HHS Public Access

Author manuscript

Mol Cell. Author manuscript; available in PMC 2022 November 18.

Published in final edited form as:

Mol Cell. 2021 November 18; 81(22): 4677–4691.e8. doi:10.1016/j.molcel.2021.09.018.

A Polycomb domain found in committed cells impairs differentiation when introduced into PRC1 in pluripotent cells

Elizabeth S. Jaensch[1,2,4], **Jinjin Zhu**[1,2], **Jesse C. Cochrane**[1,2], **Sharon K. Marr**[1], **Theresa A. Oei**[1,3], **Manashree Damle**[1], **Ethan Z. McCaslin**[1], **Robert E. Kingston**[1,2,5,*]

[1]Department of Molecular Biology, Massachusetts General Hospital Research Institute, Massachusetts General Hospital, Boston, Massachusetts, 02114, USA

[2]Department of Genetics, Harvard Medical School, Boston, Massachusetts, 02115, USA

[3]Department of Chemical Biology, Harvard University, Cambridge, Massachusetts, 02138, USA

[4]Current Address: Bioscience, Oncology R&D, AstraZeneca, Waltham, Massachusetts, 02451, USA

[5]Lead Contact

SUMMARY

The CBX family of proteins is central to proper mammalian development via key roles in Polycomb-mediated maintenance of repression. CBX proteins in differentiated lineages have chromatin compaction and phase separation activities that might contribute to maintaining repressed chromatin. The predominant CBX protein in pluripotent cells, CBX7, lacks the domain required for these activities. We inserted this functional domain into CBX7 in embryonic stem cells to test the hypothesis that it contributes a key epigenetic function. ESCs expressing this chimeric CBX7 were impaired in their ability to properly form embryoid bodies and neural progenitor cells and showed reduced activation of lineage-specific genes across differentiation. Neural progenitors exhibited a corresponding inappropriate maintenance of Polycomb binding at neural-specific loci over the course of differentiation. We propose that a switch in the ability to compact and phase separate is a central aspect of Polycomb group function during the transition from pluripotency to differentiated lineages.

INTRODUCTION

Development of multicellular organisms requires that many different cell types attain and maintain their proper identity. In each cell type, a specific group of genes must be expressed for proper function, and genes associated with other cell types must be stably repressed. The Polycomb group (PcG) proteins ensure that their target genes are appropriately repressed

*Correspondence: kingston@molbio.mgh.harvard.edu.

AUTHOR CONTRIBUTIONS

Conceptualization, E.S.J. and R.E.K.; Methodology, E.S.J; Formal Analysis, E.S.J. and M.D.; Investigation, E.S.J, J.Z., J.C.C., and T.O.; Resources, E.S.J., J.Z., S.K.M., E.M.; Writing – Original Draft, E.S.J. and R.E.K.; Writing – Reviewing & Editing, E.S.J. and R.E.K.; Funding Acquisition, E.S.J. and R.E.K.

DECLARATION OF INTERESTS

The authors declare no competing interests.

during differentiation in a cell-type specific manner, playing a vital role in development and the ability to remember specific cell states. Misregulation of Polycomb function, and therefore loss of cellular identity, can lead to developmental defects and cancer (reviewed in Chan and Morey, 2019; Pasini and Di Croce, 2016; Schuettengruber and Cavalli, 2009).

The PcG proteins form two main Polycomb Repressive Complexes: PRC1 and PRC2. PRC2 catalyzes the methylation of lysine 27 in histone H3 (H3K27me3), which both it and PRC1 can recognize and bind (Bernstein, E., et al., 2006; Kuzmichev et al., 2002; Margueron et al., 2009). Variant PRC1 (vPRC1) can efficiently catalyze the H2AK119ub mark, which also plays a role in Polycomb recruitment through binding by PRC2 (Blackledge et al., 2014; Blackledge et al., 2020; Cooper et al., 2014; Tamburri et al., 2020). Canonical PRC1 (cPRC1) functions non-enzymatically through its different subunits to physically compact chromatin, both locally and over larger genomic regions, and to drive phase separation (Boyle et al., 2020; Du et al., 2020; Grau et al, 2011; Isono et al., 2013; Kent et al., 2020; Kundu et al., 2017; Plys et al., 2019; Tatavosian et al., 2019).

Compaction and phase separation activities in mammalian cPRC1 have been characterized in the CBX subunit of the complex (Grau et al., 2011; Plys et al., 2019; Tatavosian et al., 2019). Previous work examined CBX2 and identified a positively-charged, intrinsically disordered region of the protein as being necessary for both activities (Grau et al., 2011; Plys et al., 2019). Reducing the positive charge of this region through a series of point mutations severely impairs the ability of CBX2 to perform either compaction or phase separation. We therefore refer to this functional domain as the CaPS (Compaction and Phase Separation) domain. In flies and in mammals, the domain responsible for compaction activity of cPRC1 is necessary for proper gene repression and development. In flies, mutations in the Psc subunit of PRC1 that reduce its ability to compact chromatin in vitro correlate with developmental phenotypes (King et al., 2005). Similarly, mice with charge-reducing mutations in the CBX2 CaPS domain have homeotic transformations that phenocopy the axial development phenotypes of a CBX2 null mouse (Lau et al., 2017).

In mammals, there are five Polycomb CBX homologs that differ in their predicted or demonstrated ability to compact and phase separate (Grau et al, 2011; Plys et al, 2019). These have distinct expression patterns and appear to serve non-overlapping functions in development (reviewed in Ma et al., 2014). CBX7 stands out within this family, as it lacks the entire CaPS domain, and it has been shown to lack both compaction and phase separation activities in vitro (Grau et al., 2011; Plys et al., 2019). Notably, it shows a unique expression pattern among its homologs; CBX7 is the main CBX protein expressed in embryonic stem cells (ESCs) and is highly expressed in other multipotent cells such as hematopoietic stem cells (Klauke et al., 2013; Morey et al., 2012; O'Loghlen et al., 2012). It is downregulated as cells differentiate in many lineages, while the other CBX proteins show the opposite expression pattern, with little to no expression in ESCs and upregulation upon lineage commitment.

Polycomb targets in ESCs are frequently found in domains of bivalent chromatin, originally defined as having both the repressive H3K27me3 and the active H3K4me3 histone modifications (Bernstein, B.E. et al, 2006). Bivalent chromatin is further characterized

by low levels of transcription at many of its associated genes (Brookes et al., 2012; Min et al., 2011). At the same time, these genes are marked to allow for more complete repression following lineage commitment. These genes are hypothesized to be “poised”, with the underlying theory that it is important that the repressed state of developmental genes be reversible in pluripotent cells, so more committed cell types can adopt specific transcriptional programs during differentiation (reviewed in Kuroda et al., 2020). The chromatin state at such genes must therefore allow for a balance between activation and repression to contribute to a fluid chromatin state that can be resolved in either direction upon differentiation, without requiring the de novo recruitment of either set of factors.

