

# 1    **REPRESSION BY HDAC3 AND DAX1 MEDIATES LINEAGE** 2    **RESTRICTION OF EMBRYONIC STEM CELLS**

3    Daniel Olivieri<sup>1,\*</sup>, Panagiotis Papasaikas<sup>1,2</sup>, Ilya Lukonin<sup>1</sup>, Melanie Rittirsch<sup>1</sup>, Daniel  
4    Hess<sup>1</sup>, Sébastien A. Smallwood<sup>1</sup>, Michael B. Stadler<sup>1,2</sup>, Joerg Betschinger<sup>1,\*</sup>

5  
6    <sup>1</sup> Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058  
7    Basel, Switzerland

8    <sup>2</sup> Swiss Institute of Bioinformatics, 4058 Basel, Switzerland

9  
10    \* Correspondence to: [daniel.olivieri@fmi.ch](mailto:daniel.olivieri@fmi.ch), [joerg.betschinger@fmi.ch](mailto:joerg.betschinger@fmi.ch)

11  
12    Contact for editorial correspondence: [joerg.betschinger@fmi.ch](mailto:joerg.betschinger@fmi.ch)

## 13 14    **ABSTRACT**

15    Mouse embryonic stem cells (mESCs) give rise to embryonic but not extraembryonic  
16    endoderm fates. Here, we identify the mechanism of this lineage barrier and report  
17    that the histone deacetylase Hdac3 and the corepressor Dax1 cooperatively restrict  
18    transdifferentiation of mESCs by silencing an enhancer of the extraembryonic  
19    endoderm-specifying transcription factor (TF) Gata6. This restriction is opposed by  
20    the pluripotency TFs Nr5a2 and Esrrb, which promote cell type conversion.  
21    Perturbation of the barrier extends mESC potency, and allows formation of 3D  
22    spheroids that mimic the spatial segregation of embryonic epiblast and  
23    extraembryonic endoderm in early embryos. Overall, this study shows that  
24    transcriptional repressors stabilize pluripotency by biasing the equilibrium between  
25    embryonic and extraembryonic lineages that is hardwired into the mESC TF network.

# 1 INTRODUCTION

2 Binary cell fate decisions generate two distinct daughter cell types from one common  
3 progenitor. Once specified, cells need to prevent de-differentiation and  
4 transdifferentiation in order to stay fate-committed. The processes of lineage  
5 specification and maintenance can employ different mechanisms (Holmberg and  
6 Perlmann, 2012), yet the regulatory principles underlying these differences are not  
7 well understood.

8 A well-studied developmental binary cell fate decision is the differentiation of the  
9 mouse embryonic day (E) 3.5 inner cell mass (ICM) into extraembryonic primitive  
10 endoderm (PrE) and pluripotent epiblast (EPI) (Hermitte and Chazaud, 2014). This is  
11 mediated by the lineage-specifying transcription factors (TFs) Nanog and Gata6 that  
12 are co-expressed in the ICM. Positive auto-regulation and mutual inhibition of Nanog  
13 and Gata6, modulated by extrinsic fibroblast growth factor (FGF) signaling, drives  
14 segregation into Gata6-expressing PrE and Nanog-expressing EPI at E4.5 (Simon et  
15 al., 2018).

16 Developmental potency of the EPI is captured *in vitro* by mouse embryonic stem cells  
17 (mESCs) that are restricted to predominantly give rise to embryonic cell types *in vitro*  
18 and when injected into blastocysts (Beddington and Robertson, 1989). The lineage  
19 barrier impeding PrE differentiation of mESCs is not well understood.  
20 Overexpression of Gata6 enables transdifferentiation into extraembryonic endoderm  
21 (XEN) (Shimosato et al., 2007), raising the possibility that the Nanog-Gata6  
22 antagonism that segregates EPI and PrE during mouse pre-implantation development  
23 maintains lineage separation in mESCs (Graf and Enver, 2009). Nanog is, however,  
24 dispensable for post-implantation development and mESC self-renewal (Chambers et  
25 al., 2007), suggesting operation of additional mechanisms that govern fate restriction  
26 (Holmberg and Perlmann, 2012), such as the activation of TFs that function  
27 redundantly with Nanog.

28 Different mESC states are stabilized by specific extrinsic signaling conditions (Smith,  
29 2017): A naïve pluripotent groundstate in the presence of 2 inhibitors and leukemia  
30 inhibitory factor (LIF) (2iL), and a metastable pluripotent state by foetal calf serum  
31 (S) and LIF (SL). Although both states are interconvertible, mESCs in SL  
32 heterogeneously express differentiation and pluripotency markers (Smith, 2017),

inefficiently differentiate into Epiblast-like cells (EpiLCs) *in vitro* (Hayashi et al., 2011), and are thought to recapitulate a developmentally more advanced state than mESCs in naïve 2iL conditions (Gonzalez et al., 2016; Schröter et al., 2015).

Here, we exploit the transition from the naïve to the metastable mESC state (Schröter et al., 2015) to define the mechanism of lineage restriction. We identify inhibitors of PrE transdifferentiation, define their interaction with pluripotent TFs, and describe how they enact competition between PrE and EPI fate. Our findings reveal that silencing of a Gata6 enhancer by transcriptional repressors antagonizes lineage plasticity of the mESC gene regulatory network (GRN) and secures the pluripotent lineage.

## RESULTS

### Hdac3 inhibits transdifferentiation of mESCs into PrE

While working on a putative Hdac3 interactor, we genetically deleted *Hdac3* in naïve TNG-A mESCs that express GFP under the control of the endogenous *Nanog* locus (Chambers et al., 2007) (**Figure S1A**). Compared to *wildtype* (WT) controls, *Hdac3*<sup>-/-</sup> cells expressed higher levels of the Nanog reporter in 2iL, but rapidly downregulated Nanog when converted to SL (**Figure 1A**) and were lost upon further passaging, indicating undue differentiation specifically in metastable conditions.

To determine the transcriptional changes underlying this phenotype, we performed RNA sequencing (RNAseq) of WT and *Hdac3*<sup>-/-</sup> cells in the naïve ESC state, and after 1 and 2 days (d) in SL and epiblast-like cell (EpiLC) (Hayashi et al., 2011) differentiation conditions (**Table S2**). Contrasting these results with existing datasets from the early embryo (Boroviak et al., 2015; Mohammed et al., 2017) using pairwise correlation and principal component (PC) analysis (**Figure 1B,C, S1B,C**) revealed that mESCs in 2iL are most similar to the embryonic day E3.5 ICM (Gonzalez et al., 2016). It further showed that SL and EpiLC conditions drive naïve WT mESCs into cell states resembling the embryonic E4.5 - E6.5 pre- and post-implantation EPI. In contrast, differentiating *Hdac3*<sup>-/-</sup> cells, in particular in SL, were transcriptionally most similar to primitive and visceral (VE), but not definitive endoderm (Anderson et al., 2017) (**Figure S1D**). Clustering analysis identified a class of genes (cluster 5) that was selectively induced in differentiating mutant cells and the E4.5 PrE *in vivo* (**Figure 1D**). Cluster 5 is enriched for endoderm regulators by gene ontology analysis

and includes the TFs Gata4, Gata6, and Sox17 that are required for PrE development *in vivo* (Hermitte and Chazaud, 2014) (**Figure S1E,F**). Absence of *Hdac3* therefore causes PrE transdifferentiation upon transition from naïve to metastable conditions.

Naïve mESCs engineered to express catalytically inactive Hdac3 (*Hdac3*<sup>Y118F/Y118F</sup>), or depleted of *Ncor1* and *Ncor2* by siRNA transfection and compound knockout (Sun et al., 2013) similarly induced PrE markers after SL conversion (**Figure S1G,H, Table S1**), demonstrating that Hdac3 acts as a deacetylase and in Ncor1/2 nuclear corepressor complexes.

### **Hdac3 shields mESCs from PrE conversion in response to extrinsic developmental signals**

The induction of post-implantation markers in *Hdac3*<sup>-/-</sup> cells was unperturbed (**Figure S1I**), suggesting simultaneous activation of embryonic and extra-embryonic gene expression. To test if this is due to population heterogeneity, we deleted *Hdac3* in a TNG-A-derived ESC line (G6C18) that in addition to *Nanog* reports endogenous transcription of *Gata6* (**Figure S1J**). After 3d in the presence of SL, around 10% of mutant cells were positive for the *Gata6* reporter (**Figure 1E,F, S1K**). Addition of minimal amounts (1nM) of retinoic acid (SL<sup>RA</sup>) increased this fraction to more than 30%, which required activation of LIF and FGF but not of BMP signaling. Transdifferentiation of *Hdac3*<sup>-/-</sup> cells is therefore driven by the same pathways that control progression of ICM into PrE *in vivo* (Hermitte and Chazaud, 2014; Morgani and Brickman, 2015).

To explore the ability of *Hdac3* mutants to mimic additional aspects of ICM development, we turned to a 3D culture system. After 3d in SL<sup>RA</sup>, single *Hdac3*<sup>-/-</sup> but not *WT* cells formed spatially organized spheroids (**Figure 1G, S1L**): *Nanog* reporter-positive cells were enriched in the inside, while *Gata6* reporter-positive cells co-expressing Sox17 and showing polarized distribution of the apical PrE marker Dab2 (Hermitte and Chazaud, 2014) were on the outside. This resembles formation of the polarized epithelial PrE cell layer on the surface of the epiblast in E4.5 embryos and indicates spatial lineage segregation in *Hdac3*<sup>-/-</sup> spheroids.

### **Dax1 inhibits mESCs transdifferentiation into PrE similarly to Hdac3**

To gain a more complete understanding of lineage restriction, we set out to determine the relationship of Hdac3 with two previously described inhibitors of PrE

transdifferentiation, *Dax1* and *Prdm14* (Khalfallah et al., 2009; Ma et al., 2010; Zhang et al., 2014). Naïve and SL<sup>RA</sup>-exposed *Prdm14*<sup>-/-</sup> cells generated in the G6C18 background were indistinguishable from *WT* controls. In contrast to *Prdm14*<sup>-/-</sup> and similar to *Hdac3*<sup>-/-</sup> cells, *Dax1* mutants showed increased *Nanog* reporter levels in 2iL, and approximately 30% of the cells expressed the *Gata6* reporter after 3d in SL<sup>RA</sup> (**Figure 2A, S2A-C, Table S1**).

RNAseq of *Dax1* mutants revealed that transcriptional changes in general correlated with those in *Hdac3* mutants in both, 2iL (R=0.55) and during differentiation (R=0.58) (**Figure 2B, Table S3**). Cluster 5 genes were partially deregulated in 2iLIF in both *Dax1*<sup>-/-</sup> and *Hdac3*<sup>-/-</sup> mESCs, indicating PrE-priming (**Figure 2B**). Also, the signaling pathway dependencies for PrE transdifferentiation were the same for *Dax1*<sup>-/-</sup> and *Hdac3*<sup>-/-</sup> cells (**Figure 1F, 2C, S2D**). In *Dax1* mutant 3D aggregates, however, peripheral enrichment of *Gata6* reporter-positive PrE cells was perturbed when compared to *Hdac3* mutants (**Figure 2D,E, S2E, Table S3**), suggesting that *Dax1* has additional roles in spheroid self-organization.

We note that *Dax1* knockout and knockdown mESCs in SL have been described before, reporting apart from the induction of PrE markers (Fujii et al., 2015; Niakan et al., 2006; Zhang et al., 2014) lack of viability (Yu et al., 1998), loss of pluripotency (Khalfallah et al., 2009; Niakan et al., 2006), induction of 2-cell stage specific genes (Fujii et al., 2015), and multi-lineage differentiation (Khalfallah et al., 2009). We speculate that these discrepancies arise because absence of *Dax1* in metastable, but not naïve, conditions destabilizes pluripotency and results in or exacerbates culture heterogeneity.

## **Hdac3 and Dax1 independently restrict PrE fate and antagonize Nr5a2 and Esrrb**

PrE conversion upon loss of *Dax1* and *Hdac3* was qualitatively and quantitatively highly similar, suggesting that *Dax1* and *Hdac3* may act together, potentially in a protein complex. However, affinity-purification coupled to label-free quantitation (**Figure 3A, Table S3**) revealed that *Hdac3* co-immunoprecipitated subunits of the Ncor1/Ncor2 complexes (Gps2, NCor1, NCor2, Tbl1x, Tbl1xr1), but not *Dax1*. *Vice versa*, *Dax1* formed a complex with the nuclear receptors Nr5a2 and Esrrb, but not with *Hdac3*. To test if *Hdac3* and *Dax1* mechanisms of PrE repression were truly

independent, we analyzed their genetic interaction by generating compound knock-out cell lines (**Figure S2B, Table S1**). After 3d in SL<sup>RA</sup>, PrE marker levels in *Dax1*<sup>-/-</sup>; *Hdac3*<sup>-/-</sup> double mutants were elevated two- to threefold compared to single mutants (**Figure 3B**). Notably, this additivity was due to a doubling of the fraction of cells expressing the *Gata6* reporter, reaching to more than 70% (**Figure 3C, S3A**). *Dax1* and *Hdac3* therefore act in parallel pathways that threshold the probability of single cells to exit pluripotency and activate the PrE program.

