a NovaSeq 6000 to generate > 80 million paired-end (2x50) reads (>80 million) undifferentiated and day 4 F9 samples. Sequence data quality was determined using FastQC software (Andrews, 2010) and quality based trimming and filtering (minimum quality score 30 and minimum read-length 20) was performed through TrimGalore tool (Krueger, 2012). Greater than 95% of the reads from all samples were retained after quality control and were used for the mapping. Mapping was performed against the mouse reference genome (GRCm38) using Bowtie2 (Langmead and Salzberg, 2012) with a maximum of 1 mismatch. The overall mapping rate was >97% for all samples. Bowtie2 derived BAM files were further filtered to retain the reads with minimum MAPping Quality (MAPQ) 10.

Peak-calling: Peak calling was performed using epic2 (Stovner and Saetrom, 2019) for each Input-ChIP pair using the mouse reference genome (GRCm38). The tool was run with MAPQ10 filtered BAM files and the default parameters (--falsediscovery-rate-cutoff 0.05, --binsize 200).

Peak-annotation: Peak-annotation and visualization was performed with R-package ChIPseeker (Yu et al., 2015a). Annotations were performed for all epic2 peaks called with default parameters.

Overlap of H3K4me1 peaks in F9 ECCs with those in ESCs: Overlapping LSD1 bound sites between H3K4me1 peaks in F9 ECCs and ESCs (Whyte et al., 2012). Correct overlap, was determined by converting peak coordinates from mm9 to mm10 using CrossMap tool (Zhao et al., 2014).

Pathway Analysis

IPA (Ingenuity Pathway Analysis), (IPA, QIAGEN Redwood City, www.qiagen.com/ingenuity), was used in the annotation of genes and in performing the pathway analyses.

QUANTIFICATION AND STATISTICAL ANALYSIS

methylRAD statistical analyses and software

-Methods are described in the section entitled “DNA methylation is not established at PpGe during F9 ECC differentiation”, and also in the Method Details section entitled “MethylRAD sequencing Analysis.”

-fastQC version 0.11.7 was used to check data quality before and after filtering.

-Read trimming was performed using Trimmomatic v. 0.36 to remove Illumina adapter sequences from reads. Reads containing greater than 0 N’s were discarded.

-Cutadapt v 2.2 was used to remove the first two and last two bases of each read.

-grep “CCGG|GGCC|CCAGG|GGTCC|CCTGG” *fastq > sites.fastq was used to keep only reads with FspEI sites present

-Bowtie v2.3.3 was used to map reads to the GRCm38.93 Mus Muscusus genome, removing all reads with more than 1 mismatch

-Methylated sites were identified and methylation events counted using a custom Python script, which is available at www.github.com/natallah

-A total of 1,370,254 cytosines were captured genome-wide with a cutoff of 5 reads per site minimum (1,370,254 cytosines had at least 5 reads aligned with Bowtie2). The difference in methylation was calculated by summing reads which overlap with enhancer downloaded from EnhancerAtlas 2.0.

-Differences were calculated between differentiated and undifferentiated samples by calculating the difference d = xij-xij with x equivalent to the read count (total raw reads) for enhancer i in sample j. Undifferentiated enhancer read counts were subtracted from differentiated enhancer counts. Special focus was given to the enhancers which were identified as Lsd1 bound previously (Whyte et al, 2012). Low-intermediate-high levels of methylation were calculated using the total read counts in all enhancers from the EnhancerAtlas by using the 75th percentile as the lower bound for high methylation levels and the 25th percentile as the upper bound for low methylation levels.

-BEDtools overlap was used to identify reads overlapping with specific genomic features (enhancers, Lsd1 bound enhancers, promotors)

-ChIPseeker v 1.22.0 was used to annotate site distribution across the genome.

<u>ChIP-seq statistical analyses and software</u>

-Methods are described in the section entitled “High throughput analysis of changes in H3K4me1 at PpGe” as well as in the Supplemental section entitled “ChIP-Seq Analysis”.

-FastQC v 0.11.7 was used to check read quality before and after read trimming.

-Bowtie2 v 2.3.3 was used to map to the mouse reference genome version GRCm38, removing alignments with greater than 1 mismatch. Reads with less than a MAPQ of 10 were removed.

-TrimGalore v 0.6.4 was used to trim adapter sequences and remove reads with under a read-length of 20 and with less than a quality PHRED score of 30.

-Peak calling was performed using Epic2 v 2019-Jan-03 with all defaults on each IP and input pair. The peaks were filtered selecting only peaks with false discovery rate less than or equal to 5% and with greater than or equal to a log fold-change (IP/input control) of 2. --binsize was set to 200 in epic2.

-Peaks which had greater than a z-score of 1 were counted as increased, while those with below a z-score of -1 were counted as decreased. The peaks with a z-score greater than or equal to -1 and less than or equal to 1 were annotated as not changing (see Figure 4). Z-score was calculated as how many standard deviations a certain value is above or below the mean using formulae $z = (x - \mu) / \sigma$, where x corresponds to difference between log2FC of D4 and undifferentiated samples for a specific gene, μ is mean difference across all genes, and σ is standard deviation of difference across all genes.

<u>Pathway analyses</u>

-All pathway analyses were performed using Ingenuity Pathway analysis (IPA), as described in the section entitled "DNA methylation is not established at PpGe during F9 ECC differentiation", as well as the Supplemental section entitled "Pathway Analysis" using Mus musculus as the organism, and input as genes which are low (less than or equal to the 25th percentile methylation), intermediate (greater than the 25th percentile and less than the 75th percentile for all enhancers from EnhancerAtas 2.0), or high (greater than or equal to 75th percentile methylation).

-A one-tailed Fisher's exact test was used to calculate functional enrichment, and all p-values were adjusted for multiple testing using the Benjamini-Hochberg method. The adjusted p-value cutoff for significance is padj <0.05.

-ggplot was used to generate pie charts, waterfall plots, and bar charts.

828 **DATA AND CODE AVAILABILITY**

829 -The datasets generated during this study are available at GEO under accession GSE135225
830 (ChIP-seq) and GSE135226 (methylRAD-seq). ChIP-seq peak annotations and genome
831 browser track files are included.

832 -All code and scripts use to analyze data and to generate genomic plots are at
833 www.github.com/natallah

834 **Table S4. A list of all PCR primers used in this study. Related to DNA methylation**
835 **analysis, Chromatin Immunoprecipitation and ChIP-Seq and Gene expression analysis in**
836 **Method Details**

837 A list of all PCR primers used in this study (5' to 3'), separated by technique. The presence of
838 an (AminoC6) modification at the 3' end of the antisense oligonucleotide of each adaptor is used
839 to block extension

840 **References**

841 AKHTAR-ZAIDI, B., COWPER-SAL, R., CORRADIN, O., SAIAKHOVA, A., BARTELS, C. F., BALASUBRAMANIAN,
842 D., MYEROFF, L., LUTTERBAUGH, J., JARRAR, A. & KALADY, M. F. 2012. Epigenomic enhancer
843 profiling defines a signature of colon cancer. *Science,* 336**,** 736-739.
844 ALONSO, A., BREUER, B., STEUER, B. & FISCHER, J. 1991. The F9-EC cell line as a model for the analysis of
845 differentiation. *Int J Dev Biol,* 35**,** 389-97.
846 ANDREWS, P. W., MATIN, M. M., BAHRAMI, A. R., DAMJANOV, I., GOKHALE, P. & DRAPER, J. S. 2005.
847 Embryonic stem (ES) cells and embryonal carcinoma (EC) cells: Opposite sides of the same coin.
848 *Biochemical Society Transactions,* 33**,** 1526-1530.
849 ANDREWS, S. 2010. FastQC: a quality control tool for high throughput sequence data. Available:
850 http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
851 ARAN, D. & HELLMAN, A. 2013. DNA methylation of transcriptional enhancers and cancer predisposition.
852 *Cell,* 154**,** 11-13.
853 ARAN, D., SABATO, S. & HELLMAN, A. 2013. DNA methylation of distal regulatory sites characterizes
854 dysregulation of cancer genes. *Genome biology,* 14**,** R21.
855 ATHANASIADOU, R., DE SOUSA, D., MYANT, K., MERUSI, C., STANCHEVA, I. & BIRD, A. 2010. Targeting of
856 De Novo DNA Methylation Throughout the Oct-4 Gene Regulatory Region in Differentiating
857 Embryonic Stem Cells. *Plos One,* 5.
858 BANERJI, J., RUSCONI, S. & SCHAFFNER, W. 1981. Expression of a beta-globin gene is enhanced by
859 remote SV40 DNA sequences. *Cell,* 27**,** 299-308.
860 BARON, R., BINDA, C., TORTORICI, M., MCCAMMON, J. A. & MATTEVI, A. 2011. Molecular mimicry and
861 ligand recognition in binding and catalysis by the histone demethylase LSD1-CoREST complex.
862 *Structure,* 19**,** 212-20.
863 BARSKI, A., CUDDAPAH, S., CUI, K., ROH, T.-Y., SCHONES, D. E., WANG, Z., WEI, G., CHEPELEV, I. & ZHAO,
864 K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell,* 129**,** 823-
865 837.
866 BEN-PORATH, I., THOMSON, M. W., CAREY, V. J., GE, R., BELL, G. W., REGEV, A. & WEINBERG, R. A. 2008.
867 An embryonic stem cell-like gene expression signature in poorly differentiated aggressive
868 human tumors. *Nat Genet,* 40**,** 499-507.
869 BOLGER, A. M., LOHSE, M. & USADEL, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence
870 data. *Bioinformatics,* 30**,** 2114-20.
871 BULGER, M. & GROUDINE, M. 2011. Functional and mechanistic diversity of distal transcription
872 enhancers. *Cell,* 144**,** 327-339.
873 CALO, E. & WYSOCKA, J. 2013. Modification of enhancer chromatin: what, how, and why? *Mol Cell,* 49**,**
874 825-837.
875 CAO, K., COLLINGS, C. K., MORGAN, M. A., MARSHALL, S. A., RENDLEMAN, E. J., OZARK, P. A., SMITH, E.
876 R. & SHILATIFARD, A. 2018. An Mll4/COMPASS-Lsd1 epigenetic axis governs enhancer function
877 and pluripotency transition in embryonic stem cells. *Sci Adv,* 4**,** eaap8747.
878 CARNESECCHI, J., FORCET, C., ZHANG, L., TRIBOLLET, V., BARENTON, B., BOUDRA, R., CERUTTI, C., BILLAS,
879 I. M., SERANDOUR, A. A., CARROLL, J. S., BEAUDOIN, C. & VANACKER, J. M. 2017. ERRalpha
880 induces H3K9 demethylation by LSD1 to promote cell invasion. *Proc Natl Acad Sci U S A,* 114**,**
881 3909-3914.
882 CHAPUY, B., MCKEOWN, M. R., LIN, C. Y., MONTI, S., ROEMER, M. G., QI, J., RAHL, P. B., SUN, H. H.,
883 YEDA, K. T., DOENCH, J. G., REICHERT, E., KUNG, A. L., RODIG, S. J., YOUNG, R. A., SHIPP, M. A. &
884 BRADNER, J. E. 2013. Discovery and characterization of super-enhancer-associated
885 dependencies in diffuse large B cell lymphoma. *Cancer Cell,* 24**,** 777-90.