Since CBX7 functions mainly in pluripotent or multipotent cells, we hypothesized that it plays a specific and important role in the plasticity of a poised chromatin state due to its lack of a CaPS domain. The lack of this domain in ESCs might, via this hypothesis, be essential for maintaining plasticity of cell fate. Here, we test this hypothesis by introducing the CBX2 CaPS domain into endogenous CBX7 in mouse ESCs and examining the resultant phenotype. These cells are unable to properly form embryoid bodies, and during differentiation PRC1 targets in these cells resist both activation and repression. This is echoed in neural progenitor formation, where cells with CBX7-CaPS show reduced activation of neural-associated genes and incomplete repression of genes associated with alternate lineages. This is correlated with a prolonged maintenance of PRC1 on neural-associated genes in CBX7-CaPS neural progenitors. Thus, the addition of a domain capable of compaction and phase separation to CBX7 has an impact on gene expression over the course of differentiation and may impair pluripotency through disruption of a poised chromatin state in ESCs.

RESULTS

A region of CBX2 confers *in vitro* chromatin compaction and phase separation activity on CBX7

While the CaPS domain has been demonstrated to be necessary for compaction and phase separation in CBX2, its sufficiency to impart these activities to another member of the CBX family of proteins has not been tested. To determine if the CaPS domain is sufficient to confer these functions, we designed chimeras of CBX7 that contained either the N terminal half (‘CaPS1’) or the entire domain (‘CaPS2’) and tested them in a series of *in vitro* protocols (Figure 1A). We additionally designed chimeric CBX7 constructs with mutant CaPS (mCaPS) insertions. The mutations in mCaPS reduce the positive charge of the region and have been shown to reduce *in vitro* compaction and phase separation activity in CBX2 (Grau et al., 2011; Plys et al., 2019) and to alter axial patterning in mice (Lau et al., 2017).

We examined the ability of these chimeric proteins to compact nucleosome arrays by electron microscopy. While the arrays were spread out with individual nucleosomes visible when no protein is added, the addition of CBX2 resulted in the formation of compact, globular structures (Figure 1B). As seen previously (Grau et al., 2011), compaction did not occur in the presence of CBX7. When either CBX7-CaPS1 or CBX7-CaPS2 was incubated with the array, however, the arrays became compacted. Compaction was not observed when CBX7-mCaPS was added, suggesting this activity is specific to the positive charge of the

CaPS domain, and not an artifact of increasing the size of CBX7. The degree of compaction was quantified both by determining the end-to-end distance of each array and by counting individually observable particles in each array (Francis et al., 2004). More particles reflect a more extended state, while fewer reflect compaction of the array. This quantification showed a significant difference in degree of compaction between proteins, with CBX2 and CBX7-CaPS showing fewer particles than an array with no protein, and CBX7 and CBX7-mCaPS showing a similar number of particles (Figure 1C; Figure S1A). CBX2 and CBX7-CaPS also showed a reduced array length when compared to no protein controls, as well as to CBX7 or CBX7-mCaPS arrays (Figure 1D; Figure S1B).

When proteins capable of compaction are pre-incubated with an array, they inhibit subsequent remodeling by SWI/SNF family complexes. We quantified the ability of each construct to inhibit SWI/SNF nucleosome remodeling using a restriction enzyme accessibility protocol (Grau et al., 2011). As previously described, CBX2 showed effective inhibition of remodeling at increasing protein concentrations, while CBX7 had little to no effect (Figure 1E). CBX7-CaPS1 and - CaPS2 had similar activity to CBX2 in this assay, while the equivalent CBX7-mCaPS proteins did not show significant activity.

Having shown that the CaPS region can transfer *in vitro* chromatin compaction activity to CBX7, we tested whether it could additionally transfer phase separation activity, as these activities have correlated with each other when various mutations in CBX2 have been tested (Plys et al., 2019). We tested the ability of different CBX7-CaPS constructs to form turbid solutions, a known property of phase separating proteins. CBX7-CaPS proteins formed turbid solutions at increasing protein concentrations, while CBX7 and CBX7-mCaPS proteins did not (Figure S1C). We further studied the ability of these proteins to phase separate by examining whether they formed droplets of high local protein concentration in solution after purification with GFP-RING1B. CBX7-CaPS proteins were able to form GFP-labeled droplets as observed via microscopy, while CBX7-mCaPS proteins were not (Figure 1F, Figure S1D). CBX7-CaPS proteins formed droplets over a wide range of protein and salt concentrations, even at salt concentrations that exceeded physiological levels (Figure 1G, Figure S1E). These droplets displayed sensitivity to both high and low salt concentrations, as they often appeared smaller or in reduced quantities near the extremes of investigated conditions (Figure S1F). CBX7-CaPS1 could additionally incorporate biologically relevant ligands, including nucleosomal arrays, into these droplets (Figure 1F). These data together show that the CBX2 CaPS domain confers its *in vitro* characterized chromatin compaction and phase separation activities on CBX7, extending our previous conclusions that these activities can be attributed to this region of CBX2 and may be intrinsically linked. They further demonstrate that the N-terminal half of the CaPS domain (CaPS1) is sufficient to confer both activities on CBX7.

Embryonic stem cells containing CBX7-CaPS are phenotypically similar to WT

Having confirmed that CBX7-CaPS contained a functional CaPS domain, we examined the impact of this gain-of-function mutation in a biological context. We used CRISPR-Cas9 to insert the CaPS1 region into the endogenous CBX7 locus of mouse embryonic stem cells and created two independent homozygous insertions in cell lines called CBX7-CaPS (the

line used for most of the studies described here) and CBX7-CaPS-R (a replicate line used to confirm the major findings of the initial CaPS line) (Figure 2A, Figure S2A). For use as controls, we generated cell lines with a homozygous mCaPS insertion at the endogenous CBX7 locus (CBX7-mCaPS) or with a homozygous CBX7 knock-out (KO). We noted that the protein levels of CBX7-mCaPS in the nuclear extract, in whole cell extracts, and in insoluble chromatin appeared lower than those of other CBX7 constructs (Figure 2A, S2B).

The CBX7-CaPS ESCs did not display any significant phenotypic differences from wild-type (WT) ESCs. CBX7-CaPS ESCs stained positively for alkaline phosphatase and appeared morphologically normal in culture (Figure S2C). There was no difference in their growth rate across a three-day time course (Figure S2D). Since PRC1 plays a role in gene repression in ESCs, we studied whether the addition of the CaPS domain to the complex had any effect on gene expression in this context. Measurement of RNA levels using high throughput sequencing in WT and CBX7-CaPS ESCs showed minimal changes in gene expression (Figure 2B).

Having observed no phenotypic or gene expression defects in CBX7-CaPS ESCs, we investigated whether there were any observable molecular differences in CBX7-CaPS interactions or localization. To investigate whether CBX7-CaPS interacts with proteins in a manner consistent with CBX7, or whether the presence of a CBX2 domain disrupts these protein interactions, we performed RIME, a crosslinked immunoprecipitation followed by mass spectrometry (Mohammed et al., 2016), against CBX7, CBX7-CaPS, and CBX2, which is expressed at low levels in ESCs. As expected, CBX7-CaPS interacted with PRC1 components (Table 1, Table S1). CBX7 and CBX7-CaPS additionally showed unexpected pulldown of PRC2 components, which was not mirrored in the CBX2 immunoprecipitation. While CBX7-CaPS maintains the interactions of CBX7, it does not form novel protein interactions. We conclude that despite the addition of a CBX2 domain, CBX7-CaPS maintains the protein interaction behaviors of CBX7 in mESCs.