Since *Nr5a2* and *Esrrb* have been implicated in modulating PrE gene expression (McDonald et al., 2014; Uranishi et al., 2016), we decided to investigate their role in more detail. Consistent with the specific binding to *Dax1*, deletion of *Nr5a2* completely reverted upregulation of the *Nanog* and *Gata6* reporters and of PrE marker genes in *Dax1*<sup>-/-</sup> cells, but only modestly in *Hdac3*<sup>-/-</sup> cells (**Figure 3B,C, S2B, S3A,B, Table S1**). RNAseq in 2iL and SL<sup>RA</sup> similarly showed that the majority of transcriptional changes induced by absence of *Dax1*, including those of cluster 5 genes, were reverted by co-deletion of *Nr5a2*, while loss of *Nr5a2* on its own caused comparatively minor defects (**Figure 3D, S3C,D, Table S3**). Therefore, *Nr5a2* is the key interactor mediating *Dax1* function. In contrast to *Nr5a2*, deletion of *Esrrb* partially suppressed the upregulation of the *Gata6* reporter and of PrE marker genes in both, *Dax1* and *Hdac3* mutants (**Figure 3B,C, S2B, S3A**), suggesting that *Esrrb* has non-specific functions downstream of *Dax1* and *Hdac3*. In summary, *Dax1*/*Nr5a2* and *Hdac3*/*Ncor1*/*Ncor2* form biochemically and genetically distinct complexes that antagonize PrE fate in mESCs.

## ***Gata6* enh<sup>-45</sup> is a mechanistic target of *Hdac3* and required for transdifferentiation**

The transcriptional and phenotypic similarity of *Hdac3* and *Dax1* mutants indicates that they are independently acting components of the same regulatory network. To identify the mechanism of convergence, we exploited published *Hdac3*, *Dax1* and *Nr5a2* chromatin immunoprecipitation sequencing (ChIPseq) data (Atlasi et al., 2019; Beck et al., 2014; Żylicz et al., 2018), and determined cis-regulatory elements (CRE) activity changes by performing assays for transposase-accessible chromatin sequencing (ATACseq), and H3K27ac and H4K5ac ChIPseq in *Dax1*<sup>-/-</sup> and *Hdac3*<sup>-/-</sup> cells, respectively. Focusing on 29'969 regions bound by either *Hdac3* or *Dax1* (**Table S4**) we found that *Dax1* occupancy was more similar to that of *Nr5a2*

1 (R=0.58) than of Hdac3 (R=0.11) (**Figure S4A**), which is consistent with distinct  
 2 biochemical complexes. CRE activity changes in differentiating *Dax1* and *Hdac3*  
 3 mutants, in contrast, correlated (**Figure S4B**), and clustering analysis (**Figure 4A**)  
 4 identified two clusters of loci that were enriched for Hdac3 and Dax1 binding, and  
 5 linked to CRE repression (cluster 2, decrease of accessibility, H3K27ac and H4K5ac)  
 6 and activation (cluster 4, increase of accessibility, H3K27ac and H4K5ac). Notably,  
 7 *Nr5a2* deletion reverted CRE activity changes in *Dax1* mutants (**Figure 4A**).  
 8 Dax1/Nr5a2 and Hdac3 therefore bind to and regulate shared CREs.

9 To determine the molecular features that are associated with co-regulated CREs, we  
 10 scanned underlying TF motifs (**Figure S4C,D, Table S4**). This revealed enrichment  
 11 of the GATA and OCT4-SOX2 consensus motifs at CREs that were activated and  
 12 inactivated in mutant cells, respectively. Comparison with ChIPseq confirmed binding  
 13 of Oct4, Sox2 and Nanog (Marson et al., 2008) to CREs that are ectopically silenced  
 14 in naïve and differentiating *Hdac3* and *Dax1* mutants, and binding of Gata6  
 15 (Wamaitha et al., 2015) to CREs that are ectopically activated in differentiating but  
 16 not naïve mutant cells data (**Figure 4A, S4E**). Hdac3 and Dax1 therefore converge on  
 17 regulating CREs co-bound by core pluripotency TFs, and on repressing PrE-specific  
 18 CREs that are occupied by Gata6 and activated during transdifferentiation.

19 As Gata6 overexpression is sufficient to convert mESCs into extraembryonic  
 20 endoderm stem (XEN) cells *in vitro* (Shimosato et al., 2007), we hypothesized a  
 21 causative role for Gata6 in the PrE transdifferentiation of *Dax1* and *Hdac3* mutants.  
 22 This possibility is supported by specific upregulation of Gata6 in naïve single and  
 23 compound mutants (**Figure S4F**) and by the fact that a region 45 kb upstream of the  
 24 *Gata6* gene (*enh*<sup>-45</sup>) was amongst the most strongly activated CREs in mutant cells  
 25 both during differentiation and in 2iL (**Figure 4B,C, S4G**), and decorated with  
 26 H3K27ac in the E6.5 VE but not EPI (Xiang et al., 2019) (**Figure 4C**). *enh*<sup>-45</sup> is bound  
 27 by Hdac3 and Gata6, but not Dax1, raising the possibility that it is a relevant direct  
 28 Hdac3 target. To test this, we deleted the 500bp region that is occupied by Hdac3  
 29 ( $\Delta$ *enh*<sup>-45</sup>) (**Table S1**), which fully suppressed activation of the *Gata6* reporter and  
 30 reduced PrE marker expression in  $\Delta$ *enh*<sup>-45</sup>;*Hdac3*<sup>-/-</sup> and  $\Delta$ *enh*<sup>-45</sup>;*Dax1*<sup>-/-</sup> cells to levels  
 31 observed in the respective single mutants upon SL<sup>RA</sup> transition (**Figure 4D, S4H,I**).  
 32 Reversion of *Gata6* transcription was more efficient in *Hdac3* than *Dax1* compound  
 33  $\Delta$ *enh*<sup>-45</sup> mutants, but absent in naïve mESCs (**Figure 4D**). Surprisingly, removal of



enh<sup>-45</sup> triggered downregulation of the *Nanog* reporter and of the naïve pluripotency markers *Klf4*, *Esrrb*, *Tfcp2l1* and *Prdm14* in *Hdac3* but not *Dax1* mutants in SL (Figure 4D, S4H). This indicates additional roles of Hdac3 in the stability of the pluripotent GRN that are masked by enh<sup>-45</sup> regulation and not shared with Dax1. We conclude that direct and indirect suppression of enh<sup>-45</sup> by Hdac3 and Dax1, respectively, is crucial for repressing *Gata6* transcription and, consequentially, PrE transdifferentiation upon SL transition.

## DISCUSSION

Here we show that the transcriptional repressors Hdac3 and Dax1 set the lineage barrier that prevents mESCs from transdifferentiating into PrE upon exposure to developmental signals. In naïve conditions, *Hdac3* and *Dax1* mutant mESCs are stable, allowing us to exploit the transition from 2iL to SL conditions as paradigm for PrE conversion (Schröter et al., 2015). Adaption to 3D results in formation of *Hdac3* mutant spheroids in which pluripotent and PrE lineages are spatially segregated. In *Dax1* mutants, segregation is perturbed. This setup is therefore amenable to dissecting the molecular requirement underlying PrE differentiation and sorting.

Hdac3 and Dax1 cooperatively inhibit PrE transdifferentiation. Notably, conversion of naïve *Hdac3*<sup>-/-</sup> and *Dax1*<sup>-/-</sup> mESCs occurs without overexpression of PrE-specifying TFs, such as *Sox17* (McDonald et al., 2014), *Gata4* or *Gata6* (Fujikura et al., 2002), or use of selective culture conditions that confer expanded potential (Sozen et al., 2019; Yang et al., 2017) or select the survival and/or proliferation of PrE cell types (Anderson et al., 2017), but upon transition to SL, which supports both, EPI and PrE fates. The pluripotent gene regulatory network is therefore permissive to lineage switching in the absence of ectopic reprogramming TFs or developmental reversion.

Dax1 and Hdac3 act in independent protein complexes that converge on repressing *Gata6*. We show that lineage conversion in the absence of *Dax1* and *Hdac3* requires activation of a single *Gata6* enhancer, enh<sup>-45</sup>, that is directly targeted by Hdac3. Recruitment of epigenetic regulators, such as Sin3a (Mall et al., 2017) or Groucho (Kutejova et al., 2016; Muhr et al., 2001) is associated with the repression of alternate lineages, but this study – to our knowledge for the first time – links the physical binding and enzymatic activity of a transcriptional repressor, Hdac3, at a single CRE with suppression of transdifferentiation. enh<sup>-45</sup> is occupied by *Gata6* during



transdifferentiation, indicating a positive feedforward loop that is characteristic of cell type-specifying TFs (Crews and Pearson, 2009) and can stabilize cell identity (Leyva-Díaz and Hobert, 2019). Our findings therefore suggest that the continuous silencing of autoregulated enhancers of competing master TFs is an endogenous mechanism for maintaining lineage restriction.

We find that *Hdac3* and *Dax1* repress transdifferentiation by modulating the pluripotent GRN network, specifically *Nr5a2* and *Esrrb*, which are TFs that regulate mESC self-renewal (Festuccia et al., 2020) and iPSC reprogramming (Feng et al., 2009; Guo and Smith, 2010). *Nr5a2* is a well-established interactor of *Dax1* (Sablin et al., 2008). Co-deletion of *Nr5a2* rescues PrE conversion and most transcriptional and epigenetic alterations in *Dax1* but not *Hdac3* mutants. Therefore, *Nr5a2* and *Dax1* operate as a functional unit where *Nr5a2* is the transcriptional activator and *Dax1* the repressor. The direct targets of *Nr5a2* driving PrE differentiation remain unclear, since *Nr5a2* and *Dax1* do not bind to *enh*<sup>-45</sup>. Lack of *Esrrb*, in contrast, partially blocks PrE transdifferentiation of *Dax1* and *Hdac3* mutants. Although *Esrrb* interacts and shares genomic binding with *Dax1* and *Nr5a2* (**Figure 3A, 4A, S4A**), incomplete phenotypic rescue and lack of specificity for *Dax1* suggests a role of *Esrrb* in PrE conversion beyond the *Nr5a2*/*Dax1* axis, e.g. by binding (**Figure 4C**) and activating the *Gata6* promoter (Uranishi et al., 2016) or, indirectly, by stabilizing pluripotency (Martello et al., 2012). Taken together, we propose that *Nr5a2* and *Esrrb* support both the EPI and the PrE fate within a coherent TF network that integrates extrinsic signals with transcriptional repressors to maintain lineage choice.

In summary, we uncover how the pluripotent lineage is safeguarded in mESCs. We find that both, *Gata6* and *Nanog*, are induced in naïve *Hdac3*<sup>-/-</sup> and *Dax1*<sup>-/-</sup> mESCs, and that *enh*<sup>-45</sup> is not occupied by *Nanog* (**Figure 4C**). This is inconsistent with mutual inhibition of *Gata6* by *Nanog* and indicates that maintenance of the EPI-PrE segregation is independent of direct TF antagonism. We, instead, suggest operation of a GRN that (a) epigenetically silences *Gata6*, the master TF of the competing PrE fate, and (b) balances the activities of lineage-divergent TFs, such as *Nr5a2* and *Esrrb* (**Figure S4J**). We speculate that this regulatory network is established during developmental progression of pluripotency to stabilize the initially labile lineage segregation set by antagonism between *Nanog* and *Gata6*.

# 1 REFERENCES

- 2 Anderson, K.G.V., Hamilton, W.B., Roske, F.V., Azad, A., Knudsen, T.E., Canham,  
3 M.A., Forrester, L.M., and Brickman, J.M. (2017). Insulin fine-tunes self-renewal  
4 pathways governing naive pluripotency and extra-embryonic endoderm. *Nat Cell Biol*  
5 *19*, 1164–1177.
- 6 Argelaguet, R., Clark, S.J., Mohammed, H., Stapel, L.C., Krueger, C., Kapourani, C.-  
7 A., Imaz-Rosshandler, I., Lohoff, T., Xiang, Y., Hanna, C.W., et al. (2019). Multi-  
8 omics profiling of mouse gastrulation at single-cell resolution. *Nature* *576*, 487–491.
- 9 Atlasi, Y., Megchelenbrink, W., Peng, T., Habibi, E., Joshi, O., Wang, S.-Y., Wang,  
10 C., Logie, C., Poser, I., Marks, H., et al. (2019). Epigenetic modulation of a hardwired  
11 3D chromatin landscape in two naive states of pluripotency. *Nat Cell Biol* *21*, 568–  
12 578.
- 13 Barisic, D., Stadler, M.B., Iurlaro, M., and Schübeler, D. (2019). Mammalian ISWI  
14 and SWI/SNF selectively mediate binding of distinct transcription factors. *Nature*  
15 *569*, 136–140.
- 16 Beck, S., Lee, B.-K., Rhee, C., Song, J., Woo, A.J., and Kim, J. (2014). CpG island-  
17 mediated global gene regulatory modes in mouse embryonic stem cells. *Nat Commun*  
18 *5*, 5490.
- 19 Beddington, R.S., and Robertson, E.J. (1989). An assessment of the developmental  
20 potential of embryonic stem cells in the midgestation mouse embryo. *Dev Camb Engl*  
21 *105*, 733–737.
- 22 Betschinger, J., Nichols, J., Dietmann, S., Corrin, P.D., Paddison, P.J., and Smith, A.  
23 (2013). Exit from pluripotency is gated by intracellular redistribution of the bHLH  
24 transcription factor Tfe3. *Cell* *153*, 335–347.
- 25 Boroviak, T., Loos, R., Lombard, P., Okahara, J., Behr, R., Sasaki, E., Nichols, J.,  
26 Smith, A., and Bertone, P. (2015). Lineage-Specific Profiling Delineates the  
27 Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis.  
28 *Dev Cell* *35*, 366–382.
- 29 Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A  
30 Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol*  
31 *Biology Ed Frederick M Ausubel Et Al* *109*, 21.29.1-9.
- 32 Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J.,  
33 Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and  
34 mediates germline development. *Nature* *450*, 1230–1234.
- 35 Chronis, C., Fiziev, P., Papp, B., Butz, S., Bonora, G., Sabri, S., Ernst, J., and Plath,  
36 K. (2017). Cooperative Binding of Transcription Factors Orchestrates  
37 Reprogramming. *Cell* *168*, 442-459.e20.