886 COHEN-KARNI, D., XU, D., APONE, L., FOMENKOV, A., SUN, Z., DAVIS, P. J., KINNEY, S. R., YAMADA-
887 MABUCHI, M., XU, S. Y., DAVIS, T., PRADHAN, S., ROBERTS, R. J. & ZHENG, Y. 2011. The MspJI
888 family of modification-dependent restriction endonucleases for epigenetic studies. *Proc Natl*
889 *Acad Sci U S A,* 108**,** 11040-5.
890 CREYGHTON, M. P., CHENG, A. W., WELSTEAD, G. G., KOOISTRA, T., CAREY, B. W., STEINE, E. J., HANNA,
891 J., LODATO, M. A., FRAMPTON, G. M. & SHARP, P. A. 2010. Histone H3K27ac separates active
892 from poised enhancers and predicts developmental state. *Proceedings of the National Academy*
893 *of Sciences,* 107**,** 21931-21936.
894 DE THE, H. 2018. Differentiation therapy revisited. *Nat Rev Cancer,* 18**,** 117-127.
895 DING, J., XU, H., FAIOLA, F., MA'AYAN, A. & WANG, J. 2012. Oct4 links multiple epigenetic pathways to
896 the pluripotency network. *Cell Res,* 22**,** 155-67.
897 ERNST, J. & KELLIS, M. 2010. Discovery and characterization of chromatin states for systematic
898 annotation of the human genome. *Nat Biotechnol,* 28**,** 817.
899 ERNST, J., KHERADPOUR, P., MIKKELSEN, T. S., SHORESH, N., WARD, L. D., EPSTEIN, C. B., ZHANG, X.,
900 WANG, L., ISSNER, R. & COYNE, M. 2011. Mapping and analysis of chromatin state dynamics in
901 nine human cell types. *Nature,* 473**,** 43.
902 FESKE, S. 2007. Calcium signalling in lymphocyte activation and disease. *Nat Rev Immunol,* 7**,** 690-702.
903 FUHRMANN, G., CHUNG, A. C., JACKSON, K. J., HUMMELKE, G., BANIAHMAD, A., SUTTER, J., SYLVESTER,
904 I., SCHOLER, H. R. & COONEY, A. J. 2001. Mouse germline restriction of Oct4 expression by germ
905 cell nuclear factor. *Dev Cell,* 1**,** 377-87.
906 GAO, T. S., HE, B., LIU, S., ZHU, H., TAN, K. & QIAN, J. 2016. EnhancerAtlas: a resource for enhancer
907 annotation and analysis in 105 human cell/tissue types. *Bioinformatics,* 32**,** 3543-3551.
908 GAO, X. N., YAN, F., LIN, J., GAO, L., LU, X. L., WEI, S. C., SHEN, N., PANG, J. X., NING, Q. Y., KOMENO, Y.,
909 DENG, A. L., XU, Y. H., SHI, J. L., LI, Y. H., ZHANG, D. E., NERVI, C., LIU, S. J. & YU, L. 2015.
910 AML1/ETO cooperates with HIF1alpha to promote leukemogenesis through DNMT3a
911 transactivation. *Leukemia,* 29**,** 1730-40.
912 GORDEEVA, O. & KHAYDUKOV, S. 2017. Tumorigenic and Differentiation Potentials of Embryonic Stem
913 Cells Depend on TGFβ Family Signaling: Lessons from Teratocarcinoma Cells Stimulated to
914 Differentiate with Retinoic Acid. *Stem cells international,* 2017.
915 GROSCHEL, S., SANDERS, M. A., HOOGENBOEZEM, R., DE WIT, E., BOUWMAN, B. A., ERPELINCK, C., VAN
916 DER VELDEN, V. H., HAVERMANS, M., AVELLINO, R., VAN LOM, K., ROMBOUTS, E. J., VAN DUIN,
917 M., DOHNER, K., BEVERLOO, H. B., BRADNER, J. E., DOHNER, H., LOWENBERG, B., VALK, P. J.,
918 BINDELS, E. M., DE LAAT, W. & DELWEL, R. 2014. A single oncogenic enhancer rearrangement
919 causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell,* 157**,** 369-81.
920 GUO, X., WANG, L., LI, J., DING, Z., XIAO, J., YIN, X., HE, S., SHI, P., DONG, L., LI, G., TIAN, C., WANG, J.,
921 CONG, Y. & XU, Y. 2015. Structural insight into autoinhibition and histone H3-induced activation
922 of DNMT3A. *Nature,* 517**,** 640-4.
923 GWAK, J. M., KIM, M., KIM, H. J., JANG, M. H. & PARK, S. Y. 2017. Expression of embryonal stem cell
924 transcription factors in breast cancer: Oct4 as an indicator for poor clinical outcome and
925 tamoxifen resistance. *Oncotarget,* 8**,** 36305-36318.
926 HAN, J. W., GURUNATHAN, S., CHOI, Y. J. & KIM, J. H. 2017. Dual functions of silver nanoparticles in F9
927 teratocarcinoma stem cells, a suitable model for evaluating cytotoxicity- and differentiation-
928 mediated cancer therapy. *International Journal of Nanomedicine,* 12**,** 7529-7549.
929 HEINTZMAN, N. D., STUART, R. K., HON, G., FU, Y., CHING, C. W., HAWKINS, R. D., BARRERA, L. O., VAN
930 CALCAR, S., QU, C. & CHING, K. A. 2007. Distinct and predictive chromatin signatures of
931 transcriptional promoters and enhancers in the human genome. *Nat Genet,* 39**,** 311.

932 HEINZ, S., BENNER, C., SPANN, N., BERTOLINO, E., LIN, Y. C., LASLO, P., CHENG, J. X., MURRE, C., SINGH,
933 H. & GLASS, C. K. 2010. Simple combinations of lineage-determining transcription factors prime
934 cis-regulatory elements required for macrophage and B cell identities. *Mol Cell,* 38**,** 576-589.
935 HEINZ, S., ROMANOSKI, C. E., BENNER, C. & GLASS, C. K. 2015. The selection and function of cell type-
936 specific enhancers. *Nat Rev Mol Cell Biol,* 16**,** 144-54.
937 HNISZ, D., SCHUIJERS, J., LIN, C. Y., WEINTRAUB, A. S., ABRAHAM, B. J., LEE, T. I., BRADNER, J. E. &
938 YOUNG, R. A. 2015. Convergence of developmental and oncogenic signaling pathways at
939 transcriptional super-enhancers. *Mol Cell,* 58**,** 362-70.
940 HOSSEINI, A. & MINUCCI, S. 2017. A comprehensive review of lysine-specific demethylase 1 and its roles
941 in cancer. *Epigenomics,* 9**,** 1123-1142.
942 IGLESIAS, J. M., GUMUZIO, J. & MARTIN, A. G. 2017. Linking Pluripotency Reprogramming and Cancer.
943 *Stem Cells Transl Med,* 6**,** 335-339.
944 JETER, C. R., LIU, B., LIU, X., CHEN, X., LIU, C., CALHOUN-DAVIS, T., REPASS, J., ZAEHRES, H., SHEN, J. J. &
945 TANG, D. G. 2011. NANOG promotes cancer stem cell characteristics and prostate cancer
946 resistance to androgen deprivation. *Oncogene,* 30**,** 3833-45.
947 JONES, P. A., ISSA, J.-P. J. & BAYLIN, S. 2016. Targeting the cancer epigenome for therapy. *Nature*
948 *Reviews Genetics,* 17**,** 630.
949 KASHYAP, V., AHMAD, S., NILSSON, E. M., HELCZYNSKI, L., KENNA, S., PERSSON, J. L., GUDAS, L. J. &
950 MONGAN, N. P. 2013. The lysine specific demethylase-1 (LSD1/KDM1A) regulates VEGF-A
951 expression in prostate cancer. *Mol Oncol,* 7**,** 555-66.
952 KIM, B. W., CHO, H., CHOI, C. H., YLAYA, K., CHUNG, J. Y., KIM, J. H. & HEWITT, S. M. 2015. Clinical
953 significance of OCT4 and SOX2 protein expression in cervical cancer. *BMC Cancer,* 15**,** 1015.
954 KNOECHEL, B., RODERICK, J. E., WILLIAMSON, K. E., ZHU, J., LOHR, J. G., COTTON, M. J., GILLESPIE, S. M.,
955 FERNANDEZ, D., KU, M. & WANG, H. 2014. An epigenetic mechanism of resistance to targeted
956 therapy in T cell acute lymphoblastic leukemia. *Nature genetics,* 46**,** 364.
957 KRUEGER, F. 2012. Trim Galore: A wrapper tool around Cutadapt and FastQC to consistently apply
958 quality and adapter trimming to FastQ files.
959 KUMAR, S. M., LIU, S., LU, H., ZHANG, H., ZHANG, P. J., GIMOTTY, P. A., GUERRA, M., GUO, W. & XU, X.
960 2012. Acquired cancer stem cell phenotypes through Oct4-mediated dedifferentiation.
961 *Oncogene,* 31**,** 4898-911.
962 LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods,* 9**,** 357-
963 9.
964 LANGMEAD, B., TRAPNELL, C., POP, M. & SALZBERG, S. L. 2009. Ultrafast and memory-efficient
965 alignment of short DNA sequences to the human genome. *Genome Biol,* 10**,** R25.
966 LAURENT, B., RUITU, L., MURN, J., HEMPEL, K., FERRAO, R., XIANG, Y., LIU, S., GARCIA, B. A., WU, H., WU,
967 F., STEEN, H. & SHI, Y. 2015. A specific LSD1/KDM1A isoform regulates neuronal differentiation
968 through H3K9 demethylation. *Mol Cell,* 57**,** 957-970.
969 LI, B. Z., HUANG, Z., CUI, Q. Y., SONG, X. H., DU, L., JELTSCH, A., CHEN, P., LI, G., LI, E. & XU, G. L. 2011a.
970 Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase.
971 *Cell Res,* 21**,** 1172-81.
972 LI, H., FAN, R., SUN, M., JIANG, T. & GONG, Y. 2013. Nspc1 regulates the key pluripotent Oct4-Nanog-
973 Sox2 axis in P19 embryonal carcinoma cells via directly activating Oct4. *Biochem Biophys Res*
974 *Commun,* 440**,** 527-32.
975 LI, J. Y., PU, M. T., HIRASAWA, R., LI, B. Z., HUANG, Y. N., ZENG, R., JING, N. H., CHEN, T., LI, E., SASAKI, H.
976 & XU, G. L. 2007. Synergistic function of DNA Methyltransferases dnmt3a and dnrnt3b in the
977 methylation of Oct4 and Nanog. *Molecular and Cellular Biology,* 27**,** 8748-8759.

978 LI, Q., SHI, L., GUI, B., YU, W., WANG, J., ZHANG, D., HAN, X., YAO, Z. & SHANG, Y. 2011b. Binding of the
979 JmjC demethylase JARID1B to LSD1/NuRD suppresses angiogenesis and metastasis in breast
980 cancer cells by repressing chemokine CCL14. *Cancer Res,* 71**,** 6899-908.
981 LINDROTH, A. M., PARK, Y. J., MCLEAN, C. M., DOKSHIN, G. A., PERSSON, J. M., HERMAN, H., PASINI, D.,
982 MIRO, X., DONOHOE, M. E. & LEE, J. T. 2008. Antagonism between DNA and H3K27 methylation
983 at the imprinted Rasgrf1 locus. *PLoS genetics,* 4**,** e1000145.
984 LINN, D. E., YANG, X., SUN, F., XIE, Y., CHEN, H., JIANG, R., CHUMSRI, S., BURGER, A. M. & QIU, Y. 2010. A
985 Role for OCT4 in Tumor Initiation of Drug-Resistant Prostate Cancer Cells. *Genes Cancer,* 1**,** 908-
986 16.
987 LIU, A., YU, X. & LIU, S. 2013. Pluripotency transcription factors and cancer stem cells: small genes make
988 a big difference. *Chin J Cancer,* 32**,** 483-7.
989 LIU, H., DENG, S., ZHAO, Z., ZHANG, H., XIAO, J., SONG, W., GAO, F. & GUAN, Y. 2011. Oct4 regulates the
990 miR-302 cluster in P19 mouse embryonic carcinoma cells. *Mol Biol Rep,* 38**,** 2155-60.
991 LOVEN, J., HOKE, H. A., LIN, C. Y., LAU, A., ORLANDO, D. A., VAKOC, C. R., BRADNER, J. E., LEE, T. I. &
992 YOUNG, R. A. 2013. Selective inhibition of tumor oncogenes by disruption of super-enhancers.
993 *Cell,* 153**,** 320-34.
994 LV, T., YUAN, D., MIAO, X., LV, Y., ZHAN, P., SHEN, X. & SONG, Y. 2012. Over-expression of LSD1
995 promotes proliferation, migration and invasion in non-small cell lung cancer. *PLoS One,* 7**,**
996 e35065.
997 LYKO, F. 2018. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nat Rev*
998 *Genet,* 19**,** 81-92.
999 MA, H. S., WANG, E. L., XU, W. F., YAMADA, S., YOSHIMOTO, K., QIAN, Z. R., SHI, L., LIU, L. L. & LI, X. H.
1000 2018. Overexpression of DNA (Cytosine-5)-Methyltransferase 1 (DNMT1) And DNA (Cytosine-5)-
1001 Methyltransferase 3A (DNMT3A) Is Associated with Aggressive Behavior and Hypermethylation
1002 of Tumor Suppressor Genes in Human Pituitary Adenomas. *Med Sci Monit,* 24**,** 4841-4850.
1003 MAGNANI, L., STOECK, A., ZHANG, X., LÁNCZKY, A., MIRABELLA, A. C., WANG, T.-L., GYORFFY, B. &
1004 LUPIEN, M. 2013. Genome-wide reprogramming of the chromatin landscape underlies
1005 endocrine therapy resistance in breast cancer. *Proceedings of the National Academy of Sciences***,**
1006 201219992.
1007 MAK, V. C., SIU, M. K., WONG, O. G., CHAN, K. K., NGAN, H. Y. & CHEUNG, A. N. 2012. Dysregulated
1008 stemness-related genes in gynecological malignancies. *Histol Histopathol,* 27**,** 1121-30.
1009 MANSOUR, M. R., ABRAHAM, B. J., ANDERS, L., BEREZOVSKAYA, A., GUTIERREZ, A., DURBIN, A. D.,
1010 ETCHIN, J., LAWTON, L., SALLAN, S. E., SILVERMAN, L. B., LOH, M. L., HUNGER, S. P., SANDA, T.,
1011 YOUNG, R. A. & LOOK, A. T. 2014. Oncogene regulation. An oncogenic super-enhancer formed
1012 through somatic mutation of a noncoding intergenic element. *Science,* 346**,** 1373-7.
1013 MARIKAWA, Y., TAMASHIRO, D. A., FUJITA, T. C. & ALARCON, V. B. 2011. Dual roles of Oct4 in the
1014 maintenance of mouse P19 embryonal carcinoma cells: as negative regulator of Wnt/beta-
1015 catenin signaling and competence provider for Brachyury induction. *Stem Cells Dev,* 20**,** 621-33.
1016 MARTIN, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011,*
1017 17**,** 3.
1018 MCCLELLAN, D., CASEY, M. J., BAREYAN, D., LUCENTE, H., OURS, C., VELINDER, M., SINGER, J., LONE, M.
1019 D., SUN, W., CORIA, Y., MASON, C. & ENGEL, M. E. 2019. Growth Factor Independence (GFI) 1B-
1020 mediated transcriptional repression and lineage allocation require Lysine Specific Demethylase
1021 (LSD)1-dependent recruitment of the BHC complex. *Mol Cell Biol*.
1022 MENDENHALL, E. M., WILLIAMSON, K. E., REYON, D., ZOU, J. Y., RAM, O., JOUNG, J. K. & BERNSTEIN, B.
1023 E. 2013. Locus-specific editing of histone modifications at endogenous enhancers. *Nat*
1024 *Biotechnol,* 31**,** 1133-6.