We performed immunofluorescence (IF) and ChIP-Seq to determine if the distribution of PRC1 was altered with the addition of the CaPS domain. In WT cells, PRC1 puncta containing CBX7 were visible in the nucleus by IF (Figure 2C). The formation of these puncta are unlikely to be caused by CBX7, but might form due to the action of other PRC1 components, such as PHC1 (Boyle et al., 2020; Isono et al., 2013; Kent et al., 2020; Seif et al., 2020), or associated chromatin structures (Gibson et al., 2019). When we examined CBX7 and RING1B in the CBX7-CaPS nucleus, we also observed the presence of puncta (Figure 2C). These puncta were not visually different from those in WT nuclei, but we have not yet studied whether they exhibit different characteristics beyond appearance.

ChIP-Seq in CBX7-CaPS cells further showed that both CBX7 and RING1B were present at the same loci as in WT cells, indicating that there was not mis-targeting of PRC1 (Figure 2D). Previous *in vitro* data have shown that the density of binding at loci by PRC1 might be impacted by the presence of compaction activity, as a compacted nucleosomal array contains one complex bound per three to four nucleosomes (Francis et al., 2004). We observed somewhat lowered overall levels of PRC1 at target loci, as measured by RING1B binding, while H3K27me3, the modification catalyzed by PRC2, displayed modest

but consistently higher ChIP-Seq signal in CBX7-CaPS cells (Figure 2E,F). The results of CUT&RUN against RING1B and H3K27me3 showed a similar pattern (Figure S2E). The CBX7-mCaPS line showed somewhat lower levels of PRC1 binding compared to CBX7-CaPS, as assessed by CUT&RUN (Figure S2E), consistent with the decreased expression level of CBX7-mCaPS. The CBX7 KO line showed no signal in CBX7 ChIP-Seq and a large reduction in RING1B binding (Figure S2F), which is expected since most cPRC1 in ESCs contains CBX7 (Morey et al., 2012). The reduced binding of PRC1 in the KO cell line is important for its use as a genetic control in the gain-of-function studies described below and is notably distinct from the binding characteristics of the CBX7-mCaPS cell line, indicating that this latter line remains useful as a control despite its lowered expression and PRC1 binding characteristics.

Embryoid bodies formed from CBX7-CaPS mESCs exhibit differentiation defects

Since PRC1 containing CBX2 and its CaPS domain is normally found in committed cell types, we hypothesized that it might lead to a repressive state that is more resistant to perturbation than that of PRC1 containing WT CBX7. We therefore predicted that though the CBX7-CaPS ESCs were phenotypically normal, the presence of the CaPS domain in PRC1 in these gain-of-function cells would impair their ability to efficiently change gene expression programs and differentiate. To test this, we performed a non-directed differentiation time course by allowing the ESCs to form embryoid bodies (EBs). We found that CBX7-CaPS ESCs differentiated into EBs, but they were significantly smaller over the entire differentiation time course than their WT counterparts (Figure 3A; Figure S3A). The CBX7-CaPS-R cell line also showed significant defects in EB size. In the CBX7-mCaPS control line, EBs were smaller than their WT counterparts at day 6 of differentiation, but were significantly larger than CBX7-CaPS EBs as differentiation proceeded. This is consistent with the fact that while the mCaPS mutation impairs compaction and phase separation, it does not completely eliminate these functions (Figure 1E) (Grau et al., 2011; Plys et al., 2019). CBX7 KO embryoid bodies, on the other hand, were smaller than all other cell lines over the entire time course. We further tested the ability of CBX7-CaPS ESCs to differentiate properly by quantifying the number of beating clusters that developed late in embryoid body formation, representing proper development of early cardiac cells (Movie S1). We found that CBX7-CaPS EBs formed significantly fewer beating clusters than their WT counterparts (Figure 3B). Similar to what was observed when measuring EB size, CBX7-mCaPS EBs showed a delay in beating cluster formation when compared to WT, but over time exceeded the number formed in CBX7-CaPS EBs, and CBX7 KO EBs were impaired in beating cluster formation. Thus, CBX7-CaPS and CBX7 KO cells did not differentiate properly, and while CBX7-mCaPS differentiation was slightly impaired, it was similar in behavior to WT. The beating cluster and EB size data together shows that the impairment seen in CBX7-CaPS cells is specific to its gain-of-function.

To understand how CBX7-CaPS impacted gene regulation during differentiation, we examined how the changes in gene expression during CBX7-CaPS EB formation related to changes in gene expression during WT EB formation. We broke the differentiation time course into two parts: day 0 to day 6, and day 6 to day 10. As expected, in WT cells we saw both up- and down-regulation of over 1,000 Polycomb target genes, as defined by

our ESC RING1B ChIP-Seq, over both intervals as cells committed to different lineages. In CBX7-CaPS cells, however, hundreds of target genes showed reduced changes in gene expression (Figure 3C; Figure S3B; Table S2). This resistance to change occurred both in upregulation, as with developmental genes like *Sox17* and *Foxf1* (Figure 3D), and in downregulation. While most genes clustered as "similarly regulated" with the conservative cutoffs we designed, this group also showed a significant resistance to change in gene regulation over time in both the CBX7-CaPS and CBX7-CaPS-R cell lines (Figure 3E; Figure S3C). We conclude that replacing WT CBX7 with CBX7-CaPS in ESCs impairs the changes in gene expression that accompany differentiation.

Similar analysis of the CBX7-mCaPS cell line over EB formation showed that it did not have the same resistance to changes in gene expression as the CBX7-CaPS cells (Figure 3C,E; Figure S3C). CBX7-mCaPS EBs showed some defects in early upregulation of PRC1 target genes, which correlated well with the observed phenotypic defects early in CBX7-mCaPS EB formation. However, these differences did not mirror those of CBX7-CaPS, and later in differentiation CBX7-mCaPS EBs mimicked the regulation patterns of WT. The CBX7 KO embryoid bodies also exhibited a different gene regulation pattern from CBX7-CaPS (Figure 3C,E). The results from these control lines indicate that the patterns of resistance to gene expression changes are specific to the gain-of-function CBX7-CaPS protein and are not due to a general perturbation of CBX7 function in stem cells.

We examined whether the consequences of the CaPS domain addition had differential impacts on specific germ layers during EB formation. Lineage markers of all three germ layers showed the same trend of resistance to upregulation in CBX7-CaPS differentiation compared to WT, while pluripotency markers showed a resistance to downregulation (Figure 3F). Further exploration of the genes that resist upregulation during CBX7-CaPS differentiation using Gene Set Enrichment Analysis (Mootha et al., 2003; Subramanian et al., 2005) revealed significant enrichment for many biological processes associated with differentiation and development (Figure S3D). Thus, many genes related to differentiation and cell fate commitment showed resistance to upregulation in CBX7-CaPS differentiation, and this impacted all germ layers. We conclude that resistance to change in gene expression during differentiation is a general consequence of CaPS domain introduction into CBX7.