- 1 Crews, S.T., and Pearson, J.C. (2009). Transcriptional autoregulation in development.  
2 *Curr Biol* 19, R241–R246.
- 3 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P.,  
4 Chaisson, M., and Gingeras, T.R. (2012). STAR: ultrafast universal RNA-seq aligner.  
5 *Bioinform Oxf Engl* 29, 15–21.
- 6 Feng, B., Jiang, J., Kraus, P., Ng, J.-H., Heng, J.-C.D., Chan, Y.-S., Yaw, L.-P.,  
7 Zhang, W., Loh, Y.-H., Han, J., et al. (2009). Reprogramming of fibroblasts into  
8 induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat Cell Biol* 11,  
9 197–203.
- 10 Festuccia, N., Owens, N., Chervova, A., Dubois, A., and Navarro, P. (2020). The  
11 combined action of Esrrb and Nr5a2 is essential for naïve pluripotency. *Biorxiv*  
12 2020.06.05.134999.
- 13 Fujii, S., Nishikawa-Torikai, S., Futatsugi, Y., Toyooka, Y., Yamane, M., Ohtsuka,  
14 S., and Niwa, H. (2015). Nr0b1 is a negative regulator of Zscan4c in mouse  
15 embryonic stem cells. *Sci Rep-Uk* 5, 9146.
- 16 Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki,  
17 J., and Niwa, H. (2002). Differentiation of embryonic stem cells is induced by GATA  
18 factors. *Gene Dev* 16, 784–789.
- 19 Gaidatzis, D., Lerch, A., Hahne, F., and Stadler, M.B. (2014). QuasR: quantification  
20 and annotation of short reads in R. *Bioinform Oxf Engl* 31, 1130–1132.
- 21 Gonzalez, J.M., Morgani, S.M., Bone, R.A., Bonderup, K., Abelchian, S.,  
22 Brakebusch, C., and Brickman, J.M. (2016). Embryonic Stem Cell Culture Conditions  
23 Support Distinct States Associated with Different Developmental Stages and Potency.  
24 *Stem Cell Rep* 7, 177–191.
- 25 Graf, T., and Enver, T. (2009). Forcing cells to change lineages. *Nature* 462, 587–594.
- 26 Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and  
27 correlations in multidimensional genomic data. *Bioinform Oxf Engl* 32, 2847–2849.
- 28 Guo, G., and Smith, A. (2010). A genome-wide screen in EpiSCs identifies Nr5a  
29 nuclear receptors as potent inducers of ground state pluripotency. *Development* 137,  
30 3185–3192.
- 31 Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011).  
32 Reconstitution of the Mouse Germ Cell Specification Pathway in Culture by  
33 Pluripotent Stem Cells. *Cell* 146, 519–532.
- 34 Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X.,  
35 Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-  
36 determining transcription factors prime cis-regulatory elements required for  
37 macrophage and B cell identities. *Mol Cell* 38, 576–589.

- 1 Hermitte, S., and Chazaud, C. (2014). Primitive endoderm differentiation: from  
2 specification to epithelium formation. *Philosophical Transactions Royal Soc Lond Ser*  
3 *B Biological Sci* 369, 20130537.
- 4 Hoffman, J.A., Wu, C.-I., and Merrill, B.J. (2013). Tcf7l1 prepares epiblast cells in  
5 the gastrulating mouse embryo for lineage specification. *Development* 140, 1665–  
6 1675.
- 7 Holmberg, J., and Perlmann, T. (2012). Maintaining differentiated cellular identity.  
8 *Nat Rev Genet* 13, 429–439.
- 9 Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2008). Bioinformatics enrichment  
10 tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic*  
11 *Acids Res* 37, 1–13.
- 12 Kang, M., Piliszek, A., Artus, J., and Hadjantonakis, A.-K. (2013). FGF4 is required  
13 for lineage restriction and salt-and-pepper distribution of primitive endoderm factors  
14 but not their initial expression in the mouse. *Development* 140, 267–279.
- 15 Khalfallah, O., Rouleau, M., Barbry, P., Bardoni, B., and Lalli, E. (2009). Dax-1  
16 Knockdown in Mouse Embryonic Stem Cells Induces Loss of Pluripotency and  
17 Multilineage Differentiation. *Stem Cells* 27, 1529–1537.
- 18 Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M.D.C., and Yusa, K. (2013).  
19 Genome-wide recessive genetic screening in mammalian cells with a lentiviral  
20 CRISPR-guide RNA library. *Nat Biotechnol* 32, 267–273.
- 21 Kutejova, E., Sasai, N., Shah, A., Gouti, M., and Briscoe, J. (2016). Neural  
22 Progenitors Adopt Specific Identities by Directly Repressing All Alternative  
23 Progenitor Transcriptional Programs. *Dev Cell* 36, 639–653.
- 24 Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R.,  
25 Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating  
26 genomic ranges. *Plos Comput Biol* 9, e1003118.
- 27 Leyva-Díaz, E., and Hobert, O. (2019). Transcription factor autoregulation is required  
28 for acquisition and maintenance of neuronal identity. *Development* 146, dev177378.
- 29 Lun, A.T.L., and Smyth, G.K. (2015a). csaw: a Bioconductor package for differential  
30 binding analysis of ChIP-seq data using sliding windows. *Nucleic Acids Res* 44, e45.
- 31 Lun, A.T.L., and Smyth, G.K. (2015b). From reads to regions: a Bioconductor  
32 workflow to detect differential binding in ChIP-seq data. *F1000research* 4, 1080.
- 33 Luo, J., Sladek, R., Bader, J.-A., Matthysen, A., Rossant, J., and Giguère, V. (1997).  
34 Placental abnormalities in mouse embryos lacking the orphan nuclear receptor ERR-  
35  $\beta$ . *Nature* 388, 778–782.

- 1 Ma, Z., Swigut, T., Valouev, A., Rada-Iglesias, A., and Wysocka, J. (2010).  
2 Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering  
3 extraembryonic endoderm fates. *Nat Struct Mol Biol* 18, 120–127.
- 4 Mall, M., Kareta, M.S., Chanda, S., Ahlenius, H., Perotti, N., Zhou, B., Grieder, S.D.,  
5 Ge, X., Drake, S., Ang, C.E., et al. (2017). Myt1l safeguards neuronal identity by  
6 actively repressing many non-neuronal fates. *Nature* 544, 245–249.
- 7 Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S.,  
8 Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., et al. (2008). Connecting  
9 microRNA genes to the core transcriptional regulatory circuitry of embryonic stem  
10 cells. *Cell* 134, 521–533.
- 11 Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S.,  
12 Göttgens, B., Niwa, H., and Smith, A. (2012). Esrrb is a pivotal target of the  
13 Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. *Cell Stem Cell* 11, 491–  
14 504.
- 15 Mathelier, A., Fornes, O., Arenillas, D.J., Chen, C.-Y., Denay, G., Lee, J., Shi, W.,  
16 Shyr, C., Tan, G., Worsley-Hunt, R., et al. (2015). JASPAR 2016: a major expansion  
17 and update of the open-access database of transcription factor binding profiles.  
18 *Nucleic Acids Res* 44, D110–5.
- 19 McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis  
20 of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic  
21 Acids Res* 40, 4288–4297.
- 22 McDonald, A.C.H., Biechele, S., Rossant, J., and Stanford, W.L. (2014). Sox17-  
23 Mediated XEN Cell Conversion Identifies Dynamic Networks Controlling Cell-Fate  
24 Decisions in Embryo-Derived Stem Cells. *Cell Reports* 9, 780–793.
- 25 Mohammed, H., Hernando-Herraez, I., Savino, A., Scialdone, A., Macaulay, I.,  
26 Mulas, C., Chandra, T., Voet, T., Dean, W., Nichols, J., et al. (2017). Single-Cell  
27 Landscape of Transcriptional Heterogeneity and Cell Fate Decisions during Mouse  
28 Early Gastrulation. *Cell Reports* 20, 1215–1228.
- 29 Montgomery, R.L., Potthoff, M.J., Haberland, M., Qi, X., Matsuzaki, S., Humphries,  
30 K.M., Richardson, J.A., Bassel-Duby, R., and Olson, E.N. (2008). Maintenance of  
31 cardiac energy metabolism by histone deacetylase 3 in mice. *J Clin Invest* 118, 3588–  
32 3597.
- 33 Morgani, S.M., and Brickman, J.M. (2015). LIF supports primitive endoderm  
34 expansion during pre-implantation development. *Development* 142, 3488–3499.
- 35 Muhr, J., Andersson, E., Persson, M., Jessell, T.M., and Ericson, J. (2001). Groucho-  
36 Mediated Transcriptional Repression Establishes Progenitor Cell Pattern and  
37 Neuronal Fate in the Ventral Neural Tube. *Cell* 104, 861–873.

- 1 Niakan, K.K., Davis, E.C., Clipsham, R.C., Jiang, M., Dehart, D.B., Sulik, K.K., and  
2 McCabe, E.R.B. (2006). Novel role for the orphan nuclear receptor Dax1 in  
3 embryogenesis, different from steroidogenesis. *Mol Genet Metab* 88, 261–271.
- 4 Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K.  
5 (2015). limma powers differential expression analyses for RNA-sequencing and  
6 microarray studies. *Nucleic Acids Res* 43, e47.
- 7 Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2009). edgeR: a Bioconductor  
8 package for differential expression analysis of digital gene expression data. *Bioinform*  
9 *Oxf Engl* 26, 139–140.
- 10 Sablin, E.P., Woods, A., Krylova, I.N., Hwang, P., Ingraham, H.A., and Fletterick,  
11 R.J. (2008). The structure of corepressor Dax-1 bound to its target nuclear receptor  
12 LRH-1. *Proc National Acad Sci* 105, 18390–18395.
- 13 Schröter, C., Rué, P., Mackenzie, J.P., and Arias, A.M. (2015). FGF/MAPK signaling  
14 sets the switching threshold of a bistable circuit controlling cell fate decisions in  
15 embryonic stem cells. *Development* 142, 4205–4216.
- 16 Shimosato, D., Shiki, M., and Niwa, H. (2007). Extra-embryonic endoderm cells  
17 derived from ES cells induced by GATA Factors acquire the character of XEN cells.  
18 *Bmc Dev Biol* 7, 80.
- 19 Simon, C.S., Hadjantonakis, A., and Schröter, C. (2018). Making lineage decisions  
20 with biological noise: Lessons from the early mouse embryo. *Wiley Interdiscip Rev*  
21 *Dev Biology* 7, e319.
- 22 Smith, A. (2017). Formative pluripotency: the executive phase in a developmental  
23 continuum. *Development* 144, 365–373.
- 24 Sozen, B., Cox, A.L., Jonghe, J.D., Bao, M., Hollfelder, F., Glover, D.M., and  
25 Zernicka-Goetz, M. (2019). Self-Organization of Mouse Stem Cells into an Extended  
26 Potential Blastoid. *Dev Cell* 51, 698–712.e8.
- 27 Sun, Z., Feng, D., Fang, B., Mullican, S.E., You, S.-H., Lim, H.-W., Everett, L.J.,  
28 Nabel, C.S., Li, Y., Selvakumaran, V., et al. (2013). Deacetylase-independent  
29 function of HDAC3 in transcription and metabolism requires nuclear receptor  
30 corepressor. *Mol Cell* 52, 769–782.
- 31 Trendafilov, N.T. (2010). Stepwise estimation of common principal components.  
32 *Comput Stat Data An* 54, 3446–3457.
- 33 Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M.,  
34 and Cox, J. (2016). The Perseus computational platform for comprehensive analysis  
35 of (prote)omics data. *Nat Methods* 13, 731–740.
- 36 Uranishi, K., Akagi, T., Koide, H., and Yokota, T. (2016). Esrrb directly binds to  
37 Gata6 promoter and regulates its expression with Dax1 and Ncoa3. *Biochem Bioph*  
38 *Res Co* 478, 1720–1725.

- 1 Villegas, F., Lehalle, D., Mayer, D., Rittirsch, M., Stadler, M.B., Zinner, M., Olivieri,  
2 D., Vabres, P., Duplomb-Jego, L., Bont, E.S.J.M.D., et al. (2019). Lysosomal  
3 Signaling Licenses Embryonic Stem Cell Differentiation via Inactivation of Tfe3. *Cell*  
4 *Stem Cell* 24, 257-270.e8.
- 5 Wamaitha, S.E., Valle, I. del, Cho, L.T.Y., Wei, Y., Fogarty, N.M.E., Blakeley, P.,  
6 Sherwood, R.I., Ji, H., and Niakan, K.K. (2015). Gata6 potently initiates  
7 reprogramming of pluripotent and differentiated cells to extraembryonic endoderm stem  
8 cells. *Gene Dev* 29, 1239–1255.
- 9 Xiang, Y., Zhang, Y., Xu, Q., Zhou, C., Liu, B., Du, Z., Zhang, K., Zhang, B., Wang,  
10 X., Gayen, S., et al. (2019). Epigenomic analysis of gastrulation identifies a unique  
11 chromatin state for primed pluripotency. *Nat Genet* 52, 95–105.
- 12 Yang, J., Ryan, D.J., Wang, W., Tsang, J.C.-H., Lan, G., Masaki, H., Gao, X.,  
13 Antunes, L., Yu, Y., Zhu, Z., et al. (2017). Establishment of mouse expanded  
14 potential stem cells. *Nature* 550, 393–397.
- 15 Yu, R.N., Ito, M., Saunders, T.L., Camper, S.A., and Jameson, J.L. (1998). Role of  
16 Ahch in gonadal development and gametogenesis. *Nat Genet* 20, 353–357.
- 17 Zhang, J., Liu, G., Ruan, Y., Wang, J., Zhao, K., Wan, Y., Liu, B., Zheng, H., Peng,  
18 T., Wu, W., et al. (2014). Dax1 and Nanog act in parallel to stabilize mouse  
19 embryonic stem cells and induced pluripotency. *Nat Commun* 5, 5042.
- 20 Żylicz, J.J., Bousard, A., Žumer, K., Dossin, F., Mohammad, E., Rocha, S.T. da,  
21 Schwalb, B., Syx, L., Dingli, F., Loew, D., et al. (2018). The Implication of Early  
22 Chromatin Changes in X Chromosome Inactivation. *Cell* 176, 182-197.e23.

