Loss of MLL3/4 decouples enhancer H3K4 monomethylation, H3K27 acetylation, and gene activation during ESC differentiation

Ryan M. Boileau[1,2,3], Kevin X. Chen[1,2,3], Robert Blelloch[1,2,3,#]

[1]The Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, Center for Reproductive Sciences, University of California San Francisco, San Francisco, USA

[2]Department of Urology, University of California San Francisco, San Francisco, USA

[3]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA

[#]Corresponding Author

Contact Information
Ryan M. Boileau, Ryan.Boileau@ucsf.edu
Kevin X. Chen, Kevin.Chen4@ucsf.edu
Robert Blelloch, Robert.Blelloch@ucsf.edu

Abstract

Enhancers are essential in defining cell fates through the control of cell type specific gene expression. Enhancer activation is a multi-step process involving chromatin remodelers and histone modifiers including the monomethylation of H3K4 (H3K4me1) by MLL3 (KMT2C) and MLL4 (KMT2D). MLL3/4 are thought to be critical for enhancer activation and cognate gene expression including through the recruitment of acetyltransferases for H3K27. Here we test this model by evaluating the impact of MLL3/4 loss on chromatin and transcription during early embryonic stem cell differentiation. We find that MLL3/4 activity is required at most if not all sites that gain or lose H3K4me1 but is largely dispensable at sites that remain stably methylated during this transition. This requirement extends to H3K27 acetylation (H3K27ac) at most transitional sites. However, many sites gain H3K27ac independent of MLL3/4 or H3K4me1 including enhancers regulating key factors in early differentiation. Furthermore, despite the failure to gain active histone marks at thousands of enhancers, transcriptional activation of nearby genes is largely unaffected, thus uncoupling the regulation of these chromatin events from transcriptional changes during this transition. These data challenge current models of enhancer activation and imply distinct mechanisms between stable and dynamically changing enhancers. Collectively, our study highlights gaps in knowledge about the steps and epistatic relationships of enzymes necessary for enhancer activation and cognate gene transcription.

Keywords

Enhancer, epistasis, pluripotency, MLL3, MLL4, H3K4me1, P300, H3K27ac, transcription, differentiation

83 **Background**

84 Gene regulatory networks that drive cell fate are regulated spatiotemporally by cell type
85 specific transcription factors (TFs). The critical functions of TFs in development are coupled to
86 their target genes through TF binding of *cis*-regulatory elements such as enhancers. Target gene
87 regulation is mediated by general chromatin regulators which are recruited to enhancers bound
88 by TFs. As such, chromatin regulators are an essential part of enhancer function and their
89 dysregulation has severe consequences in development and disease. For example, MLL3 and
90 MLL4 are functionally redundant histone methyltransferases within the COMPASS complex which
91 deposit the histone modification H3K4me1 primarily at enhancers(1,2). Moreover, mutations in
92 MLL3 and MLL4 cause developmental defects such as Kabuki Disease and are in the top ten of
93 frequently mutated genes in cancer(3–5). Despite the importance of chromatin regulators to
94 human health, we lack fundamental insight into how general chromatin regulators such as MLL3/4
95 coordinate enhancer function.

96 The interdependence of different histone modifications and their depositing enzymes at
97 enhancers has been widely studied, although their mechanistic roles remain controversial. A
98 prevailing model of enhancer activation is that DNA bound TFs recruit MLL3/4 which deposit
99 H3K4me1(6–9). MLL3/4 in turn recruits the acetyltransferases P300/CBP which deposit
100 H3K27ac. An activated enhancer then promotes expression of its cognate gene, most often one
101 that is a nearby neighbor(10,11). However, it is unclear how generalizable this model is. Indeed,
102 recent studies have begun to challenge the importance of histone modifications and even
103 chromatin regulators in gene regulation by enhancers(12–16).

104 Therefore, we decided to revisit the prevailing model including the epistatic relationship
105 between MLL3/4 to H3K4me1, H3K27ac, and gene transcription in the context of ESCs
106 transitioning from the naive to formative pluripotent states. The naive to formative transition is a
107 rapid, homogenous, and well-characterized transition that faithfully recapitulates the early
108 embryonic cell fate transition in epiblast cells of the mouse between E4.5 and E5.5 at the

109 transcriptional and epigenomic level(Fig1A)(17–20). Importantly, this developmental window
110 immediately precedes previously described cell fate specification and migrational phenotypes
111 seen in gastrulation with the loss of MLL3/4 *in vivo*(6,21). Therefore, the naive to formative
112 transition provides an ideal model to study relationships in enhancer mechanics and
113 transcriptional regulation.

114 Using state-of-the-art epigenomics techniques to profile genetic deletion of both MLL3 and
115 MLL4 in the naive to formative transition we make several surprising discoveries. First, MLL3/4 is
116 only required at a small subset of sites that maintain stable H3K4me1 levels during the transition
117 but is essential for all sites that gain or lose H3K4me1. This demonstrates the existence of distinct
118 mechanisms of H3K4me1 deposition at most stable vs dynamic sites. Second, loss of MLL3/4
119 reduces H3K27ac levels at most dynamic sites, but often does so independent of underlying
120 H3K4me1 levels. Finally, despite dramatic loss of active histone modifications at many enhancers
121 in the formative state, few loci can be linked to changes in nearby gene expression with these
122 changes being relatively minor. Taken together, these results demonstrate that MLL3/4 protein
123 and H3K4me1 deposition can be uncoupled from H3K27ac deposition and also from changes in
124 gene expression at most enhancer sites during early ESC differentiation. Therefore, our data
125 challenge the prevailing model of enhancer activation and suggests that the functional role of
126 MLL3/4 protein in cell fate transitions is more context dependent than previously appreciated.

127

128 **Results**

129 **MLL3/4 is dispensable for transcriptional activation of much of the formative program**

130 To study the role of MLL3 and MLL4 in stem cell self-renewal and differentiation we
131 acquired MLL3$^{-/-}$; MLL4$^{fl/fl}$ ESCs (hereafter called MLL3KO)(22). We integrated Cre-ERT2 into the
132 *Rosa26* locus and induced Cre-recombination with tamoxifen to generate MLL3$^{-/-}$; MLL4$^{-/-}$ double
133 knockout clones (DKO)(Fig.S1A). Western blotting confirmed loss of MLL3 protein alone in the
134 parental line and both MLL3 and MLL4 in the DKO cells compared to a wildtype (WT) control

135 (Fig.1B). Under naive pluripotency culture conditions (Serum+LIF+2i), the DKO cells formed
136 colonies that were less compact and grew at a slower rate compared to their WT or MLL3KO
137 counterparts (Fig.1C, Fig.S1B). Next, we removed LIF+2i from the media to induce the formative
138 state after ~63 hours in WT cells (23). As expected, WT ESC colonies flattened out during this
139 transition. Similar phenotypic changes were seen in MLL3KO and DKO, although in both cases
140 cells appeared more dispersed. Proliferation was reduced in MLL3KO and DKO cells relative to
141 WT, with the greatest impact on transitioning DKO (Fig.S1B). All cell lines expressed OCT4 and
142 NANOG protein in naive culture conditions and similarly downregulated NANOG in formative
143 conditions as indicated by Western blot (Fig.S1C). These data show that while there are
144 morphological and proliferative differences, the knockout cells self-renew under naive culture
145 conditions, have a pluripotent-like cell identity, and show evidence of a cell state transition in
146 formative conditions.

147 To determine the transcriptional effects of loss of MLL3/4 we performed RNA-seq on WT,
148 MLL3KO, and DKO lines in naive and formative conditions. Differential gene expression analysis
149 between WT naive and formative samples uncovered 887 downregulated genes (naive enriched
150 genes) and 1602 upregulated genes (formative enriched genes) (Fig.1D). After subsetting for all
151 naive and formative enriched genes we conducted UMAP analysis of both the WT and mutant
152 cells under both conditions and found samples separated predominately by developmental state
153 rather than by genotype (Fig. 1E). Correlation analysis of the gene expression changes seen in
154 MLL3KO vs WT and DKO vs. WT showed Pearson correlations of 0.8 and 0.6 respectively
155 showing that transitional changes in gene regulation still occur in the absence of MLL3/4
156 (Fig.1F,G). The dynamic range of gene expression changes was reduced in the MLL3/4 DKO as
157 evidenced by a reduced slope in the linear model. This reduction was largely driven by diminished
158 expression of naïve enriched genes in the naïve state (Fig.S1D). Upregulation of the formative
159 program was largely unimpacted by MLL3/4 loss (Fig.S1D). qPCR analysis validated the
160 successful upregulation of known formative markers (Fgf5, Otx2, Dnmt3b)(Fig.S1E). These

161 results suggested a relatively normal overall transition in gene expression changes between naive
162 and formative even in the absence of both MLL3 and MLL4.

163 In contrast, direct comparison of MLL3/4 DKO vs. WT identified hundreds of misregulated
164 genes in both the naive and formative states. In the naive state, DKO cells had 783 upregulated
165 genes and 929 downregulated genes (P.adj < 0.05 and Log2CPM > 1)(Fig.1H). We categorized
166 these genes that were up or down by whether they are normally naive enriched, formative
167 enriched, or unchanged between naive and formative (Fig.1I). This analysis showed a similar
168 proportion of formative genes to be prematurely expressed (275 of 783, 35%) as naive enriched
169 genes that were reduced in the naive state (348 of 928, 39%). In the formative state we found
170 646 genes were upregulated and 461 genes were downregulated in DKO cells compared to
171 WT(Fig.1J). Surprisingly, fewer genes were down in the DKOs in the formative state than the
172 naive state (461 vs. 929). Focusing on genes that normally change during the transition, 83 naive
173 enriched genes were abnormally up and 283 formative enriched genes were abnormally down in
174 the formative state (Fig.1K). These numbers were overall low relative to normal gene expression
175 changes seen with naive to formative transition (887 down, 1602 up) showing that MLL3/4 only
176 regulates a minor subset of genes among those that are normally gained with the transition from
177 naive to formative. Given that the WT and DKO ESCs represent slightly different strain
178 backgrounds (see methods), we also compared MLL3/4DKO to its MLL3KO parental line. This
179 comparison showed even fewer impacted genes (Fig.S1F,G).

180

181 **MLL3/4 is required for all dynamic H3K4me1 deposition during pluripotent transition**

182 Though expression of a small number of genes were impacted it was unclear if this was
183 due to direct or indirect effects of MLL3/4 loss. To gain a better understanding of the direct effects
184 of MLL3/4 loss on transcription changes during the naive to formative transition we evaluated the
185 monomethylation of H3K4, the primary known role of MLL3/4 and a surrogate marker of their
186 binding. We performed CUT&RUN(24) to profile H3K4me1 in naive and formative samples for

187 WT, MLL3KO, and DKO cells. Differential peak analysis using DiffBind(25) on WT naive and
188 formative samples with peaks called by SEACR(26) uncovered approximately ~13000 peaks
189 each that significantly gained (formative H3K4me1 peaks) or lost H3K4me1 (naive H3K4me1
190 peaks) during the normal pluripotent transition (FDR < 0.01, Log2 FC > 1). An additional ~94000
191 peaks did not significantly change between cell states (shared H3K4me1 peaks, FDR > 0.1, Log2
192 FC < 0.7)(Fig.2A,B). Naïve and formative enriched peaks were almost completely lost in the DKO
193 cells, but only slightly reduced in MLL3KO cells (Fig.2C,D). Reanalysis by Diffbind uncovered 0
194 total peaks that showed a significant change between the two states in the DKO background while
195 we found 29088 significantly changing peaks in MLL3KO compared to 27458 total peaks in WT
196 (Fig.2E, Fig.S2A). In contrast, shared H3K4me1 peaks (shown in gray) were not obviously
197 affected. To test whether the loss of dynamic H3K4me1 was a direct effect of MLL3/4 enzymatic
198 activity, we performed and evaluated RNA-seq and H3K4me1 CUT&RUN data using ESCs with
199 point mutations in the SET domains of both MLL3 and 4 that render them catalytically dead (dCD
200 cells)(12). Consistent with previous reports(12,27), we found a comparable reduction in global
201 H3K4me1 in westerns of acid-extracted histones between DKO and dCDs when compared to WT
202 while seeing few transcriptional differences between dCD and WT cells in the naive and formative
203 states (Fig.S2B-E). Moreover, H3K4me1 CUT&RUN on dCD cells largely recapitulated the
204 H3K4me1 defects observed in DKOs (Fig.2F). Diffbind uncovered only 272 total differential peaks
205 between naive and formative in the catalytic mutant background (Fig.2G).

206 The majority of H3K4me1 sites, whether shared, naive-enriched, or formative-enriched
207 were located at distal elements (intergenic or intronic) and not promoters (Fig.2H). To quantify
208 dependency of MLL3/4 H3K4me1 at the various sites, we set cutoffs for dependent (Log2 FC
209 naive DKO/WT < -1 or Log2 FC formative DKO/WT < -1) and independent H3K4me1 (Log2 FC
210 naive DKO/WT > 0.7 & Log2 FC form DKO/WT < 0.7). Of the distal H3K4me1 sites, only 15% of
211 shared sites (8632 of 64494) showed dependency on MLL3/4 for H3K4me1, while all naive and
212 formative specific sites were dependent. Heatmap visualization validated these categories and

213 similar behavior at dynamic sites during the transition was observed with analysis of published
214 naive and formative H3K4me1 ChIP-seq datasets (Fig.2J and Fig.S2F)(23,28). Furthermore,
215 evaluation of published naive MLL3/4 ChIP-seq data showed enrichment of the proteins at
216 H3K4me1 sites whether naïve-enriched or shared (12). Together, these data show that MLL3/4
217 are essential for a subset of H3K4me1 peaks including all dynamic sites and a small fraction of
218 shared sites.

219 Next, we asked about the impact of MLL3/4 loss on the transcription of genes closest to
220 the sites representing the different MLL3/4 dependent categories (shared, naive, and formative
221 H3K4me1 sites) (Fig. 2K). Genes near MLL3/4 dependent shared H3K4me1 sites showed a clear
222 upregulation in expression in the transition from naive to formative in WT cells, suggesting these
223 enhancers may be primed in the naive state and became activated in the formative state. The
224 loss of MLL3/4 had no impact on the upregulation of these genes with the transition to the
225 formative state, although there was a slight reduction in expression in the naïve state. This later
226 finding is consistent with previous findings showing downregulation of genes near MLL3/4 peaks
227 in steady state naïve DKO ESCs(12), something we also saw with our expression data (Fig.S2G).
228 In contrast, genes nearby MLL3/4 dependent naive-specific H3K4me1 sites were not impacted in
229 the naïve DKO ESCs. Most surprising, while genes nearby MLL3/4 dependent formative-specific
230 H3K4me1 showed a strong gain in expression during the naive to formative transition, this gain
231 was not diminished by the loss of MLL3/4. Together, these results show that while MLL3/4 is
232 required for all dynamic H3K4me1 and a subset of shared H3K4me1 between the naive and
233 formative state, this requirement has minimal impact on gain of nearby gene transcription with the
234 transition to the formative state.

235

236 **MLL3/4 dependent and independent distal H3K27ac deposition**

237 MLL3/4 is thought to be important for the recruitment of histone acetyltransferase
238 P300/CBP and thus histone acetylation including at H3K27 at enhancers. To test this general

239 model at the genomic level, we conducted CUT&TAG(29) for H3K27ac in WT and DKO cells in
240 the naive and formative states. We performed Diffbind at all peaks called by SEACR, intersected
241 these with ATAC-seq peaks from naive and formative WT cells, and subset for distal sites to
242 further enrich for active enhancers. A total of 12802 H3K27ac peaks were significantly down
243 (naive H3K27ac peaks) and 8924 peaks were significantly up (formative H3K27ac peaks) during
244 the transition (Fig.3A,B). Another 12205 peaks did not show significant change in H3K27ac
245 between the two cell states (shared H3K27ac peaks). Loss of MLL3/4 resulted in a dramatic
246 decrease in the dynamic changes in H3K27ac normally seen in the WT cells (Fig. 3C). However,
247 unlike with H3K4me1, many sites still showed some change including sites that appeared to be
248 unique to each pluripotent state. Diffbind using DKO samples uncovered 4295 naive-enriched,
249 1935 formative-enriched, and 23454 shared peaks, a reduction compared to WT, but still
250 suggestive of cell state dependent dynamic acetylation (Fig. 3D). In both WT or DKO cells, the
251 majority of H3K27ac was located at distal sites, including both shared and naive/formative-
252 enriched peaks. Most of the cell-state specific H3K27ac sites were dependent on MLL3/4 (Log2
253 FC DKO/WT < -1 or Log2 FC DKO/WT > -0.7 respectively) with MLL3/4 dependent sites
254 accounting for 22% of shared (2661 of 12205), 57% at naive-enriched (5717 of 9963), and 62%
255 at formative-enriched sites (4495 of 7205) (Fig.3F). Heatmap visualization confirmed these
256 findings (Fig. 3G,H, Fig. S3A). Interestingly though, metagene analysis identified a low-level
257 persistent signal and premature H3K27ac signal at MLL3/4 independent naive and formative
258 distal sites respectively. Analysis of published ChIP-seq for H3K27ac and the acetyltransferase
259 P300 in naive and formative cells, showed commensurate binding at the sites with correlated
260 increases and decreases in signal with H3K27ac signal in WT cells (Fig.S3B)(23,28). These data
261 uncover both MLL3/4 dependent and independent mechanisms of P300 acetylation of H3K27 at
262 cell-type specific enhancer sites.