We performed CUT&RUN in Day 10 EBs to see if these differences in gene regulation were associated with differences in PRC1 binding. We found no significant differences in RING1B occupancy between WT and CBX7-CaPS EBs (Figure S3E,F). This indicates that even though PRC1 is occupying the same genomic locations in the WT and CBX7-CaPS lines in both EBs and in ESCs (Figure 2D,E), the gain of function version has a different effect on gene expression than WT during differentiation. It is worth noting, however, that EBs are inherently heterogeneous, so it is possible that there are differences in binding in the distinct cell types that were obscured by analyzing the bulk population.

Neural progenitor differentiation is disrupted in CBX7-CaPS cells

We investigated whether CBX7-CaPS would also impair development in a directed differentiation system by differentiating cells down a neuronal lineage to form neural

progenitor cells (NPCs) (Bibel et al., 2007). Both WT and CBX7-CaPS ESCs formed aggregates of NPCs in this system, with no obvious morphological differences (Figure S4A).

At the molecular level, however, CBX7-CaPS cells showed substantive misregulation of developmental genes as they progressed down the differentiation time course (Figure 4A). We performed gene ontology analysis on misregulated genes and found that on both day 6 and day 8 of differentiation, CBX7-CaPS NPCs showed lower expression of genes associated with neural differentiation and higher expression of genes associated with alternate lineages when compared to WT NPCs (Figure 4B). This is consistent with CBX7-CaPS cells failing to adopt a neural specific gene expression program when pushed down that specific lineage.

We analyzed changes in PRC1 target gene expression over time in NPCs as we had in EBs. In this system, CBX7-CaPS NPCs again showed reduced activation and repression of target genes (Figure 4C,D; Table S3). As observed with EB formation, most differences in gene regulation were due to resistance to changes in expression in CBX7-CaPS cells (Figure S4B). Analysis of genes that resisted activation in CBX7-CaPS NPC formation showed an enrichment for biological processes associated with a neuronal lineage (Figure S4C). This is consistent both with the general failure of these cells to activate a neural expression program, and with a resistance to activation contributing to the misregulation of genes identified in Figures 4A and B. We examined expression of CBX2, CBX4 and CBX8, the predominant differentiation specific family members, to determine whether their expression was altered in a manner that might contribute to the molecular phenotypes observed and found that they all maintained similar expression patterns in WT and the CBX7-CaPS (Table S3).

In contrast to EB formation, the mCaPS and CaPS lines showed similar gene expression patterns across the NPC differentiation course (Figure S4D), which might reflect the residual compaction and phase separation activity in mCaPS noted previously that could suffice to generate a molecular phenotype in this differentiation protocol. The gene expression data for NPC formation from CBX7-CaPS ESCs thus indicates a resistance to change in gene expression of PRC1 targets that impairs the ability of cells to adopt a specific differentiated identity.

Changes in gene expression in CBX7-CaPS NPCs are linked to differences in Polycomb binding

The phenotypes we report above for the CBX7-CaPS gain-of-function mutation are consistent with the possibility that PRC1 in these mutant cells might retain its presence on genes that normally lose PRC1 binding during differentiation, thereby leading to maintained repression of those genes. To test this hypothesis, we investigated the binding status of PRC1 in NPCs by measuring RING1B ChIP-Seq signal because RING1B provides a common measure for PRC1 binding between the distinct CBX7 cell lines. We expected that if CBX7-CaPS PRC1 was maintained longer than WT PRC1 on target genes, there would be fewer changes in PRC1 occupancy in CBX7-CaPS cells over the course of differentiation. This expectation was met; from day 0 to day 6 of NPC formation, fewer peaks changed in RING1B binding intensity in CBX7-CaPS NPCs than changed in WT NPCs (Figure

5A). Later in differentiation, CBX7-CaPS NPCs showed more peaks losing PRC1 than their WT counterparts, potentially reflecting a delay in the differentiation process since many of these peaks had already lost PRC1 by day 6 in WT differentiation (Figure S5A). Genes that showed a loss of RING1B in WT differentiation but not in CBX7-CaPS were highly enriched for biological processes related to neural development, consistent with the impairment in activation of these genes in the CBX7-CaPS line (Figure 5B).

To determine whether the differences in PRC1 binding correlated with the differences in gene expression in CBX7-CaPS NPCs, we identified gene sets that were either downregulated relative to WT, expressed similarly to WT, or upregulated relative to WT based on the data in Figure 4A. We tested the prediction that genes that were normally upregulated during WT NPC differentiation would be maintained in a repressed state by continued binding by CBX7-CaPS PRC1. Consistent with this prediction, the enrichment of RING1B in CBX7-CaPS relative to WT was highest on genes that showed lower expression in CBX7-CaPS relative to WT and was lowest on genes that were more highly expressed in CBX7-CaPS (Figure 5C). The same pattern was evident when H3K27me3 levels were assessed. We conclude that there is a correlation between the greater presence of the repressive Polycomb group, as represented both by RING1B and H3K27me3, and the misregulation of developmentally regulated genes during differentiation in CBX7-CaPS cells.

We examined this correlation further by analyzing how a resistance to change in gene expression related to the maintenance of PRC1 on individual neural loci. Investigation of neural marker genes such as *Pax6* and *Nestin* showed a loss of PRC1 binding over WT differentiation, which as expected resulted in upregulation of these genes by day 8 (Figure 5D; Figure S5B). However, these same targets showed sustained PRC1 binding in CBX7-CaPS differentiation, and repression of these genes continued through day 8. We confirmed these findings in the CBX7-CaPS-R line using ChIP-qPCR, which demonstrated sustained RING1B binding to neural loci across differentiation when compared to WT (Figure 5E).

To further explore the impact of the CBX7-CaPS mutant on PRC1 binding to genes that normally become up regulated, we examined whether this trend extended to the supervised clusters of upregulated genes in WT cells that we identified in Figure 4C. In these clusters, 1074 PRC1 target genes were similarly upregulated in both WT and CBX7-CaPS cells (designated ‘U’) and 226 PRC1 targets were resistant to upregulation in CBX7-CaPS differentiation (designated ‘R’). We expected that genes that were upregulated in WT cells would have decreased RING1B occupancy after 6 days of differentiation, and we verified that trend (Figure 5F, compare lanes 1 and 3 with lanes 5 and 7). If resistance to upregulation in CBX7-CaPS cells was a result of prolonged retention of PRC1 binding during differentiation, we predicted that we would see greater RING1B occupancy on the resistant genes in CBX7-CaPS NPCs compared to WT NPCs at day 6. This pattern was observed (Figure 5F, compare lane 7 to lane 8). As was noted in EB formation, genes that we classified as ‘similarly’ upregulated in NPC differentiation still showed some resistance to activation when compared to their WT counterparts, but in a manner that was below our conservative cut-off (Figure S5C, compare lane 1 to lane 2). Therefore, these genes would also be predicted to have increased binding by PRC1, which was observed (Figure 5F,

compare lane 5 to lane 6), demonstrating a general sustained maintenance of PRC1 binding over time in CBX7-CaPS cells. Finally, we expected genes that resisted upregulation in the CBX7-CaPS cells to retain more RING1B than genes that were upregulated more normally in CBX7-CaPS cells, and this pattern was also validated (Fig. S5D, compare lane 2 to lane 4). We conclude that CBX7-CaPS cells maintain higher PRC1 occupancy on many PRC1 target genes that normally become upregulated in WT cells and that this effect is increased on the genes that most strongly resist upregulation, consistent with the repressive function of PRC1.