23



# 1    **ACKNOWLEDGMENTS**

2    We thank H.Kohler (FMI) for cell sorting, Karoline Guja for technical support, V.  
3    Iesmantavicius (FMI) for help with computational analysis of mass spectrometry data,  
4    Min Jia and Jeff Chao (FMI) for providing LIF, and Austin Smith (Wellcome-MRC  
5    Cambridge Stem Cell Institute) for providing TNG-A cells. We are grateful to F.  
6    Mohn, A.H.F.M. Peters, D. Schuebeler (FMI) and S. Stricker (University Munich) for  
7    comments on the manuscript. D.O. was supported by EMBO (ALTF 1632-2014) and  
8    Marie Curie Actions (LTFCOFUND2013, GA-2013-609409). Research in the lab of  
9    J.B. is supported by the Novartis Research Foundation.

# 10    **AUTHOR CONTRIBUTIONS**

11    D.O. conceived the study and performed experiments, P.P. and J.B. bioinformatical  
12    analysis, M.R. RT-qPCRs, D.H. mass spectrometry, and I.L. spheroid image analysis.  
13    S.S. supervised sequencing library generation and sequencing, and M.B.S.  
14    bioinformatical analysis. D.O. and J.B. wrote the paper.

# 15    **COMPETING INTERESTS**

16    The authors declare no conflict of interest.

# 1 MAIN FIGURES

## 2 **Figure 1: Naïve *Hdac3*<sup>-/-</sup> mESCs transdifferentiate to PrE in response to FGF,** 3 **LIF, and RA.**

4 (A) *Nanog*>GFP intensity of *WT* and *Hdac3*<sup>-/-</sup> TNG-A cells in 2iL and at d2 and d5 of  
5 SL exposure.

6 (B,C) Pairwise Pearson correlations (B) and common PC analysis (C) of *in vitro*  
7 TNG-A and *in vivo* embryo (Boroviak et al., 2015) RNAseq samples. Morula  
8 (MOR).

9 (D) k-means clustering of gene expression changes relative to naïve *WT* TNG-A  
10 ESCs. (E) *Nanog*>GFP and *Gata6*::mCherry fluorescence intensities of *WT* and  
11 *Hdac3*<sup>-/-</sup> cells in indicated conditions.

12 (F) Percentage of *Gata6*::mCherry positive cells after 3d of exposure in indicated  
13 conditions. Jak(i) blocks LIF, LDN19 BMP4 and PD03 FGF signaling. Average and  
14 SD of three independent clones.

15 (G) Immunofluorescence and reporter expression in spheroids derived from single  
16 *WT* and *Hdac3*<sup>-/-</sup> naïve mESCs. DNA counterstain is Hoechst. The embryonic  
17 *Gata6*::mRFP negative part is indicated. Scalebar: 10µM.

## 18 **Figure 2: *Dax1* is required for lineage restriction and sorting.**

19 (A,C) Fraction of *Gata6*::mCherry positive cells in indicated genotypes and  
20 conditions. Average and SD of at least two independent clones.

21 (B) Scatter plots of log2 fold change (FC) gene expression changes in *Hdac3*<sup>-/-</sup> and  
22 *Dax1*<sup>-/-</sup> mutants relative to *WT* cells in naïve mESCs (top) and after 2 d of  
23 differentiation (bottom); Cluster 5 genes are colored and selected PrE TFs labeled.

24 (D,E) Quantification of spatial lineage segregation in spheroids. Representative  
25 spheroid segmentations of indicated genotypes (D). Z-score-normalized fluorescence  
26 distributions in mutants compared to *WT* spheroids (E). Positive Z-scores indicate  
27 peripheral, outside, and negative Z-scores central, inside, enrichment. Scalebar:  
28 25µM.

## 29 **Figure 3: *Hdac3* and *Nr5a2/Dax1* are genetically and biochemically distinct** 30 **pathways repressing PrE differentiation.**

1 (A) Z-scores of high-confidence interactors coIPed by Hdac3, Dax1, Nr5a2, and Esrrb  
2 (indicated in red) in naïve mESCs.

3 (B,C) Expression of PrE markers relative to *Dax1*<sup>-/-</sup> cells (B), and fraction of  
4 *Gata6*::mCherry positive cells after 3d in SL<sup>RA</sup> (C) in indicated genotypes. Average  
5 and SD of three independent clones.

6 (D) Gene expression changes in *Dax1*<sup>-/-</sup>, *Nr5a2*<sup>-/-</sup> and *Dax1*<sup>-/-</sup>;*Nr5a2*<sup>-/-</sup> cells compared  
7 to WT cells in 2iLIF (upper) and after 2d in SL<sup>RA</sup> (lower).

8 **Figure 4: Hdac3 and Dax1 co-repress Gata6-bound CREs, including the essential**  
9 **Gata6 enh<sup>-45</sup>.**

10 (A) k-means clustered heatmap showing enrichment of TFs, and changes of chromatin  
11 marks or accessibility in indicated mutants relative to WT at Hdac3- and/or Dax1-  
12 bound regions. Repressed and induced clusters 2 and 4 are marked.

13 (B) Scatterplot of log2 fold accessibility changes in d2\_SL<sup>RA</sup> *Dax1*<sup>-/-</sup> cells and of log2  
14 fold H3K27ac ChIP signal changes in d2\_SL *Hdac3*<sup>-/-</sup> cells relative to WT controls at  
15 Hdac3- and/or Dax1-bound regions. Regions colored in red are associated with Gata6.

16 (C) Genome-browser view of the Gata6 locus showing accessibility and chromatin  
17 mark changes in indicated mutants, and of ChIPseq signals of indicated TFs and  
18 H3K27ac in E6.5 VE and EPI.

19 (D) FC of PrE and naïve pluripotency markers relative to WT cells in 2iL of indicated  
20 genotypes in indicated conditions. Average and SD of at least three independent  
21 clones.

# **MATERIALS AND METHODS**

## **Mouse ESCs**

Male TNG-A mESCs where a GFP-IRES-Puromycin-N-acetyltransferase is knocked into one of the Nanog alleles (Chambers et al., 2007) are a gift from Austin Smith (Wellcome – MRC Cambridge Stem Cell Institute and Living Systems Institute, University of Exeter).

## **Cell culture**

mESCs were cultured on gelatin-coated plates in N2B27 medium (DMEM/F12 medium (Life Technologies) and Neurobasal medium (Gibco) 1:1, N2 supplement 1/200 (homemade), B-27 Serum-Free Supplement 1/100 (Gibco), 2mM L-glutamine (Gibco), 0.1mM 2-mercaptoethanol (Sigma)), 1μM PD0325901, 3μM CHIR99021 (Steward lab, Dresden), and mLIF (Smith lab, Cambridge; Chao lab, FMI). For medium switch and differentiation experiments cells were washed with PBS, detached in Accutase (Sigma), centrifuged at 300g for 3 minutes (min) in DMEM/F12 with 0.1% BSA (Gibco), resuspended in the appropriate medium, counted using Fast Read 102 counting chambers (Kova), diluted and plated. Cells were plated in Serum LIF (SL) medium (GMEM (Sigma), 10% fetal bovine serum (Sigma), 1mM sodium pyruvate (Gibco), 2mM L-glutamine (Gibco), 0.1mM non-essential amino acids (Gibco), 0.1mM 2-mercaptoethanol (Sigma), and mLIF), plus Retinoic acid (1nM, Sigma), JAK Inhibitor I (1μM, Calbiochem) and LDN193189 (100nM, Sigma), PD0325901, where specified, at 10'000 cells/cm<sup>2</sup> on gelatin-coated plates, or in EpiLC medium (N2B27 medium, 20 ng/ml activin A and 12 ng/ml bFGF (Smith lab, Cambridge), and 1% KSR (Life Technologies)) at 25'000 cells/cm<sup>2</sup> on fibronectin-coated plates. To generate spheroids, 1200 cells were plated as described above in 1.5 mL SL plus 1nM Retinoic acid in a well of an AggreWell 400 plate (Stemcell Technologies), prepared with anti-adherence rinsing solution (Stemcell Technologies) according to manufacturer's instructions, and incubated for 3 days.

TNG-A mESCs stably transfected with pPB-LR51-EF1a-bsd2ACas9 (modified from (Koike-Yusa et al., 2013)) were reverse transfected with U6>sgRNA plasmids (George Church, Addgene plasmid #41824) as follows: 3μL of Lipofectamin2000 (Life Technologies) were mixed with 250μL of OPTIMEM (Gibco) and incubated for 5 min at RT, 350ng of sgRNA plasmid diluted in 250μL of OPTIMEM were added to

the Lipofectamin2000 mix and incubated for 30 min at room temperature (RT), the transfection mix was added to freshly resuspended 200'000 cells in 2mL of medium in a well of a 6-well-plate. The next day medium was changed. 3 days after transfection, single cells were sorted in 96-well plates for clonal isolation. Parental cell line, sgRNA sequences, genotyping strategy, and sequencing results/western blots for every knock-out clone used in this study are detailed in **Table S1** and supplementary figure panels. The Hdac3Y118F knock-in cell lines were generated by reverse transfection as above of 400ng Hdac3\_gRNA7 plasmid and 10pmol of oligo Hdac3YF\_ki\_f (for clone 1034-15) or oligo Hdac3YF\_ki\_r (for clone 1035-6). The Gata6::mCherry knock-in cell line G6C18 was generated by reverse transfection as above of 300ng Gata6-C gRNA#1 plasmid and 800ng of Gata6\_3xGS\_mCherry homologous recombination template. sgRNAs, oligonucleotide sequences, genotyping strategies, and sequencing results for knock-in clones are detailed in **Table S1**.

#### **siRNA transfections**

Cells were reverse transfected in 2iLIF as follows: 2.4μL of LipofectamineRNAiMAX (Life Technologies) were mixed with 200μL of OPTIMEM (Gibco) and incubated for 5 min at RT, 2.4μL of 20μM Ncor1 siRNA mix and Ncor2 siRNA mix or AllStars Neg. Control siRNA (FlexiTube GeneSolution, Qiagen) diluted in 200μL of OPTIMEM were added to the LipofectamineRNAiMAX mix and incubated for 30 min at RT, the transfection mix was added to freshly resuspended 100'000 cells in 1.6mL of medium in a well of a 12-well-plate. The next day the cells were detached and transferred to fresh 2iLIF or Serum LIF. siRNA sequences are displayed in **Table S1**.

#### **Western blotting**

Cells were lysed in 1x RIPA buffer (50mM Tris pH7.5, 150mM NaCl, 1% IGEPAL, 0.5% Na Deoxycholate, 0.1% SDS, 2mM EDTA), protein concentration was determined by the BCA protein assay (Pierce), and 10μg of protein per sample were loaded on 10% polyacrylamide gels. Primary antibodies used: anti-HDAC3 antibody, ab7030, by abcam, 1/5000; DAX-1/NR0B1 antibody (mAb), Clone: 1DA-2F4, Active Motif, Catalog No:39983, 1/10'000; Human ERR beta/NR3B2 Antibody, PP-H6705-0, by R&D Systems, 1/10'000.

#### **Quantitative PCR (qPCR)**

1 RNA was purified with the RNeasy Mini Kit with on-column DNase digest (Qiagen).  
 2 1µg of RNA was reverse transcribed using SuperScript III Reverse Transcriptase  
 3 (LifeTechnologies) and qPCR was performed with the TaqMan Fast Universal PCR  
 4 Master Mix (Thermo Fisher) and the following TaqMan probes: GAPDH  
 5 (4352339E), Gata6 (Mm00802636\_m1), Gata4 (Mm00484689\_m1), Pdgfra  
 6 (Mm00440701\_m1), Sox7 (Mm00776876\_m1), Sox17 (Mm00488363\_m1), Oct4  
 7 (Mm03053917\_g1), Esrrb (Mm00442411\_m1), Klf4 (Mm00516104\_m1), Tfcp2l1  
 8 (Mm00470119\_m1), Prdm14 (Mm01237814\_m1).

## 9 **Co-immunoprecipitations (IPs)**

10 Endogenous Hdac3 was IP'ed from nuclear extract of G6C18 cells in 2iLIF using  
 11 anti-HDAC3 antibody (ab7030, by abcam) and rabbit normal IgGs (Santa Cruz) as  
 12 control. N-3xFLAG-3GS-Dax1, Nr5a2-3GS-3xFLAG-C, and Esrrb-3GS-3xFLAG-C  
 13 cloned in pPB-CAG-DEST-pgk-hph (Betschinger et al., 2013) were expressed in  
 14 G6C18 cells in 2iLIF and IP'ed from nuclear extracts using anti-FLAG M2 antibody  
 15 (Sigma) using un-transfected cells as control. Briefly, cells from a confluent T-75  
 16 flask were harvested and nuclei isolated by hypotonic buffer (10mM TrisHCl pH7.5,  
 17 10mM KCl, 1mM dithiothreitol (DTT), 0.5% IGEPAL) for 20 min on ice. Nuclei  
 18 were lysed by rotation for 90 min at 4°C in 1mL of lysis buffer (20mM TrisHCl  
 19 pH7.5, 100mM KCl, 1.5mM MgCl<sub>2</sub>, 1mM DTT, 10% glycerol, 1x Protease Inhibitor  
 20 Tablet (Roche), 2x Phosphatase Inhibitor Tablets (Roche), 0.5% Triton X-100, 250  
 21 units/mL Benzonase (Sigma)). Clarified lysates were incubated with 1.2µg of  
 22 antibody pre-bound to 5µL protein Dynabeads (Life Technologies) for 6 hours (h) at  
 23 4°C. Beads were washed 4 times with wash buffer (20mM TrisHCl pH 7.5, 150mM  
 24 NaCl, 0.5% IGEPAL) and 3 times with wash buffer without detergent, and digested  
 25 with 0.2µg Lys-C (WAKO) and 0.2µg modified porcine trypsin (Promega) before  
 26 subjecting to mass spectrometry as described before (Villegas et al., 2019).