263 Given that H3K27ac is thought to be a marker of active enhancers, we next asked whether
264 sites with MLL3/4 dependent H3K27ac had impacted transcription at neighboring genes upon

265 loss of MLL3/4. We identified the nearest neighbor transcription start site (TSS) to all distal sites
266 where H3K27ac was dependent on MLL3/4 and separated these sites by whether they were
267 normally shared, naive-enriched, or formative-enriched. We then evaluated expression of genes
268 associated with the TSSs (Fig. 3I). Shared peaks were associated with genes that normally
269 increased in expression during the transition; this increase was not impacted by MLL3/4 loss.
270 Naive-enriched peaks were associated with genes that show a slight reduction during the
271 transition and that reduction occurred prematurely in the MLL3/4 DKO cells. Strikingly, there was
272 a strong upregulation in gene expression nearby sites that normally show MLL3/4 dependent
273 increases in H3K27ac, but this upregulation was not impacted by MLL3/4 loss. These data show
274 that the upregulation of gene expression associated with a gain of H3K27ac at nearby enhancer
275 sites is largely independent of that acetylation, while maintenance of gene expression associated
276 with naive enriched H3K27ac sites is at least in part dependent on maintenance of that
277 acetylation.
278
279 **Enhancer Activation Can Occur Independently of MLL3/4**
280 The data above suggested similar trends in terms of the effects of MLL3/4 on H3K4me1,
281 H3K27ac and nearby gene expression, consistent with a partial dependency on MLL3/4
282 recruitment for H3K27ac deposition. To investigate this dependency, we focused on all distal sites
283 that show changes in H3K27ac in the context of their H3K4me1 status. We first separated all
284 chromatin accessible H3K4me1 sites (ATAC+) into those that lose, gain, or have shared
285 methylation during the naive to formative transition (Fig.S4A). Reanalyzing nearest neighbor
286 transcription with these sites did not appreciably change our results obtained with all H3K4me1
287 sites (Fig.S4B vs. Fig.2K). Next, we subdivided the H3K4me1 categories into ones that lose, gain,
288 maintain, or never have coincident H3K27ac (Fig.4A). In general, most H3K4me1 sites never
289 showed H3K27ac. However, among the 5960 sites that lose H3K4me1 during the transition, 1370
290 also lose H3K27ac, and 913 maintain acetylation. Among 5583 sites that gain H3K4me1, 1172

291 sites also gain H3K27ac. We identified few sites (<150) for the remaining categories of H3K27ac
292 at sites that either gain or lose H3K4me1. Among 37589 sites maintaining H3K4me1, 2264 sites
293 gain acetylation, 3498 sites lose acetylation, 4353 sites show maintained acetylation. This
294 suggests, although sites with dynamic H3K4me1 tended to have dynamic H3K27ac in the same
295 direction (gain or loss), most H3K27ac dynamics during the naive to formative transition typically
296 occur at sites with preexisting H3K4me1.

297 To understand how these subgroups were impacted by the loss of MLL3/4, we visualized
298 the data as heatmaps and filtered for MLL3/4 independent or dependent H3K27ac (Log2FC
299 DKO/WT > -0.7 or Log2FC DKO/WT < -1 respectively). Several important findings were
300 uncovered by these efforts. Among the sites that are enriched for both H3K4me1 and H3K27ac
301 (H3K4me1+/H3K27ac+) in the naive state, the loss of MLL3/4 resulted in a consistent loss of
302 monomethylation as seen across all naive enriched H3K4me1 sites (Fig. 4B). However, only 61%
303 of these sites lost H3K27ac (891 of 1452 sites). Among sites that gain H3K4me1 and H3K27ac
304 during the transition from naive to formative, again all H3K4me1 is dependent on MLL3/4, while
305 77% also lost H3K27ac (901 out of 1172 sites)(Fig. 4C). To confirm that the naive and formative
306 H3K4me1+/H3K27ac+ sites were direct MLL3/4 targets, we analyzed published naive ChIP-seq
307 of MLL3/4 and generated CUT&RUN for MLL4 in the formative state (Fig. 4B,C). Naive MLL3/4
308 ChIP-seq signal was enriched at naive sites and MLL4 CUT&RUN was enriched specifically at
309 formative sites. Together these analyses show H3K27ac maintenance and de novo deposition
310 can occur at enhancer sites independent of MLL3/4 including at sites requiring MLL3/4 for
311 H3K4me1 deposition.

312 Next, we analyzed the sites that normally retain monomethylation during the transition
313 (H3K4me1 shared sites), which included a mix of sites that did or did not require MLL3/4 for
314 H3K4me1 (Fig. 2J). Among the MLL3/4 independent H3K4me1 sites, 2039 sites normally gained
315 H3K27ac, 2357 sites lost H3K27ac, and 3611 sites had H3K27ac that did not change. An
316 additional 18135 sites never showed H3K27ac. For sites that gained H3K27ac, 48% required

317 MLL3/4 to do so even though the same enzymes were not required for the monomethylation at
318 those sites (Fig. 4D). The remaining 52% showed strong H3K27ac in formative cells independent
319 of MLL3/4 and most showed premature H3K27ac deposition in naive DKO cells. MLL3/4 signal
320 was low at all these sites consistent with the unaffected H3K4me1 which suggested the loss of
321 acetylation at these sites was due to indirect effects. Similar to the formative H3K27ac sites, naive
322 H3K27ac at shared H3K4me1 sites also demonstrated mixed dependency on MLL3/4 with only
323 844 of the sites losing H3K27ac (Fig.S4B). Among the remaining 1320 sites, which retained high
324 levels of naive H3K27ac, the reduction in acetylation upon differentiation was incomplete in DKOs.
325 In contrast to the formative sites, MLL3/4 signal was strong at the naive dependent sites. In
326 MLL3/4 dependent shared H3K4me1 sites (2113 sites) few sites gained or had shared H3K27ac
327 (225 and 742, respectively). Despite losing H3K4me1 these sites showed mixed dependency of
328 H3K27ac on MLL3/4. Notably, sites with naive enriched H3K27ac comprised the largest group
329 (1141 sites) and most of these sites (973 of 1141 sites, 85%) were dependent on MLL3/4 for
330 H3K27ac. Finally, H3K4me1 independent shared sites showed no appreciable MLL3/4 binding
331 despite a loss of H3K27ac in DKO cells at 992 of 3297 sites, implying indirect effects (Fig.S4D).
332 Taken together, these data show that H3K27ac deposition at cell-type specific sites is not
333 necessarily coupled with binding or activity of MLL3/4. Indeed, we identified almost as many
334 exceptions as examples for the canonical model coupling MLL3/4 and H3K27ac deposition at
335 enhancers.

336 Next, we asked how the loss of MLL3/4 impacted H3K4me1 and H3K27ac at well-known
337 functionally validated enhancers regulating genes important for pluripotent cell identity including
338 Sox2, Nanog, Fgf5, and Oct6(28,30–34). Sox2 and Nanog expression normally decrease, while
339 Fgf5 and Oct6 expression increase during the naive to formative transition. In the MLL3/4 DKOs,
340 Sox2 expression failed to decrease during the transition, while the other genes showed normal
341 patterns of expression (Fig.1G, Fig.S4E). However, the impact of MLL3/4 loss on H3K4me1 and
342 H3K27ac at associated enhancers for each of these genes varied (Fig. 4E). The Sox2 enhancer

343 cluster showed stable H3K4me1 and naive specific H3K27ac, both of which were dependent on
344 MLL3/4. Nanog has at least two known enhancers, all normally showing stable H3K4me1 and a
345 loss of H3K27ac during the transition; none of which were impacted by MLL3/4 loss. Fgf5 has 4
346 known enhancers, all normally showing increases in H3K4me1 and H3K27ac during the
347 transition. In the DKO cells, all four enhancers showed reduced H3K4me1, with two showing
348 normal upregulation of H3K27ac and two showing reduced H3K27ac in the formative state. Oct6
349 has one known enhancer, which normally shows a gain in H3K4me1 and H3K27ac during the
350 transition; H3K4me1 was reduced, but H3K27ac was unchanged in DKO cells. These examples
351 of well-known enhancers demonstrate the variable nature of dependency on MLL3/4 for
352 H3K4me1 and H3K27ac, including a surprising dispensability for de novo activation of enhancers
353 for Fgf5 and Oct6 expression.

354 We hypothesized that dependency of H3K4me1/H3K27ac on MLL3/4 might vary based
355 on the TFs bound at each enhancer. Therefore, we performed differential motif analysis using
356 Gimmemotifs(35) to determine TF binding motifs enriched at enhancer sites dependent on
357 MLL3/4 for H3K4me1 and/or H3K27ac (Fig.S5A). This analysis identified strong enrichment for
358 the binding motifs for TCF7L2, GRHL, and ZIC1 at sites requiring MLL3/4 for the gain of both
359 H3K4me1 and H3K27ac during the naive to formative transition. Expression of each of these TFs
360 is maintained or increased during the transition, and GRHL2 is essential for the gain of H3K4me1
361 and H3K27ac at hundreds of formative specific enhancers(23). At MLL3/4 dependent naive
362 peaks, we identified a strong enrichment for multiple TF binding motifs including RXRA, ZBTB20,
363 ZNF257, MYC, ESR2, and NR6A1, all of which are either expressed or have closely related TFs
364 that are expressed in naive cells (Fig.S5B). Interestingly, the motifs for the critical TF regulators
365 of the naive and/or formative states such as Oct4, Sox2, Nanog, KLFs, Oct6, and Otx2, were
366 generally not uncovered. These analyses suggest that TFs differ in their dependency on MLL3/4
367 to maintain or establish H3K27ac.
368

369 **Distal H3K4me1 and H3K27ac are not functionally coupled with formative transcriptional**
370 **activation**

371 We next asked whether focusing on sites of H3K27ac change in the context of dynamic
372 H3K4me1 peaks might enrich for sites near genes whose expression was impacted by MLL3/4.
373 Similar to analysis of H3K27ac alone, genes nearby naive H3K4me1+/H3K27ac+ sites with
374 MLL3/4 dependent H3K27ac showed reduced expression in naive cells (Fig.4B,Fig.S6A) while
375 genes nearby MLL3/4 independent H3K27ac were unchanged, supporting a role for acetylation
376 in maintaining expression at these sites. In contrast, in the formative state there was little impact
377 on transcription of genes nearby either MLL3/4 independent or dependent H3K27ac
378 (Fig.4C,Fig.S6B). This suggested gain of both H3K4me1 and H3K27ac at formative sites was not
379 associated with formative transcriptional activation.

380 To expand on this surprising finding, we used ChromHMM as an unbiased approach to
381 look at all potential combinations of H3K4me1, H3K27ac in WT and DKO cells in the naive and
382 formative states. We modeled 16 chromatin states to segment all possible combinations of these
383 modifications, and subset for distal peaks overlapping WT ATAC peaks separately in naive and
384 formative cells (Fig.5A,C). The sites were then mapped to nearby TSS, and the impact of MLL3/4
385 on expression of these genes was measured (Fig.5B,D). In the naive state, genes near sites that
386 lost H3K27ac showed highly significant downregulation whether or not they also lost H3K4me1
387 (States 1,15). These sites also showed strong MLL3/4 signal supporting a direct role for the
388 enzymes in coordinating both acetylation and transcription. In the formative state, only genes near
389 distal sites that lost both H3K4me1 and H3K27ac showed a significant reduction in gene
390 expression (State 2) (Fig. 5D). This state represented a relatively small number of enhancers
391 (1040 sites) linked to only 154 genes. Therefore, this unbiased approach at defining chromatin
392 states again showed that even though MLL3/4 loss dramatically alters the landscape of enhancer
393 activity states, the impact on gene expression changes that occur during the naive to formative
394 transition is relatively minor.

395 Genes can be regulated by multiple enhancers(36,37). Thus, we hypothesized that
396 MLL3/4 independent enhancers may compensate for loss of MLL3/4 dependent enhancers
397 explaining the underwhelming impact on expression. To test this hypothesis, we compared the
398 change in expression (naive or formative DKO/WT) relative to the number of enhancers impacted
399 by loss of MLL3/4 for each gene and their associated enhancers. In naive cells, this approach
400 showed a correlation with the change in nearby gene expression (Fig.5E). In formative cells, only
401 a very small effect was seen and only for genes with the greatest numbers of impacted enhancers.
402 We obtained similar results by analyzing the proportion of enhancers that were lost per gene in
403 each pluripotent state (Fig.5F). The contrast between naive and formative was even more striking
404 when only considering sites within 10kb of the TSS (Fig.S6C,D). Therefore, compensation by
405 non-MLL3/4 dependent enhancers does not appear to be a major factor rescuing gene expression
406 in formative cells.
407
408 **Gene-centric analysis reveals a subset of distal loci associate with MLL3/4 dependent**
409 **formative genes**
410 Our data showed a requirement for MLL3/4 for H3K4 monomethylation at all sites that normally
411 gain or lose the mark during the naive to formative transition (Fig. 2D,J,Fig.6A). They also showed
412 a requirement for MLL3/4 for H3K27ac at 66% of sites that gain and 66% that lose acetylation
413 during the transition (Fig3G,H, Fig.6A, Fig.S7A). However, only 13% of formative genes that
414 normally increase in expression in the transition and only 39% naive genes that are normally
415 downregulated were impacted by loss of MLL3/4. Among the naive genes that were impacted,
416 there was a strong enrichment for cytokine pathways (Fig.S7B). Among the formative genes that
417 were impacted, there was an enrichment for development related pathways including
418 regionalization, migration, and morphogenesis (Fig.6B). This later finding suggested a role for
419 initiating transcription of genes involved in a later stage of development, specifically gastrulation.

420 Our previous analyses utilized an enhancer-centric approach to identify changes in nearby
421 genes. We next asked if impacted genes could be connected to nearby enhancers that showed
422 loss of H3K27ac. To take this gene-centric approach, we categorized all genes that were normally
423 upregulated during the transition to the formative state into those that were not impacted
424 (independent) vs. impacted (dependent) by MLL3/4 loss. Among formative genes, this resulted in
425 820 independent formative genes and 131 dependent formative genes which could be linked to
426 2693 and 402 nearby H3K4me1/H3K27ac+ enhancers respectively. As a background control, we
427 linked 12506 expressed genes from our RNA-seq to an associated 37362 H3K4me1/H3K27ac+
428 peaks. Both the dependent and independent formative genes normally showed an increase in
429 H3K27ac at nearby enhancer sites during the naive to formative transition (Fig.6C). This increase
430 in acetylation was mostly lost in DKO cells, whether or not the nearby genes failed to increase
431 gene expression. Similarly, both dependent and independent naive genes (normally
432 downregulated during transition) showed a decrease in nearby enhancer H3K27ac during the
433 naive to transition. However, H3K27ac at enhancers nearby dependent naive genes was more
434 impacted than those nearby independent genes in the DKO cells (Fig.S7C). Therefore, the gene-
435 centric approach was consistent with enhancer-centric approach showing a more prevalent role
436 of MLL3/4 regulated chromatin changes on the maintenance of naïve gene expression than the
437 gain of formative gene expression.

438 To look deeper, we investigated the chromatin states nearby the small number of genes
439 that were dependent on MLL3/4 for upregulation of transcript expression during the naive to
440 formative transition by creating genome tracks (Fig.6D). These dependent genes included
441 important gastrulation genes such as *T*/*Brachyury*, *Fgf8*, and *Lefty1*, the latter of which is a
442 previously identified target of MLL3/4 (Wang et al. 2016). All three of these genes had one or
443 more nearby enhancers that showed a loss of H3K4me1 and H3K27ac in MLL3/4 DKO formative
444 cells. Thus, the expression of these key regulators relies on enhancers which require MLL3/4 for
445 proper spatiotemporal regulation. However, they represent a very small subset of the thousands

446 of enhancers that lose active marks upon the loss of MLL3/4. All together, these data show that
447 from both a gene-centric or enhancer-centric perspective, H3K4me1 and/or H3K27ac at distal
448 sites are largely uncoupled from target gene activation during the naïve to formative transition
449 with the loss of MLL3/4.
450

451 **Discussion**

452 Enhancers are thought to play a central role in gene regulation during cell fate transitions.
453 Yet, the molecular epistasis of enhancer activation is incompletely understood. The prevailing
454 model proposes that H3K4me1 deposited by MLL3/4 precedes H3K27ac deposition by P300/CBP
455 and these events stimulate transcription. By combining state-of-the-art CUT&RUN/CUT&TAG
456 technologies and knockouts of both MLL3 and MLL4, we evaluated this model during the well-
457 defined transition from naive to formative pluripotency, where hundreds of genes are both
458 silenced and activated. Surprisingly, the majority of these gene expression changes including de
459 novo expression of many genes were unaffected by MLL3/4 loss. In contrast, loss of MLL3/4 led
460 to large effects on the active histone modification signatures at thousands of enhancers. These
461 results suggest that MLL3/4 and its orchestration of post-translational histone marks, play a
462 relatively minor role in gene regulation during early ESC differentiation.

463 MLL3/4 are thought to be the major H3K4 monomethyltransferases in mammals(1,2).
464 However, we find that the majority of H3K4me1 is unaffected by MLL3/4 loss in both the naive
465 and formative states. Consistent with our findings, similar results have been previously described
466 in naive ESCs(12). By investigating the transition to the formative state, we were able to make
467 several novel discoveries though. We discovered that all distal sites that normally show dynamic
468 changes in H3K4me1, either gained or lost, were fully dependent on MLL3/4 for monomethylation.
469 In contrast, only a small fraction of the sites that normally do not change during the transition
470 showed any requirement for enzymes in maintaining monomethylation levels. Therefore, there
471 appear to be at least two distinct types of cis-regulatory elements with regard to MLL3/4

472 dependency. This raises important follow-up questions such as what enzymes are responsible
473 for the predominant group of H3K4me1 sites that do not require MLL3/4 for maintenance? MLL2
474 has been suggested as an alternative monomethyltransferase in naive ESCs, but only at a small
475 number of sites that co-exist with H3K27m3, a repressive mark(12,38). MLL1 represents another
476 potential candidate. Interestingly, MLL1 has been shown to inhibit the reprogramming of primed
477 pluripotent (EpiSCs) to naive ESCs(39). However, its role in regulating H3K4me1 in naive cells
478 during differentiation has not been evaluated. Other candidates include SMYD1 and SMYD2
479 which are SET domain containing proteins that have previously been shown to be required for
480 H3K4me1 and or H3K4m2 at de novo enhancers during macrophage activation(40). Additionally,
481 whether there are distinct roles for the MLL3/4 dependent vs. independent distal sites remains
482 unclear. Given most enhancer activation in the absence of MLL3/4 occurs at these independent
483 H3K4me1 sites they may represent a collection of primed sites distinct from MLL3/4. Surprisingly
484 though, previous studies have shown that at least in the case of MLL3/4, the proteins, but not
485 their enzymatic activity are important in CBP/P300 recruitment and H3K27ac(12). Therefore, it
486 will be important to understand the link between the factors resulting in methylation of these
487 MLL3/4 independent sites and those that deposit H3K27ac at these sites.