We also examined the supervised clusters (Fig. 4C) containing genes that were downregulated in WT cells but resisted this change in CBX7-CaPS cells and genes that were downregulated in both WT and CBX7-CaPS cells. Both groups of downregulated genes showed attenuated reduction of RING1B signal in CBX7-CaPS NPCs compared to WT (Figure S5E, compare lanes 1 and 3 to lanes 2 and 4). The genes that resisted downregulation showed a greater reduction in RING1B binding than those that were effectively repressed (Figure S5E, compare lane 2 to lane 4). The modest loss of RING1B signal on genes that resist downregulation might reflect a failure at these targets to transition to a new PRC1 complex containing different CBX proteins, and a corresponding failure to effect stable repression, or might reflect a role for PRC1 in activation of this subset of genes (Cohen et al., 2018; Loubiere et al., 2020).

DISCUSSION

Our findings advance the understanding of how chromatin modifying complexes play essential roles in governing the plasticity of cell fate. While stably differentiated cells must have mechanisms to maintain appropriate repressed and active states, those mechanisms must be tempered in pluripotent cells so that they can differentiate efficiently. We show that the ability of ESCs to properly differentiate is disrupted when they contain a cPRC1 that incorporates a CaPS domain. This domain functions to compact nucleosomal templates and to drive phase separation *in vitro*, two capabilities that lead to plausible mechanisms for maintenance of gene repression (Larson et al., 2017; Strom et al., 2017). This is consistent with the idea that bivalent chromatin in ESCs contributes to a fluid, “poised” state of gene regulation, and provides an explanation for why the CBX family of proteins in mammals has evolved to contain one member (CBX7) that lacks a positively charged and disordered domain that generates these key functions while the others, all primarily expressed in differentiated lineages, contain a domain with these characteristics. This work connects the function of the CBX2 CaPS domain to phenotypic outcome and illustrates the importance of this domain in the regulation of cellular plasticity.

In both directed and non-directed differentiation, CBX7-CaPS ESCs were impaired in their ability to differentiate, as shown phenotypically and by a resistance to change in gene expression in PRC1 targets. During neural differentiation, PRC1 was not redistributed appropriately and misregulated genes showed differences in PRC1 binding levels that correlated with their expression patterns, suggesting the defects in gene regulation were Polycomb dependent. Thus, the addition of the CBX2 CaPS domain to the PRC1 complex

in embryonic stem cells presents a hurdle for changing developmental gene expression programs involving both activation and repression.

We propose a model in which wild-type CBX7 plays a role in recruiting cPRC1 to target genes, via binding to the H3K27me3 modification, but does not allow cPRC1 to fully maintain the repressed state. This is consistent with recent work showing a larger effect of vPRC1 and its associated H2AK119 ubiquitylation activity on repression in ESCs than cPRC1 (Blackledge et al., 2020; Fursova et al., 2019; Tamburri et al., 2020). The functions of the CaPS domain might play key roles in cPRC1 mediated repression and/or memory later in development, while in pluripotent and multipotent cells that express CBX7 the absence of this domain might allow for a chromatin state that is poised for efficient activation upon receiving specific differentiation cues (Figure 6). By this model, replacement of WT CBX7 with the CBX7-CaPS protein may disrupt this poised state not by directly impacting ESC gene expression, but by impairing the ability of genes to change expression pattern when differentiation initiates.

The changes that occur when CBX7-CaPS is present include defects in both activation and repression of PRC1 target genes (Figure 6). CBX7-CaPS might stabilize cPRC1 on genes in ESCs, thereby occluding binding of activating transcription factors for lineage specific genes. Defects in genes achieving a fully repressed state might occur due to occluded binding by other repressive factors, via contributions of CBX7-CaPS to structural changes that promote activation (Loubiere et al., 2020), or due to failed redistribution of PRC1 because it is maintained on genes that are normally activated. Both compaction and phase separation increase the local concentration of cPRC1. These increases in local concentration might contribute to maintenance of CBX7-CaPS PRC1 complexes on target genes where WT CBX7 is normally removed during differentiation (Keenan et al., 2021; Strom et al., 2017). Such a mechanism would provide justification for a PRC1 complex that lacks these activities in pluripotent cells; in this stage of development, it is specifically important that the PcG repressed state not be remembered. Instead, it must be amenable to dynamic changes, including activation.

Why is the CaPS domain important to PRC1 function when other components of PRC1, such as PHC, generate phase separated structures and contribute to compaction of PcG targeted regions? For example, PRC1 with WT CBX7 in pluripotent cells is found in condensates due to other proteins, and perhaps nucleic acids, that drive condensate formation. An important mechanistic question concerns whether the presence of a CaPS domain in the CBX component of PRC1 alters the characteristics of condensates and/or the nature of the compacted state in a specific manner that contributes to the gain-of-function phenotypes described here or the loss-of-function phenotypes described previously with mutations in the CBX2 CaPS domain. Both phase separation and compaction can generate regions of increased concentration of components that have distinct characteristics. For example, different phase separated condensates are expected to differ in client molecules that partition into the condensate, hydration of the condensate and dynamics of the components of the condensate (Riback et al., 2020; Statt et al., 2020; reviewed in Shin and Brangwynne, 2017). The CaPS domain of the CBX family offers an opportunity to understand how different biophysical properties might contribute to biological function.

This work tested the hypothesis that it is important that CBX7 have limited function in pluripotent ESCs for differentiation to occur efficiently, and therefore that increasing the ability of cPRC1 in stem cells to compact chromatin and drive phase separation might be detrimental to differentiation. In studies of CBX2, which is primarily expressed in differentiated lineages, loss of function of the CaPS domain has substantive impact on axial patterning in mice that mirror the axial patterning defects of CBX2 null mice (Lau et al., 2017). While such work demonstrates the importance of CaPS domain activities in development, the studies we present here on the phenotypes of a gain-of-function mutation in CBX7 indicate that the addition of the CaPS domain might also be harmful in development and differentiation if present in inappropriate contexts.

Limitations of the study

In this work, we demonstrate that addition of a protein domain with known *in vitro* chromatin compaction and phase separation activities has an impact on gene regulation. However, we did not directly explore the effects of this domain on chromatin dynamics within stem cells or within cells undergoing differentiation. It would be interesting to study how the biophysical properties of ESC condensates, as shown in Figure 2, are altered with the addition of the CaPS domain. Similarly, in this work we do not directly investigate the degree of chromatin compaction within ESCs. Though there are technologies to investigate accessibility (e.g., ATAC-seq), these have not been optimized to study degrees of compacted chromatin, as they preferentially target accessible chromatin. As the field advances, it would be interesting to learn how observable measures of chromatin accessibility are impacted by the CaPS domain. Such experiments would help translate our *in vitro* characterizations of CaPS domain activities into its *in vivo* context and differentiate between how chromatin compaction and phase separation activities each contribute to regulation of gene expression.

STAR METHODS

Resource availability

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Robert Kingston (kingston@molbio.mgh.harvard.edu).