1025 METZGER, E., WISSMANN, M., YIN, N., MULLER, J. M., SCHNEIDER, R., PETERS, A. H., GUNTHER, T.,
1026 BUETTNER, R. & SCHULE, R. 2005. LSD1 demethylates repressive histone marks to promote
1027 androgen-receptor-dependent transcription. *Nature,* 437**,** 436-9.
1028 ONG, C.-T. & CORCES, V. G. 2011. Enhancer function: new insights into the regulation of tissue-specific
1029 gene expression. *Nature Reviews Genetics,* 12**,** 283.
1030 OOI, S. K., QIU, C., BERNSTEIN, E., LI, K., JIA, D., YANG, Z., ERDJUMENT-BROMAGE, H., TEMPST, P., LIN, S.
1031 P., ALLIS, C. D., CHENG, X. & BESTOR, T. H. 2007. DNMT3L connects unmethylated lysine 4 of
1032 histone H3 to de novo methylation of DNA. *Nature,* 448**,** 714-7.
1033 OTANI, J., NANKUMO, T., ARITA, K., INAMOTO, S., ARIYOSHI, M. & SHIRAKAWA, M. 2009. Structural basis
1034 for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-
1035 DNMT3L domain. *EMBO Rep,* 10**,** 1235-41.
1036 PALMIERI, S. L., PETER, W., HESS, H. & SCHOLER, H. R. 1994. Oct-4 transcription factor is differentially
1037 expressed in the mouse embryo during establishment of the first two extraembryonic cell
1038 lineages involved in implantation. *Dev Biol,* 166**,** 259-67.
1039 PARDO, M., LANG, B., YU, L., PROSSER, H., BRADLEY, A., BABU, M. M. & CHOUDHARY, J. 2010. An
1040 expanded Oct4 interaction network: implications for stem cell biology, development, and
1041 disease. *Cell Stem Cell,* 6**,** 382-95.
1042 PATEL, D., SHIMOMURA, A., MAJUMDAR, S., HOLLEY, M. C. & HASHINO, E. 2018. The histone
1043 demethylase LSD1 regulates inner ear progenitor differentiation through interactions with Pax2
1044 and the NuRD repressor complex. *PLoS One,* 13**,** e0191689.
1045 PENG, S., MAIHLE, N. J. & HUANG, Y. 2010. Pluripotency factors Lin28 and Oct4 identify a sub-population
1046 of stem cell-like cells in ovarian cancer. *Oncogene,* 29**,** 2153-9.
1047 PETELL, C. J., ALABDI, L., HE, M., SAN MIGUEL, P., ROSE, R. & GOWHER, H. 2016. An epigenetic switch
1048 regulates de novo DNA methylation at a subset of pluripotency gene enhancers during
1049 embryonic stem cell differentiation. *Nucleic Acids Res,* 44**,** 7605-17.
1050 PLANK, J. L. & DEAN, A. 2014. Enhancer function: mechanistic and genome-wide insights come together.
1051 *Mol Cell,* 55**,** 5-14.
1052 RADA-IGLESIAS, A., BAJPAI, R., SWIGUT, T., BRUGMANN, S. A., FLYNN, R. A. & WYSOCKA, J. 2011. A
1053 unique chromatin signature uncovers early developmental enhancers in humans. *Nature,* 470**,**
1054 279.
1055 SAMPIERI, K. & FODDE, R. 2012. Cancer stem cells and metastasis. *Semin Cancer Biol,* 22**,** 187-93.
1056 SCHOENHALS, M., KASSAMBARA, A., DE VOS, J., HOSE, D., MOREAUX, J. & KLEIN, B. 2009. Embryonic
1057 stem cell markers expression in cancers. *Biochem Biophys Res Commun,* 383**,** 157-62.
1058 SCHÜBELER, D. 2015. Function and information content of DNA methylation. *Nature,* 517**,** 321.
1059 SHI, Y., LAN, F., MATSON, C., MULLIGAN, P., WHETSTINE, J. R., COLE, P. A., CASERO, R. A. & SHI, Y. 2004.
1060 Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell,* 119**,** 941-53.
1061 SHI, Y. J., MATSON, C., LAN, F., IWASE, S., BABA, T. & SHI, Y. 2005. Regulation of LSD1 histone
1062 demethylase activity by its associated factors. *Mol Cell,* 19**,** 857-64.
1063 SPERANZINI, V., CIOSSANI, G., MARABELLI, C. & MATTEVI, A. 2017. Probing the interaction of the p53 C-
1064 terminal domain to the histone demethylase LSD1. *Arch Biochem Biophys,* 632**,** 202-208.
1065 STOVNER, E. B. & SAETROM, P. 2019. epic2 efficiently finds diffuse domains in ChIP-seq data.
1066 *Bioinformatics*.
1067 TABERLAY, P. C., STATHAM, A. L., KELLY, T. K., CLARK, S. J. & JONES, P. A. 2014. Reconfiguration of
1068 nucleosome-depleted regions at distal regulatory elements accompanies DNA methylation of
1069 enhancers and insulators in cancer. *Genome research***,** gr. 163485.113.
1070 TIERLING, S., SCHMITT, B. & WALTER, J. 2018. Comprehensive Evaluation of Commercial Bisulfite-Based
1071 DNA Methylation Kits and Development of an Alternative Protocol With Improved Conversion
1072 Performance. *Genet Epigenet,* 10**,** 1179237X18766097.

1073 VAN DEN BERG, D. L. C., SNOEK, T., MULLIN, N. P., YATES, A., BEZSTAROSTI, K., DEMMERS, J.,
1074 CHAMBERS, I. & POOT, R. A. 2010. An Oct4-Centered Protein Interaction Network in Embryonic
1075 Stem Cells. *Cell Stem Cell,* 6**,** 369-381.
1076 VINYARD, M. E., SU, C., SIEGENFELD, A. P., WATERBURY, A. L., FREEDY, A. M., GOSAVI, P. M., PARK, Y.,
1077 KWAN, E. E., SENZER, B. D., DOENCH, J. G., BAUER, D. E., PINELLO, L. & LIAU, B. B. 2019. CRISPR-
1078 suppressor scanning reveals a nonenzymatic role of LSD1 in AML. *Nat Chem Biol,* 15**,** 529-539.
1079 WANG, C., LEE, J. E., LAI, B., MACFARLAN, T. S., XU, S., ZHUANG, L., LIU, C., PENG, W. & GE, K. 2016.
1080 Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. *Proc Natl Acad*
1081 *Sci U S A,* 113**,** 11871-11876.
1082 WANG, J., HEVI, S., KURASH, J. K., LEI, H., GAY, F., BAJKO, J., SU, H., SUN, W., CHANG, H., XU, G., GAUDET,
1083 F., LI, E. & CHEN, T. 2009a. The lysine demethylase LSD1 (KDM1) is required for maintenance of
1084 global DNA methylation. *Nat Genet,* 41**,** 125-9.
1085 WANG, J., TELESE, F., TAN, Y., LI, W., JIN, C., HE, X., BASNET, H., MA, Q., MERKURJEV, D., ZHU, X., LIU, Z.,
1086 ZHANG, J., OHGI, K., TAYLOR, H., WHITE, R. R., TAZEARSLAN, C., SUH, Y., MACFARLAN, T. S.,
1087 PFAFF, S. L. & ROSENFELD, M. G. 2015a. LSD1n is an H4K20 demethylase regulating memory
1088 formation via transcriptional elongation control. *Nat Neurosci,* 18**,** 1256-64.
1089 WANG, S., LV, J., ZHANG, L., DOU, J., SUN, Y., LI, X., FU, X., DOU, H., MAO, J., HU, X. & BAO, Z. 2015b.
1090 MethylRAD: a simple and scalable method for genome-wide DNA methylation profiling using
1091 methylation-dependent restriction enzymes. *Open Biol,* 5.
1092 WANG, S. P., TANG, Z., CHEN, C. W., SHIMADA, M., KOCHE, R. P., WANG, L. H., NAKADAI, T., CHRAMIEC,
1093 A., KRIVTSOV, A. V., ARMSTRONG, S. A. & ROEDER, R. G. 2017. A UTX-MLL4-p300 Transcriptional
1094 Regulatory Network Coordinately Shapes Active Enhancer Landscapes for Eliciting Transcription.
1095 *Mol Cell,* 67**,** 308-321 e6.
1096 WANG, X. Q., ONGKEKO, W. M., CHEN, L., YANG, Z. F., LU, P., CHEN, K. K., LOPEZ, J. P., POON, R. T. &
1097 FAN, S. T. 2010. Octamer 4 (Oct4) mediates chemotherapeutic drug resistance in liver cancer
1098 cells through a potential Oct4-AKT-ATP-binding cassette G2 pathway. *Hepatology,* 52**,** 528-39.
1099 WANG, Y., ZHANG, H., CHEN, Y., SUN, Y., YANG, F., YU, W., LIANG, J., SUN, L., YANG, X., SHI, L., LI, R., LI,
1100 Y., ZHANG, Y., LI, Q., YI, X. & SHANG, Y. 2009b. LSD1 is a subunit of the NuRD complex and
1101 targets the metastasis programs in breast cancer. *Cell,* 138**,** 660-72.
1102 WANG, Y. D., CAI, N., WU, X. L., CAO, H. Z., XIE, L. L. & ZHENG, P. S. 2013. OCT4 promotes tumorigenesis
1103 and inhibits apoptosis of cervical cancer cells by miR-125b/BAK1 pathway. *Cell Death Dis,* 4**,**
1104 e760.
1105 WANG, Z., ZANG, C., ROSENFELD, J. A., SCHONES, D. E., BARSKI, A., CUDDAPAH, S., CUI, K., ROH, T.-Y.,
1106 PENG, W. & ZHANG, M. Q. 2008. Combinatorial patterns of histone acetylations and
1107 methylations in the human genome. *Nature genetics,* 40**,** 897.
1108 WEI, F., SCHOLER, H. R. & ATCHISON, M. L. 2007. Sumoylation of Oct4 enhances its stability, DNA
1109 binding, and transactivation. *J Biol Chem,* 282**,** 21551-60.
1110 WEN, J., PARK, J. Y., PARK, K. H., CHUNG, H. W., BANG, S., PARK, S. W. & SONG, S. Y. 2010. Oct4 and
1111 Nanog expression is associated with early stages of pancreatic carcinogenesis. *Pancreas,* 39**,**
1112 622-6.
1113 WHYTE, W. A., BILODEAU, S., ORLANDO, D. A., HOKE, H. A., FRAMPTON, G. M., FOSTER, C. T., COWLEY,
1114 S. M. & YOUNG, R. A. 2012. Enhancer decommissioning by LSD1 during embryonic stem cell
1115 differentiation. *Nature,* 482**,** 221-5.
1116 YANG, F., ZHANG, J. & YANG, H. 2018. OCT4, SOX2, and NANOG positive expression correlates with poor
1117 differentiation, advanced disease stages, and worse overall survival in HER2(+) breast cancer
1118 patients. *Onco Targets Ther,* 11**,** 7873-7881.
1119 YEGNASUBRAMANIAN, S., WU, Z., HAFFNER, M. C., ESOPI, D., ARYEE, M. J., BADRINATH, R., HE, T. L.,
1120 MORGAN, J. D., CARVALHO, B. & ZHENG, Q. 2011. Chromosome-wide mapping of DNA

1121 methylation patterns in normal and malignant prostate cells reveals pervasive methylation of
1122 gene-associated and conserved intergenic sequences. *BMC genomics,* 12**,** 313.
1123 YU, G., WANG, L. G. & HE, Q. Y. 2015a. ChIPseeker: an R/Bioconductor package for ChIP peak
1124 annotation, comparison and visualization. *Bioinformatics,* 31**,** 2382-3.
1125 YU, Z., XIAO, Q., ZHAO, L., REN, J., BAI, X., SUN, M., WU, H., LIU, X., SONG, Z., YAN, Y., MI, X., WANG, E.,
1126 JIN, F. & WEI, M. 2015b. DNA methyltransferase 1/3a overexpression in sporadic breast cancer is
1127 associated with reduced expression of estrogen receptor-alpha/breast cancer susceptibility gene
1128 1 and poor prognosis. *Mol Carcinog,* 54**,** 707-19.
1129 ZARET, K. S. & CARROLL, J. S. 2011. Pioneer transcription factors: establishing competence for gene
1130 expression. *Genes Dev,* 25**,** 2227-41.
1131 ZENTNER, G. E., TESAR, P. J. & SCACHERI, P. C. 2011. Epigenetic signatures distinguish multiple classes of
1132 enhancers with distinct cellular functions. *Genome Res*.
1133 ZHANG, X. M., LU, F., WANG, J., YIN, F., XU, Z. S., QI, D. D., WU, X. H., CAO, Y. W., LIANG, W. H., LIU, Y. Q.,
1134 SUN, H., YE, T. & ZHANG, H. 2013. Pluripotent Stem Cell Protein Sox2 Confers Sensitivity to LSD1
1135 Inhibition in Cancer Cells. *Cell Reports,* 5**,** 445-457.
1136 ZHAO, H., SUN, Z. F., WANG, J., HUANG, H. J., KOCHER, J. P. & WANG, L. G. 2014. CrossMap: a versatile
1137 tool for coordinate conversion between genome assemblies. *Bioinformatics,* 30**,** 1006-1007.
1138 ZHENG, Y., COHEN-KARNI, D., XU, D., CHIN, H. G., WILSON, G., PRADHAN, S. & ROBERTS, R. J. 2010. A
1139 unique family of Mrr-like modification-dependent restriction endonucleases. *Nucleic Acids Res,*
1140 38**,** 5527-34.
1141 ZHU, B., LIU, T., HU, X. & WANG, G. 2013a. Developmental toxicity of 3,4-dichloroaniline on rare minnow
1142 (Gobiocypris rarus) embryos and larvae. *Chemosphere,* 90**,** 1132-9.
1143 ZHU, Y., SUN, L., CHEN, Z., WHITAKER, J. W., WANG, T. & WANG, W. 2013b. Predicting enhancer
1144 transcription and activity from chromatin modifications. *Nucleic Acids Res,* 41**,** 10032-43.
1145 ZIBETTI, C., ADAMO, A., BINDA, C., FORNERIS, F., TOFFOLO, E., VERPELLI, C., GINELLI, E., MATTEVI, A.,
1146 SALA, C. & BATTAGLIOLI, E. 2010. Alternative splicing of the histone demethylase LSD1/KDM1
1147 contributes to the modulation of neurite morphogenesis in the mammalian nervous system. *J*
1148 *Neurosci,* 30**,** 2521-32.