488 MLL3/4 are thought to be important recruiters of CBP/P300, whose catalytic byproduct
489 H3K27ac is a key marker of enhancer activation. In our data, we identify that more than half of
490 H3K27ac distal sites lose acetylation with MLL3/4 loss, with a greater proportion at dynamic sites
491 than unchanging sites. Many sites that lose H3K27ac do so independent of H3K4me1 loss.
492 Whether this represents an indirect role or a direct role for MLL3/4 in the recruitment of CBP/P300
493 at sites that also retain H3K4me1 is unclear. However, the presence of MLL3/4 binding at many
494 of these sites supports a direct role. In addition to the MLL3/4 dependent sites, there is also a
495 large number of MLL3/4 independent H3K27ac sites, including many occurring at regions that do
496 require MLL3/4 for H3K4me1. This indicates multiple mechanisms for CBP/P300 recruitment to
497 enhancers that are differentially used depending on the specific enhancer and chromatin state.

498 How different enhancers recruit CBP/P300 using distinct mechanisms is unclear but is likely in
499 part due to the underlying transcription factors binding to those enhancers. Enhancer
500 subcategories consisting of varying types of TFs have recently been shown to require specific
501 chromatin complexes for enhancer activity(41). Consistent with the role for different TFs
502 underlying these requirements, TF motif enrichment analysis uncovered several TFs whose
503 motifs were highly enriched at each distinct category of MLL3/4 dependent and independent
504 enhancers. CBP/P300 could either be directly recruited by these TFs and/or be recruited through
505 other factors compensating for loss of MLL3/4. Future studies are necessary to fully elucidate the
506 requirements for recruitment of specific epigenetic complexes by distinct TFs.

507 Ultimately, it is thought that enhancer activation as defined by the gain of H3K27ac drives
508 increased expression of its target gene. Surprisingly though, while we observe major impacts on
509 the chromatin state of sites with dynamic H3K4me1 and H3K27ac during the naive to formative
510 transition, there are relatively modest changes in the gain and loss of gene expression. The
511 activation of the formative transcriptional program is largely unaffected, with only a few exceptions
512 such as Lefty1 and Fgf8. The key markers of the formative state including Otx2, Oct6, and Fgf5
513 are all normally upregulated. This unexpected finding raised the question whether there was any
514 specific MLL3/4 dependent enhancer phenotype in formative cells that could be linked to a failure
515 to activate gene expression. Analysis of genes either nearby MLL3/4 dependent formative
516 H3K4me1 sites or MLL3/4 dependent formative H3K27ac sites during the transition did not
517 identify a reduction in gene activation. Most surprising, interrogation of genes nearby enhancers
518 that fail to gain both H3K4me1 and H3K27ac also showed normal increases in gene expression.
519 Only when considering the small number of sites where both H3K4me1 and H3K27ac were
520 depleted, independent of their dynamics, was there a small but significant effect for a handful of
521 genes. Consistent with our data, recent work mutating H3 and/or H3.3 Lysine 27 sites to Arginine,
522 which cannot be acetylated did not impact either naive gene expression or the activation of the
523 formative program in differentiating ESCs(14,15). However, these findings are interpreted as a

524 lack of a requirement for H3K27ac rather than an inability to recruit enhancer CBP/P300 at large
525 where we also expect a loss in other active enhancer histone acetylations deposited by CBP/P300
526 such as H2B acetylation(42). Either way, it is unclear whether many of the sites dependent on
527 MLL3/4 for active histone marks are dispensable for gene expression changes and or these sites
528 are functional enhancers without an active histone signature. Of note, loss of MLL3/4 can also
529 impact the recruitment of chromatin remodelers (i.e. the BAF complex) and enhancer-promoter
530 looping, making the minimal impact MLL3/4 loss has on gene transcription during this transition
531 even more surprising (9,43,44).

532 MLL4 KO embryos die shortly following gastrulation but show defects starting in early
533 gastrulation(6,21). In contrast, MLL3 KO embryos die perinatally due to developmental defects in
534 the lung. MLL3/4 double KO embryos phenocopy the MLL4 KO consistent with the MLL4 being
535 able to compensate for MLL3 loss in early development. The direct molecular basis for these
536 phenotypes is poorly understood. However, it is important to note that even though MLL3/4 loss
537 had little impact on gene activation during the ESC transition from naive to formative pluripotency,
538 many genes were dysregulated in steady state naive ESCs and remained so in the formative
539 state. In the case of persisting naive genes (Fig. 1K), this is consistent with previous reports that
540 MLL4 KOs fail to repress naive genes in LIF cultured cells after withdrawal of 2i from naive (LIF+2i)
541 culture conditions(45). Some misregulated genes are likely direct targets given the
542 downregulation of genes nearby MLL3/4 dependent enhancers upon loss of the enzymes. In
543 addition, there are likely indirect effects potentially including additional unknown non-transcription
544 related functions for MLL3/4. It is also important to note that other studies have shown important
545 roles for MLL3/4 in the upregulation of gastrulation markers in the setting of embryoid body
546 differentiation, a model for a later stage in early development than studied here(22). Even later in
547 development, MLL3/4 is essential for developmental programs in a variety of tissues including
548 cardiac, neural crest, muscle, and adipose specific programs(6,7,46,47). The differences between
549 these studies and ours suggest a cell-context specific function of MLL3/4. A potential reason is

550 the compendium of TFs that require or do not require MLL3/4 is different for each transition. For
551 instance, PPARγ and CEBPα are major adipogenic factors that require MLL3/4 to function, while
552 at least some of the pluripotency factors and the drivers of formative state are MLL3/4
553 independent based on our data. One possibility may be that transcriptional roles of MLL3/4
554 proteins are limited to determinants of differentiation pathways post-pluripotency and is most
555 important in specifying somatic cell fates. Interestingly, a recent preprint suggests that MLL3/4
556 enzymatic activity is also important for a subset of cell fates such as extraembryonic endoderm
557 and trophectoderm but not germ layer formation(48). Taken together, these data suggest
558 fundamental and currently underappreciated differences in transcriptional control between
559 pluripotent and somatic cell types.
560

561 **Conclusions**

562 In sum, our data demonstrates that cell-type specific enhancer activation is not uniquely
563 performed by MLL3 and MLL4. Instead, gene regulation appears to be partitioned during the naive
564 to formative transition. MLL3/4 is required for a small subset of genes critical for gastrulation and
565 later developmental processes, but formative transcriptional activation is mediated by other
566 mechanisms including MLL3/4 independent enhancers (Fig.6E). By identifying rules and
567 exceptions to the current model of enhancer activation we highlight current gaps in knowledge
568 concerning the molecular steps regulating these processes. Further investigation into the
569 molecular interplay and steps governing enhancer function will advance our understanding of
570 gene regulation in health and gene dysregulation in disease.
571

572 **Methods**

573 **Cell Culture**

574 Mouse ESCs were cultured in Knockout DMEM (Thermo Fisher, CAT#10829018) supplemented
575 with 15% FBS, L-Glutamine, Penicillin/Streptamycin, NEAA, LIF (1000U/mL), and 2i (1uM MEK
576 inhibitor PD0325901 and 3uM GSK3 inhibitor CHIR99021). Wildtype and dCD cells were R1
577 ESCs from a 129 strain background. MLL3$^{-/-}$; MLL4$^{fl/fl}$ ESCs were a mixed background of
578 C57BL/6J and 129 strains. Formative cells were generated by removal of LIF and 2i. Briefly, 5e4
579 ESCs were plated per well of a 6 well plate on day −1 in LIF+2i media. To initiate differentiation,
580 LIF and 2i were removed 24 hours after seeding (Day 0). Formative cells were collected on day
581 3 of differentiation, 63 hours after removal of LIF and 2i. To overcome proliferation defects 7.5e4
582 MLL3/4 DKO cells were plated per well of a 6 well. Naive cells were passaged and staged
583 appropriately for simultaneous harvest. The expected amino acid substitutions in MLL3/4 dCD
584 cells were validated using Sanger sequencing on amplicons generated from genomic PCR. To
585 generate DKOs from MLL3$^{-/-}$; MLL4$^{fl/fl}$ ESCs we transfected with Rosa26-CRE-ERT2 plasmid
586 linearized by EcoRI, selected with puromycin, and genotyped using genomic PCR. We then
587 added tamoxifen for 3 days and picked clones to validate knockout of MLL4 floxed allele. Primers
588 for genotyping listed in the supplemental table. Lines consistently tested negative for mycoplasm.
589

590 **Crystal Violet Assay**

591 To assess proliferation cells were plated in 24 well plates at approximately 12.5e4 cells per well.
592 After 24, 48, and 72 hours, cells were then washed with PBS and 200ul Crystal Violet (0.2%
593 Crystal Violet, 2% ethanol in dH2O) was added to the plate for 10 minutes at room temperature.
594 Cells were then washed twice by gently submerging plate in tap water before 600ul 1% SDS was
595 added to the well to solubilize the stain. Plate was placed on a shaker until color was uniform in
596 the well and 200ul was transferred to a 96 well plate for reading absorbance at 570nm with a plate
597 reader (SpectraMax M5). Absorbance for each time point was normalized to a blank well.
598

599 **qPCR and analysis**

600 To perform qPCR we first extracted RNA by adding Trizol directly to plates. After adding
601 chloroform, an isopropanol precipitation with GlycoBlue was performed followed by ethanol
602 washes. RNA pellets were resuspended in RNAse-free water and quantified using a NanoDrop.
603 200ug of RNA was used for cDNA synthesis using Maxima First Strand Synthesis Kit (Thermo
604 Fisher, CAT#K1672) with half reactions according to the manufacturer's protocol. qPCR on cDNA
605 was performed using SYBRgreen master mix (Applied Biosystems, CAT#A25742) using 6ul final
606 volume on a QuantStudio5 qPCR machine (Applied Biosystems). qPCR primers for targets are
607 listed in the supplemental table. Target Ct values were normalized to GAPDH internally for each
608 sample and then set relative to the naive WT negative control.

609

610 **Protein extraction and Westerns**

611 To harvest protein for western assays, cells were trypsinized, washed once with ice cold PBS
612 before adding RIPA buffer with protease inhibitors (Sigma CAT#P8340). After 15 minutes on ice,
613 lysed cells were centrifuged at 16000g for 10 minutes and the supernatants representing whole
614 cell fractions were collected and snap frozen in liquid nitrogen. Protein quantification was
615 conducted using a Micro BCA protein assay kit (Thermo CAT#23235). 40ug of protein was loaded
616 per well of SDS PAGE gels for western. For histone purification we utilized the acid extraction
617 protocol exactly as described in Shechter et al. 2007 using 5e6 cells as input(49).

618

619 Westerns were typically conducted using a Bio-Rad system with Tris-Glycine gels purchased from
620 Bio-Rad and transferred to methanol activated PVDF membranes. Membranes were blocked, and
621 stained with primaries and secondaries using Li-Cor Odyssey Blocking Buffer (Li-Cor, CAT#927-
622 60001) mixed 1:1 with TBS. Primary antibodies for western were Rabbit anti-MLL3 (provided by
623 Kai Ge), Rabbit anti-MLL4 (provided by Kai Ge), Mouse anti-OCT4 (BD Biosciences,

624 CAT#611202), Rabbit anti-NANOG (Cell Signaling, CAT#61419), Rabbit anti-H3K4me1 (Abcam,
625 CAT#ab8895), and Mouse anti-H3 (Cell Signaling, CAT#3738). Westerns for histones were
626 conducted similarly to other targets except gels were transferred to PVDF membranes using
627 CAPS buffer (500mM CAPS, adjusted to pH 10.5 with NaOH).
628
629 High molecular weight westerns for MLL3 and MLL4 were conducted using a NuPage SDS-PAGE
630 and Transfer system with an XCell electrophoresis unit according to manufacturer's protocols with
631 several modifications (Thermo Fisher, CAT#E10002). In brief, 40ug of whole cell protein lysate
632 was incubated with LDS loading buffer (Thermo Fisher, CAT#NP0007, added BME to 1% final
633 concentration) and incubated at 70C for 10min. Samples were loaded into 3-8% Tris-acetate gels
634 (Thermo Fisher, CAT#EA03752BOX) and ran in sample running buffer (Thermo Fisher,
635 CAT#LA0041) supplemented with NuPage Antioxidant (Thermo Fisher, CAT#NP0005). Samples
636 were run at 80V for 30min and then 120V for 120min. Samples were then transferred to PVDF
637 membranes using NuPage Transfer Buffer (Thermo Fisher, CAT#NP0006, with added 10%
638 Methanol, NuPage Antioxidant, 0.01% SDS) at 30V in a cold room for either 2 hours or overnight.
639
640 **RNA sequencing**
641 Total RNA was extracted and purified from cells using Trizol followed by ethanol precipitation.
642 RNA-seq libraries were generated using the QuantSeq 3' mRNA-Seq Library Prep Kit FWD for
643 Illumina (Lexogen, CAT#A01172) according to their protocol using 200ng of total RNA for input.
644 We utilized the PCR Add-on kit for Illumina (Lexogen, CAT#M02096) to determine an appropriate
645 number of PCR cycles to amplify libraries. Amplified libraries were quantified using Agilent
646 Tapestation 4200. Libraries were pooled and sequenced using a HiSeq 4000 to obtain single end
647 50bp reads. At least 10 million mapped reads or more per sample were obtained.
648

649 **RNA sequencing processing and analysis**

650 Adapters for sample FastQ files were trimmed using cutadapt v2.5 followed by alignment to the
651 mm10 genome using STAR align 2.7.2b. Featurecounts from Subread v1.6.4 was used to
652 generate a counts matrix of reads per gene. FastQC v0.11.8 and MultiQC were used to validate
653 quality sequencing and mapping.

654

655 The gene count matrix was converted to CPM and filtered for genes greater than 1 cpm in at least
656 two total samples. Samples were normalized using TMM. Next, Log2 CPM averages were
657 calculated for replicates of each sample for scatterplot visualization and nearest neighbor TSS
658 analysis. All transcriptomics analyses were conducted using CPM values from TMM normalization
659 of all samples except for nearest neighbor analysis where WT and DKO were TMM normalized
660 together. To conduct differential gene expression we performed DESeq2 v1.34.0 analysis using
661 the raw gene count matrix as input for each desired comparison. Gene ontology was performed
662 using ClusterProfiler 4.2.2. Custom R code for other downstream transcriptomics analyses and
663 visualization provided on Github.

664

665 **CUT&RUN sequencing and processing**

666 CUT&RUN was conducted using the protocol from Skene et al. 2018 with the following
667 modifications: Freshly trypsinized cells were bound to activated Concanavalin A beads (Bang
668 Laboratories, #BP531) at a ratio of 2e5 cells/10ul beads in CR Wash buffer (20mM HEPES,
669 150mM NaCl, 0.5mM Spermidine with protease inhibitors added) at room temp. An input of 2e5
670 cells were used per target. Bead-bound cells were then incubated rotating overnight at 4C in CR
671 Antibody buffer (CR Wash with 0.05% Digitonin, 2mM EDTA) containing primary antibody. We
672 used the following antibodies for CUT&RUN: 1:100 Rabbit anti-H3K4me1 (Abcam, CAT#ab8895),
673 Rabbit anti-MLL4 1:100 (provided by Kai Ge), and 1:100 Rabbit IgG isotype control (Abcam,

674 CAT#ab171870). After primary, we washed 3 times 5 minutes each with cold CR Dig-wash buffer
675 (CR Wash with 0.05% Digitonin) and incubated with pA-MNase (1:100 of 143ug/mL provided by
676 Steve Henikoff) for 1 hour rotating at 4C. After MNase binding, we washed 3 times 5 each with
677 cold CR Dig-wash buffer, and chilled cells down to 0C using a metal tube rack partially submerged
678 in an ice water slurry. MNase digestion was induced by adding CaCl2 at a final concentration of
679 2mM. After 30 minutes of digestion, the reaction was quenched using Stop Buffer containing
680 340mM NaCl, 20mM EDTA, 4mM EGTA, 0.05% Digitonin, 100ug/mL RNAse A, 50ug/mL
681 Glycogen, and approximately 2pg/mL Yeast spike-in DNA (provided by Steve Henikoff). The
682 digested fragments for each sample were then extracted using a phenol chloroform extraction.
683 Library preparation on samples was conducted using manufacturer's protocols for NEBNext Ultra
684 II Dna Library Prep Kit (New England BioLabs, CAT#E7645) and NEB Multiplex Dual Index oligos
685 (New England BioLabs, CAT#E7600, #E7780) with the following modifications. We input
686 approximately 10ng of sample for half reactions, we diluted the NEBNext Illumina adaptor 1:25,
687 we used the following PCR cycling conditions: 1 cycle of Initial Denaturation at 98C for 30
688 seconds, 12+ cycles of Denaturation at 98C for 10 seconds then Annealing/Extension at 65C for
689 10 seconds, and 1 final cycle of extension at 65C for 5 minutes. Following library preparation,
690 double size selection was performed using Ampure beads and quality and concentration of
691 libraries were determined by an Agilent 4200 Tapestation with High-Sensitivity D1000 reagents
692 before pooling for sequencing.
693
694 Fastq files for CUT&RUN samples were processed using Nextflow(50) and the nf-core CUT&RUN
695 pipeline v1.0.0 beta. In brief, adapters were trimmed using Trim Galore. Paired-end alignment
696 was performed using Bowtie2 and peaks were called using SEACR with a peak threshold of 0.05
697 using spike in calibration performed using the E.coli genome K12. However, all downstream
698 analysis was performed using CPM normalized samples. To reduce high background observed