Materials availability—Materials generated in this study are available with a Material Transfer Agreement.

Data and code availability

- Sequencing data have been deposited on the GEO repository with code GEO: GSE151901 and will be publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Experimental Model and Subject Details

J1 mouse embryonic stem cells (male) were cultured at 37°C on 1% gelatin with PMEF-N feeder mouse embryonic fibroblasts (MEFs) (EMD Millipore) in DMEM media supplemented with 15% HyClone FBS (GE Healthcare), GlutaMax (Thermo Fisher Scientific), penicillin/streptomycin (Thermo Fisher Scientific), non-essential amino acids (Thermo Fisher Scientific), 1000U/mL Leukemia Inhibitory Factor (LIF), and 2-b-mercaptoethanol (Sigma). Protocols and reagents used to differentiate stem cells are described below.

Method Details

Chimeric protein design—The CaPS region insertions were selected based on previous work (Grau et al., 2011; Plys et al., 2019). The CaPS1 insertion contained CBX2 AA (amino acid) 93 through AA 240 and CaPS2 contained AA 93 through AA 415 (CaPS2). These regions were inserted at AA 80 of CBX7, after its chromodomain, mimicking the placement of the CaPS region in Cbx2.

Protein expression and purification—Chimeric CBX7 constructs were created using overlap PCR and cloned into the pFastBac plasmid with an N-terminal Flag tag. Plasmids were transfected into DH10Bac *E. coli* according to the Bac-to-Bac system (Thermo Fischer Scientific) and resulting bacmids were transfected using Cellfectin (Thermo Fischer Scientific) into SF9 cells. Virus was collected in the supernatant and used to infect SF9 cells to produce protein. The CBX proteins and chimeras were purified using a Flag-tag on the N-terminus of the protein and were stabilized by co-expression with RING1B, the direct CBX binding partner (Grau et al., 2011). SF9 cells expressing protein were pelleted by centrifugation and nuclear extracts were prepared as previously described (Abmayr et al., 2006). Cells were pelleted by centrifugation at 3000rpm for 10 minutes and washed once with PBS. Washed pelleted were rapidly resuspended in hypotonic buffer () at 5x packed cell volume (pcv) and centrifuged for 5 minutes at 3000 rpm, then resuspended in hypotonic buffer again to 3x the original pcv. Cells swelled on ice for 10 minutes, then were homogenized with 25 strokes in a glass Dounce homogenizer with a type B pestle. Cells were spun down for 15 minutes at 4000rpm. Nuclei were resuspended in ½ packed nuclear volume (pnv) low-salt buffer (20mM HEPES pH 7.9, 25% glycerol, 1.5mM $MgCl_2$, 1.2M KCl, 0.2mM EDTA, 0.2mM PMSF, 0.5mM DTT). While vortexing, the same volume of high-salt buffer (20mM HEPES pH 7.9, 25% glycerol, 1.5mM $MgCl_2$, 1.2M KCl, 0.2mM EDTA, 0.2mM PMSF, 0.5mM DTT) was added dropwise. Nuclei extracted with continuous gentle mixing at 4°C for 30 minutes. Nuclei were pelleted at 14,500rpm for 30 minutes, and supernatant containing nuclear extract was collected.

To purify protein, nuclear extracts were incubated with anti-Flag M2 resin (Sigma) for 2 hours at 4°C with rocking. Resin was collected by centrifugation and washed five times with BC600 buffer (20mM HEPES pH 7.6, 20% glycerol, 0.2mM EDTA, 600mM KCl) followed by four washes with BC300 buffer (20mM HEPES pH 7.6, 20% glycerol, 0.2mM EDTA, 300mM KCl). Protein was eluted using 0.8mg/mL Flag peptide solution. Fractions were examined on Coomassie stained gels, and those containing appropriate bands were pooled and dialyzed into BC300 buffer with 1mM DTT.

Preparation of histone octamers—HeLa chromatin was a kind gift from Michael Marr. Chromatin extraction was based on the method by Dignam et al. (1983). Pelleted HeLa cells were washed twice with PBS containing 5mM $MgCl_2$ and pelleted in a JA-14 rotor at 3000rpm for 10 minutes. Cells were resuspended in 10mM Tris pH 8, 10mM KCl, 1.5mM MgCl2, 1mM DTT, 1x protease inhibitor (Roche) and incubated on ice for 30 minutes. Sample was homogenized in a Dounce homogenizer with pestle B and pelleted in JA-17 rotor at 3500 rpm for 10 minutes. Nuclei were resuspended in 20mM Tris pH 7.5, 20% glycerol, 10% sucrose, 420mM KCl, 5mM $MgCl_2$, 0.1mM EDTA, 1mM DTT, 1x protease inhibitor. Sample was stirred on a magnetic stirrer for 30 minutes at 4°C. Lysate was centrifuged in a SW28 rotor at 100,000g for 1 hour. The chromatin pellet was resuspended in 5mL of resuspension buffer (50mM Tris pH 7.5, 0.1mM EDTA, 100mM NaCl, 0.1% Triton X-100, 5% glycerol, 1x protease inhibitor). Benzonase (10μL) was added and incubated on ice for 30 minutes. One-tenth volume of 5M NaCl was added and sample was sonicated until homogenous. Insoluble debris was pelleted at 15,000g for 30 minutes at 4°C. Concentrated sulfuric acid was added to the supernatant at a final concentration of 0.1M and histones were acid extracted with continuous mixing overnight at 4°C. Insoluble debris was pelleted at 15,000g for 30 minutes. Supernatant was neutralized by the addition of 1/5 volume of 1M Tris base. Extract was loaded onto a 5mL HiTrap SP HP column (Cytiva) equilibrated in 100mM NaCl in Buffer A (50mM Tris pH 8.0, 2mM EDTA, 5mM 2-mercaptoethanol), washed with 600mM NaCl in Buffer A, and eluted with a linear gradient from 600mM-2000mM NaCl in Buffer A. Histone-containing fractions were dialyzed in refolding buffer (2M NaCl, 10mM Tris, pH 7.5, 1mM EDTA, 5mM 2-mercaptoethanol). Histones were applied to a HiLoad 26_600 Superdex 200 gel filtration column (Cytiva) equilibrated in refolding buffer. Octamer fractions were combined and stored at 4°C.

Recombinant histones were purified based on the method by Dyer et al. (1999). Recombinant xenopus histones were expressed in *E. coli* from the following plasmids: pET28-H2A, pET28-H2B, pET28-H3, pET28-H4. Plasmids were transformed into BL21 DE3 cells and grown at 37°C in 2YT containing 50μg/ml antibiotic (kanamycin for pET28 plasmids, streptomycin for pCDF plasmid). Proteins were induced with 1mM IPTG and harvested after 3 hours. Cell pellets were resuspended and sonicated in wash buffer (50mM Tris pH 7.5, 100mM NaCl, 1mM PMSF, 1mM 2-mercaptoethanol). Inclusion bodies were pelleted at 48,000g. Pellet was resuspended in wash buffer containing 1% Triton X-100 and pelleted as before. Wash was repeated twice. Pellet was resuspended in SAUDE 200 (7M urea, 20mM sodium acetate pH 5.2, 200mM NaCl, 5mM 2-mercaptoethanol, 1mM EDTA). Sample was centrifuged at 48,000g for 15 minutes and supernatant was filtered through a 0.22 micron syringe filter. Sample was loaded onto a 5 ml HiTrap SP HP column (Cytiva) equilibrated in SAUDE 200. The column was washed with 10% SAUDE 1000 (7M urea, 20mM sodium acetate pH 5.2, 1000mM NaCl, 5mM 2-mercaptoethanol, 1mM EDTA) and protein was eluted with a linear gradient from 10–50% SAUDE 1000. Histone fractions were dialyzed overnight into 5mM 2-mercaptoethanol and then lyophilized.