1149 ANDREWS, S. 2010. FastQC: a quality control tool for high throughput sequence data. Available:
1150 http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
1151 BOLGER, A. M., LOHSE, M. & USADEL, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence
1152 data. *Bioinformatics,* 30**,** 2114-20.
1153 GAO, T. S., HE, B., LIU, S., ZHU, H., TAN, K. & QIAN, J. 2016. EnhancerAtlas: a resource for enhancer
1154 annotation and analysis in 105 human cell/tissue types. *Bioinformatics,* 32**,** 3543-3551.
1155 KRUEGER, F. 2012. Trim Galore: A wrapper tool around Cutadapt and FastQC to consistently apply
1156 quality and adapter trimming to FastQ files.
1157 LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods,* 9**,** 357-
1158 9.
1159 LANGMEAD, B., TRAPNELL, C., POP, M. & SALZBERG, S. L. 2009. Ultrafast and memory-efficient
1160 alignment of short DNA sequences to the human genome. *Genome Biol,* 10**,** R25.
1161 MARTIN, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011,*
1162 17**,** 3.
1163 STOVNER, E. B. & SAETROM, P. 2019. epic2 efficiently finds diffuse domains in ChIP-seq data.
1164 *Bioinformatics*.
1165 WHYTE, W. A., BILODEAU, S., ORLANDO, D. A., HOKE, H. A., FRAMPTON, G. M., FOSTER, C. T., COWLEY,
1166 S. M. & YOUNG, R. A. 2012. Enhancer decommissioning by LSD1 during embryonic stem cell
1167 differentiation. *Nature,* 482**,** 221-5.

1168 YU, G., WANG, L. G. & HE, Q. Y. 2015. ChIPseeker: an R/Bioconductor package for ChIP peak annotation,
1169 comparison and visualization. *Bioinformatics,* 31**,** 2382-3.
1170 ZHAO, H., SUN, Z. F., WANG, J., HUANG, H. J., KOCHER, J. P. & WANG, L. G. 2014. CrossMap: a versatile
1171 tool for coordinate conversion between genome assemblies. *Bioinformatics,* 30**,** 1006-1007.

1172

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-cMyc antibody	Millipore	MABE282
anti-SSEA-1	Millipore	MAB430
and Alexa Fluor 555 nm	Life Technologies	A21422
anti-Oct4	Abcam	ab181557
anti-CHD4	Abcam	ab72418
anti-HDAC1	Abcam	ab7028
anti-Lsd1	Abcam	ab17721
anti-HBO1	Abcam	ab124993
anti-Oct4	Santa Cruz	sc-8628
anti-H3K4me2	Abcam	ab32356
anti-H3K27Ac	Abcam	ab4729
anti-H3K4me1	Abcam	ab8895
anti-H3K27me3	Abcam	ab8898
anti-β Actin	Santa Cruz	sc8628c
Anti-Rabbit IgG	Jackson Immunoresearch	111-035-003

Anti-Mouse IgG	Jackson Immunoresearch	115-035-003
Biological Samples		
pCAG-Myc-Oct4	Addgene	13460
Chemicals, Peptides, and Recombinant Proteins		
TRIzol	Invitrogen	15596026
*Fsp*EI	NEB	
Protein A magnetic beads	Life Technologies	10002D
Protein G magnetic beads	Life Technologies	10004D
RNase	Roche	11119915001
Proteinase K	Worthington	LS004222
Quick-RNA™ MiniPrep Plus Kit	ZymoReseach	R1057
Verso One-Step RT-qPCR kit	Thermo Scientific	AB-4104A
Alkaline phosphatase staining kit	Sigma	AB0300
GST-Lsd1	Sigma	SRP0122
Oct4	Abcam	ab134876
Oct4	Abcam	ab169842

Bulk Histones	Sigma	H9250
EpiTect Fast Bisulfite Conversion Kit	Qiagen	59802
Lipofectamine 2000	Thermofischer	11668019
Critical Commercial Assays		
Epigenase™ kit	Epigentek	P-0379
Deposited Data		
Raw and analyzed data	This paper	GSE135225
Experimental Models: Cell Lines		
P19 embryonal carcinoma cells	ATCC	CRL-182
F9 embryonal carcinoma cells	ATCC	CRL-1720
E14Tg2A Embryonic stem cells	MMRRC	015890-UCD
Oligonucleotides		
List of primers used	This paper	Table S4
Software and Algorithms		

Trimmomatic v 0.36	Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics*, btu170.	https://github.com/timflutre/trimmomatic
Bowtie2 v.2.3.3	Langmead B, Salzberg S. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. 2012, 9:357-359.	http://bowtie-bio.sourceforge.net/bowtie2/manual.shtml
Cutadapt v 2.2	MARTIN, Marcel. Cutadapt removes adapter sequences from high-throughput sequencing reads. **EMBnet.journal**, [S.l.], v. 17, n. 1, p. pp. 10-12, may 2011. ISSN 2226-6089. Available at: <http://journal.embnet.org/index.php/embnetjournal/article/view/200>. Date accessed: 11 dec. 2019. doi:https://doi.org/10.14806/ej.17.1.200.	https://cutadapt.readthedocs.io/en/stable/installation.html
ENSEMBL Mus musculus reference genome version GRCm38.93	GRCm38.p2 (Genome Reference Consortium Mouse Reference 38), INSDC Assembly GCA_000001635.4, Jan 2012	ftp://ftp.ensembl.org/pub/release-98/fasta/mus_musculus/dna/
Custom python scripts	https://github.com/natallah/methylRad	https://github.com/natallah/methylRad
Bismark v 0.18.2	Felix Krueger, Simon R. Andrews, Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications, *Bioinformatics*, Volume 27, Issue 11, 1 June 2011, Pages 1571–1572, https://doi.org/10.1093/bioinformatics/btr167	https://www.bioinformatics.babraham.ac.uk/projects/bismark/
BEDTools v 2.29.0	Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic	https://bedtools.readthedocs.io/en/latest/

	features. *Bioinformatics*. 2010;26(6):841–842. doi:10.1093/bioinformatics/btq033	
Enhancer Atlas v 2.0	Tianshun Gao and Jiang Qian. *EnhancerAtlas 2.0: an updated resource with enhancer annotation in 586 tissue/cell types across nine species.* *Nucleic Acids Res.* 2019.	http://enhanceratlas.org/
TrimGalore v 0.6.4	Krueger F: Trim Galore!. [http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/]	https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
epic2 2019-Jan-03	Endre Bakken Stovner, Pål Sætrom, epic2 efficiently finds diffuse domains in ChIP-seq data, *Bioinformatics*, Volume 35, Issue 21, 1 November 2019, Pages 4392–4393, https://doi.org/10.1093/bioinformatics/btz232	https://bioepic.readthedocs.io/en/latest/
FastQC v. 0.11.7	Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
ChIPseeker v 1.22.0	Yu G, Wang L, He Q (2015). "ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization." *Bioinformatics*, **31** (14), 2382-2383. doi: 10.1093/bioinformatics/btv145.	https://bioconductor.org/packages/release/bioc/html/ChIPseeker.html
Ingenuity Pathway Analysis (IPA)	Qiagen	https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/?gclid=Cj0KCQiAuefvBRDXARIsAFEOQ9GUaRrR3QZy42lJtb5QHvlHsmCrtw5fOeyRo0d2GEgfWxKTIPwd8w8aArVwEALw_wcB

Table S4. Primers used in this study.

A list of all PCR primers used i

	Primer
Trim28 out HSO3 F	ATGGATGTTTATGGAAGTAGTAGAAATA
Trim28 out HSO3 R	ACATCTAAATACTACCCAAAACCATTAC
Trim28 in HSO3 F	AATAGTTTAGGTTTTATTTTTTTTAAGATT
Trim28 in HSO3 R	AAACCATTACTTCAAAATAACTTTAAATTA
Sall4 out HSO3 F	AGGGATTATAATTTTTTGAGTTTTAGTTTATA
Sall4 out HSO3 R	AAAACCTCTAAAAAAACAATCAATACTCTTAA
Sall4 in HSO3 F	TATATAGAGAGGTTTAAATAAAGGGTTTTT
Sall4 in HSO3 R	TTTAAAACCACTAAACATATTAAAACATAA
Sox2 out HSO3 F	AGAAAATTGAGTTATTAAGGTAGTAATTATTT
Sox2 out HSO3 R	AAACCAAAAACCTTAACTACCAAACATAA
Sox2 in HS03 F	TAAAATTTTTATAGTTTTAATTGTTAAATA
Sox2 in HSO3 R	TATTATATCTAAAACCAACTAACAATATTAT
Lefty2 out HSO3 F	TTAGAAGTTTTTGGGGGAGAGGTTTGATTTA
Lefty2 out HSO3 R	TCAAAAATCATAACTCTTCCCACACCTCAAA
Lefty2 in HSO3 F	ATAATATGAGGGAGAGGTTTAGTTTTT
Lefty2 in HSO3 R	CACCTCAAACTCTATCTACTACTAACTTTA
Esrrb out HSO3 F	TTTGGAGAGGAAATATGTTAATTTTGAATA
Esrrb out HSO3 R	AAATCAACACACAAAATTCACTAAAAAAACA
Esrrb in HSO3 F	AATAGGGATTTTTTTTGGGATAGAAAT
Esrrb in HSO3 R	ATTCACTAAAAAAAAAAAAAATCTCAAA
Lefty1 out HSO3 F	AAATAAGGAGGTAGGGGTAGAGAATATTTGA
Lefty1 out HSO3 R	AAAAAACAATCTCCCTCCCACCTAACA
Lefty1 in HSO3 F	TTTAGAGGAGAAGTTAAGTTTAGTATAGAGAATA
Lefty1 in HSO3 R	ACACCTAATCAAACCCATTATACAAAAT
Oct3/4 out HSO3 F	TAATGGGATTTTTAGATTGGGTTTAGAAAA
Oct3/4 out HSO3 R	TAACCCTAAACAAATACTCAACCCTTAAAT
Oct3/4 in HSO3 F	TTTGAGGGTTATTTTTTTGTAAAGATAA
Oct3/4 in HSO3 R	AAAAAAAATATCTAACTTCAAATTCAAA
H19 out HSO3 F	GAGTATTTAGGAGGTATAAGAATT
H19 out HSO3 R	ATCAAAAACTAACATAAACCCT
H19 in HSO3 F	GTAAGGAGATTATGTTTATTTTTGG
H19 in HSO3 R	CCTCATTAATCCCATAACTAT
For Bisulfite PCR (promoters):	
Sox2 out HSO3 F	AAGGTTTTTTTTGAAATTGTTATTATTAAA
Sox2 out HSO3 R	AATATTAACTCTCTCTTAAATACATAA
Sox2 in HS03 F	AAAAATTTTTTTGTGTTAGGGTTGGGAGTTA
Sox2 in HSO3 R	ACTATCCAACTAATATTTCAAAAAACTAACAAA
Esrrb out HSO3 F	AATTTGTAAAAGATAGGGTTTTTGATTTGAAGTTTA
Esrrb out HSO3 R	TATAAACCAAAACCCCCACCCAAAACCTAAAT
Esrrb in HSO3 F	AATTTAAAGTTTTTTAGAATTTTGAAGT

Esrrb in HSO3 R	AATCCTCCATAATACCACTAATACAATCCT
Trim28 out HSO3 F	TTTATTTTGTTTTTGTGTTTTTGGTATTTTAAAG
Trim28 out HSO3 R	TCTAATCTAAAATAAAAAATAACAAAATCAAAT
Trim28 in HSO3 F	ATTATATTTTTTTTTTTTGAGGTAGGAAGT
Trim28 in HSO3 R	TTTCATCTCAAAAACAAAATCAATAT
Lefty1 out HSO3 F	ATAATAAAGGGTGGTTTTGTTGGAGAGTTAGGGAATA
Lefty1 out HSO3 R	TATCCTTTCCCTTTCTTACTATAAAAAAATT
Lefty1 in HSO3 F	TTTTATAATTTATTATGGTTTTTTTATTTAT
Lefty1 in HSO3 R	AAATCTACCCCATCCCAATCTCTAAACCAAAATACA
Lefty2 out HSO3 F	ATTTATTAAGTATTAAGATTTAGTTGTTGTAAA
Lefty2 out HSO3 R	AACTACAACTATATATCACTATACCAAACCTATTTAA
Lefty2 in HSO3 F	AGATTTGGTTTTTGGATGGTTTTTGTTTTTAA
Lefty2 in HSO3 R	ATTCTCCATAAACCAATAAAAATCTAAATTT
Sall4 out HSO3 F	TAAAGGAATAGAATTATATTATATAATGTTTTTTTTAA
Sall4 out HSO3 R	ATCTTTCTACTTTACATTTTCTTAATTAACTATTT
Sall4 in HSO3 F	TTTTTTGAGTTGTAAATTGTATATTTTTTTTAGATTT
Sall4 in HSO3 R	AATAAATCAACTAAACCACAAATTTAATAT
For RT-qPCR:	**Primer**
Oct3/4 F	TCTTTCCACCAGGCCCCCGGCTC
Oct3/4 R	TGCGGGCGGACATGGGGAGATCC
Sox2 F	ATGATGGAGACGGAGCTGAA
Sox2 R	TTGCTGATCTCCGAGTTGTG
Esrrb F	CCTCATCAACTGGGCCAAGC
Esrrb R	TACACGATGCCCAAGATGAGAATCT
Nanog F	ACCTGAGCTATAAGCAGGTTAAGAC
Nanog F	GTGCTGAGCCCTTCTGAATCAGAC
Trim28 F	GAGATGGAGAGCGAACAGTCTAC
Trim28 R	TGTCACAGCTCTCACAGAACAG
Prdm14 F	ACAGCCAAGCAATTTGCACTAC
Prdm14 R	TTACCTGGCATTTTCATTGCTC
Lefty1 F	CTACAACACAGCCATGCCAG
Lefty1 R	CTCCATTCCGAACACTAGCAGGT
Lefty2 F	GCAAGAGGTTCAGCCAGAATTT
Lefty2 R	GCTGCTCCATTCCGAACACTA
Olig2 F	CGCAGCGAGCACCTCAAATCTAA
Olig2 R	CCCAGGGATGATCTAAGCTCTCGAA
Foxa2 F	CTCTTCCGTGAGCAACATGA
Foxa2 R	GCGCCCACATAGGATGAC
Gata6 F	CTTGCGGGCTCTATATGAAACTCCAT
Gata6 R	TAGAAGAAGAGGAAGTAGGAGTCATAGGGACA
Gata4 F	CTCTATCACAAGATGAACGGCATCAAC
Gata4 R	TCTGGCAGTTGGCACAGGAGAG
Cxc4 F	GAAGTGGGGTCTGGAGACTAT
Cxcr4 R	TTGCCGACTATGCCAGTCAAG
Fgf5 F	GAAAAGACAGGCCGAGAGTG
Fgf5 R	GAAGTGGGTGGAGACGTGTT