699 with MLL4 antibody we subtracted MLL4 CUT&RUN in DKO cells away from WT to generate
700 formative MLL4 CUT&RUN bigWigs for use in heatmaps.
701
702 **CUT&TAG sequencing and processing**
703 CUT&TAG was conducted using the protocol from Kaya-Okur et al. 2020 with the following
704 modifications: freshly trypsinized cells were bound to Concanavalin A beads at a ratio of 2e5
705 cells/7ul beads in CR wash at room temp. We used 2e5 cells as input per sample. Bead-bound
706 cells were then incubated rotating overnight at 4C in CT Antibody buffer (CR Wash with 0.05%
707 Digitonin, 2mM EDTA, 1mg/mL BSA) containing primary antibody. We used the following primary
708 antibodies for CUT&TAG: 1:100 Rabbit anti-H3K27ac (Abcam, ab4729) and 1:100 Rabbit IgG
709 isotype control (Abcam, ab171870). After primary, samples were washed 3 times for 5 minutes
710 each using CR Dig-wash buffer and resuspended in 1:100 secondary antibody (Guinea pig anti-
711 rabbit, Antibodies Online #ABIN101961) in CR Dig-wash buffer at 4C for 1 hour rotating at 4C.
712 Samples were then incubated for 1 hour at 4C with 50ul of approximately 25nM homemade pA-
713 Tn5 in CT Dig300 wash buffer (20mM HEPES, 300mM NaCl, 0.01% Digitonin, 0.5mM Spermidine
714 with Roche cOmplete protease inhibitors added). Recombinant Tn5 was purified and loaded with
715 adapters as previously described (Kaya-Okur et al. 2019). After Tn5 incubation, samples were
716 washed 3 times for 5 minutes each with CT Dig300 wash buffer. Tagmentation was then initiated
717 for 1hr at 37C in a thermocycler by adding MgCl2 to 10mM final concentration in 50uL volume.
718 The tagmentation reaction was quenched immediately afterwards by adding 1.6ul of 0.5M EDTA,
719 1ul of 10mg/mL Proteinase K, and 1ul of 5% SDS. Samples were then incubated at 55C for 2
720 hours in a thermocycler to denature Tn5 and solubilize tagmented chromatin. After incubation,
721 samples were magnetized and the supernatant was transferred to new wells where SPRI bead
722 purification was performed using homemade beads to select all DNA fragment lengths larger than
723 100bp. Samples were eluted in 0.1X TE and approximately half of each sample was used for
724 library preparation using NEBNext HIFI Polymerase with custom indices synthesized by IDT. An

725 appropriate number of cycles for each target was chosen to prevent overamplification bias. After
726 amplification, libraries were purified with 1.2x homemade SPRI beads to select for fragments
727 >250bp and eluted in 0.1X TE. Quality and concentration of libraries were determined by an
728 Agilent 4200 Tapestation with D1000 reagents before pooling for sequencing.
729
730 CUT&TAG samples were processed similarly to CUT&RUN samples using the same Nextflow
731 pipeline.
732
733 **ATAC-seq sequencing and processing**
734 ATAC-seq libraries were generated using the Active Motif commercial kit following the provided
735 protocol with kit components (Active Motif, CAT#53150): Specifically, freshly trypsinized cells
736 were washed with ice cold PBS using 75K cells as input for each sample. Cells were lysed using
737 ice cold ATAC Lysis Buffer and we added 50ul of Tagmentation Master Mix to each sample. We
738 incubated tagmentation reactions at 37C for 30min in a thermomixer set to 800 rpm. We
739 immediately transferred samples to a new tube and performed DNA column purification. After
740 purification we amplified the libraries using the following PCR conditions: 1 cycle at 72C for 5
741 minutes, 1 cycle at 98C for 30 seconds, and finally 10 cycles of 98C for 10 seconds, 63C for 30
742 seconds, 72C for 1 minute. Amplified libraries were purified using 1.2x SPRI bead solution. Quality
743 and concentration of libraries were determined by an Agilent 4200 Tapestation with D1000
744 reagents before pooling for sequencing.
745
746 Fastq files for ATAC-seq samples were processed using Nextflow and the nf-core ATAC-seq
747 pipeline v1.2.1. In brief, adapters were trimmed using Trim Galore. Paired-end alignment was
748 performed using BWA and peaks were called using MACS2 in broad peak mode with a cutoff
749 value of 0.1. BAM files were converted to bigWigs. Using read counts from consensus peaks of

750 WT and DKO samples we performed TMM normalization using EdgeR to generate scale factors
751 used during bigWig generation of samples.
752
753 **ChIP-seq processing**
754 Fastq files for ChIP-seq from previously published studies were processed using Nextflow and
755 the nf-core ChIP-seq pipeline v1.2.1. In brief, adapters were trimmed using Trim Galore. Paired-
756 end alignment was performed using BWA, duplicates were removed, and peaks were called using
757 MACS2 in broad peak mode with a cutoff value of 0.1. We subtracted the MLL3/4 ChIP-seq in
758 DKO cells away from WT samples to generate naive MLL3/4 ChIP-seq bigWigs for use in
759 heatmaps.
760
761 **Peak analysis**
762 We performed differential signal enrichment analysis using Diffbind on SEACR peaks derived
763 from CUT&RUN or CUT&TAG samples. We then used the output from Diffbind, containing the
764 union of input SEACR peaks, for two packages: 1) We used Deeptools multiBigwigsummary to
765 count reads at each peak for the different targets/samples and 2) we used HOMER
766 annotatePeaks to identify nearest gene and annotate genomic features. These files were then
767 imported into R and joined together by peak coordinates to make a dataframe which we used to
768 filter for phenotypes and calculate fold changes. After genomic feature annotation by Homer we
769 classify intergenic and intronic peaks as “Distal”, TSS as “Promoter”, and everything else as
770 “Other” including: exon, transcriptional termination sites (TTS), UTR regions, and those without
771 assigned features. To calculate Log2CPM peak densities we took counts for each peak from
772 multibigwigsummary multiplied them by a constant (1e6), added 1, and then performed Log2(x).
773 Finally, we divided by the width in basepairs of the peak. Fold changes were conducted using
774 these Log2CPM peak density values. H3K27ac CUT&TAG data was subset by H3K27ac peaks

775 that overlapped at least 75% of an WT ATAC-seq peak using bedops. We also used bedops on
776 ATAC subset H3K27ac sites to identify H3K27ac peaks that had overlapping H3K4me1 peaks to
777 generate our H3K4me1/H3K27ac+ peak lists. ChromHMM segmentations were similarly filtered
778 for chromatin states that overlapped at least 75% of an WT ATAC-seq peak.
779
780 **Declarations**
781 **Ethics approval and consent to participate**
782 Not applicable
783
784 **Consent for publication**
785 Not applicable
786
787 **Availability of data and materials**
788 The data generated and their analysis in the current study are available in the GEO Database
789 using the identifier GSExxxxxx. Code generated for analysis will be hosted on Github or otherwise
790 available by request.
791
792 **Competing interests**
793 The authors declare that they have no competing interests.
794
795 **Funding**
796 This work was made possible by funding from the National Institute of General Medical Sciences
797 (grant no. R01GM122439). Additionally, R.M.B. was supported by the ARCS foundation and the
798 UCSF Discovery Fellowship.
799
800 **Author's Contributions**

Conceptualization: RMB, RB; Methodology: RMB; Validation: RMB; Formal Analysis: RMB; Investigation: RMB, KC; Resources: RB; Data Curation: RMB; Writing - original draft: RMB; Writing - review and editing: RMB, RB; Visualization: RMB; Supervision: RMB, RB; Project Administration: RB; Funding Acquisition: RB.

Acknowledgements

We thank Kai Ge for providing the MLL3KO;MLL4 floxed ESC line and antibodies for MLL3 and MLL4. We thank Joanna Wysocka for providing the MLL3/4 catalytic dead cell line. We thank Steve Henikoff for sending purified pA-MNase. We would like to thank the Nextflow community for helpful advice on processing of sequencing samples including Chris Cheshire and Charlotte West for providing access to a beta for nf-core CUT&RUN pipeline. We would like to thank Bryan Marsh, Deniz Gökbuget, and Brian Deveale for helpful discussions throughout the project including proofreading the original manuscript. We would like to thank Dan Lim, Yin Shen, and Elphege Nora for their thoughtful comments and constructive feedback on the manuscript.

References

1. Hu D, Gao X, Morgan MA, Herz HM, Smith ER, Shilatifard A. The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. Mol Cell Biol. 2013 Dec;33(23):4745–54.

2. Sze CC, Shilatifard A. MLL3/MLL4/COMPASS Family on Epigenetic Regulation of Enhancer Function and Cancer. Cold Spring Harb Perspect Med. 2016 Nov;6(11):a026427.

3. Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, et al. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. Nat Genet. 2010 Sep;42(9):790–3.

4. Lavery WJ, Barski A, Wiley S, Schorry EK, Lindsley AW. KMT2C/D COMPASS complex-associated diseases [KCDCOM-ADs]: an emerging class of congenital regulopathies. Clin Epigenetics. 2020 Jan 10;12(1):10.

5. Mendiratta G, Ke E, Aziz M, Liarakos D, Tong M, Stites EC. Cancer gene mutation frequencies for the U.S. population. Nat Commun. 2021 Oct 13;12(1):5961.

830 6. Lee JE, Wang C, Xu S, Cho YW, Wang L, Feng X, et al. H3K4 mono- and di-
831 methyltransferase MLL4 is required for enhancer activation during cell differentiation. eLife.
832 2013 Dec 24;2:e01503.

833 7. Lai B, Lee JE, Jang Y, Wang L, Peng W, Ge K. MLL3/MLL4 are required for CBP/p300
834 binding on enhancers and super-enhancer formation in brown adipogenesis. Nucleic Acids
835 Res. 2017 Jun 20;45(11):6388–403.

836 8. Wang SP, Tang Z, Chen CW, Shimada M, Koche RP, Wang LH, et al. A UTX-MLL4-p300
837 Transcriptional Regulatory Network Coordinately Shapes Active Enhancer Landscapes for
838 Eliciting Transcription. Mol Cell. 2017 Jul 20;67(2):308-321.e6.

839 9. Park YK, Lee JE, Yan Z, McKernan K, O'Haren T, Wang W, et al. Interplay of BAF and
840 MLL4 promotes cell type-specific enhancer activation. Nat Commun. 2021 Mar
841 12;12(1):1630.

842 10. Moore JE, Pratt HE, Purcaro MJ, Weng Z. A curated benchmark of enhancer-gene
843 interactions for evaluating enhancer-target gene prediction methods. Genome Biol. 2020
844 Jan 22;21(1):17.

845 11. Nasser J, Bergman DT, Fulco CP, Guckelberger P, Doughty BR, Patwardhan TA, et al.
846 Genome-wide enhancer maps link risk variants to disease genes. Nature. 2021
847 May;593(7858):238–43.

848 12. Dorighi KM, Swigut T, Henriques T, Bhanu NV, Scruggs BS, Nady N, et al. Mll3 and Mll4
849 facilitate enhancer RNA synthesis and transcription from promoters independently of H3K4
850 monomethylation. Mol Cell. 2017 May 18;66(4):568-576.e4.

851 13. Rickels R, Herz HM, Sze CC, Cao K, Morgan MA, Collings CK, et al. Histone H3K4
852 monomethylation catalyzed by Trr and mammalian COMPASS-like proteins at enhancers is
853 dispensable for development and viability. Nat Genet. 2017 Nov;49(11):1647–53.

854 14. Zhang T, Zhang Z, Dong Q, Xiong J, Zhu B. Histone H3K27 acetylation is dispensable for
855 enhancer activity in mouse embryonic stem cells. Genome Biol. 2020 Feb 21;21(1):45.

856 15. Sankar A, Mohammad F, Sundaramurthy AK, Wang H, Lerdrup M, Tatar T, et al. Histone
857 editing elucidates the functional roles of H3K27 methylation and acetylation in mammals.
858 Nat Genet. 2022 Jun;54(6):754–60.

859 16. Martire S, Nguyen J, Sundaresan A, Banaszynski LA. Differential contribution of p300 and
860 CBP to regulatory element acetylation in mESCs. BMC Mol Cell Biol. 2020 Jul 20;21(1):55.

861 17. Smith A. Formative pluripotency: the executive phase in a developmental continuum. Dev
862 Camb Engl. 2017 Feb 1;144(3):365–73.

863 18. Gökbuget D, Blelloch R. Epigenetic control of transcriptional regulation in pluripotency and
864 early differentiation. Dev Camb Engl. 2019 Sep 25;146(19):dev164772.

865 19. Yang P, Humphrey SJ, Cinghu S, Pathania R, Oldfield AJ, Kumar D, et al. Multi-omic
866 Profiling Reveals Dynamics of the Phased Progression of Pluripotency. Cell Syst. 2019 May
867 22;8(5):427-445.e10.

868 20. Krishnakumar R, Chen AF, Pantovich MG, Danial M, Parchem RJ, Labosky PA, et al.
869 FOXD3 Regulates Pluripotent Stem Cell Potential by Simultaneously Initiating and
870 Repressing Enhancer Activity. Cell Stem Cell. 2016 Jan 7;18(1):104–17.

871 21. Ashokkumar D, Zhang Q, Much C, Bledau AS, Naumann R, Alexopoulou D, et al. MLL4 is
872 required after implantation, whereas MLL3 becomes essential during late gestation. Dev
873 Camb Engl. 2020 Jun 17;147(12):dev186999.

874 22. Wang C, Lee JE, Lai B, Macfarlan TS, Xu S, Zhuang L, et al. Enhancer priming by H3K4
875 methyltransferase MLL4 controls cell fate transition. Proc Natl Acad Sci U S A. 2016 Oct
876 18;113(42):11871–6.

877 23. Chen AF, Liu AJ, Krishnakumar R, Freimer JW, DeVeale B, Blelloch R. GRHL2-dependent
878 enhancer switching maintains a pluripotent stem cell transcriptional subnetwork after exit
879 from naïve pluripotency. Cell Stem Cell. 2018 Aug 2;23(2):226-238.e4.

880 24. Skene PJ, Henikoff JG, Henikoff S. Targeted in situ genome-wide profiling with high
881 efficiency for low cell numbers. Nat Protoc. 2018 May;13(5):1006–19.

882 25. Stark R, Brown G. DiffBind: Differential Binding Analysis of ChIP-Seq Peak Data [Internet].
883 Bioconductor version: Release (3.15); 2022 [cited 2022 Aug 24]. Available from:
884 https://bioconductor.org/packages/DiffBind/

885 26. Meers MP, Tenenbaum D, Henikoff S. Peak calling by Sparse Enrichment Analysis for
886 CUT&RUN chromatin profiling. Epigenetics Chromatin. 2019 Jul 12;12(1):42.

887 27. Bleckwehl T, Crispatzu G, Schaaf K, Respuela P, Bartusel M, Benson L, et al. Enhancer-
888 associated H3K4 methylation safeguards in vitro germline competence. Nat Commun. 2021
889 Oct 1;12:5771.

890 28. Buecker C, Srinivasan R, Wu Z, Calo E, Acampora D, Faial T, et al. Reorganization of
891 enhancer patterns in transition from naive to primed pluripotency. Cell Stem Cell. 2014 Jun
892 5;14(6):838–53.

893 29. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, et al. CUT&Tag
894 for efficient epigenomic profiling of small samples and single cells. Nat Commun. 2019 Apr
895 29;10(1):1930.

896 30. Zhou HY, Katsman Y, Dhaliwal NK, Davidson S, Macpherson NN, Sakthidevi M, et al. A
897 Sox2 distal enhancer cluster regulates embryonic stem cell differentiation potential. Genes
898 Dev. 2014 Dec 15;28(24):2699–711.

899 31. Blinka S, Reimer MH, Pulakanti K, Rao S. Super-Enhancers at the Nanog Locus
900 Differentially Regulate Neighboring Pluripotency-Associated Genes. Cell Rep. 2016 Sep
901 27;17(1):19–28.

902 32. Thomas HF, Kotova E, Jayaram S, Pilz A, Romeike M, Lackner A, et al. Temporal
903 dissection of an enhancer cluster reveals distinct temporal and functional contributions of
904 individual elements. Mol Cell. 2021 Mar 4;81(5):969-982.e13.

905 33. Romeike M, Spach S, Huber M, Feng S, Vainorius G, Elling U, et al. Transient upregulation
906 of IRF1 during exit from naive pluripotency confers viral protection. EMBO Rep. 2022 Jul
907 19;e55375.

908 34. Agrawal P, Blinka S, Pulakanti K, Reimer MH, Stelloh C, Meyer AE, et al. Genome editing
909 demonstrates that the -5 kb Nanog enhancer regulates Nanog expression by modulating
910 RNAPII initiation and/or recruitment. J Biol Chem. 2021 Jun;296:100189.

911 35. Bruse N, Heeringen SJ van. GimmeMotifs: an analysis framework for transcription factor
912 motif analysis [Internet]. bioRxiv; 2018 [cited 2022 Aug 24]. p. 474403. Available from:
913 https://www.biorxiv.org/content/10.1101/474403v1

914 36. Moorthy SD, Davidson S, Shchuka VM, Singh G, Malek-Gilani N, Langroudi L, et al.
915 Enhancers and super-enhancers have an equivalent regulatory role in embryonic stem cells
916 through regulation of single or multiple genes. Genome Res. 2017 Feb;27(2):246–58.

917 37. Osterwalder M, Barozzi I, Tissières V, Fukuda-Yuzawa Y, Mannion BJ, Afzal SY, et al.
918 Enhancer redundancy provides phenotypic robustness in mammalian development. Nature.
919 2018 Feb 8;554(7691):239–43.

920 38. Mas G, Blanco E, Ballaré C, Sansó M, Spill YG, Hu D, et al. Promoter bivalency favors an
921 open chromatin architecture in embryonic stem cells. Nat Genet. 2018 Oct;50(10):1452–62.

922 39. Zhang H, Gayen S, Xiong J, Zhou B, Shanmugam AK, Sun Y, et al. MLL1 Inhibition
923 Reprograms Epiblast Stem Cells to Naïve Pluripotency. Cell Stem Cell. 2016 Apr
924 7;18(4):481–94.

925 40. Kaikkonen MU, Spann N, Heinz S, Romanoski CE, Allison KA, Stender JD, et al.
926 Remodeling of the enhancer landscape during macrophage activation is coupled to
927 enhancer transcription. Mol Cell. 2013 Aug 8;51(3):310–25.

928 41. Neumayr C, Haberle V, Serebreni L, Karner K, Hendy O, Boija A, et al. Differential cofactor
929 dependencies define distinct types of human enhancers. Nature. 2022 Jun;606(7913):406–
930 13.

931 42. Narita T, Higashijima Y, Kilic S, Liebner T, Walter J, Choudhary C. A unique H2B acetylation
932 signature marks active enhancers and predicts their target genes. bioRxiv. 2022 Jan
933 1;2022.07.18.500459.