## 27 **Flow cytometry**

28 Cells were washed with PBS, detached in Accutase, centrifuged at 300g for 3 min in  
 29 DMEM/F12 with 0.1% BSA, resuspended in DMEM/F12 with 0.1% BSA, run on an  
 30 LSRII SORP Analyzer (Becton Dickinson), and analyzed using FlowJo (FlowJo,  
 31 LLC). Quantification for Gata6::mCherry positive cells (**Figure 1F, 2A, 2B, 3C, S4I**)  
 32 were performed on experimental triplicates, with gates shown in fluorescence

intensity plots. Nanog>GFP intensities (**Figure S2C, S3B**) are the average of geometric GFP mean intensities of experimental triplicates.

### **High-throughput imaging of spheroids**

To fix the spheroids, 1mL of medium was removed from each well and 500μL of 8% PFA (Alfa Aesar) in PBS were added to the remaining 500μL of medium and incubated for 20 min at 37°C. The spheroids were then harvested with a 1000μL pipette, washed twice with PBST (PBS-0.1% Tween-20 (Sigma)) and once with PBST-Hoechst 33342 (1/5000, Life Technologies), transferred in PBS to a μCLEAR 96-well plate (Greiner Bio-One), and imaged on a Yokogawa CV7000s high throughput confocal microscope at 40x magnification. Images were acquired in confocal mode as z-stack multiplane images over a z distance of 100 μm with 5 μm step width. Subsequently, images were stitched to generate a single image per z-plane, channel and well, which was used for object segmentation and feature extraction.

### **Antibody staining of spheroids**

Fixed spheroids were incubated in blocking solution (3% donkey serum (Sigma), 1% BSA in PBST) for 1 h at RT, stained overnight at 4°C with antibodies (anti-Dab2 monoclonal antibody (D709T), #12906S Cell Signaling Technology, 1/200; Human SOX17 Antibody, #AF1924 R&D Systems, 1/200) in blocking solution, washed 3 times with PBST, stained with secondary antibodies (goat anti-rabbit Alexa Fluor 647 and donkey anti-goat Alexa Fluor 647, 1/500, ThermoFisher) in PBST for 1h at RT, washed twice with PBST and once with PBST-Hoechst 33342, transferred in PBS to a μCLEAR 96-well plate (Greiner bio-one), and imaged on a Zeiss LSM710.

### **Native histone ChIPseq**

Nuclear lysates were prepared from 10 million cells as for coIPs above. Clarified lysates were incubated with 4μg of primary antibody for 2h and then SDS concentration was brought to 0.1% and EDTA to 5mM to inactivate benzonase. Lysates were clarified again and 40μL of protein G Dynabeads were added and rotated for 6h at 4°C. Beads were washed twice with RIPA buffer, four times with RIPA buffer with 500mM NaCl, and once with 10mM Tris-HCl pH8.0 with 1mM EDTA. DNA was eluted by proteinase K (Macherey-Nagel) treatment in 0.5% SDS for 3h at 55°C and purified with Qiagen MinElute PCR Purification Kit. Primary antibodies used: H4K5ac (Merck-Millipore, 07-327), H3K27ac (Active Motif,



39135). Sequencing libraries for biological duplicates of each assayed histone modification and input per condition were prepared using the ChIP-Seq NEB Ultra (Dual Indexes) Kit (New England Biolabs) and sequenced on an Illumina HiSeq2500 (50 bp single-end reads). In total ~1.05 billion reads (accounting to ca 45 million reads per replicate) were generated. After demultiplexing, reads were aligned against the mouse genome (GRCb38/mm10) using the QuasR Bioconductor library (Gaidatzis et al., 2014) with default parameters. The overall alignment rate was ~95%.

## ATACseq

ATAC-seq was performed as describe before (Buenrostro et al., 2015). Briefly, freshly isolated nuclei from 50'000 cells were subjected to transposition using the Nextera Tn5 Transposase for 30' at 37°C. DNA was purified with the Qiagen MinElute PCR Purification Kit and amplified for a total of 12 cycles with the Q5 PCR mix (NEB). Libraries of experimental triplicates were purified with AMPure-XP beads (Beckman Coulter) and sequenced on an Illumina NextSeq500 machine (75 bp paired-end reads). In total ~0.96 billion reads (accounting to ca 80 million reads or 40 million read-pairs per replicate) were generated. After demultiplexing, reads were aligned against the mouse genome (GRCb38/mm10) using the QuasR Bioconductor library (Gaidatzis et al., 2014) with default parameters. The overall alignment rate was ~94%.

## RNAseq

RNAseq of biological triplicates was performed using x cells. RNA was purified with the RNeasy Mini Kit with on-column DNase digest (Qiagen), RNase treatment etc. Libraries were prepared using TruSeq mRNA Library preparation kit (Illumina) and sequenced on an Illumina HiSeq2500 machine (50 bp single-end reads). In total ~1.6 billion reads (accounting for ca 30 million reads per replicate) were generated. After demultiplexing, reads were aligned against the mouse genome and guided by transcriptome annotation (GRCb38/mm10, GENCODE release M4) using STAR (Dobin et al., 2012) version 2.5.0 with command line parameters:

```
--outSJfilterReads Unique --outFilterType BySJout --outFilterMultimapNmax 10 --
alignSJoverhangMin 6 --alignSJDBoverhangMin 3 --outFilterMismatchNoverLmax
0.1 --alignIntronMin 20 --alignIntronMax 1000000 --outFilterIntronMotifs
RemoveNoncanonicalUnannotated --seedSearchStartLmax 50 --twopassMode Basic
```

1 The overall alignment rate was consistently ~90%.

## 2 BIOINFORMATICS

### 3 Integration of external embryonic RNAseq datasets and projection on common 4 subspaces

accession	RNAseq description	reference
E-MTAB-2958	mouse embryonic development	(Boroviak et al., 2015)
GSE100597	mouse embryonic development	(Mohammed et al., 2017)
GSE77783	endoderm from naïve pluripotency	(Anderson et al., 2017)

5 GRCm38 alignments from (Boroviak et al., 2015) were downloaded (E-MTAB-2958)  
6 and exonic gene counts were obtained on the GENCODE M4 transcriptome  
7 annotation using the GenomicFeatures Bioconductor package (Lawrence et al., 2013).  
8 Single-cell RNAseq filtered count matrices from (Mohammed et al., 2017) were  
9 downloaded (GSE100597) and gene annotation was lifted from ENSEMBL to  
10 GENCODE M4 discarding ambiguous mappings. Assignments of individual cells to  
11 specific lineages were obtained through personal communication with the study  
12 authors and bulk gene counts for each stage and lineage were calculated by  
13 aggregating the corresponding single-cell data. Gene counts from all datasets were  
14 library normalized and converted to CPMs. Microarray data from (Anderson et al.,  
15 2017) were downloaded (GSE77783) and analyzed using the limma Bioconductor  
16 package (Ritchie et al., 2015). Pairwise correlation values (**Figure 1B, S1B, S1D**)  
17 between individual samples of two datasets were calculated after gene-wise mean  
18 normalization of the log2 gene expression values.

19 For the common subspace projection including datasets from different labs (**Figure**  
20 **1C, S1C**), we converted gene expression values  $^d_i S$  of a sample  $i$  in dataset  $d$  to log2  
21 fold gene expression changes ( $S'$ ) relative to a matching reference sample  $R$  (WT 2iL  
22 and ESC (Boroviak et al., 2015), and WT d2\_EpiLC and E6.5\_EPI (Mohammed et al.,  
23 2017)):

$$^d_i S' = \log_2(^d_i S) - \log_2(^d_R S)$$

24 We used the stepwise estimation of common principal components (Trendafilov,  
25 2010), implemented in the CRAN cpca package to obtain a common subspace  
26 projection for each pairwise comparison. cpca takes as input the sample covariance  
27 matrices of each dataset and performs a joint dimension reducing transformation  
28 under the assumption that the eigenvector space is identical across datasets. In each

1 pairwise comparison only the intersection of the approximately top 5000 most  
2 variable genes, according to a within-dataset mean-variance trend fit for RNAseq  
3 experiments and a log2FC-based selection for microarray experiments, was used  
4 (provided in **Table S2**) for sample covariance matrix estimation.

## 5 **Differential gene expression analysis and gene clustering**

6 Differential gene expression (**Figure 2B, 3D, S3C**) was calculated using the  
7 Bioconductor edgeR package (McCarthy et al., 2012; Robinson et al., 2009)  
8 (provided in **Table S3**). Briefly, for each experimental design a common negative  
9 binomial dispersion parameter was estimated using the *estimateGLMCommonDisp*  
10 function and a negative binomial generalized log-linear model was fit per gene using  
11 the *glmFit* function with a prior count of 1 in order to shrink log-fold change (LFC)  
12 effect sizes of lowly expressed genes towards zero. Reported p-values are the  
13 Benjamini & Hochberg adjusted estimates for multiple comparisons.

14 Clusters of genes with consistent expression profiles (**Figure 1D**) were identified with  
15 k-means clustering using as input features the gene expression log2FC of *WT* and  
16 *Hdac3*<sup>-/-</sup> TNG-A derived clones in d2\_SL and d2\_EpiLC relative to *WT* TNG-A 2iL  
17 cells, focusing on the approximately top 5000 most variable genes as described above,  
18 and applying a pseudo-count of 1 to shrink effect sizes of lowly expressed genes.  
19 Only genes with a maximum absolute log2FC >1.5 in any of the four contrasts were  
20 considered (**provided in Table S2**). The calculation was carried out using the base R  
21 *kmeans* function implementation with *centers*=8 and 200 random starts. Analyses of  
22 enriched gene sets (**Figure S1F**) was performed using DAVID (Huang et al., 2008)  
23 for GO terms of biological processes.

## 24 **Peak calling**

accession	ChIPseq description	reference
SRR8058260	Hdac3	(Żylicz et al., 2018)
SRR8058264	Hdac3, control	
SRR2141930	Dax1, rep1	(Beck et al., 2014)
SRR2141931	Dax1, rep2	
SRR2141933	Dax1, control	
SRR5110917	Nr5a2	(Atlasi et al., 2019)
SRR5110908	Nr5a2, input	
SRR5077736	Esrrb	(Chronis et al., 2017)
SRR5077675	Esrrb, input	

SRR2043315	Gata6, rep1	(Wamaitha et al., 2015)
SRR2043317	Gata6, rep2	
SRR2043316	Gata6, control1	
SRR2043318	Gata6, control2	
SRR713340	Oct4	(Marson et al., 2008)
SRR713341	Sox2	
SRR713342	Nanog	

1 ChIP-seq data were realigned to GRCb38/mm10 using the QuasR Bioconductor  
2 library (Gaidatzis et al., 2014) with default parameters. Dax1 and Hdac3 peaks of  
3 significantly high read density, taken as a proxy of protein binding, were called using  
4 the csaw software (Lun and Smyth, 2015a) and according to the workflow outlined in  
5 (Lun and Smyth, 2015b). Briefly, the average fragment length was computed  
6 (function *correlateReads*) and reads were counted in windows of width w=78bp both  
7 for the precipitates and the input (where available). Windows of significant  
8 enrichment were filtered against a local background in regions of 5Kbp centered  
9 around the 78bp windows (3-fold enrichment, function *filterWindowsLocal*), and  
10 against the matching control sample (3-fold enrichment, function  
11 *filterWindowsControl*). Finally, nearby identified peaks, closer than 750nt apart, were  
12 merged (function *mergeWindows*). Blacklisted regions that often yield artifactual high  
13 signal in ChIP-seq experiments were obtained from the ENCODE project  
14 (<https://www.encodeproject.org/files/ENCFF547MET/>) and excluded from all read-  
15 counting operations (read parameter *discard*). Peaks were assigned to their proximal-  
16 most genes and distances to transcription start-sites (TSSs) were calculated according  
17 to the GENCODE release M4 annotation. A peak was considered as a “promoter  
18 peak” if it was within 1kbp distance of an annotated TSS.

## 19 Differential ChIPseq/ATACseq enrichment

20 ChIPseq and ATACseq signal was quantified in 1kb regions at the 29969 peaks that  
21 are the union of Hdac3 (24249) and Dax1 (10671) peaks (**provided in Table S4**). In  
22 all cases, differential ChIP-seq or ATACseq enrichments (**Figure 4A, 4B, S4A-C,**  
23 **S4E, S4G**) between condition are stated, the values refer to changes in library  
24 normalized read counts, averaged, when available, over biological replicates. enr:  
25 denotes enrichment over control or input.

26 Clusters of peaks (**Figure 4A**) were identified with k-means clustering using as input  
27 features the enrichment of Hdac3 and Dax1 ChIPseq over controls, respectively, and

differential H3K27ac ChIPseq and ATACseq signals in *Hdac3* and *Dax1* mutants, respectively, compared to *WT* cells after 2d of differentiation in SL. The calculation was carried out using the base R *kmeans* function implementation with *centers*=6 and 200 random starts. The Bioconductor ComplexHeatmap package (Gu et al., 2016) was used to plot enrichments in 10kb, peak-centered windows.