Sall4 RT F	GTGCTCCAGTGAACTCCCC
Sall4 RT R	ACAGCATTTGTTGCAGATGTGA
Dnmt3l F	ATGGACAATCTGCTGCTGACTG
Dnmt3l R	CGCATAGCATTCTGGTAGTCTCTG
Dnmt3a F	GTCACACAGAAGCATATCCAGGAG
Dnmt3a R	GTTGACAATGGAGAGGTCATTGC
Lsd1 RT F	GATGGGATTTGGCAACCTTAACAAG
Lsd1 RT R	AAACAAATTGACACTTGGGTCCC
Gapdh RT F	CAAAATGGTGAAGGTCGGTGTGAA
Gapdh RT R	CAACAATCTCCACTTTGCCACTG
For ChIP-qPCR:	**Primer**
Sox2 F	TGGCGAGTGGTTAAACAGAG
Sox2 R	TAGCGAGAACTAGCCAAGCA
Esrrb F	CGAGCTTCAGCTGGCTATTT
Esrrb R	GAGCTCCAGATCCCCTACAC
Trim28 F	GGTCTGCAATTGAAGGAAGG
Trim28 R	TTAAACAGCAGGGGGTAAGG
Lefty1 F	GTAGCCAGCAGACAGGACAA
Lefty1 R	ATCCCCAATCCACATTCACT
Lefty2 F	AGGCCTAGCTTTTGCATCAC
Lefty2 R	TCTCCCAGAGTCGATCTTCC
Sall4 F	GAAATAAACATCTGGGAGAAGGA
Sall4 R	GGAAACCCCAGATTGAGAGA
Oct3/4 F	TATAGGTGTGGCATTCCGCATC
Oct3/4 R	TGCCACAAACCACCTGTATTTTAG
Ctrl F	ACCTAAACCTCATAAAGACACAACA
Ctrl R	TGACGTGTTCTTGGATTCAGT
MethylRAD-seq adaptor and index primers	
Adaptor-1 sense	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
Adaptor-1 antisense	NNNNAGATCGGAAGAGC(AminoC6)
Adaptor-2 sense	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Adaptor-2 antisense	NNNNAGATCGGAAGAGC(AminoC6)
P1	ACACTCTTTCCCTACACGACGCT
P2	GTGACTGGAGTTCAGACGTGTGCT
i5- UDA5073	ATGATACGGCGACCACCGAGATCTACACGATAACAAGTACACTCTTTCCCTACACGACGC
i5- UDA5074	ATGATACGGCGACCACCGAGATCTACACAGCGGTGGACACACTCTTTCCCTACACGACGC
i5- UDA5075	ATGATACGGCGACCACCGAGATCTACACGGTTATGCTAACACTCTTTCCCTACACGACGC
i5- UDA5076	ATGATACGGCGACCACCGAGATCTACACAACCGCATCGACACTCTTTCCCTACACGACGC
i7- UDA7120	CAAGCAGAAGACGGCATACGAGATGAACTTCCTTGTGACTGGAGTTCAGACGTGT
i7- UDA7119	CAAGCAGAAGACGGCATACGAGATAGGTCCTTCCGTGACTGGAGTTCAGACGTGT
i7- UDA7118	CAAGCAGAAGACGGCATACGAGATTATGGAGATTGTGACTGGAGTTCAGACGTGT
i7- UDA7117	CAAGCAGAAGACGGCATACGAGATCGCAAGAGCCGTGACTGGAGTTCAGACGTGT