934 43. Local A, Huang H, Albuquerque CP, Singh N, Lee AY, Wang W, et al. Identification of
935 H3K4me1-associated proteins at mammalian enhancers. Nat Genet. 2018 Jan;50(1):73–82.

936 44. Yan J, Chen SAA, Local A, Liu T, Qiu Y, Dorighi KM, et al. Histone H3 lysine 4
937 monomethylation modulates long-range chromatin interactions at enhancers. Cell Res.
938 2018 Feb;28(2):204–20.

939 45. Cao K, Collings CK, Morgan MA, Marshall SA, Rendleman EJ, Ozark PA, et al. An
940 Mll4/COMPASS-Lsd1 epigenetic axis governs enhancer function and pluripotency transition
941 in embryonic stem cells. Sci Adv. 2018 Jan;4(1):eaap8747.

942 46. Ang SY, Uebersohn A, Spencer CI, Huang Y, Lee JE, Ge K, et al. KMT2D regulates specific
943 programs in heart development via histone H3 lysine 4 di-methylation. Dev Camb Engl.
944 2016 Mar 1;143(5):810–21.

945 47. Shpargel KB, Mangini CL, Xie G, Ge K, Magnuson T. The KMT2D Kabuki syndrome histone
946 methylase controls neural crest cell differentiation and facial morphology. Dev Camb Engl.
947 2020 Jul 17;147(21):dev187997.

948 48. Xie G, Lee JE, Senft AD, Park YK, Chakraborty S, Thompson JJ, et al. MLL3/MLL4
949 methyltransferase activities control early embryonic development and embryonic stem cell
950 differentiation in a lineage-selective manner. bioRxiv. 2022 Jan 1;2020.09.14.296558.

951 49. Shechter D, Dormann HL, Allis CD, Hake SB. Extraction, purification and analysis of
952 histones. Nat Protoc. 2007;2(6):1445–57.

953 50. Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, et al. The nf-core framework
954 for community-curated bioinformatics pipelines. Nat Biotechnol. 2020 Mar;38(3):276–8.

955

956

957 **Figure Legends**

958

959 **Figure 1: MLL3/4 is dispensable for transcriptional activation of much of the formative**
960 **program**. A) Schematic of naïve to formative transition *in vivo* and *in vitro*. B) Western validation
961 of MLL3 and MLL4 knockout cell lines. C) Brightfield microscopy of cell lines in naive and
962 formative conditions, 20x, scale bar = 100um. D) Differential gene expression analysis (DGE) on
963 WT naive and formative RNA-seq samples (significant genes colored, P.adj < 0.05, Log2
964 Foldchange (Log2FC) >1). E) UMAP analysis on all RNA-seq samples subset by all naive and
965 formative genes from WT DGE. F,G) Fold change (formative/naive) of all genes for either
966 MLL3KO and DKO compared to WT. WT formative and naive genes from WT DGE analysis
967 colored. R, Pearson's coefficient. Dashed line and linear equation represent linear model of all
968 genes. H) DGE analysis on WT naive and DKO naive samples (significant genes colored, P.adj
969 < 0.05 and Log2FC > 1). I) Naive state DKO misexpressed genes categorized by whether they
970 are also WT naive or formative genes. J,K) Same as H,I respectively but comparing within the
971 formative state.

972

973 **Figure 2: MLL3/4 is required for all dynamic H3K4me1 deposition during pluripotent**
974 **transition.** A) Differential signal enrichment between naive and formative WT H3K4me1 peaks
975 identifies significant peaks (Black, FDR <0.05, Log2 Foldchange (Log2FC) > 1). B) Scatterplot of
976 H3K4me1 peak intensities between naive and formative WT samples. (gray, shared peaks, FDR
977 > 0.1 & Log2FC < 0.7, red or blue, FDR < 0.05 & Log2FC > 1). C,D) H3K4me1 signal in naive
978 and formative MLL3KO or DKO samples at peaks called from WT. Colors correspond to peak
979 categories derived from WT. E) Differential signal enrichment of naive and formative DKO
980 samples. F) Same as C,D but with naive and formative dCD cells. G) Differential signal enrichment
981 of naive and formative dCD samples. H) Feature annotation of WT peaks stratified by peak

982 category including intergenic/intronic (Distal), Promoter, and all else (Others). I) Distal H3K4me1
983 peak categories further stratified by MLL3/4 independent (DKO/WT Log2FC > -0.7) or dependent
984 (DKO/WT Log2FC < -1). J) Heatmap of MLL3/4 independent and dependent H3K4me1 peak
985 categories, rows sorted on "S" columns. All heatmap values and range are in CPM. For metagene
986 analysis the range in CPM is the same as shown in heatmap for each factor. K) nearest neighbor
987 TSS analysis of expression levels in Log2CPM for each RNA-seq dataset near H3K4me1 peak
988 categories. Multi-comparison paired Wilcoxon Rank-Sum Test, Benjamini-Hochberg corrected.
989 *p<0.05 **p<0.01 ***p<0.001 ****p<0.0001
990
991 **Figure 3: MLL3/4 dependent and independent distal H3K27ac deposition.** A) Differential
992 signal enrichment between naive and formative WT H3K27ac peaks identifies significant peaks
993 (Black, FDR <0.05, Log2 Foldchange (Log2FC) > 1). B) Scatterplot of H3K27ac peak intensities
994 between naive and formative WT samples. (gray, shared peaks, FDR > 0.1 & Log2FC < 0.7, red
995 or blue, FDR < 0.05 & Log2FC > 1). C) H3K27ac signal in naive and formative DKO samples at
996 peaks called from WT. Colors correspond to peak categories derived from WT. D) Feature
997 annotation of WT or DKO peaks stratified by peak category including intergenic/intronic (Distal),
998 Promoter, and all else (Others). E) Distal H3K27ac peak categories further stratified by MLL3/4
999 independent (DKO/WT Log2FC > -0.7) or dependent (DKO/WT Log2FC < -1). G,H) Heatmap of
1000 MLL3/4 independent and dependent H3K27ac peak categories for naive or formative specific
1001 peaks, rows sorted on "S" columns. All heatmap values and range are in CPM. For metagene
1002 analysis the range in CPM is the same as shown in heatmap for each factor. I) nearest neighbor
1003 TSS analysis of expression levels Log2CPM near H3K27ac peak categories. Multi-comparison
1004 paired Wilcoxon Rank-Sum Test, Benjamini-Hochberg corrected. *p<0.05 **p<0.01 ***p<0.001
1005 ****p<0.0001
1006

1007 **Figure 4: Enhancer Activation Can Occur Independently of MLL3/4.** A) Pie charts showing
1008 fraction of coincident changes in H3K27ac within the different categories of changing H3K4me1
1009 following MLL3/4. B) Heatmap of sites that normally lose H3K4me1 and H3K27ac during naive to
1010 formative transition clustered by either MLL3/4 dependent or independent H3K27ac. Rows sorted
1011 on "S" columns. Y-axis scale of metagene profiles above is same as signal ranges of heatmap,
1012 values in CPM. C) Same as B but using sites that normally gain H3K4me1 and H3K27ac during
1013 naive to formative transition clustered by either MLL3/4 dependent or independent H3K27ac. D)
1014 Same as B but using sites that have unchanging H3K4me1 but gain H3K27ac during naive to
1015 formative transition clustered by either MLL3/4 dependent or independent H3K27ac. All heatmap
1016 values and range are in CPM. For metagene analysis the range in CPM is the same as shown in
1017 heatmap for each factor. E) Naive and formative examples of functionally validated enhancers
1018 that are either MLL3/4 dependent or independent. Naive MLL3/4 dependent Sox2 enhancer
1019 cluster, Naive MLL3/4 independent Nanog enhancers, Formative independent and dependent
1020 enhancers of Fgf5, formative independent enhancer of Pou3f1/Oct6.
1021
1022 **Figure 5: Distal H3K4me1 and H3K27ac are not functionally coupled with formative**
1023 **transcriptional activation.** A) Naive ChromHMM emissions for 16 states yields all combinations
1024 of H3K4me1 and H3K27ac in WT and DKO. Heatmaps are clustered by each chromatin state.
1025 Black arrows denote states with MLL3/4 dependent H3K27ac. B) Gene expression changes of
1026 nearest TSS for all sites in each given state. Monte Carlo permutation sampling of "all genes
1027 control" to perform Mann-Whitney U Test for each state. Benjamini-Hochberg corrected. Selected
1028 statistics shown, additional statistics provided in supplemental table. C) Same as A with Formative
1029 samples. D) Same as B with Formative RNA-seq. All heatmap values and range are in CPM. For
1030 metagene analysis the range in CPM is the same as shown in heatmap for each factor.
1031 ChromHMM states without any H3K4me1 or H3K27ac emission probability are excluded. E) For
1032 all H3K4me1/H3K27ac+ sites in either Naive (top panel) or Formative (bottom panel), the total

1033 number of MLL3/4 dependent enhancers for a gene compared with the fold change of RNA levels
1034 DKO/WT in Log2CPM. Each dot represents one gene. Blue line represents generalized linear
1035 model, gray 95% confidence interval. F) Boxplots of relative expression DKO/WT of RNA levels
1036 for all genes associated with any H3K4me1/H3K27ac+ peak in either the naïve (top panel) or
1037 formative state (bottom panel). Each gene is binned by the percentage of their associated
1038 enhancer loss in DKOs. *p<0.05 **p<0.01 ***p<0.001 ****p<0.0001
1039
1040 **Figure 6: Gene-centric analysis reveals a subset of distal loci associate with MLL3/4**
1041 **dependent formative genes.** A) Pie charts showing fraction of normally gained H3K4me1,
1042 H3K27ac, or gene expression that are dependent on MLL3/4 or not. 215 formative genes
1043 represent those of the 283 formative down genes from Fig.1K without preexisting naive
1044 expression defects (Log2CPM naive DKO/WT > -1). B) Clusterprofile analysis of Biological
1045 Processes Gene Ontology for 215 MLL3/4 dependent formative genes. C) Fold change Log2CPM
1046 of H3K27ac density for H3K4me1/H3K27ac+ enhancers associated with formative genes that
1047 gained expression during transition in an either MLL3/4 independent or dependent fashion. Only
1048 131 out of 215 dependent genes and 820 out of 1387 independent genes were able to be
1049 associated with an enhancer. D) Genome tracks for known and predicted formative targets of
1050 MLL3/4 Lefty1, Brachyury/T, and Fgf8. E) Model for the role of MLL3/4 independent and
1051 dependent mechanisms in the regulation of gained gene expression during the naïve to formative
1052 transition.

1053 **Supplementary Figure Legends**

1054

1055 **Figure S1: MLL3/4 is dispensable for transcriptional activation of much of the formative**
1056 **program**. A) Confirmation of multiple DKO clones using genomic PCR following tamoxifen
1057 treatment. B) Proliferation measurements using crystal violet assays. C) Westerns on whole cell
1058 protein fractions of OCT4, NANOG and αTUBULIN loading control. Quantifications below are fold
1059 changes relative to naïve WT after normalizing OCT4 or NANOG for TUBULIN levels in each
1060 lane. D) qPCR of formative enriched markers Fgf5, Otx2, Dnmt3b. For each marker selected
1061 comparisons are shown resulting from a Two-way ANOVA, Tukey's multiple comparison test for
1062 all samples. Mean and S.E.M. in black. Additional statistics provided in supplemental table. E)
1063 Boxplots of transcript levels in Log2CPM for unchanging, naïve enriched or formative enriched
1064 genes. F) DGE analysis on MLL3KO naive and DKO naive samples (significant genes colored,
1065 P.adj < 0.05 and Log2FC > 1). G) Same as F but comparing the formative state. *p<0.05 **p<0.01
1066 ***p<0.001 ****p<0.0001

1067

1068 **Figure S2: MLL3/4 is required for all dynamic H3K4me1 deposition during pluripotent**
1069 **transition.** A) Diffbind analysis of H3K4me1 signal in MLL3KO cells at WT peaks. B)
1070 Quantifications of westerns on histones purified by acid-extraction. H3K4me1 levels are relative
1071 to naïve WT levels after first normalizing by H3 signal in each lane. Black dot represents mean
1072 (n=2). C) mRNA Log2CPM of naive MLL3/4 dCD cells vs. WT cells. Significant genes highlighted.
1073 D) same as B with formative samples. E) Foldchange (formative/naive) of all genes for MLL3/4
1074 dCD compared to WT. WT formative and naive genes from WT DGE analysis colored. R,
1075 Pearson's coefficient. Dashed line and linear equation represent linear model of all genes. F)
1076 Heatmap of ChIP-seq for H3K4me1 in naive and formative cells from two published studies. All
1077 heatmap values and range are in CPM. For metagene analysis the range in CPM is the same as
1078 shown in heatmap for each factor. G) Nearest neighbor TSS analysis of expression levels in

1079 Log2CPM for each RNA-seq dataset near naive MLL3/4 ChIP-seq peak categories. Multi-
1080 comparison paired Wilcoxon Rank-Sum Test, Benjamini-Hochberg corrected. *p<0.05 **p<0.01
1081 ***p<0.001 ****p<0.0001
1082
1083 **Figure S3: MLL3/4 dependent and independent distal H3K27ac deposition.** A) Heatmap of
1084 shared H3K27ac sites clustered by MLL3/4 independent or dependent H3K27ac. B) Heatmaps of
1085 ChIP-seq for H3K27ac in naive and formative cells from two published studies. All heatmap values
1086 and range are in CPM. For metagene analysis the range in CPM is the same as shown in heatmap
1087 for each factor.
1088
1089 **Figure S4: Enhancer Activation Can Occur Independently of MLL3/4.** A) Peak count of
1090 H3K4me1 categories from Fig2. after filtering for distal H3K4me1 overlapping WT ATAC peaks.
1091 B) Nearest neighbor TSS analysis of expression levels in Log2CPM for each RNA-seq dataset
1092 near H3K4me1 peak categories after subsetting for ATAC peak overlap. Multi-comparison paired
1093 Wilcoxon Rank-Sum Test, Benjamini-Hochberg corrected. C) Heatmaps for shared H3K4me1 that
1094 overlaps with naive enriched H3K27ac, clustered by MLL3/4 independent or dependent H3K27ac.
1095 D) Heatmaps for MLL3/4 dependent H3K4me1 overlapped with sites that have shared, naive, or
1096 formative H3K27ac. Clustered additionally by MLL3/4 independent or dependent H3K27ac. E)
1097 Heatmaps for shared H3K4me1 that overlaps with shared H3K27ac, clustered by MLL3/4
1098 independent or dependent H3K27ac. All heatmap values and range are in CPM. For metagene
1099 analysis the range in CPM is the same as shown in heatmap for each factor. F) Log2CPM of Sox2
1100 and Nanog in naive and formative samples comparing WT and DKO cells. Sox2 is significantly
1101 up in the formative state. *p<0.05 **p<0.01 ***p<0.001 ****p<0.0001
1102
1103 **Figure S5: Enhancer Activation Can Occur Independently of MLL3/4.** A) Clusters from
1104 differential motif analysis using Gimmemotifs on formative enriched MLL3/4 independent and

1105 dependent enhancers marked by both H3K4me1 and H3K27ac (Sites from Figure 4C and 4D).
1106 Motifs derived from JASPAR 2020 database, k-means clustering used. B) Same as A but with
1107 naive enriched independent and dependent enhancers (using sites from Figure 4B and Figure
1108 S5B).
1109
1110 **Figure S6: Distal H3K4me1 and H3K27ac are not functionally coupled with formative**
1111 **transcriptional activation.** A) Nearest neighbor TSS analysis of expression levels in Log2CPM
1112 near naive H3K4me1 dependent sites that are either have MLL3/4 independent (left panel) or
1113 dependent H3K27ac (right panel)(clusters from Fig.4B). Multi-comparison paired Wilcoxon Rank-
1114 Sum Test, Benjamini-Hochberg corrected. B) Same as A but at formative H3K4me1 dependent
1115 sites (clusters from Fig.4C). C) The relative expression DKO/WT of RNA levels for all genes
1116 associated with any H3K4me1/H3K27ac+ peak in either naive or formative state compared with
1117 the number of dependent enhancers for each gene. Only dependent enhancers within 10kb of
1118 the nearest TSS are considered. Each dot represents one gene. Blue line represents generalized
1119 linear model, gray 95% confidence interval. D) Boxplots of relative expression DKO/WT of RNA
1120 levels for all genes associated with any H3K4me1/H3K27ac+ peak in either the naïve (left) or
1121 formative state (right). Each gene is binned by the percentage of their associated enhancer loss
1122 in DKOs using only enhancers within 10kb of the nearest TSS. *p<0.05 **p<0.01 ***p<0.001
1123 ****p<0.0001
1124
1125 **Figure S7: Gene-centric analysis reveals a subset of distal loci associate with MLL3/4**
1126 **dependent formative genes.** A) A major percentage of naive enriched peaks fail to maintain
1127 distal H3K4me1 and H3K27ac with moderate consequences on naive gene expression. B)
1128 Clusterprofile analysis of Biological Processes Gene Ontology for 348 MLL3/4 dependent naive
1129 genes. C) Fold change Log2CPM of H3K27ac density for H3K4me1/H3K27ac+ enhancers that

1130 are associated with naive genes that gained expression during transition in an either MLL3/4

1131 independent or dependent fashion.