**Figure 4C** was assembled using the UCSC genome browser interface (<https://genome.ucsc.edu>).

## Motif enrichment analysis

Motif enrichment (**Figure S4D**) was performed as described before (Barisic et al., 2019). Briefly, the 29969 *Hdac3*/*Dax1* peak union set was grouped based on differential H3K27ac and ATAC signal in mutants compared to *WT* cells after 2d of differentiation in SL into 9 bins with 1319 peaks each (**Figure S4C**). Peaks with absolute log2FCs less than 0.5 were grouped into a single bin (unchanged CRE activity). Positional weight matrices (PWMs) for 519 vertebrate transcription factors available from the JASPAR database through the Bioconductor JASPAR2016 package (Mathelier et al., 2015) were used to determine motif enrichment in each bin with the findMontifsGenome.pl script of the HOMER package (Heinz et al., 2010) using all other bins as background and parameters *-nomotif -size 500*. A cutoff for the expression of TFs in ESCs and during differentiation, and for the enrichment significance of their motifs was employed. Reported enrichment significance values are FDRs corrected for multiple testing.

## Mass spectrometry

Relative protein quantifications of three independent affinity-purifications (1: experimental duplicates of anti-Dax1, -Esrrb and -Nr5a2 coIPs, 2: experimental triplicates of anti-Dax1, -Esrrb, -Nr5a2 and -Hdac3 coIPs, 3: experimental triplicates of anti-Hdac3 coIPs) were merged and missing values filled in with a 1.8 fold downshift and a 0.3 fold distribution width of the actual distribution of detected proteins, as described (Tyanova et al., 2016). Fold-change enrichments were calculated over the respective experimental negative controls, and Z-scores per each coIP and experiment determined (provided in **Table S3**). Only proteins with Z-scores>2 in both independent affinity-purifications are considered (**Figure 3A**).

## Segmentation and quantification

Object segmentation was performed using image processing libraries available for Python 3. To identify the middle plane of spheroids segmentation masks were first created for the maximum intensity projections (MIPs) of the DNA channel by automated thresholding with 2-class Otsu algorithm. Entire z-stacks were then cropped to segmentation outlines obtained from the MIP segmentation. Subsequently, sum intensity of DNA staining in each plane was measured and used to identify the z-plane with maximum sum intensity of DNA signal, corresponding to the middle plane of the object. Segmentation masks were then adjusted in-plane by automated thresholding and used for subregion analysis. Each segmentation mask of the object middle plane was subdivided in 2 regions: A 50 pixel-wide concentric region along the object periphery and a central region, ranging from the object center to the border of the peripheral region. To exclude debris, objects where the central subregion was less than 30% of the object area were eliminated from the analysis. Areas and mean intensities of the Nanog>GFP and Gata6::mCherry channels were then separately calculated for the two regions, and used for quantitative analysis. Distribution of intensity between the peripheral and central regions was calculated as the ratio of mean intensity in the peripheral and central region (**Figure S2E**, provided in **Table S3**). Extracted features describing object area and intensity were normalized to the WT controls within corresponding biological replicates using z-score transformation and unified into a cross-comparable dataset (**Figure 2E**, provided in **Table S3**).

## Quantification and Statistical analysis

Details for quantification and statistical analysis are specified in the figure legends, including number of biological replicates. Data is presented as the average and standard deviation.

## DATA AND SOFTWARE AVAILABILITY

All sequencing data have been deposited at ArrayExpress (E-MTAB-9446, E-MTAB-944, E-MTAB-0450 and E-MTAB-9453).

# 1 SUPPLEMENTAL FILES

## 2 **Figure S1: Transcriptomics of *Hdac3*<sup>-/-</sup> cells.** Related to Figure 1.

3 (A) Anti-Hdac3 western-blot of naïve *Hdac3*<sup>-/-</sup> TNG-A clones. Specific Hdac3 band is  
4 indicated.

5 (B,D) Similar to **Figure 1B**, but using different embryo RNAseq samples  
6 (Mohammed et al., 2017) (B) or *in vitro* cell types (Anderson et al., 2017) (D).  
7 Primitive streak (PS) Extraembryonic endoderm cell states (nEnd, XEN), anterior  
8 definitive endoderm (ADE)

9 (C) Similar to **Figure 1C**, but using different embryo RNAseq samples (Mohammed  
10 et al., 2017).

11 (E) Similar to **Figure 1D** with a magnified view of cluster 5.

12 (F) p-values of GO terms enriched in cluster 5.

13 (G) Indicated PrE marker mRNA expression relative to negative control siRNA  
14 transfection of cells transfected with indicated siRNAs after 2d in SL.

15 (H) *Gata6* mRNA levels relative to *WT* cells of indicated genotypes after 2d in SL.  
16 Average and standard deviation (SD) of at least two independent clones.

17 (I) Similar to **Figure 1D** with a magnified view of a panel of naïve and general  
18 pluripotency, and post-implantation markers.

19 (J) Anti-Hdac3 western-blot of naïve *Hdac3*<sup>-/-</sup> G6C18 clones. Specific Hdac3 band is  
20 indicated.

21 (K) Representative *Nanog*>GFP and *Gata6*::mCherry fluorescence intensity plots  
22 used for quantification in **Figure 1F**.

23 (L) Similar to **Figure 1G**.

## 24 **Figure S2: Characterization of *Dax1*<sup>-/-</sup> cells.** Related to Figure 2.

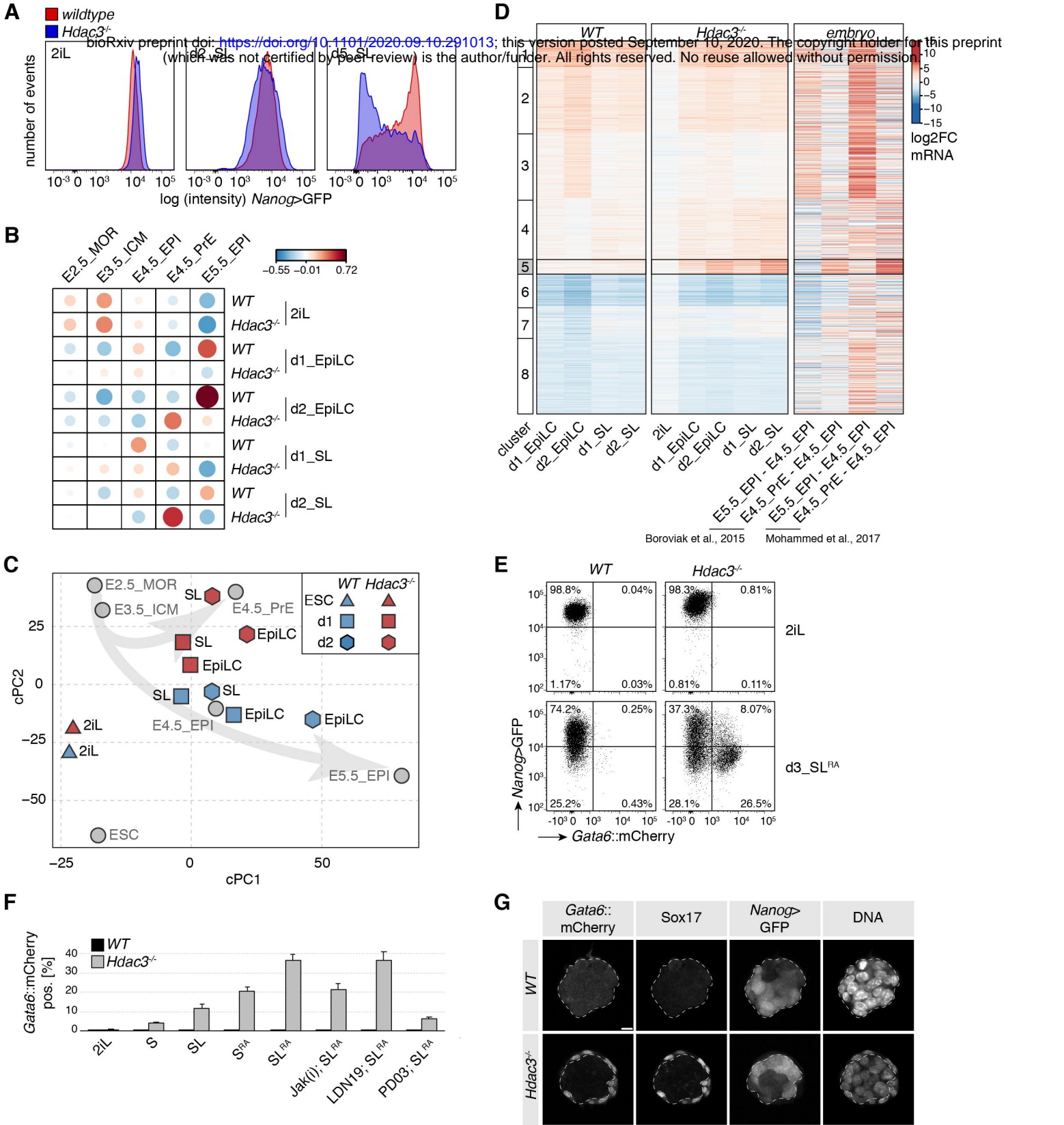
25 (A,D) Representative *Nanog*>GFP and *Gata6*::mCherry intensity plots in indicated  
26 genotypes and conditions.

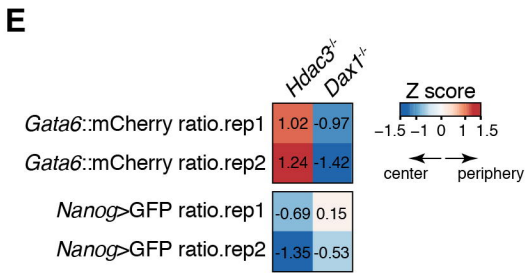
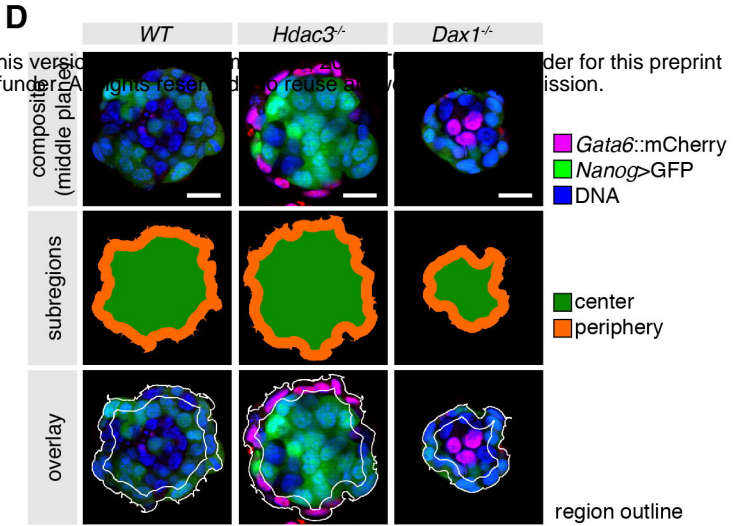
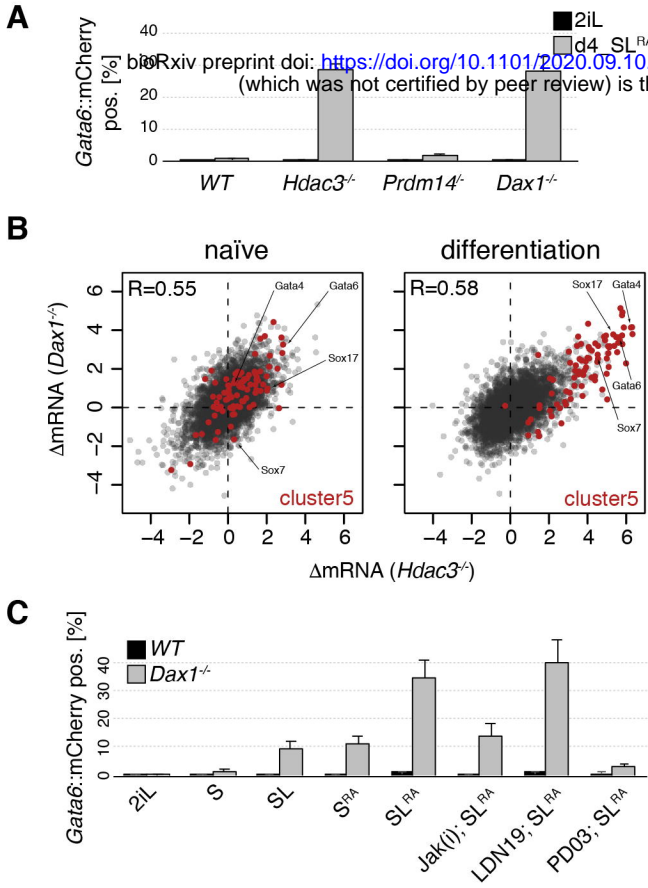
27 (B) Anti-Dax1, -Esrrb, and -Hdac3 western blots for the genotyping of G6C18 mutant  
28 clones. Migration behavior of targeted proteins is indicated.