:T
:T
:T
:T

Figure 1

Click here to access/download;Figure;Figure 1.tif

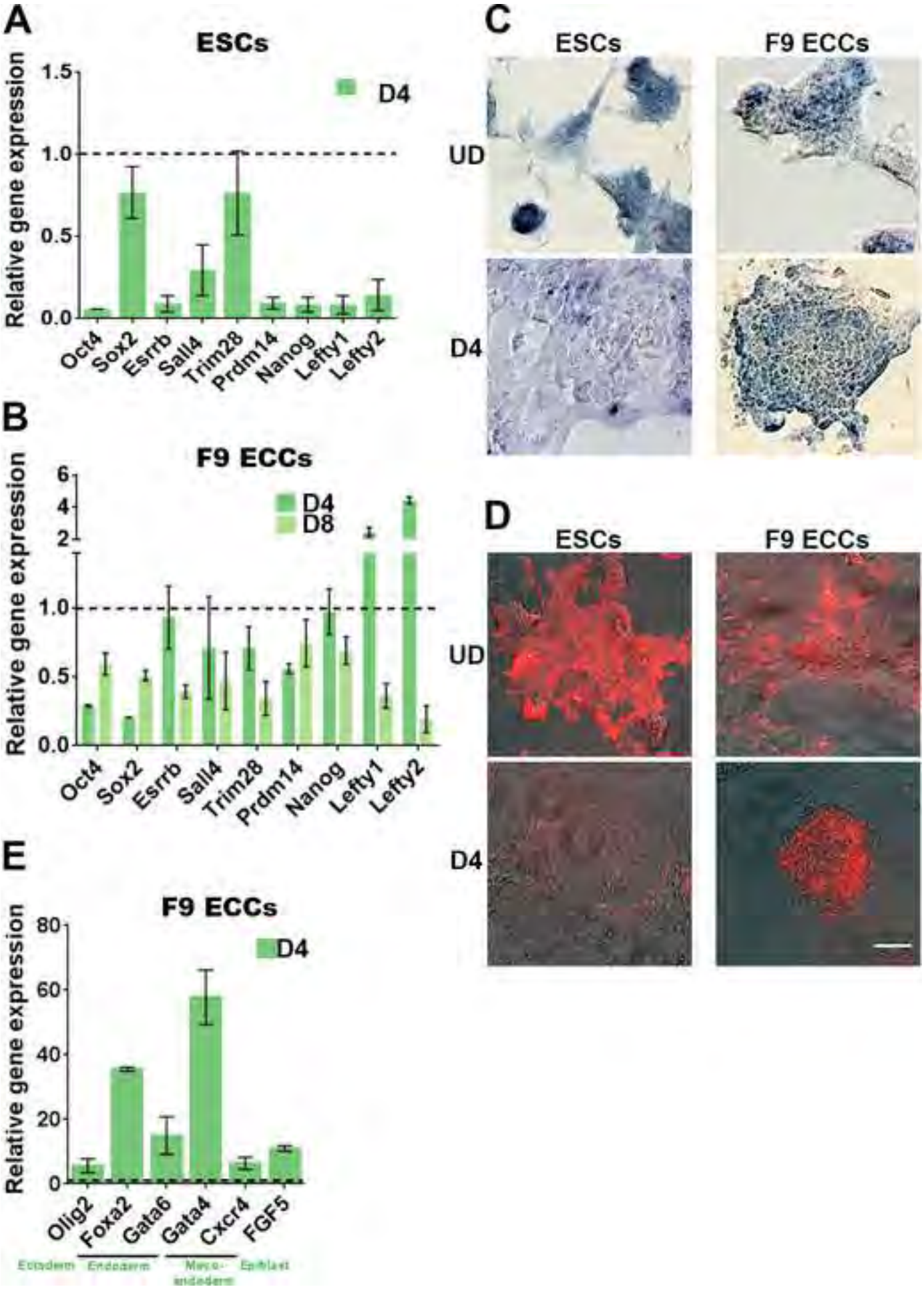


Figure 2

Click here to access/download;Figure;Figure 2.tif

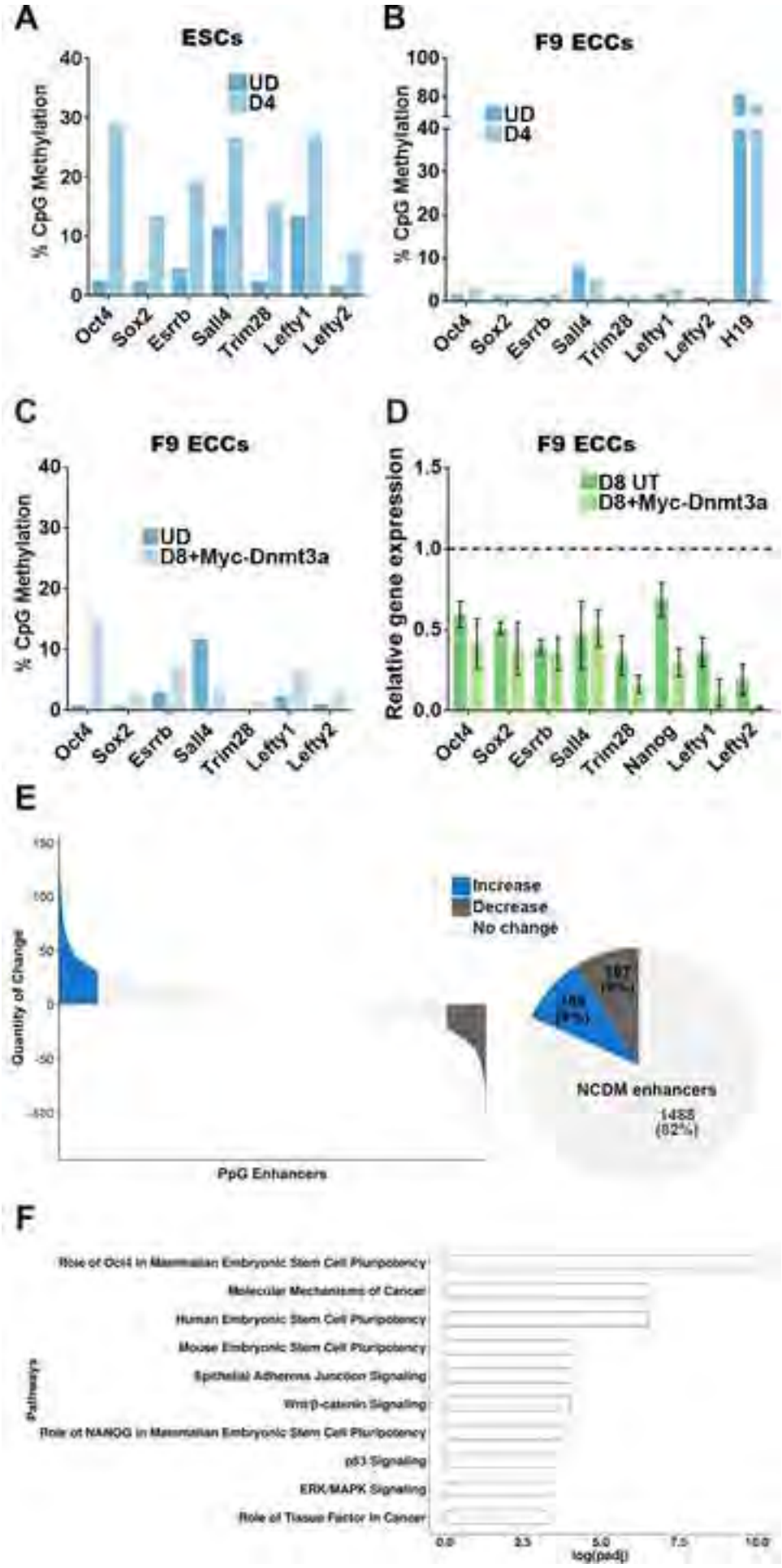


Figure 3

Click here to access/download;Figure;Figure 3-2.tif

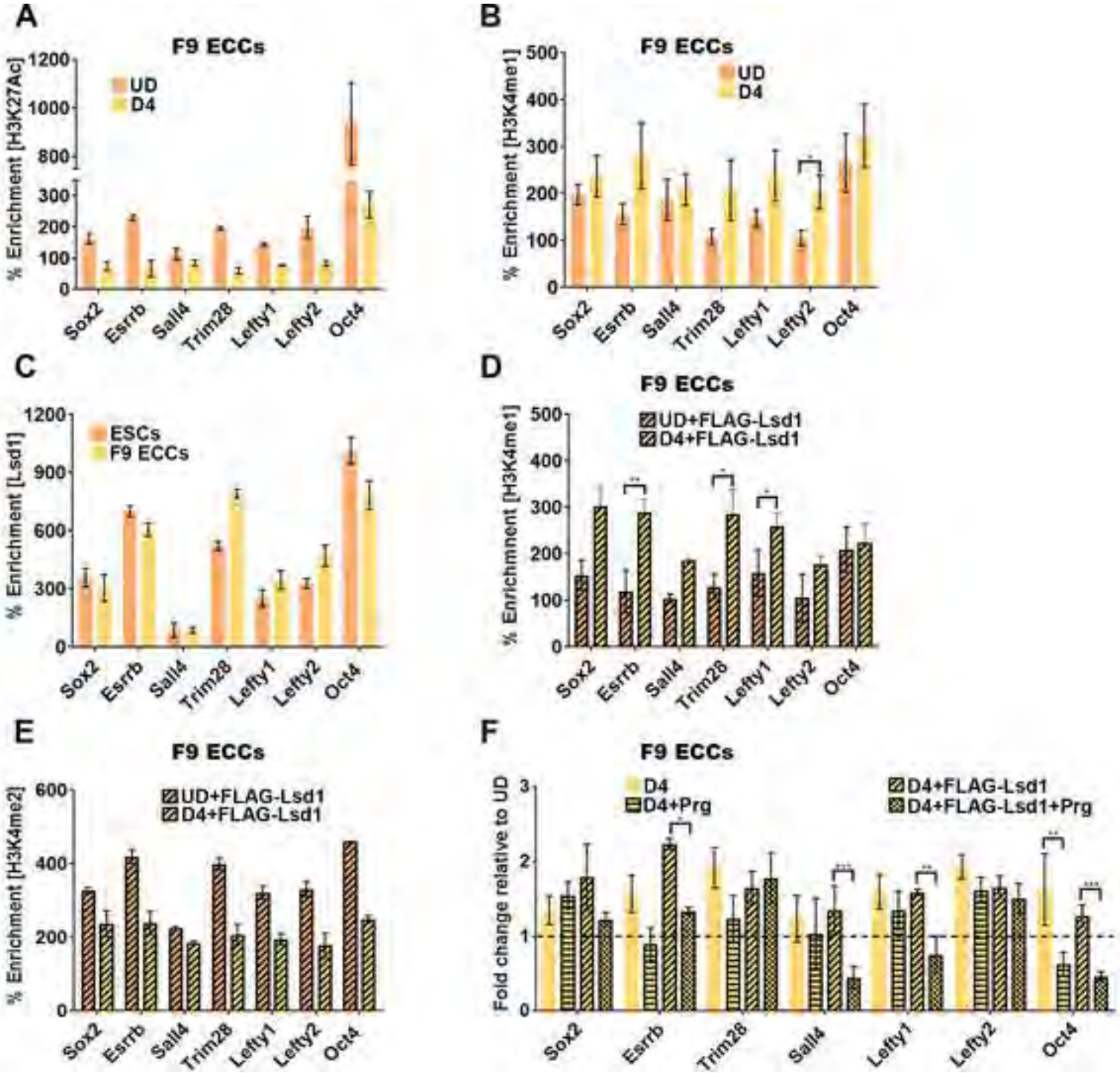


Figure 4

Click here to access/download;Figure;Figure 4.tif

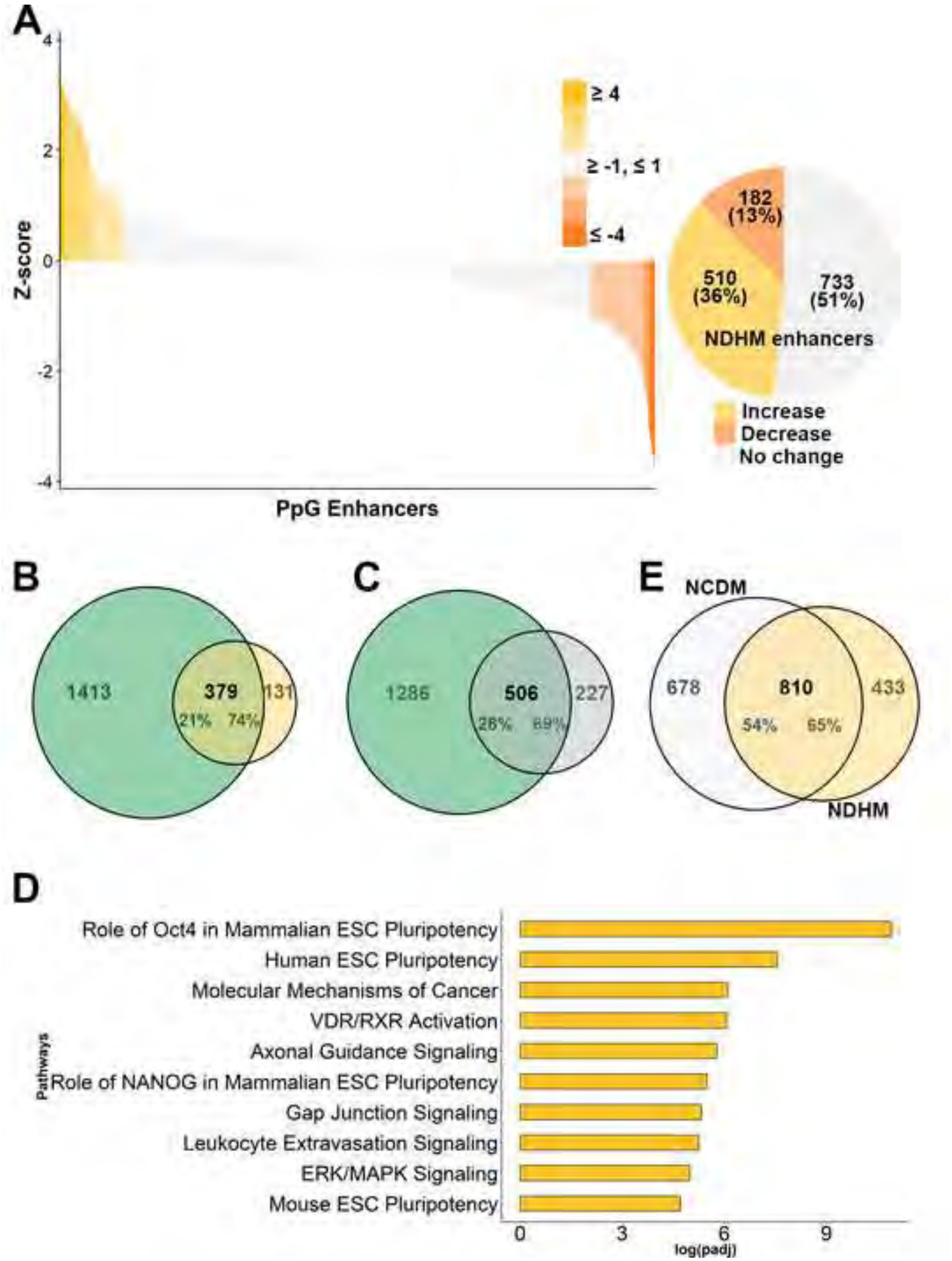


Figure 5

Click here to access/download;Figure;Figure 5.tif

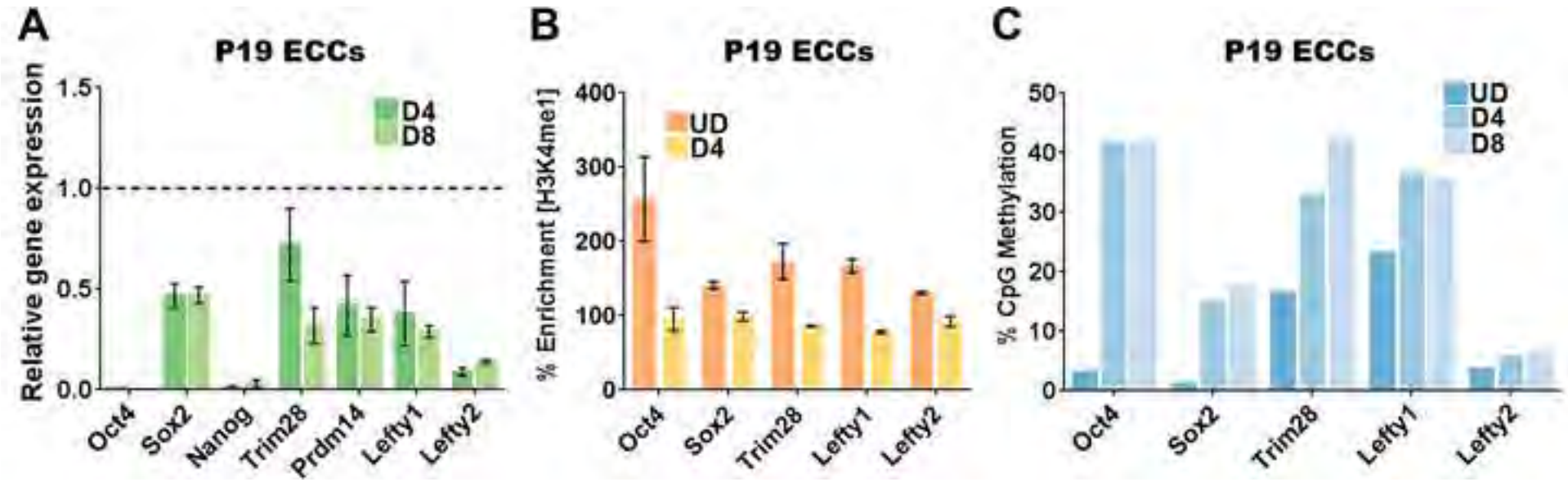


Figure 7

Click here to access/download;Figure;Figure 7.tif

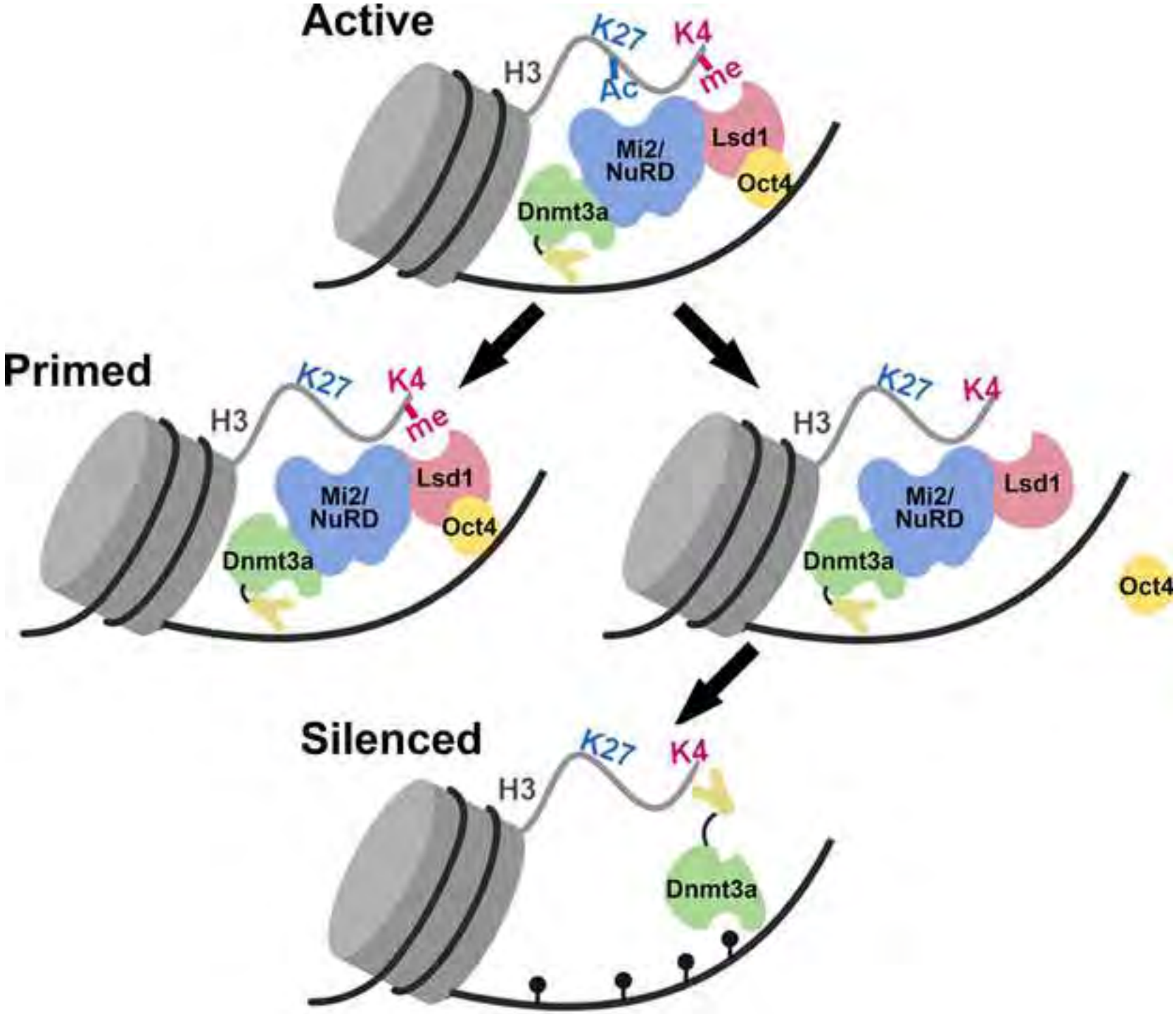
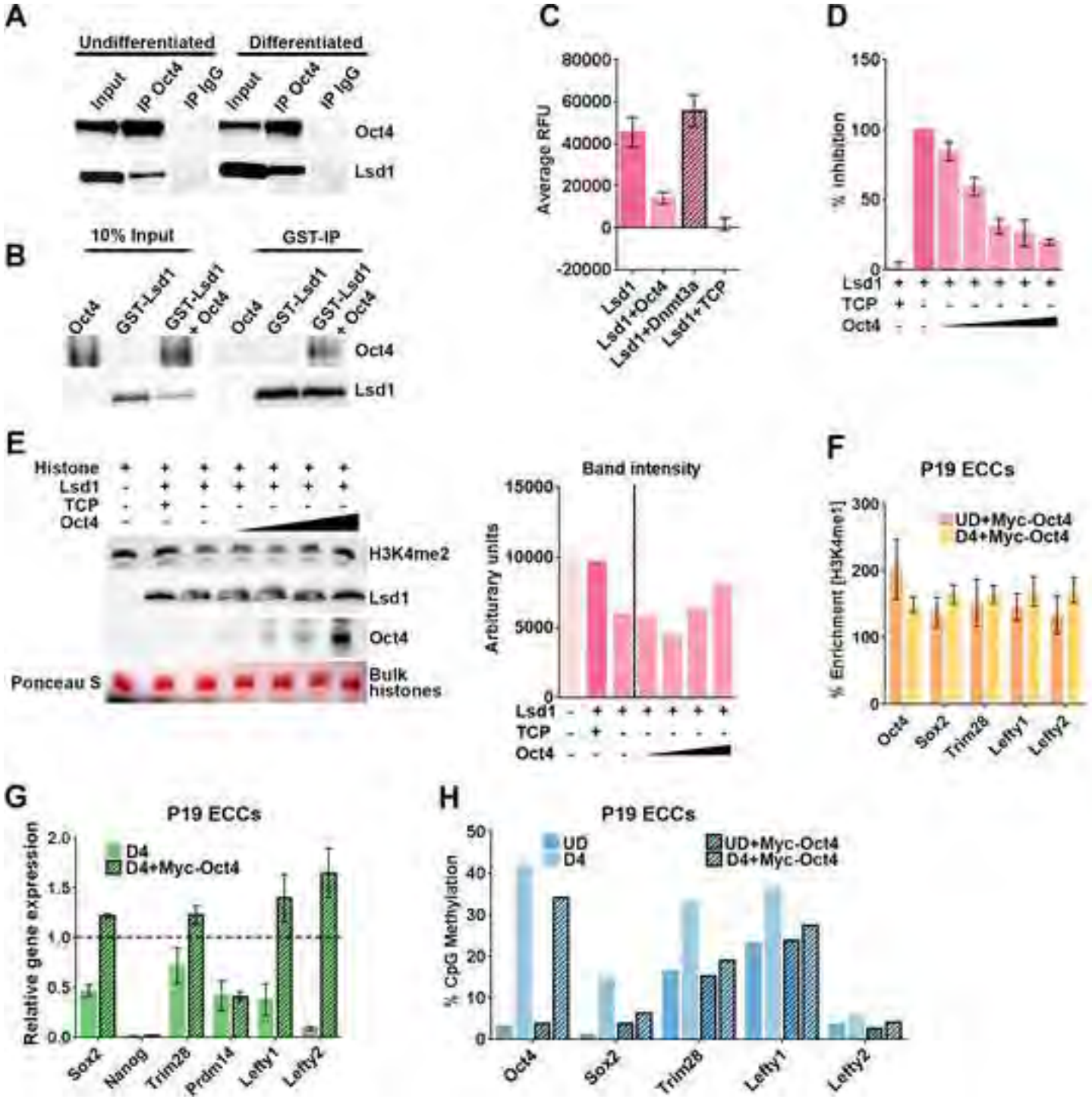