Figure 1

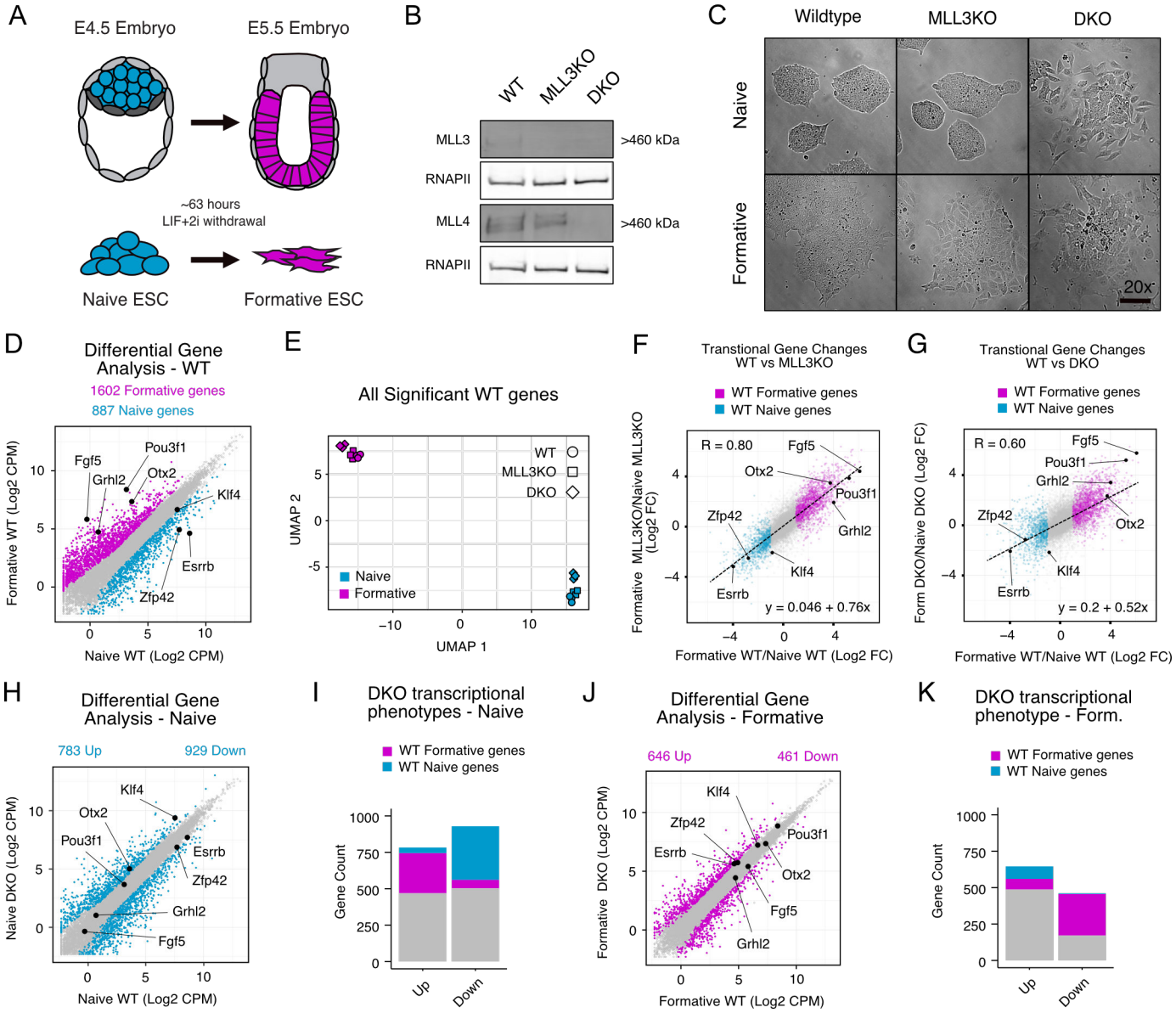

A
E4.5 Embryo
E5.5 Embryo
~63 hours
LIF+2i withdrawal
Naive ESC
Formative ESC
B
WT
MLL3KO
DKO
MLL3
>460 kDa
RNAPII
MLL4
>460 kDa
RNAPII
C
Wildtype
MLL3KO
DKO
Naive
Formative
20x
D
Differential Gene Analysis - WT
1602 Formative genes
887 Naive genes
Formative WT (Log2 CPM)
Naive WT (Log2 CPM)
Fgf5
Grhl2
Pou3f1
Otx2
Klf4
Esrrb
Zfp42
E
All Significant WT genes
WT
MLL3KO
DKO
Naive
Formative
UMAP 1
UMAP 2
F
Transtional Gene Changes WT vs MLL3KO
WT Formative genes
WT Naive genes
R = 0.80
y = 0.046 + 0.76x
Formative MLL3KO/Naive MLL3KO (Log2 FC)
Formative WT/Naive WT (Log2 FC)
G
Transtional Gene Changes WT vs DKO
WT Formative genes
WT Naive genes
R = 0.60
y = 0.2 + 0.52x
Form DKO/Naive DKO (Log2 FC)
Formative WT/Naive WT (Log2 FC)
H
Differential Gene Analysis - Naive
783 Up
929 Down
Naive DKO (Log2 CPM)
Naive WT (Log2 CPM)
I
DKO transcriptional phenotypes - Naive
WT Formative genes
WT Naive genes
Gene Count
Up
Down
J
Differential Gene Analysis - Formative
646 Up
461 Down
Formative DKO (Log2 CPM)
Formative WT (Log2 CPM)
K
DKO transcriptional phenotype - Form.
WT Formative genes
WT Naive genes
Gene Count
Up
Down

Figure 2

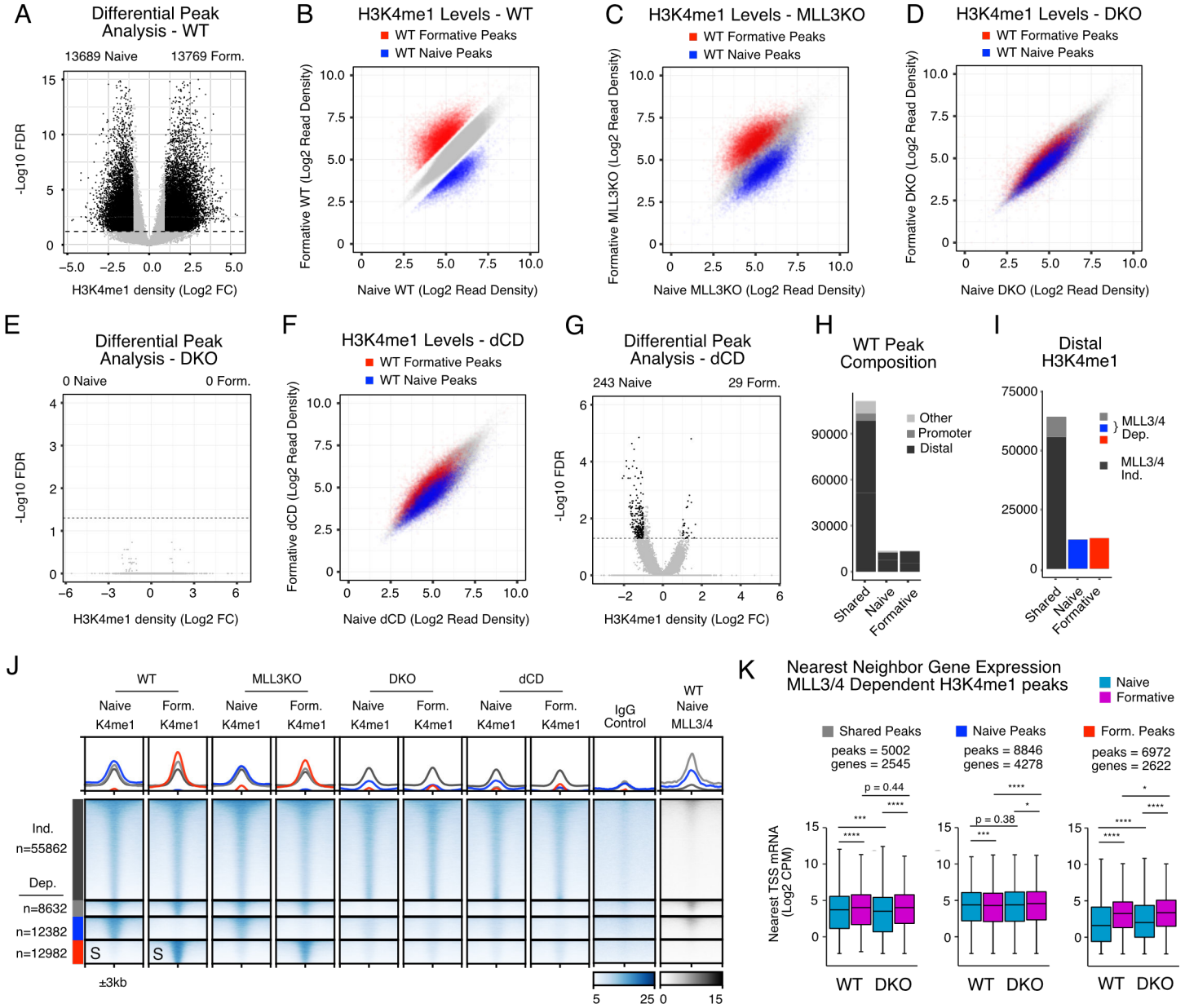

A
Differential Peak Analysis - WT
13689 Naive
13769 Form.
-Log10 FDR
H3K4me1 density (Log2 FC)
B
H3K4me1 Levels - WT
WT Formative Peaks
WT Naive Peaks
Formative WT (Log2 Read Density)
Naive WT (Log2 Read Density)
C
H3K4me1 Levels - MLL3KO
WT Formative Peaks
WT Naive Peaks
Formative MLL3KO (Log2 Read Density)
Naive MLL3KO (Log2 Read Density)
D
H3K4me1 Levels - DKO
WT Formative Peaks
WT Naive Peaks
Formative DKO (Log2 Read Density)
Naive DKO (Log2 Read Density)
E
Differential Peak Analysis - DKO
0 Naive
0 Form.
-Log10 FDR
H3K4me1 density (Log2 FC)
F
H3K4me1 Levels - dCD
WT Formative Peaks
WT Naive Peaks
Formative dCD (Log2 Read Density)
Naive dCD (Log2 Read Density)
G
Differential Peak Analysis - dCD
243 Naive
29 Form.
-Log10 FDR
H3K4me1 density (Log2 FC)
H
WT Peak Composition
Other
Promoter
Distal
Shared
Naive
Formative
I
Distal H3K4me1
MLL3/4 Dep.
MLL3/4 Ind.
Shared
Naive
Formative
J
WT
MLL3KO
DKO
dCD
Naive K4me1
Form. K4me1
IgG Control
WT Naive MLL3/4
Ind.
n=55862
Dep.
n=8632
n=12382
n=12982
S
±3kb
K
Nearest Neighbor Gene Expression
MLL3/4 Dependent H3K4me1 peaks
Naive
Formative
Shared Peaks
peaks = 5002
genes = 2545
Naive Peaks
peaks = 8846
genes = 4278
Form. Peaks
peaks = 6972
genes = 2622
p = 0.44
p = 0.38
Nearest TSS mRNA (Log2 CPM)
WT
DKO

Figure 3

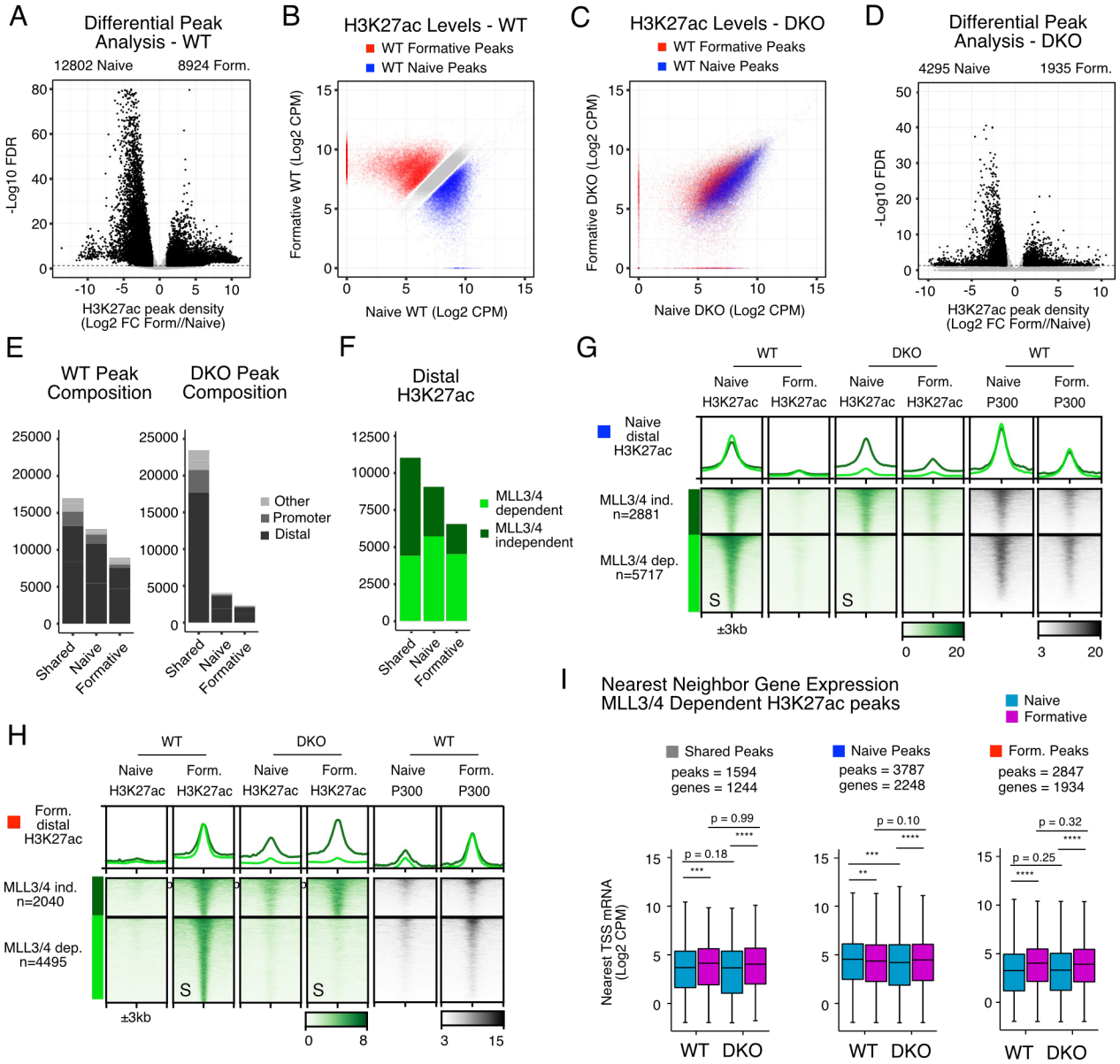

A
Differential Peak Analysis - WT
12802 Naive
8924 Form.
-Log10 FDR
H3K27ac peak density (Log2 FC Form//Naive)
B
H3K27ac Levels - WT
WT Formative Peaks
WT Naive Peaks
Formative WT (Log2 CPM)
Naive WT (Log2 CPM)
C
H3K27ac Levels - DKO
WT Formative Peaks
WT Naive Peaks
Formative DKO (Log2 CPM)
Naive DKO (Log2 CPM)
D
Differential Peak Analysis - DKO
4295 Naive
1935 Form.
-Log10 FDR
H3K27ac peak density (Log2 FC Form//Naive)
E
WT Peak Composition
DKO Peak Composition
Other
Promoter
Distal
Shared
Naive
Formative
F
Distal H3K27ac
MLL3/4 dependent
MLL3/4 independent
Shared
Naive
Formative
G
WT
DKO
WT
Naive H3K27ac
Form. H3K27ac
Naive H3K27ac
Form. H3K27ac
Naive P300
Form. P300
Naive distal H3K27ac
MLL3/4 ind. n=2881
MLL3/4 dep. n=5717
S
±3kb
H
WT
DKO
WT
Naive H3K27ac
Form. H3K27ac
Naive H3K27ac
Form. H3K27ac
Naive P300
Form. P300
Form. distal H3K27ac
MLL3/4 ind. n=2040
MLL3/4 dep. n=4495
S
±3kb
I
Nearest Neighbor Gene Expression
MLL3/4 Dependent H3K27ac peaks
Naive
Formative
Shared Peaks
peaks = 1594
genes = 1244
Naive Peaks
peaks = 3787
genes = 2248
Form. Peaks
peaks = 2847
genes = 1934
p = 0.99
p = 0.18
p = 0.10
p = 0.32
p = 0.25
Nearest TSS mRNA (Log2 CPM)
WT
DKO