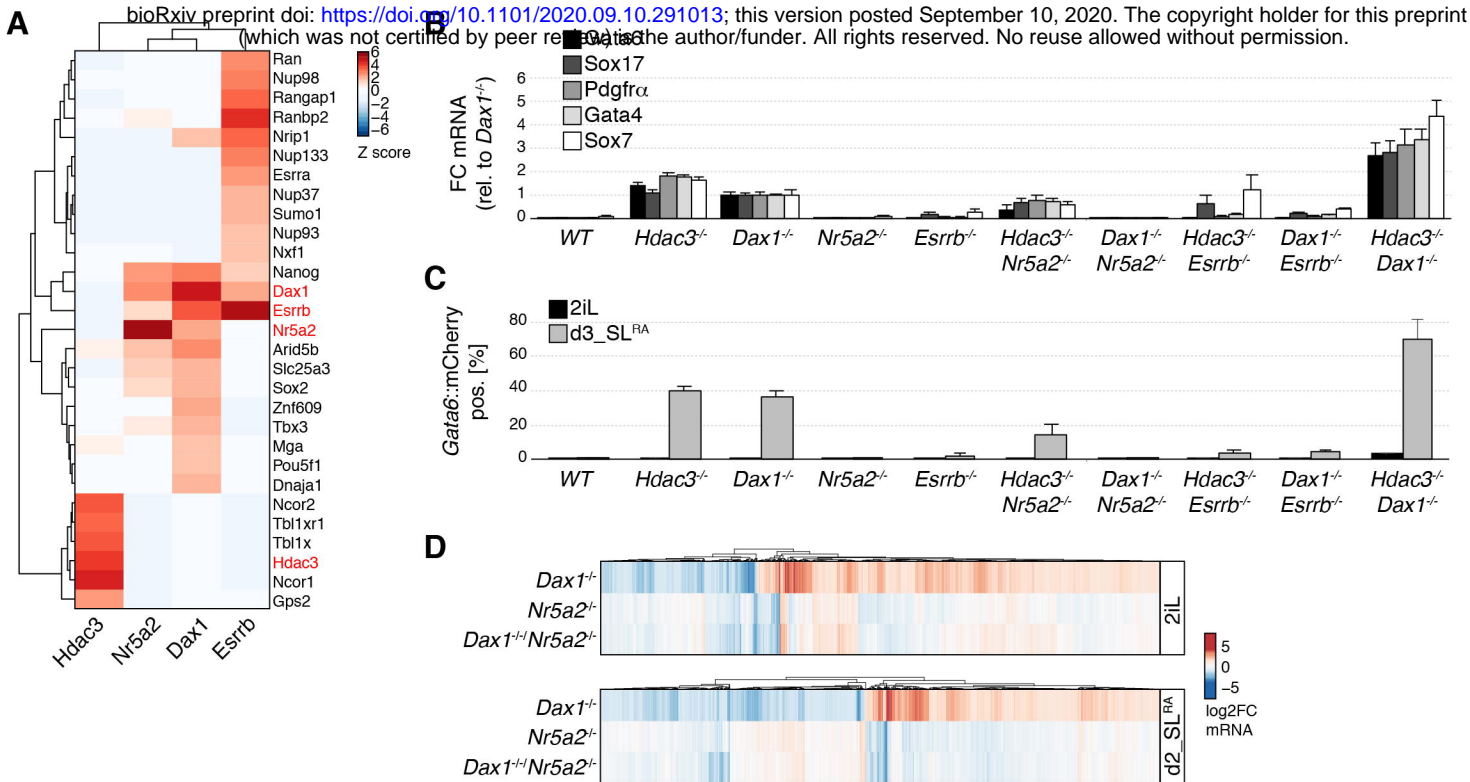


- 1 (C) Quantification of *Nanog*>GFP geometric mean intensities in 2iL. Average and SD
- 2 of at least two independent clones.
- 3 (E) Ratios of *Nanog*>GFP and *Gata6*::mCherry intensity (periphery/center) in
- 4 spheroids of indicated genotypes. Dashed lines indicate the mean in *WT* cells.
- 5 **Figure S3: Characterization of compound mutant cells.** Related to Figure 3.
- 6 (A) Representative *Nanog*>GFP and *Gata6*::mCherry intensity plots of genotypes and
- 7 conditions quantified in **Figure 3C**.
- 8 (B) Geometric mean intensity of *Nanog*>GFP reporter in 2iLIF. Average and SD of
- 9 three independent clones.
- 10 (C) Similar to **Figure 3D**, but focusing on cluster 5 genes.
- 11 (D) Similar to **Figure 1D** with a magnified view of a panel of naïve and general
- 12 pluripotency, and post-implantation markers in indicated genotypes and conditions
- 13 relative to naïve *WT* cells in 2iL.
- 14 **Figure S4: Epigenomic analysis and *enh*<sup>-45</sup> characterization.**
- 15 (A) Pairwise Pearson correlation of Hdac3, Dax1, Nr5a2, Esrrb occupancy at regions
- 16 bound by Hdac3 and/or Dax1. Individual correlation coefficients are indicated.
- 17 (B) Pairwise Pearson correlation of changes in H3K27ac, H4K5ac or accessibility in
- 18 mutant relative to *WT* cells and in indicated conditions at regions bound by Hdac3
- 19 and/or Dax1. Individual correlation coefficients are indicated.
- 20 (C) Similar to **Figure 4B**. CRE bins according to activation and repression in *Dax1*
- 21 and *Hdac3* mutants during differentiation are colored.
- 22 (D) Heat-map of high-confidence TF motifs, enrichments and false discovery rates
- 23 (FDRs) at these bins.
- 24 (E) Similar to **Figure 4B** with individual regions colored by ChIPseq signal for
- 25 indicated TFs (upper panels). The same scatterplots with log2 fold H3K27ac changes
- 26 in naïve *Hdac3* mutants instead of accessibility changes in differentiating *Dax1*<sup>-/-</sup> cells
- 27 plotted on the y-axis (lower panels).
- 28 (F) Expression changes of PrE markers in naïve ESCs of indicated genotypes relative
- 29 to *WT* cells. Average and SD of three independent clones.

- 1    **(G)** Scatterplots of log2 fold changes in H4K5ac and H3K27ac in *Hdac3* mutants at
- 2    regions bound by Dax1 and/or Hdac3 in naïve and differentiation conditions. Regions
- 3    colored in red are associated with Gata6.
- 4    **(H,I)** Representative *Nanog*>GFP and *Gata6*::mCherry intensity plots (H) and
- 5    quantification of *Gata6*::mCherry positive cells (I) in indicated genotypes after 3d in
- 6    SL<sup>RA</sup>. Average and SD of at least three independent clones.
- 7    **(J)** Model. In *WT* cells, Hdac3 and Dax1 form a regulatory network that inhibits
- 8    lineage conversion. Hdac3/NCor1/NCor2 bind and silence enh<sup>-45</sup>, while Dax1
- 9    interacts with and antagonizes Nr5a2, and indirectly influences enh<sup>-45</sup>. Derepression
- 10   of *Gata6* in *Dax1* and *Hdac3* mutants causes Gata6 to associate with Hdac3- and
- 11   Dax1/Nr5a2-bound CREs, together with Esrrb forming a positive feed-forward loop
- 12   for PrE transdifferentiation.
- 13   **Table S1: Oligonucleotides, siRNAs and antibodies, and sequencing information**
- 14   **of genome-engineered clones used in this study.**
- 15   **Table S2: Data presented in Figures 1 and S1.**
- 16   **Table S3: Data presented in Figures 1, 2, S1 and S2.**
- 17   **Table S4: Data presented in Figures 4 and S4.**