Figure 6

Click here to access/download;Figure;Figure 6-11.tif



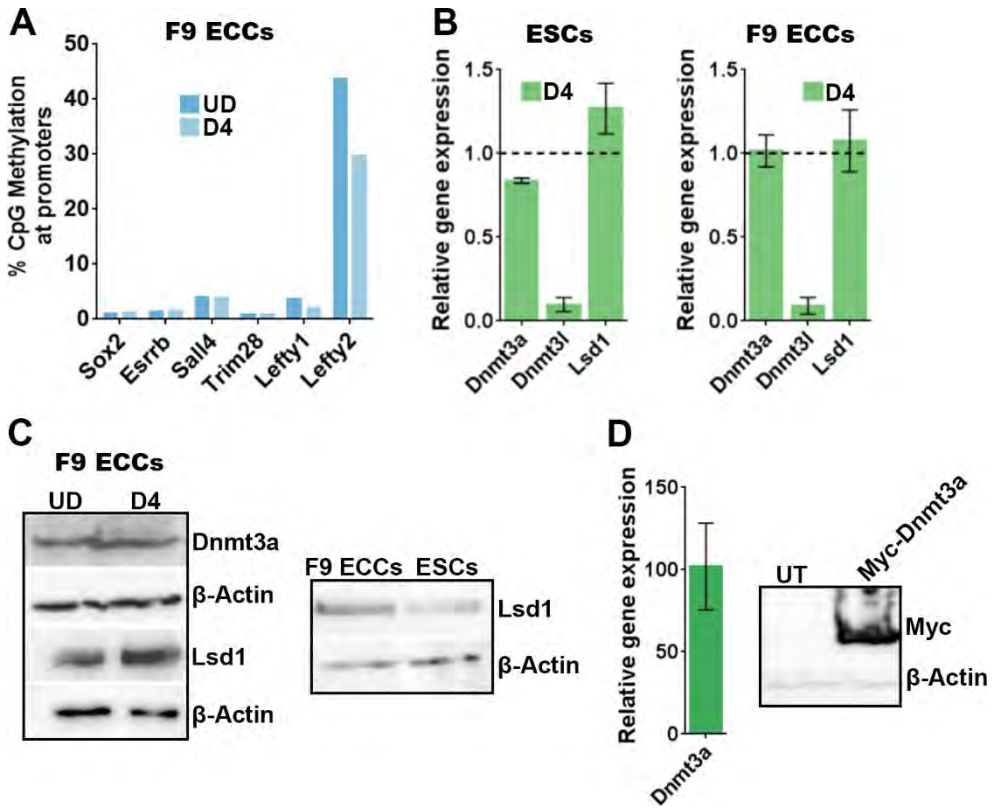


Figure S1. Expression of epigenetic effectors in embryonal carcinoma cells. (**Related to Figure 1)**

UD: undifferentiated, D4: Days post-differentiation. ESCs: embryonic stem cells and F9 ECCs: F9 embryonal carcinoma cells; PpGe: pluripotency gene enhancers.

(A) Bis-Seq analysis of DNA methylation at PpG promoters in F9 ECCs pre- and post-differentiation. DNA methylation at PpG promoters remained under 10% except *Lefty2*, which shows very high methylation in the UD state that is reduced post-differentiation.

(B) RT-qPCR comparing the changes in expression of DNA methyltransferase Dnmt3a and histone demethylase Lsd1 in ESCs with that in F9 ECCs pre- and post-differentiation. The C_t values are normalized to *Gapdh* and represented relative to expression in undifferentiated cells (dotted line). In both ESCs and ECCs, Lsd1 and Dnmt3a expression is maintained post-differentiation.

(C) Western blot. A total of 50 µg of total protein from undifferentiated and differentiated cells was loaded in each well. Left panel confirms comparable expression of Dnmt3a and Lsd1 pre- and post-differentiation in F9 ECCs. Right panel compares Lsd1 expression in F9 ECCs with ESCs showing very similar levels in these cells. β-Actin is the loading control.

(D) Gene expression analysis by RT-qPCR and Western blot confirming recombinant Myc-Dnmt3a overexpression in F9 ECCs 24 hr post-differentiation (48 hr post-transfection). The data are normalized to a *Gapdh* control and shown relative to untransfected that is set to 1. β-Actin is used as loading control for Western blot.

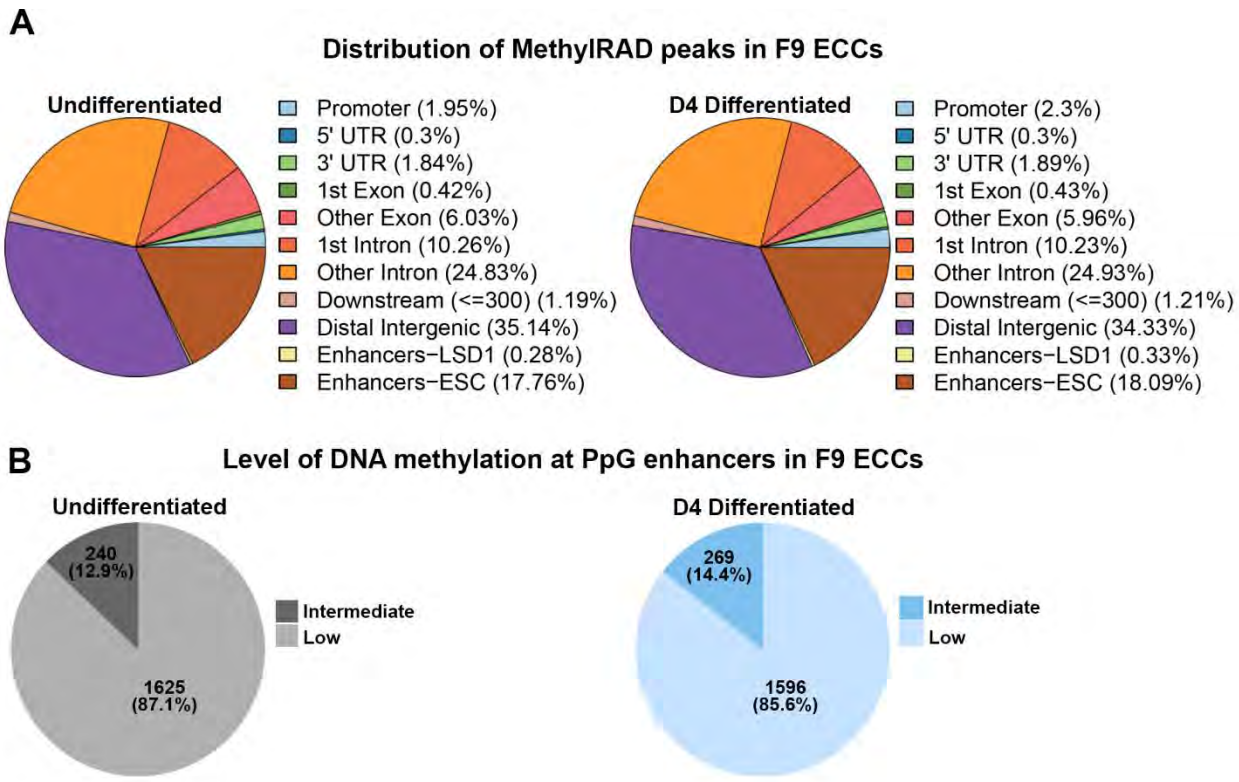


Figure S2. Distribution of MethylRAD peaks. (Related to Figure 2)

(A) Fractional distribution of MethylRAD peaks in undifferentiated and differentiated F9 ECCs across regulatory regions of the genome.

(B) Pie charts show the level of methylation at PpGe in undifferentiated (grey) and differentiated day 4 (blue) samples. Most PpGe had low levels of methylation or to a lesser extent, intermediate levels of methylation, with none having high CpG methylation.

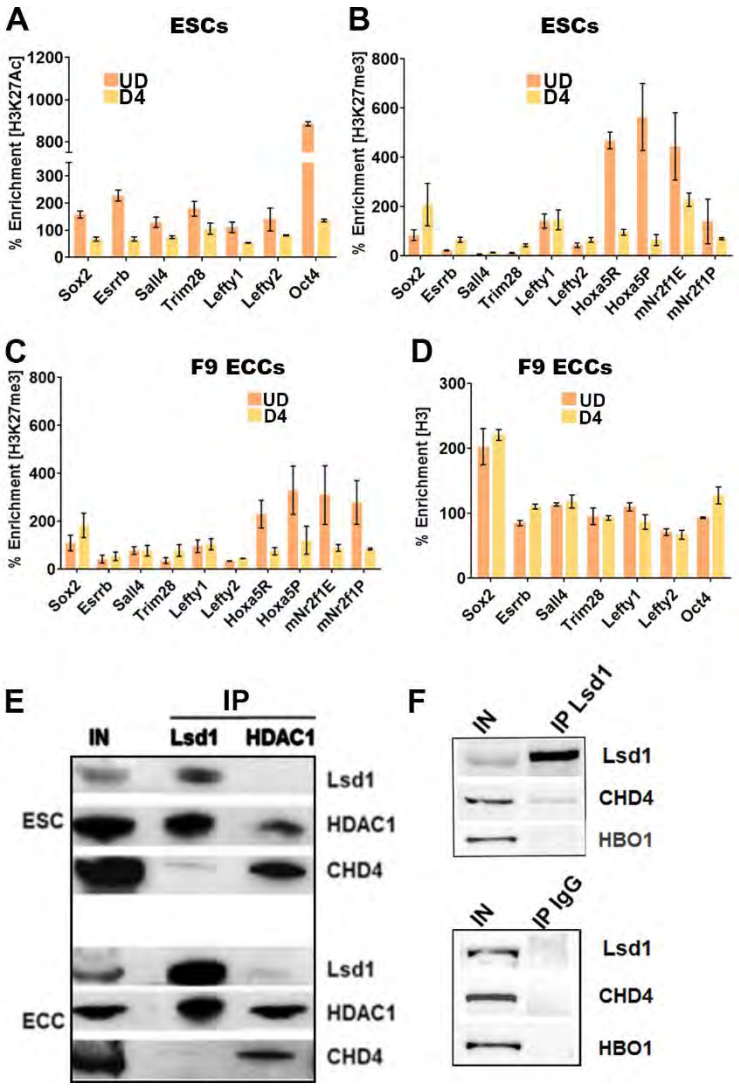


Figure S3. H3K27 modification of pluripotency gene enhancers in embryonic stem cells and embryonal carcinoma cells. (Related to Figure 3)

UD: undifferentiated; D4: Days post-induction of differentiation; IP: immunoprecipitation

(A, B, C and D) ChIP-qPCR was used to determine the enrichment of histone modifications at PpGe. (A) H3K27Ac in ESCs (B) H3K27me3 in ESCs (C) H3K27me3 and (D) H3 in F9 ECCs pre- and post-differentiation. Whereas deacetylation of PpGe is observed as a decrease in H3K27Ac signal post-differentiation in ESCs, there is no gain of H3K27me3 at these sites,

neither in ESCs nor in F9 ECCs. As previously reported, we observed a decrease in H3K27me3 at the enhancers and promoters of *Hoxa5* and *mNr2f1* genes, consistent with their transcriptional activation in response to differentiation (Laursen, Mongan et al. 2013). No change was observed in H3 occupancy between UD and D4 at PpGe. % Enrichment = Fold enrichment over input X100

(E, F) Co-IP was performed with anti-Lsd1 or anti-HDAC1 and a control IgG using whole cell extracts from undifferentiated ESCs and F9 ECCs. 20% of the input and eluate from Co-IP were probed for Lsd1-Mi2/NuRD subunits (Lsd1, HDAC1 and CHD4) on Western blot and HBO1 was used as a negative control.