Figure 4

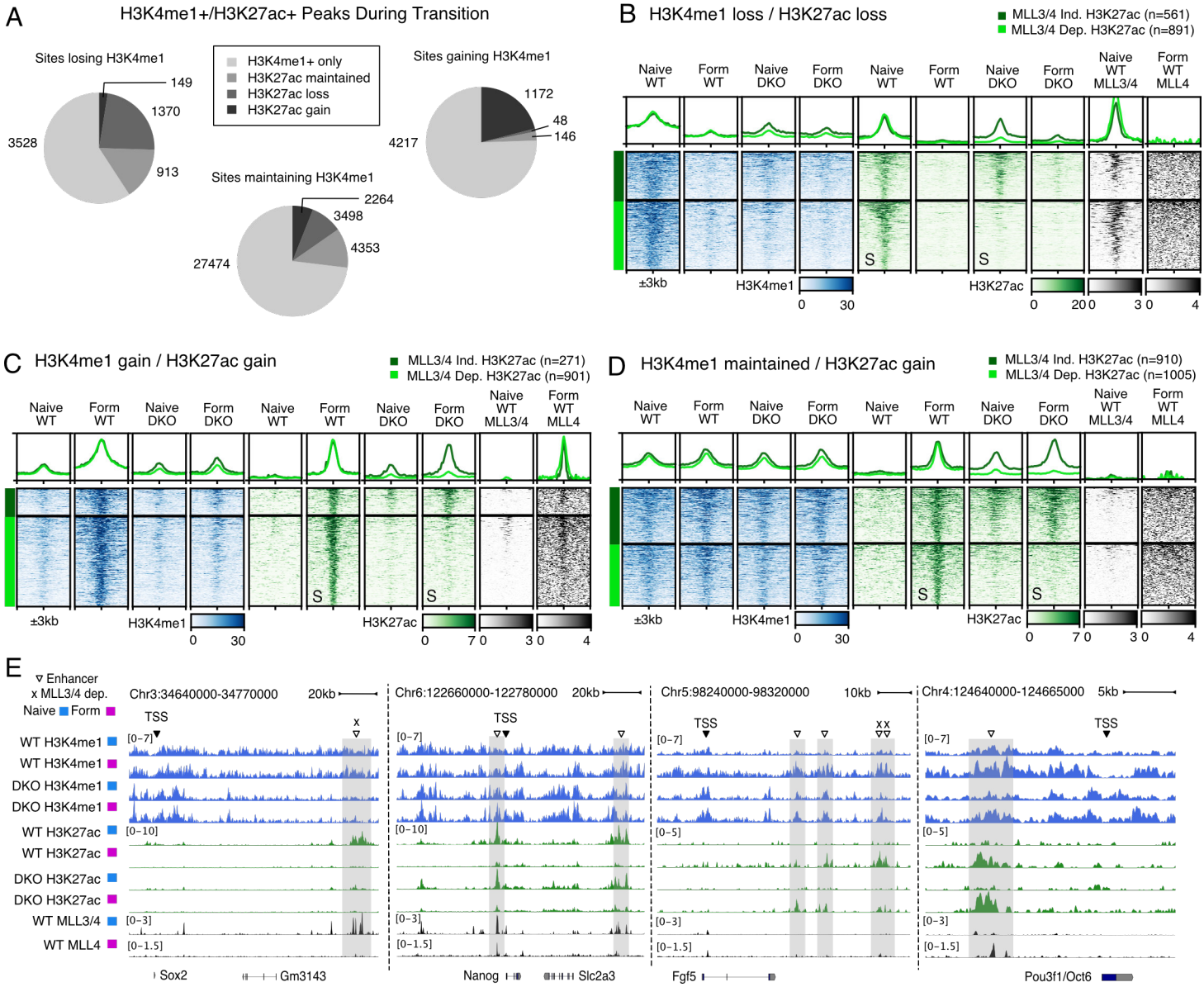

A
H3K4me1+/H3K27ac+ Peaks During Transition
Sites losing H3K4me1
149
1370
3528
913
H3K4me1+ only
H3K27ac maintained
H3K27ac loss
H3K27ac gain
Sites gaining H3K4me1
1172
48
146
4217
Sites maintaining H3K4me1
2264
3498
4353
27474
B
H3K4me1 loss / H3K27ac loss
MLL3/4 Ind. H3K27ac (n=561)
MLL3/4 Dep. H3K27ac (n=891)
Naive WT
Form WT
Naive DKO
Form DKO
Naive WT
Form WT
Naive DKO
Form DKO
Naive WT MLL3/4
Form WT MLL4
S
S
±3kb
H3K4me1
0
30
H3K27ac
0
20
0
3
0
4
C
H3K4me1 gain / H3K27ac gain
MLL3/4 Ind. H3K27ac (n=271)
MLL3/4 Dep. H3K27ac (n=901)
Naive WT
Form WT
Naive DKO
Form DKO
Naive WT
Form WT
Naive DKO
Form DKO
Naive WT MLL3/4
Form WT MLL4
S
S
±3kb
H3K4me1
0
30
H3K27ac
0
7
0
3
0
4
D
H3K4me1 maintained / H3K27ac gain
MLL3/4 Ind. H3K27ac (n=910)
MLL3/4 Dep. H3K27ac (n=1005)
Naive WT
Form WT
Naive DKO
Form DKO
Naive WT
Form WT
Naive DKO
Form DKO
Naive WT MLL3/4
Form WT MLL4
S
S
±3kb
H3K4me1
0
30
H3K27ac
0
7
0
3
0
4
E
Enhancer
x MLL3/4 dep.
Naive
Form
WT H3K4me1
WT H3K4me1
DKO H3K4me1
DKO H3K4me1
WT H3K27ac
WT H3K27ac
DKO H3K27ac
DKO H3K27ac
WT MLL3/4
WT MLL4
Chr3:34640000-34770000
20kb
TSS
x
[0-7]
[0-10]
[0-3]
[0-1.5]
Sox2
Gm3143
Chr6:122660000-122780000
20kb
TSS
[0-7]
[0-10]
[0-3]
[0-1.5]
Nanog
Slc2a3
Chr5:98240000-98320000
10kb
TSS
xx
[0-7]
[0-5]
[0-3]
[0-1.5]
Fgf5
Chr4:124640000-124665000
5kb
TSS
[0-7]
[0-5]
[0-3]
[0-1.5]
Pou3f1/Oct6

Figure 5

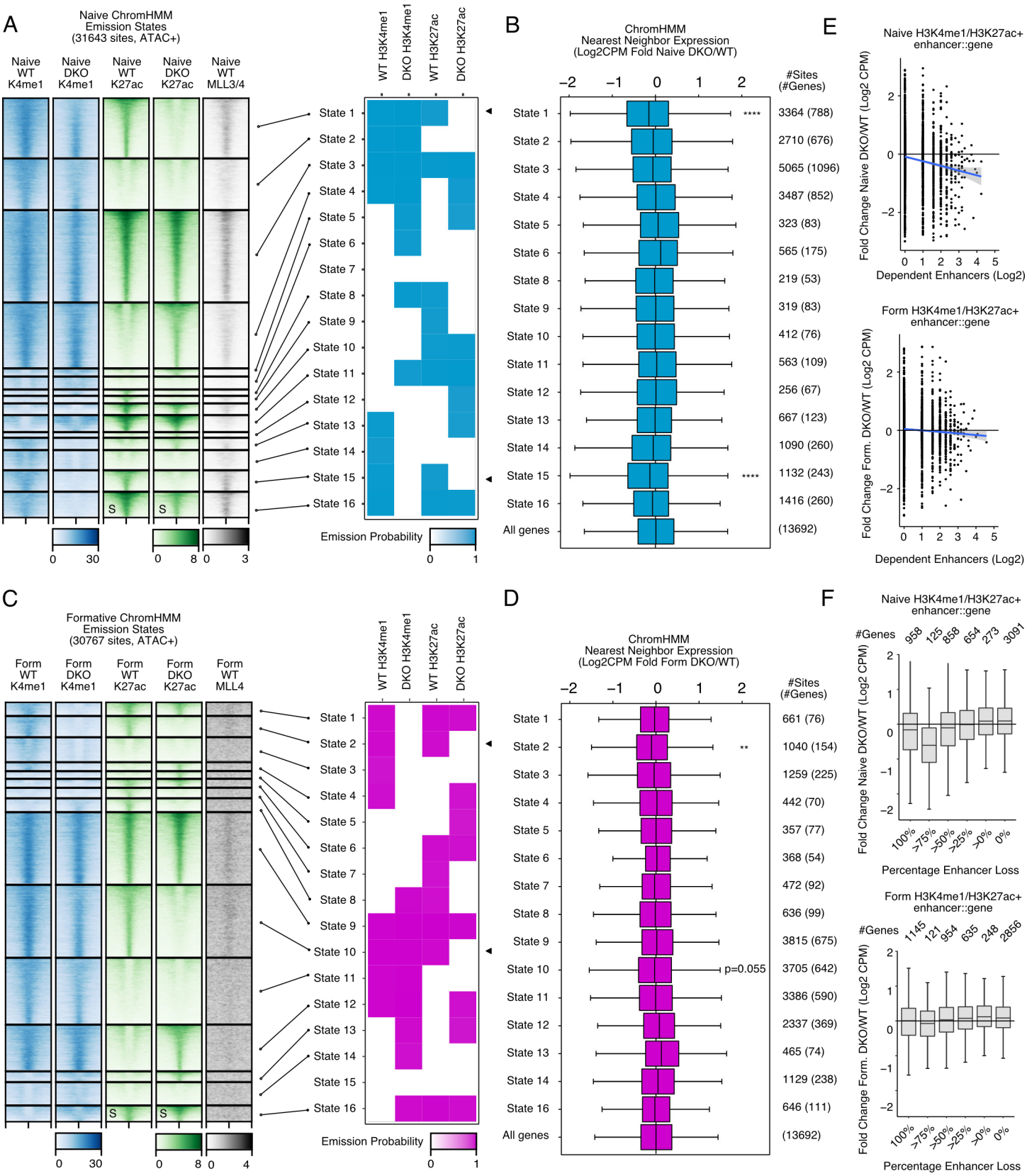
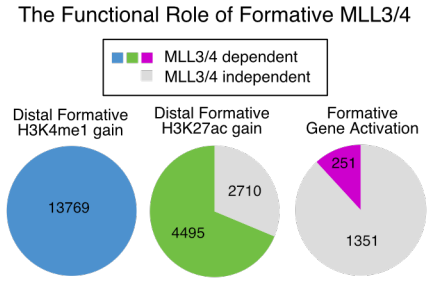
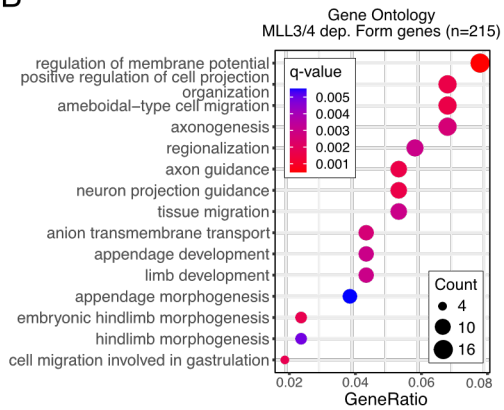

A
Naive ChromHMM
Emission States
(31643 sites, ATAC+)
Naive WT K4me1
Naive DKO K4me1
Naive WT K27ac
Naive DKO K27ac
Naive WT MLL3/4
WT H3K4me1
DKO H3K4me1
WT H3K27ac
DKO H3K27ac
State 1
State 2
State 3
State 4
State 5
State 6
State 7
State 8
State 9
State 10
State 11
State 12
State 13
State 14
State 15
State 16
Emission Probability
B
ChromHMM
Nearest Neighbor Expression
(Log2CPM Fold Naive DKO/WT)
#Sites (#Genes)
3364 (788)
2710 (676)
5065 (1096)
3487 (852)
323 (83)
565 (175)
219 (53)
319 (83)
412 (76)
563 (109)
256 (67)
667 (123)
1090 (260)
1132 (243)
1416 (260)
(13692)
All genes
E
Naive H3K4me1/H3K27ac+ enhancer::gene
Fold Change Naive DKO/WT (Log2 CPM)
Dependent Enhancers (Log2)
Form H3K4me1/H3K27ac+ enhancer::gene
Fold Change Form. DKO/WT (Log2 CPM)
C
Formative ChromHMM
Emission States
(30767 sites, ATAC+)
Form WT K4me1
Form DKO K4me1
Form WT K27ac
Form DKO K27ac
Form WT MLL4
D
ChromHMM
Nearest Neighbor Expression
(Log2CPM Fold Form DKO/WT)
661 (76)
1040 (154)
1259 (225)
442 (70)
357 (77)
368 (54)
472 (92)
636 (99)
3815 (675)
3705 (642)
3386 (590)
2337 (369)
465 (74)
1129 (238)
646 (111)
(13692)
p=0.055
F
Naive H3K4me1/H3K27ac+ enhancer::gene
#Genes
958 125 858 654 273 3091
Percentage Enhancer Loss
Form H3K4me1/H3K27ac+ enhancer::gene
1145 121 954 635 248 2856