Recombinant histone octamer was assembled as described previously (Dyer et. al., 2004), using histones prepared as described above. 4mg of each histone were dissolved in 2mL

unfolding buffer (6M guanidinium chloride, 20mM Tris pH 7.5, 5mM DTT). Equimolar amounts of each histone were combined and unfolding buffer was added to a final volume of 12 mL. Sample was dialyzed overnight at 4°C against three changes of 1L refolding buffer (2M NaCl, 10mM Tris, pH 7.5, 1mM EDTA, 5mM 2-mercaptoethanol). Precipitated protein was removed by centrifugation at 48,000g, 15 minutes. Supernatant was filtered through a 0.22 micron syringe filter. Sample was loaded onto a HiLoad 26_600 Superdex 200 gel filtration column (Cytiva) and equilibrated in refolding buffer. Octamer fractions were combined and stored at 4°C.

H3K27me3 Methyl-Lysine Analog (MLA) histones were modified as previously described (Simon et al., 2007). As described above for purification of recombinant histones, histone H3 K27C/C110A was expressed in *E. coli* from plasmid pCDF1b-6xHis H3 K27C, C110A. The histone protein was induced and purified as described above (Dyer et al., 1999). Once purified, 10mg of histone H3 K27C/C110A were dissolved in 1mL alkylation buffer (1M HEPES pH 7.8, 4M guanidinium chloride, 10mM D/L-methionine, 20mM DTT) in black tubes for 45 minutes at 37°C. 100mg (2-bromoethyl) trimethyl ammonium bromide was added and incubated at 50°C for 2.5 hours with intermittent mixing. Additional 10μL of 1M DTT was added and incubation continued for 2.5 hours. The reaction was quenched with 50μL 2-mercaptoethanol and incubated at room temperature for 30 minutes. Protein was dialyzed overnight into 5mM 2-mercaptoethanol, lyophilized and stored frozen at −80°C.

The 6xHistidine tag was cleaved from the H3Kc27me3 histone prior to octamer assembly. 4mg H3Kc27me3 and 4mg H4 were each dissolved in 2mL unfolding buffer (6M guanidinium chloride, 20mM Tris pH 7.5, 5mM DTT). H3/H4 dimer was dialyzed 4 hours at 4°C in 500mL cleavage buffer (20 mM Tris pH 8, 100 mM NaCl, 1mM EDTA, 1 mM EDTA, 5 mM 2-mercaptoethanol, 10% glycerol). Dialysis buffer was changed and 30μL TEV protease (1mg/mL) was added to dimer. Following overnight digestion at 4°C, an additional 30μL of TEV protease was added and incubated at room temperature for 1 hour. Cleaved H3Kc27me3/H4 dimer was dialyzed overnight at 4°C against three changes of 1L 5mM 2-mercaptoethanol. 4mg aliquots of dimer were lyophilized. H3Kc27me3 octamer was assembled by combining 8mg H3Kc27me3/H4 dimer with 4mg each of H2A and H2B dissolved in unfolding buffer. Dialysis and S200 gel filtration proceeded as described above for assembly of recombinant histone octamer.

Preparation of Cy5-labeled ligands—The G5E4 array was excised from pG5E4 by restriction digest using Asp718, ClaI, DdeI, and DraII. The resulting fragment was purified by PEG precipitation and end labeled using Klenow Fragment (New England Biolabs) to incorporate Cy5-dCTP.

Nucleosomal arrays were assembled using purified HeLa and H3K27me3 MLA core histones and a Cy5-labeled G5E4 nucleosome-positioning array by salt dialysis as previously described (Lee and Narlikar, 2001). Briefly, 25μL G5E4 DNA and 20μg of histone octamer were added to 2M KCl in Slide-A-Lyzer MINI Dialysis Devices (ThermoFisher Scientific). Samples were placed in high-salt buffer (20mM Tris-HCl pH 7.7, 2M KCl, 1mM EDTA, 1mM DTT), and a “rabbit” pump was set up to pump low-salt buffer (20mM Tris-HCl pH 7.7, 250mM KCl, 1mM EDTA, 1mM DTT) into the dialysis

beaker. Samples were dialyzed at 4°C for 40 hours, then dialyzed against zero-salt buffer (20mM Tris-HCl pH 7.7, 1mM EDTA, 1mM DTT) for 3 hours. Arrays were digested by EcoRI to assess the extent of nucleosome occupancy on the positioning sequences.

Restriction Enzyme Accessibility assay—Reactions took place in a final volume of 20μL with 16mM HEPES pH 7.5, 80mM KCl, 2mM $MgCl_2$, 1mM $MnCl_2$, 2mM ATP, 2.5μg BSA, 8%w/v glycerol, and 2mM DTT. CBX proteins were added as serial dilutions at indicated final concentrations. 240ng of Cy5-labeled nucleosome arrays that were 85–95% assembled were added, and the reaction was incubated at 30°C for 30 minutes. 4μL HhaI (20U/μL) was added, with or without SWI/SNF at a final concentration of 5nM. The reaction was incubated at 30°C for 1 hour. 5μL stop buffer (1.5μg/μL Proteinase K, 10mM Tris-HCl pH 7.5, 35mM EDTA, 1% SDS, 0.1% orange G) was added to the reactions, and they were incubated at 55°C for 1 hour. Reactions were run on 1% TAE gel, imaged on a Typhoon scanner, and quantified using ImageJ software. Apparent inhibition of remodeling was determined by the equation:

$$\frac{\%\ uncut\ with\ PCG\ and\ SWI/SNF - \%\ uncut\ with\ SWI/SNF\ only}{\%\ uncut\ without\ SWI/SNF - \%\ uncut\ with\ SWI/SNF\ only} * 100$$

Curves on graphical representation of inhibition data were generated using best fit polynomial equation in R.

Electron microscopy—G5E4 nucleosomal arrays were assembled as described above, but dialyzed into 20mM HEPES pH 7.9, 0.1mM EDTA pH 8.0. For each protein tested, 3.3nM nucleosomal array (~40nM total nucleosomes, assuming 12 nucleosomes per G5E4 array) was mixed with 10nM protein in a buffer containing 20mM Hepes pH 7.9, 30mM KCl 10% glycerol, 0.1mM EDTA and 0.001% NP-40. Each reaction was crosslinked with 0.1% glutaraldehyde on ice for 1 hour and then quenched with 100 mM Tris. The reactions were then dialyzed against 20mM HEPES pH 7.9, 0.1mM EDTA and 50mM NaCl overnight.