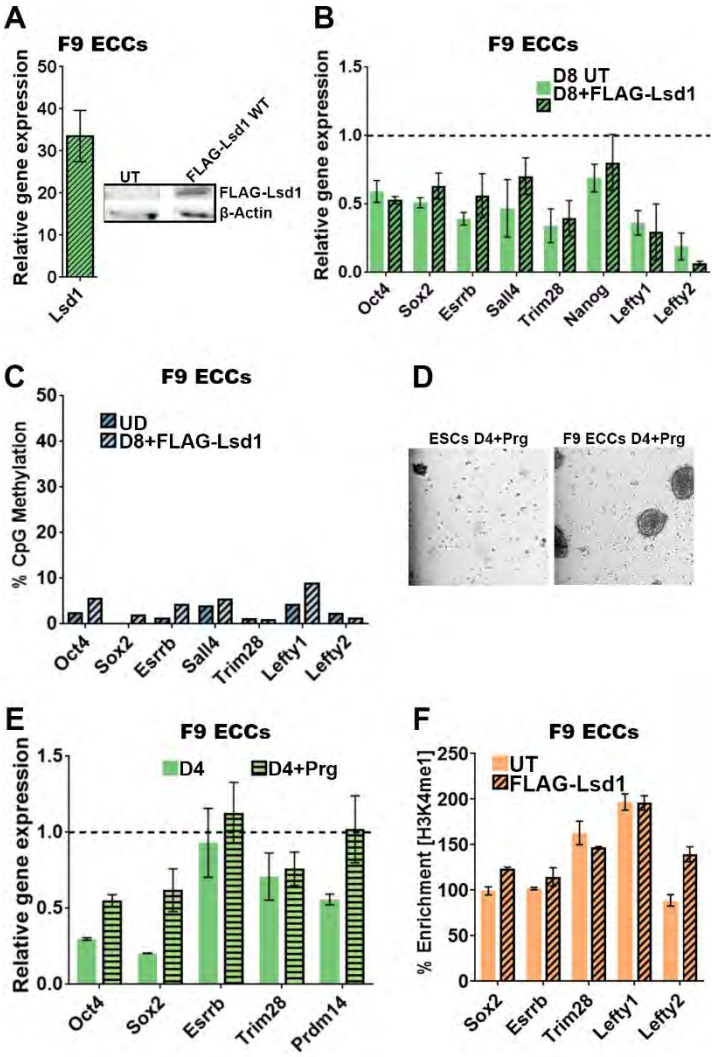


Figure S4. Effect on Lsd1 on differentiation and pluripotency gene enhancer silencing in embryonal carcinoma cells. (Related to Figure 3)

UD: undifferentiated; D8: Days post-induction of differentiation; Prg: pargyline.

(A) Gene expression analysis by RT-qPCR and Western blot examining the expression of recombinant FLAG-Lsd1 in F9 ECCs 24 hr post-differentiation (48 hr post-transfection). The C_t values are normalized to *Gapdh* and represented relative to expression in untransfected cells (set to 1). β-Actin is used as loading control for Western blot.

(B) Gene expression analysis by RT-qPCR of PpGs in F9 ECCs expressing FLAG-Lsd1. Similar to our observation in untransfected WT F9 ECCs, PpGs are partially repressed in these cells, showing no effect of recombinant Lsd1 on PpG repression (*p*-value >0.1).

(C) Bis-Seq analysis of DNA methylation at PpGe in F9 ECCs expressing recombinant FLAG-Lsd1 showed no gain, corresponding to no loss of H3K4me1 at these sites. Data are an average and SEM of two biological replicates.

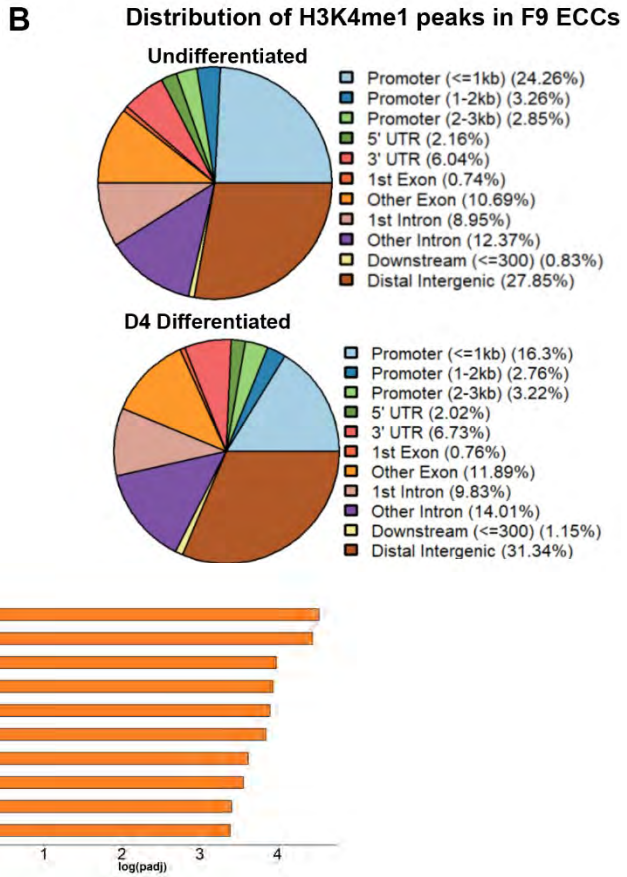
(D) Bright field microscopy of ESCs and ECCs differentiated for 4 days in the presence of the Lsd1 inhibitors Prg. 80-90% cell death is observed in ESCs. Lsd1 inhibitor has no effect on F9 ECC differentiation as shown by a normal morphology of embryoid bodies. Scale bar is 100 µm.

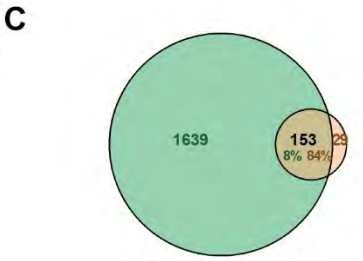
(E) Gene expression analysis of PpGs by RT-qPCR in differentiating F9 ECCs untreated or treated with Prg. A slight derepression of some PpGs was observed in inhibitor treated cells post-differentiation. Gene expression was normalized to *Gapdh* and represented as relative change to gene expression in undifferentiated (dotted line).

(F) Undifferentiated F9 ECCs were transfected with FLAG-Lsd1 and cultured for 72 hr. H3K4me1 enrichment was determined using ChIP-qPCR. No change in H3K4me1 levels was observed compared to untransfected (UT) cells. % Enrichment = Fold enrichment over input X100

A

With mm10 converted coordinates			
Sample	Epic2_peaks	LSD1 Bound	
		Single peak nearest to the gene	Within 1KB of Whyte et al Peaks
F9-undifferentiated-H3K4me1	25,382	1,837	965
F9-D4 differentiated-H3K4me1	44,237	2,002	1,103





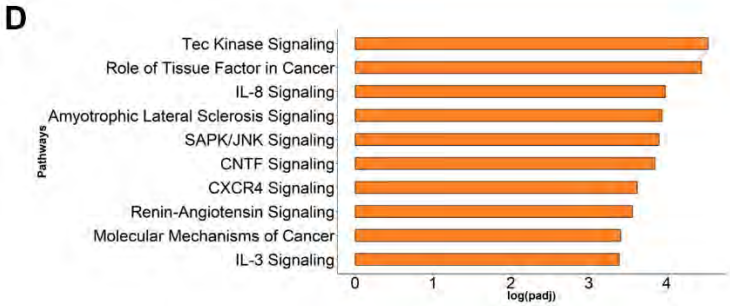


Figure S5. Distribution of H3K4me1 ChIP-Seq peaks. (Related to Figure 4)

(A) Summary of the overlapping LSD1 bound sites between H3K4me1 peaks in F9 ECCs and ESCs. To determine correct overlap, peak coordinates from Whyte.et al were converted from mm9 to mm10 using CrossMap tool.

(B) Fractional distribution of H3K4me1 peaks in undifferentiated and differentiated F9 ECCs throughout the genome.

(C) Overlap of the 182 sites that show decrease in H3K4me1 post-differentiation in F9 ECCs with those in ESCs (See Supplementary methods).

(D) Top ten most statistically significant enriched canonical pathways amongst the genes associated with decrease in F9 ECCs.

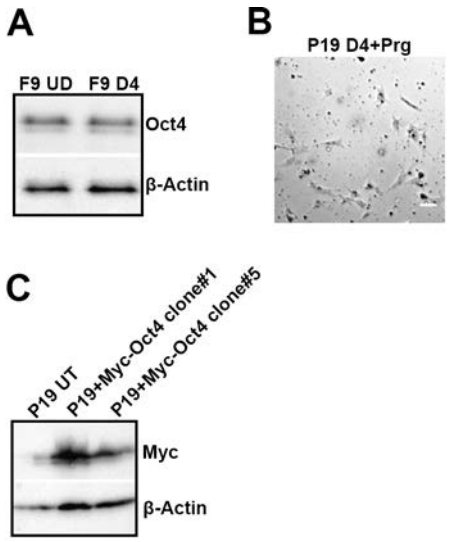


Figure S6. P19 embryonal carcinoma cells are sensitive to Lsd1 inhibition. Related to Figure 5 and Figure 6

D4: Days post-induction of differentiation; Prg: pargyline; UT: untransfected control.

(A) Western blot showing protein levels of Oct4 in undifferentiated and D4 differentiated F9 ECCs. A continued expression of Oct4 is seen in differentiated cells although it is lower than that observed in the undifferentiated cells. β-Actin was used as a loading control.

(B) Bright field microscopy of P19 ECCs differentiated for 4 days in presence of Lsd1 inhibitor, Prg. 80-90% cell death is observed in P19 ECCs indicating that Lsd1 activity at PpGe is required for differentiation. Scale bar is 100 µm.

(C) Expression analysis by Western blot confirming recombinant Myc-Oct4 overexpression in two independent clones of P19 ECCs. β-Actin was used as a loading control for Western blot.

Supplementary Tables

Genomic element		**No. of CpGs**
Enhancer	Lefty1	4
	Lefty2	6
	Sall4	4
	Sox2	10
	Trim28	6
	Esrrb	6
	Oct4	4
Promoter	Lefty1	6
	Lefty2	9
	Sall4	10
	Sox2	18
	Trim28	19
	Esrrb	29
Imprinted region	H19	11

Table S1. Number of CpGs at each site used for Bis-Seq DNA methylation analysis. Related to DNA methylation analysis in Method Details

The number of CpG sites used to compute percent methylation within the H19 imprinted region, PpG enhancers, and promoters.

Sample	*Lefty1*	*Lefty2*	*Sall4*	*Sox2*	*Trim28*	*Esrrb*	*Oct4*	***Total***
ESCs UD	28465	21981	2861	5826	17973	6034	1902	**85042**
ESCs D4	401	275	216	881	269	5381	6322	**13745**
F9 ECCs UD	57536	49939	27259	40357	14261	22395	6322	**232670**
F9 ECCs D4	62876	38534	29204	32570	13304	21569	8865	**214866**
F9 ECCs D6	62185	38590	30187	28239	10163	13059	6980	**182422**
F9 ECCs+Myc-Dnmt3a UD	3733	2692	1308	2612	1249	24671	311	**38969**
F9 ECCs+Myc-Dnmt3a D8	2734	1637	1135	2665	1097	4316	350	**24705**
F9 ECCs+FLAG-Lsd1 UD	2513	2267	816	4204	820	24412	1791	**43102**
F9 ECCs+FLAG-Lsd1 D8	2819	1950	1946	4530	1169	21576	1601	**42022**
P19 ECCs UD	26651	18427	10541	31377	9437	46154	7044	**175098**
P19 ECCs D4	40376	30470	18198	27196	11267	56736	14506	**241010**
P19 ECCs D8	4713	8696	5880	28944	6475	24128	4498	**96640**
P19 ECCs+Myc-Oct4 UD	23029	11988	11538	9324	7735	28746	7510	**130029**
P19 ECCs+Myc-Oct4 D4	35014	20519	14300	18022	13227	43559	14989	**210333**
Total	**353045**	**247965**	**156838**	**236746**	**108446**	**342735**	**76669**	**1522445**

Table S2. Number of reads for each enhancer used for Bis-Seq analysis. Related to DNA methylation analysis in Method Details

The total number of reads from Wide-Seq runs that were used for data presented in this study. The number of reads were calculated for each sample and enhancer, along with the overall total number of reads.

Sample	*Lefty1*	*Lefty2*	*Sall4*	*Sox2*	*Trim28*	*Esrrb*	*H19*	***Total***
F9 ECCs UD	51	726	486	9556	7575	4191	11062	**33647**
F9 ECCs D4	99	1590	1400	17431	16307	13865	9376	**60068**
Total	**150**	**2316**	**1886**	**26987**	**23882**	**18056**	**20438**	**93715**

Table S3. Number of reads for promoters and imprinted regions used for Bis-Seq analysis. Related to DNA methylation analysis in Method Details

The total number of reads from Wide-Seq runs that were used in this study. The number of reads were calculated for each sample, promoter site, and imprinted locus, as well as the overall total number of reads.