Figure 6

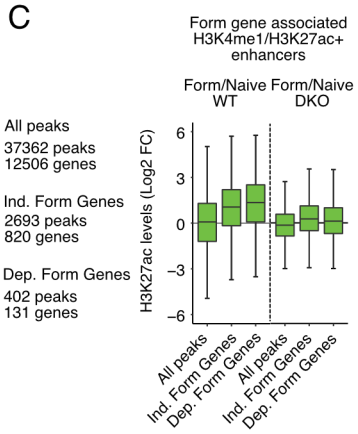
A



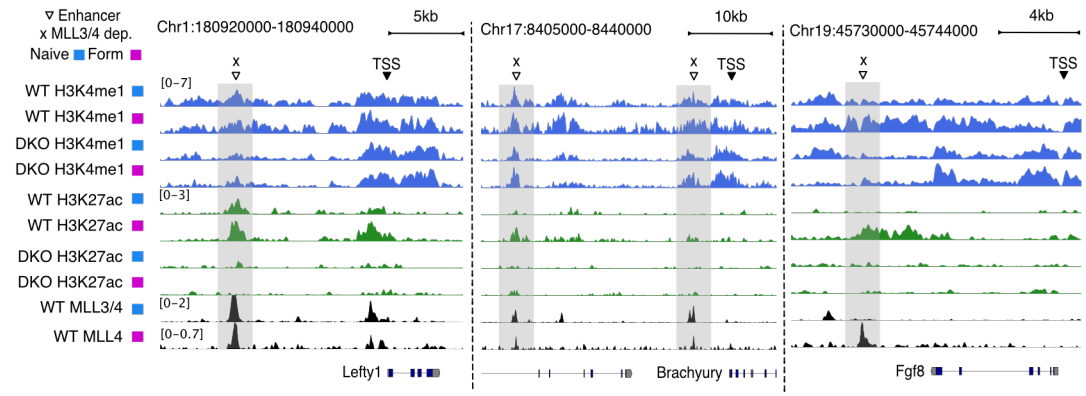
B



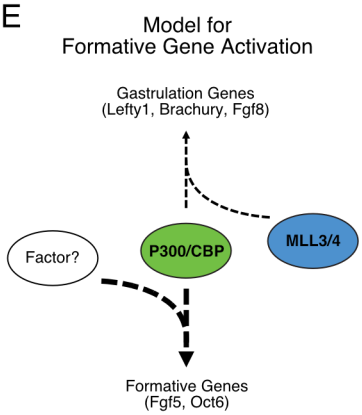
C



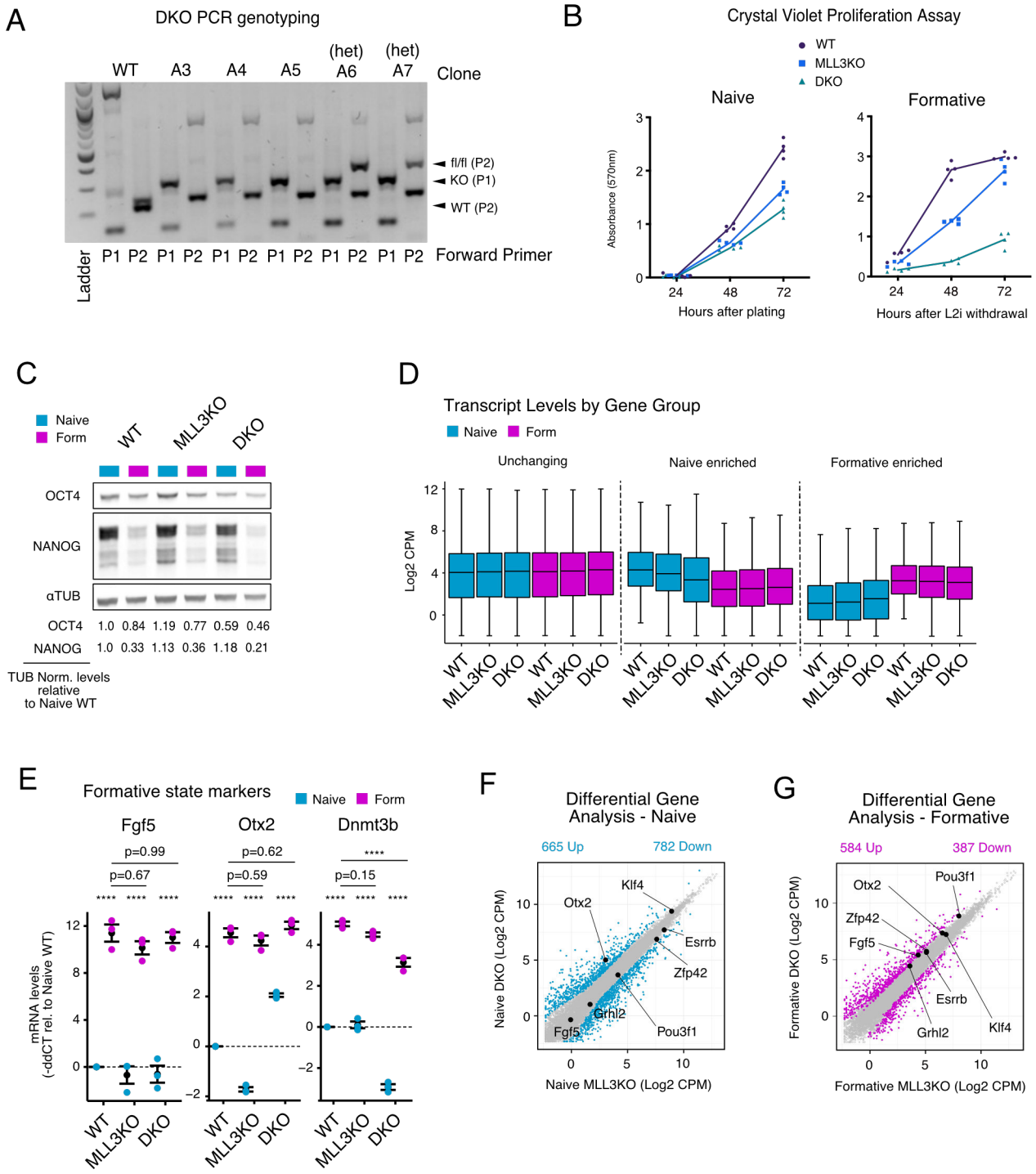
D



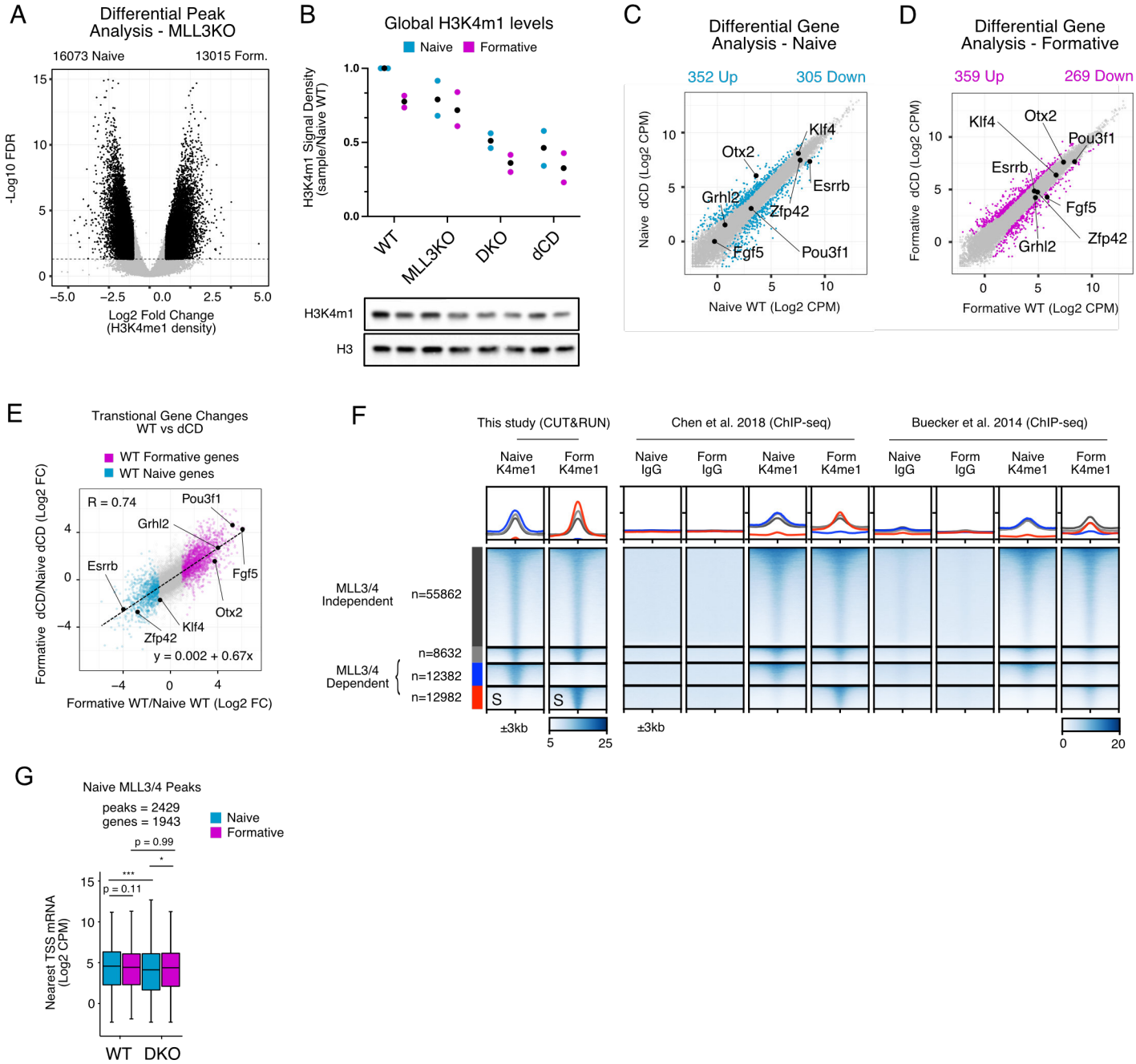
E



Supplemental Figure 1

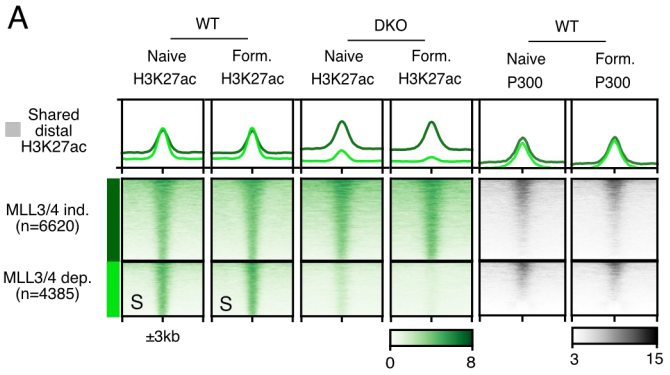


Supplemental Figure 2

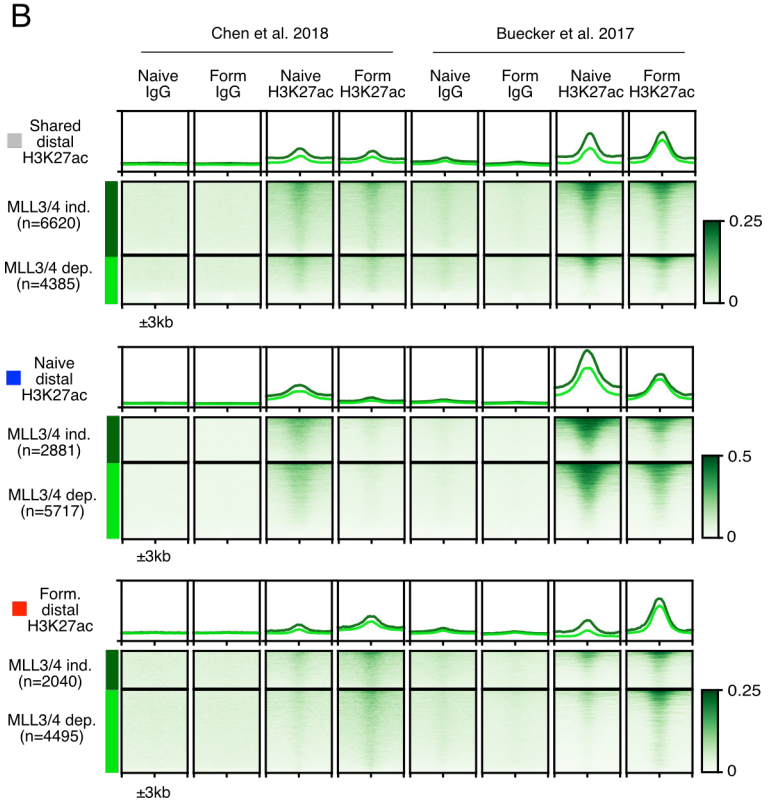


Supplemental Figure 3

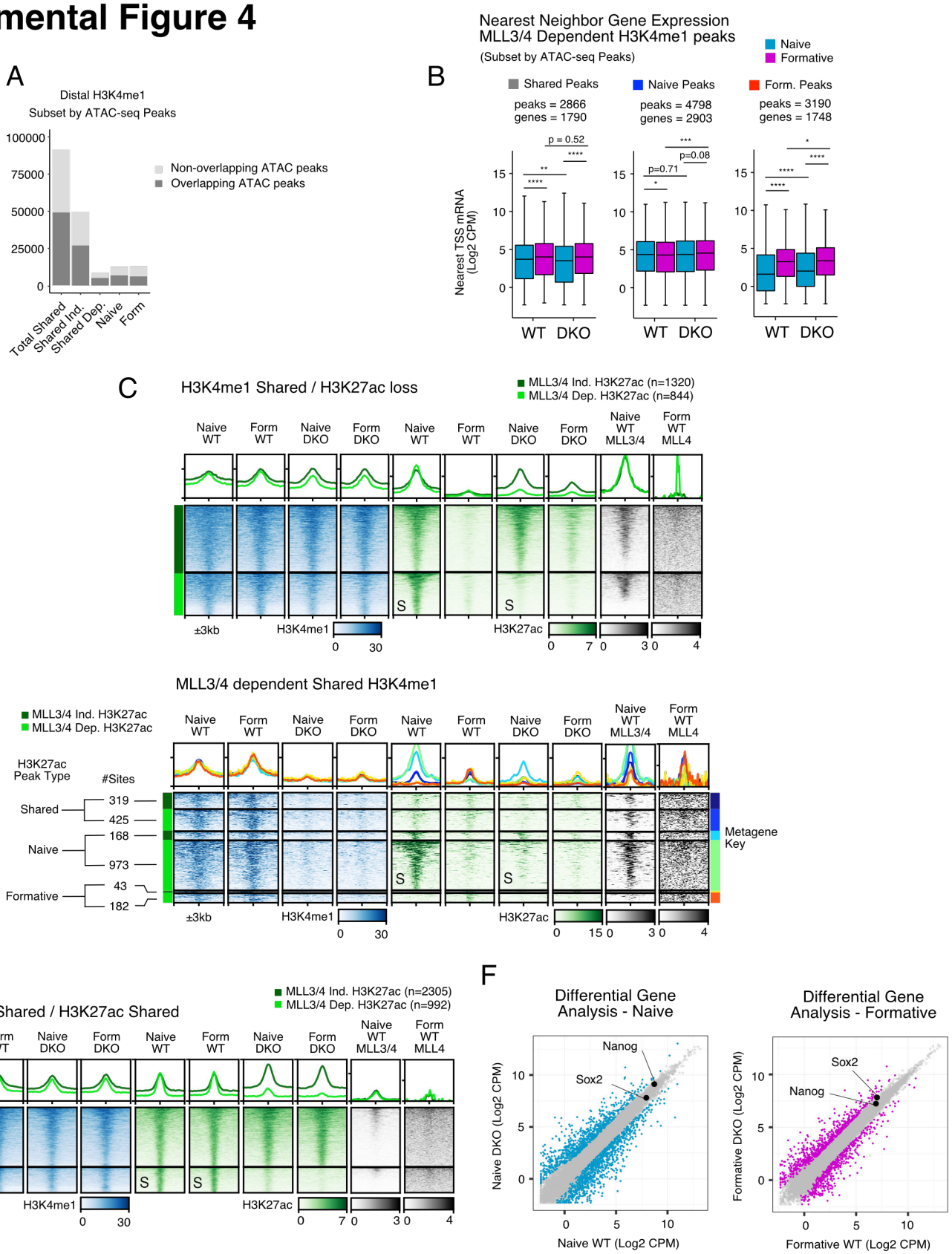
A



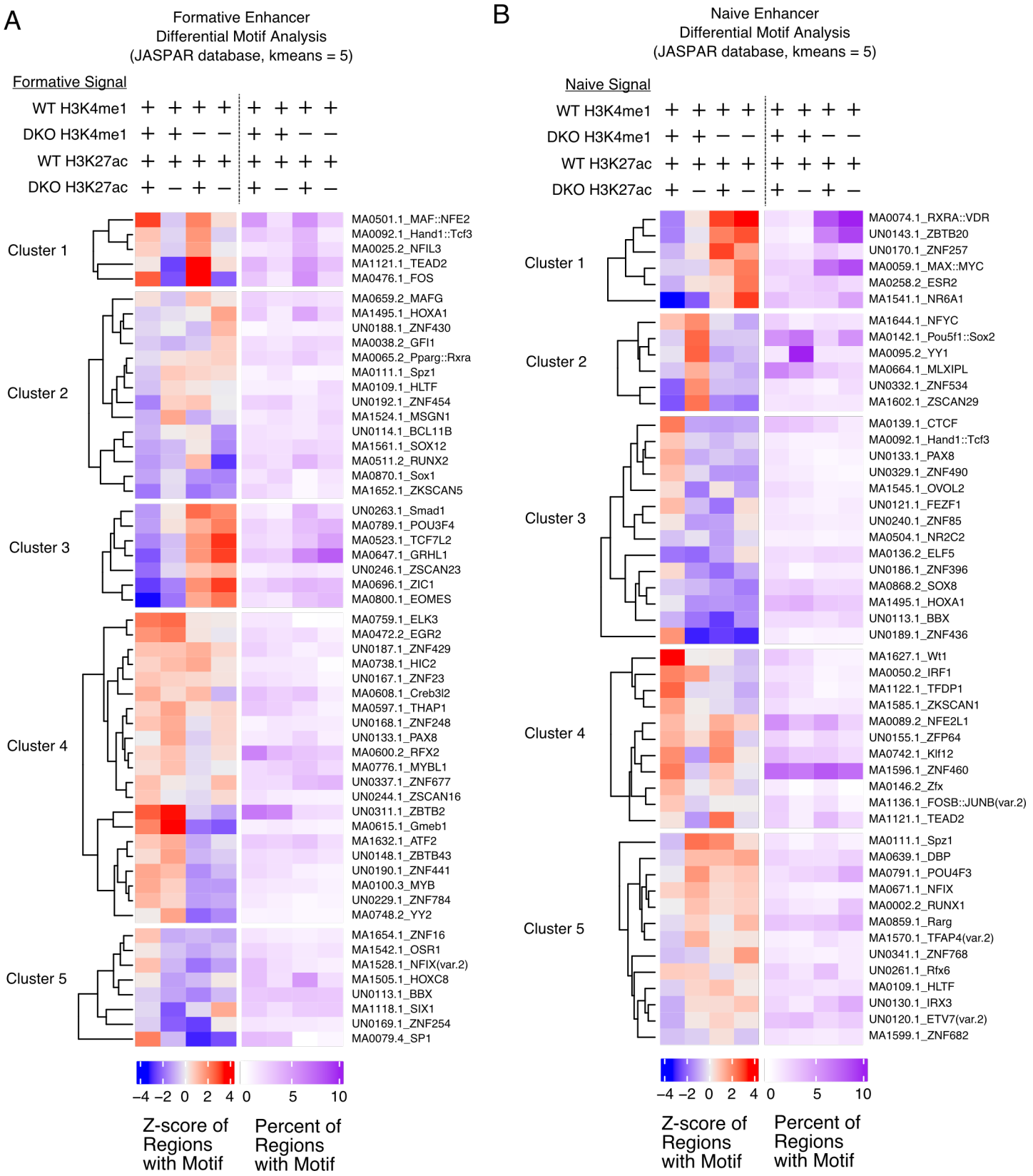
B



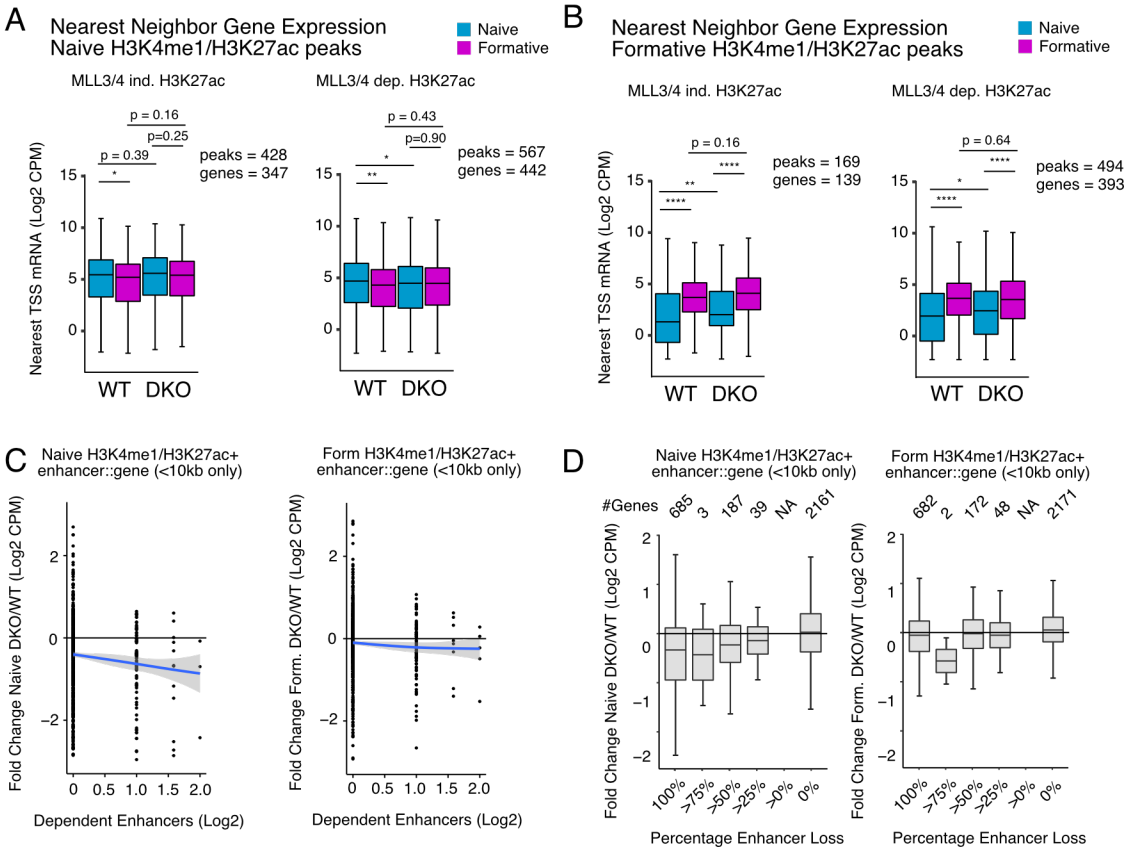
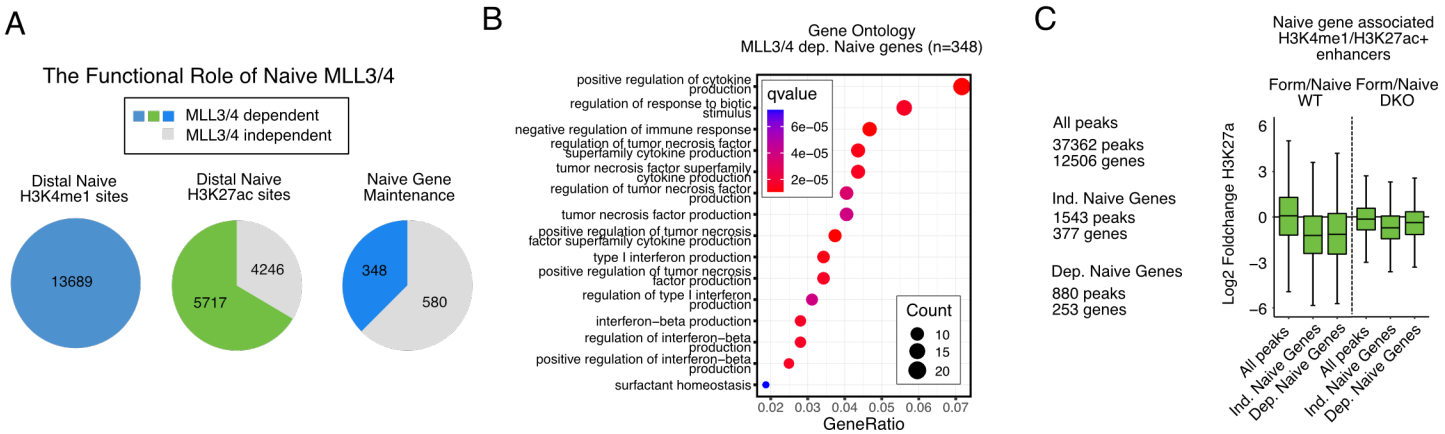
Supplemental Figure 4



Supplemental Figure 5



Supplemental Figure 6



Supplemental Figure 7