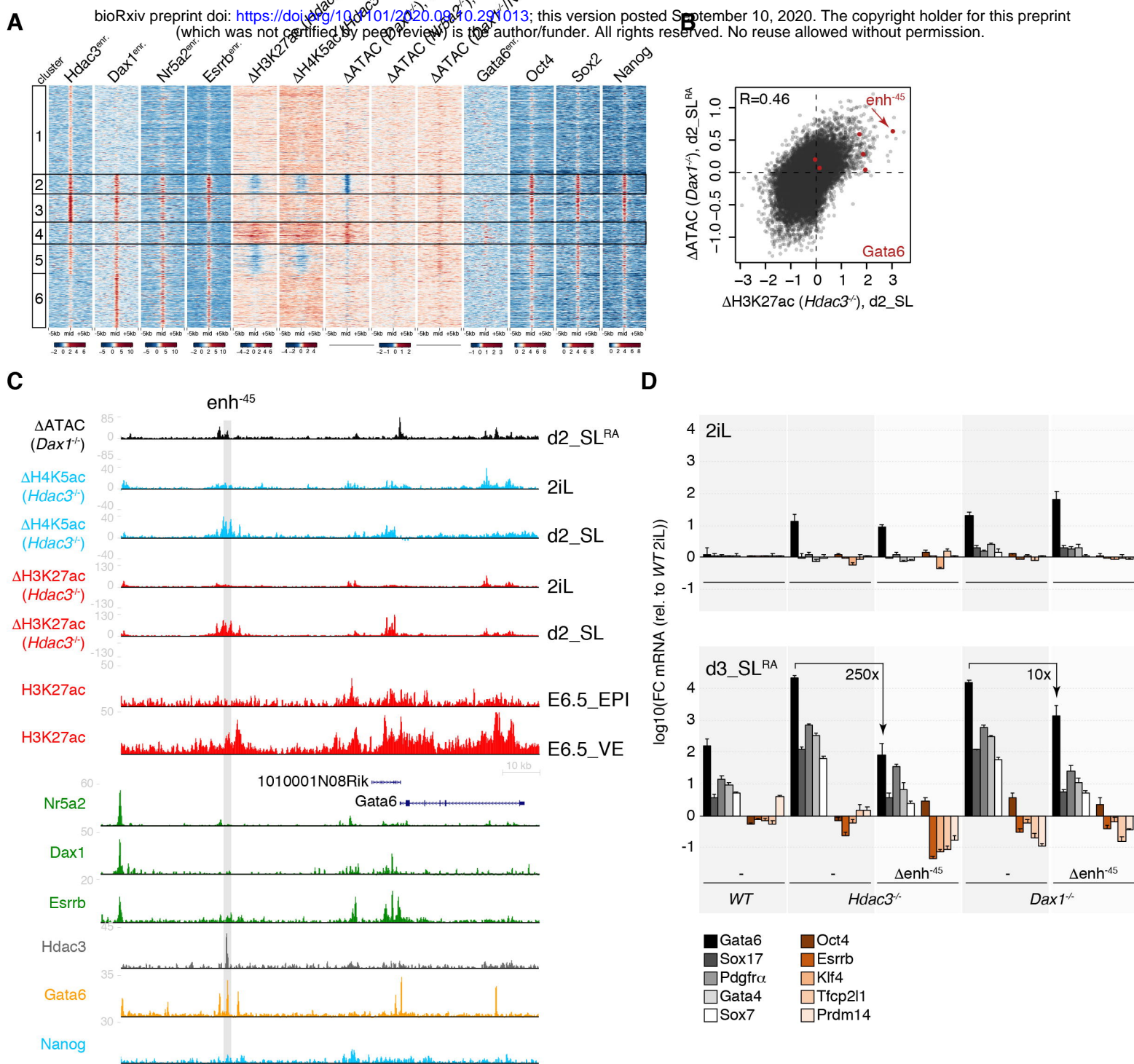


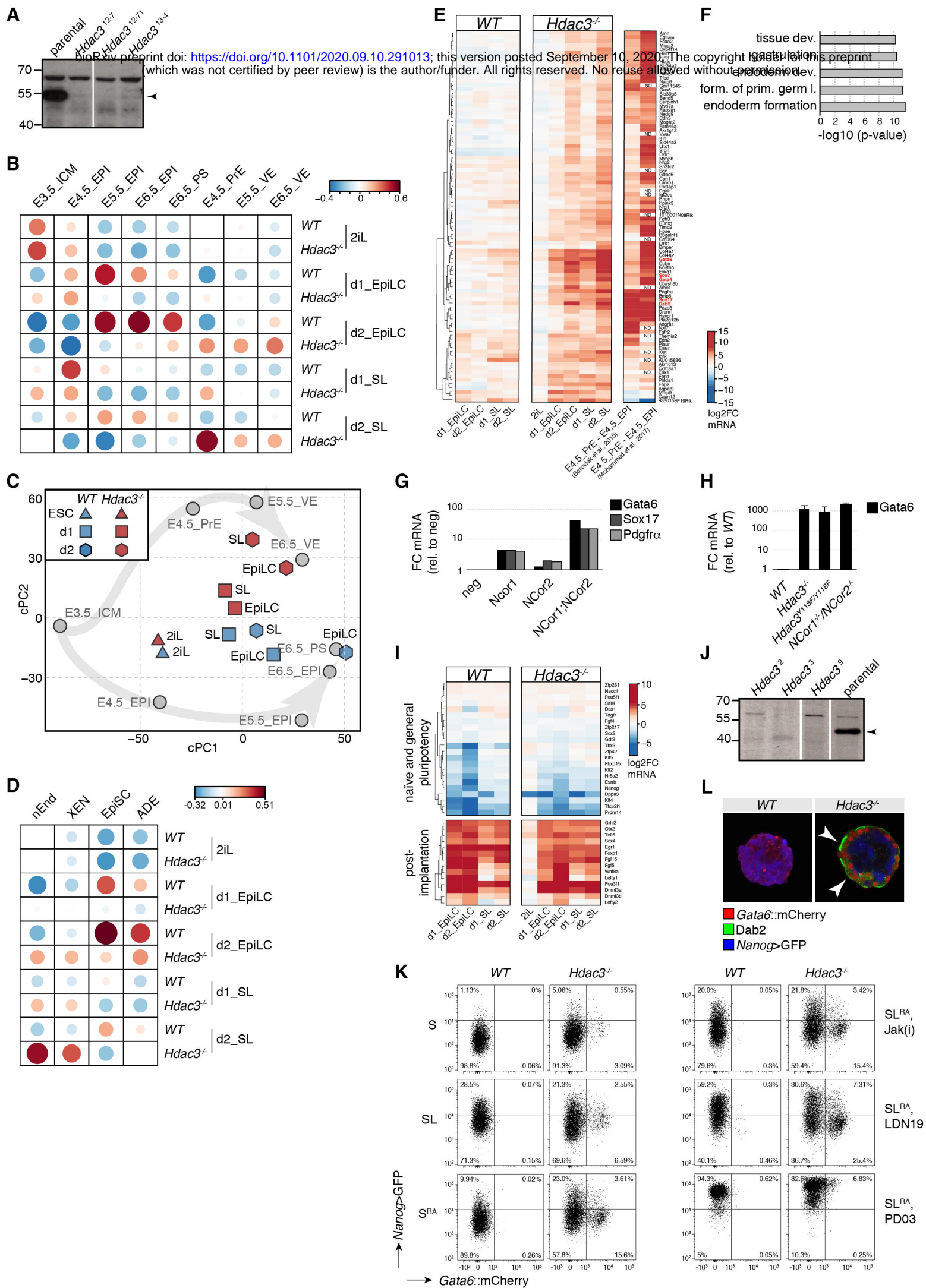




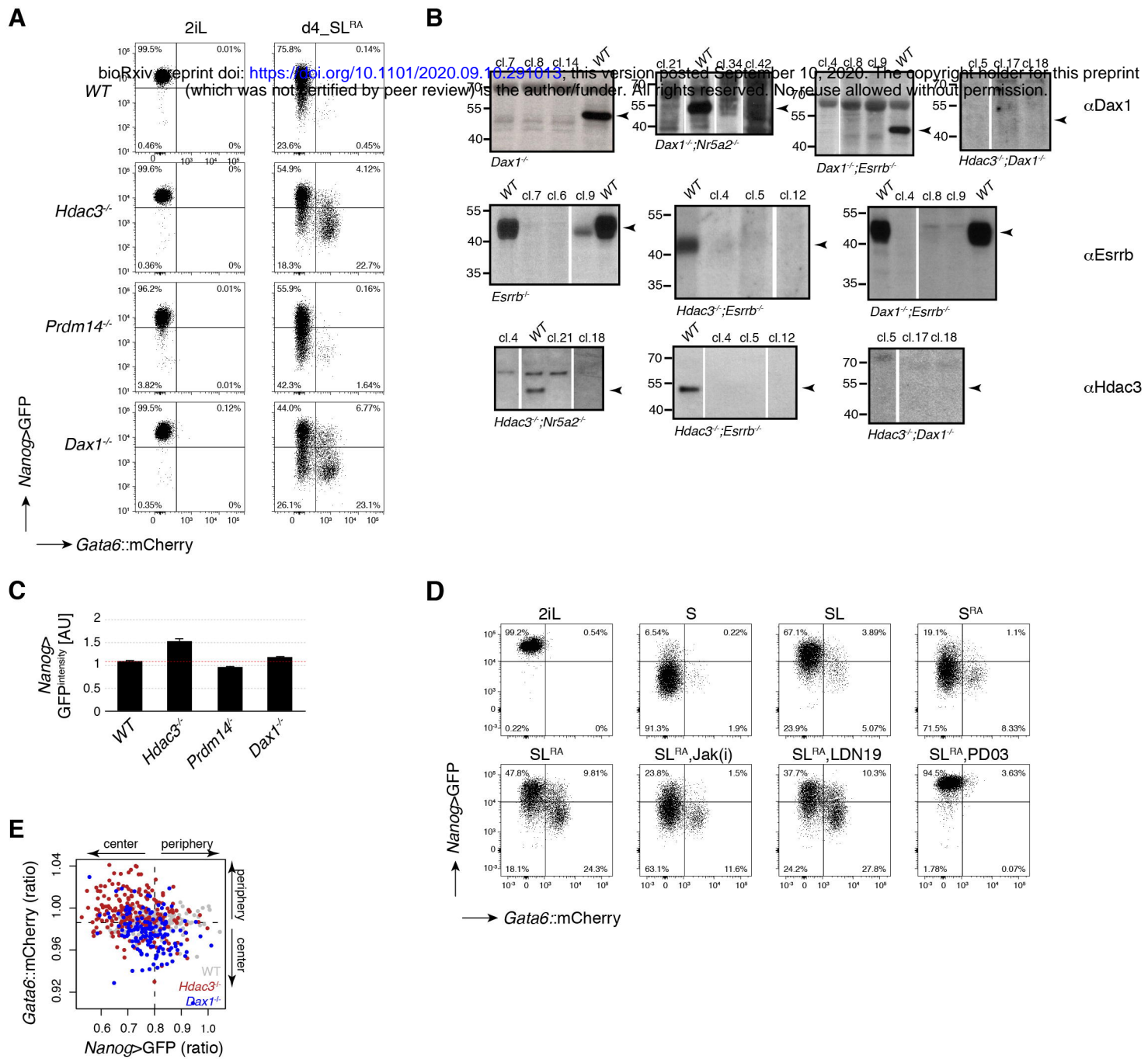


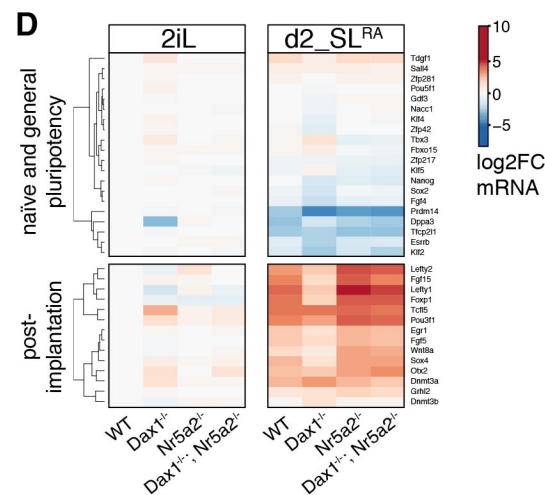
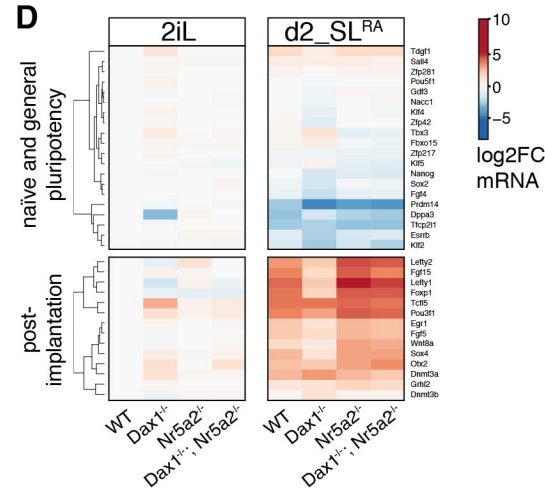
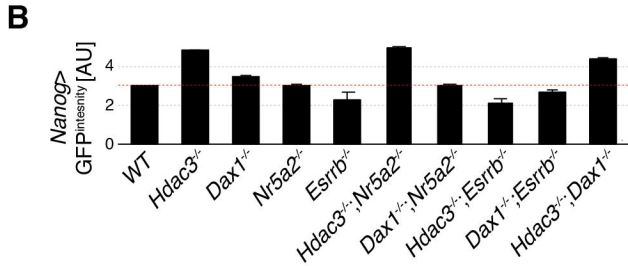
bioRxiv preprint doi: <https://doi.org/10.1101/2020.09.10.291013>; this version posted September 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



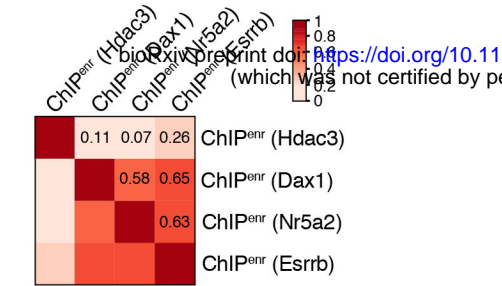




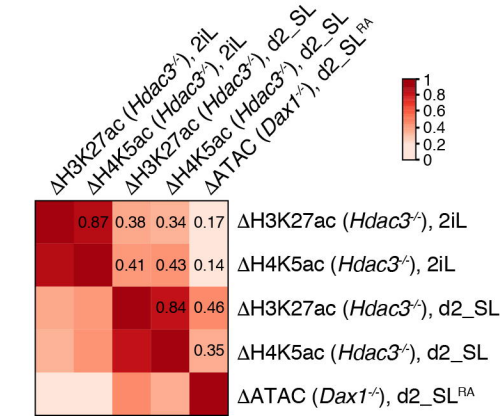




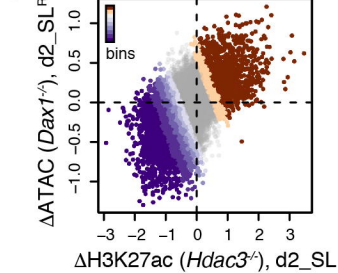
**A**



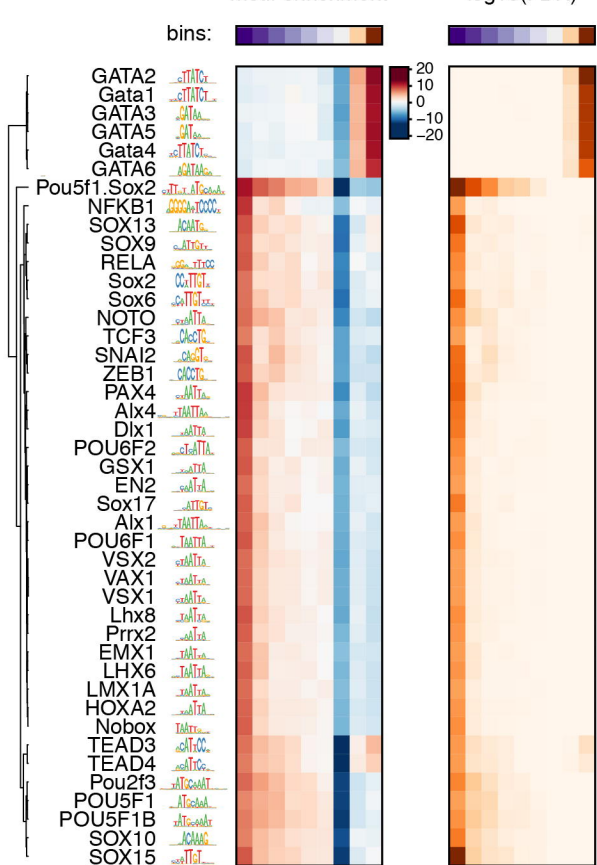
**B**



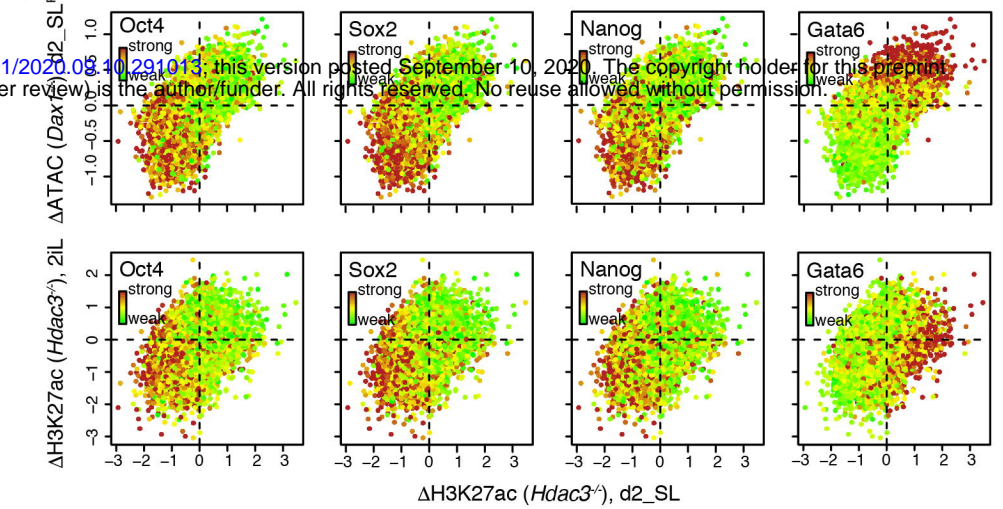
**C**



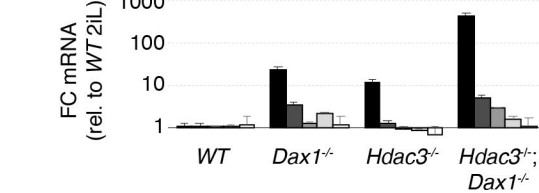
**D**



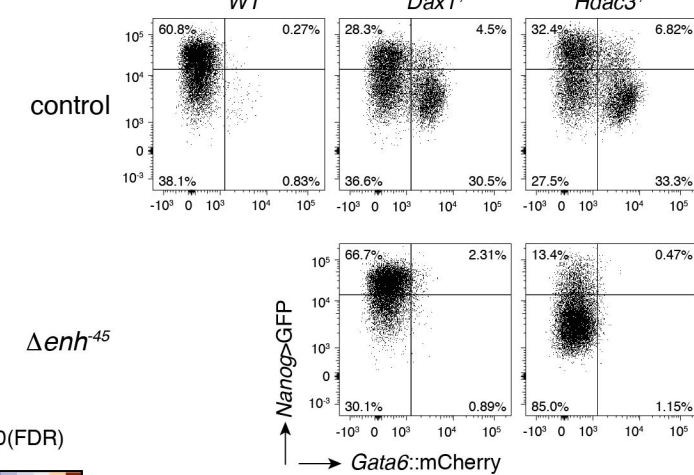
**E**



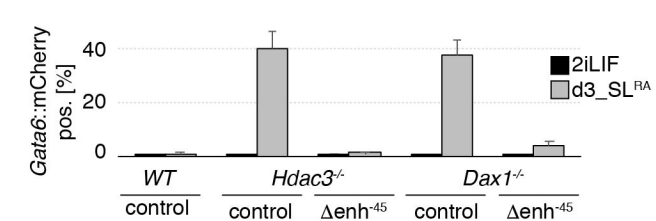
**F**



**H**



**I**



**J**