Continuous carbon grids were glow discharged using an Emitech K100X glow discharge system for 30 seconds at −20mA. 3 μL of sample was applied to the glow discharged side for 1 minute, blotted, washed with water twice and then stained with 2% uranyl acetate for 1 minute. Grids were imaged on an FEI Morgagni 268 microscope operating at 80 keV with a nominal magnification of 51,000 and a pixel size of 4.66 Å.

For each sample, approximately 50 well defined arrays were imaged. The length of the array was determined in ImageJ by measuring diameter of a circle encompassing the entire array. The number of particles in each array was determined by counting distinct entities while blinded to the identity of the samples. Array length and particle number were determined by separate researchers. Statistical significance was determined in R using a Wilcoxon test.

Turbidity assay—Proteins were diluted preliminarily in 300mM KCl buffer. Buffer was loaded into a clear-bottom 384-well plate (Corning) to result in a final concentration of 20 mM HEPES pH 7.9, 100 mM KCl, and 1 mM $MgSO_4$ after addition of proteins. Protein

was added to the buffer using a multi-channel pipette to minimize differences in incubation time. Solutions were incubated for 5 minutes at room temperature. To assay for turbidity, absorbance at 405nm was measured using a Spectramax M3 plate reader.

Droplet formation assay—Proteins were diluted to specified concentrations into a buffer containing final concentrations of 20 mM HEPES pH 7.9, 100 mM KCl, and 1 mM $MgSO_4$. In instances when ligands were included in the assay, they were added to a final concentration of 100nM. Ligands were added to buffer before the addition of protein. The proteins were loaded onto glass slides with coverslips and imaged with a Nikon 90i Eclipse epifluorescence microscope equipped with an Orca ER camera (Hamamatsu) using a 100X oil objective and Volocity software (Perkin Elmer). Images in figures were prepared using Fiji software.

Cell Culture and Differentiation—J1 mouse embryonic stem cells were cultured on 1% gelatin with PMEF-N feeder mouse embryonic fibroblasts (MEFs) (EMD Millipore) in DMEM media supplemented with 15% HyClone FBS (GE Healthcare), GlutaMax, penicillin/streptomycin, non-essential amino acids, 1000U/mL Leukemia Inhibitory Factor (LIF), and 2-b-mercaptoethanol. To test for alkaline phosphatase activity, ESCs were stained two days after passaging. Staining was performed according to manufacturer instructions using Stemgent® Alkaline Phosphatase Staining Kit II (Reprocell). For cellular growth assays, 3×10^5 ESCs were seeded per well in 6-well plates on feeder cells. Cells were dissociated from individual wells 24-, 48-, and 72-hours after initial plating and counted using an automated cell counter (Bio-Rad).

Neural progenitors were formed as previously described (Bibel et al., 2007). Briefly, ESCs were de-MEFed over two passages, then 4×10^6 cells were plated in low-adherent 10cm dishes in 15mL CA medium (DMEM with 10% FCS, 2 mM L-glutamine, non-essential amino acids, penicillin/streptomycin, and 2-b-mercaptoethanol). Media was changed once after 2 days. Four days after the initial transfer to CA media, media was changed again and retinoic acid was added to a final concentration of 5μM. Media was changed again after two more days (at which point day 6 NPCs were collected), once again adding retinoic acid, and on day 8 cells were collected. To form embryoid bodies, embryonic stem cells were de-MEFed, and resuspended in differentiation media (IMDM supplemented with 15% HyClone FBS, GlutaMax, and penicillin/streptomycin) at a concentration of 8,400 cells/mL. Drops were set at 50μL per drop on the bottom of a non-adherent tissue culture plate, and plates were turned upside down. After 4 days, embryoid bodies were pooled together and grown in suspension in non-adherent plates, changing media every day. EBs were collected for analysis on days 6 and 10 of differentiation. EBs were imaged using an Axio Zoom microscope (Zeiss) at 50x magnification. EB size was determined by measuring the area of individual EBs from these images using ImageJ, and statistical significance of differences in size was determined using a Wilcoxon test in R. For the beating heart assay, on day 4 of differentiation, one drop each containing an embryoid body was placed in the wells of 24-well plates coated with 0.2% gelatin. Media was changed daily, and beating clusters were counted on days 12, 13, and 14 of differentiation. Statistical significance was determined using a paired T test.

CRISPR generation of cell lines—CRISPR gRNA for CaPS knock-in was designed by identifying the closest PAM sequence (5'-NGG) to the intended insertion site and using the 20bp sequence immediately upstream of it as the guide, as described in Ran et al. (2013). CRISPR gRNAs for CBX7 knock-out cell lines were generated using predesigned guide RNAs from IDT, selecting the two highest-ranking guides from two different exons. Alt-R® gRNAs were ordered from IDT, along with universal tracrRNA. For CBX7-CaPS knock-in cloning using homology-directed repair (HDR), a donor plasmid was used with 900bp homology arms.

To assemble the CRISPR-RNP complex (Kim et al., 2014), 10μL each of 200μM gRNA and tracrRNA were hybridized by gradual cooling from 95°C. 5μL of hybridized RNA was added to 50μg of Cas9 protein and incubated at room temperature for 25 minutes. Before transfection, 100ng linear puromycin marker and 1μg HDR donor plasmid (where necessary) were added to the solution. During CRISPR-RNP complex assembly, J1 ESCs were de-MEFed. ESCs were collected by centrifugation, and 1×10^6 cells per transfection were resuspended in 100μL Mouse ES Cell Nucleofector® Solution (Amaxa) with CRISPR-RNP solution. Cells were nucleofected using Amaxa Nucleofector® II program A-30 and plated on puromycin-resistant DR4 MEFs (Fischer Scientific). Cells were selected using 1μg/mL puromycin for two days, then grown in ES cell media. Individual clones were picked, dissociated with Trypsin, and plated into individual wells of 24-well plates seeded with MEFs. Clones were screened for CaPS insertion using PCR, and confirmed using Sanger sequencing and Western blot. CBX7 knock-out clones were screened by Western blot and confirmed by next-generation amplicon sequencing.

Nuclear protein extraction—For Western blot analysis, proteins were extracted from nuclei. ESCs were resuspended in 5mL buffer containing 15mM Tris-HCl pH 8.0, 15mM NaCl, 60mM KCl, 1mM EDTA pH 8.0, 0.5mM EGTA, 0.5mM spermidine, 0.3mM spermine, 1mM DTT, and Protease Inhibitor Complex (PIC) (Roche). To lyse cells, 5mL of the same buffer but with 0.08% NP-40 was added dropwise and the solution was mixed by inversion and incubated on ice for 10 minutes. Nuclei were collected by centrifugation, counted, and washed once in the original resuspension buffer. Nuclei were resuspended in protein extraction buffer (50mM Tris-HCl pH 7.5, 300mM NaCl, 0.5mM EDTA pH 8.0, 1% Triton X-100, PIC, 1mM DTT) at a volume of 20μL buffer per 10^6 nuclei and incubated on ice for 5 minutes. Extracts were centrifuged at 12,000g for 10 min at 4°C and supernatant was collected for Western blot analysis.

Immunofluorescence—ESCs were seeded on glass coverslips with gelatin in 12-well plates overnight. Coverslips were washed in wells with PBS, then fixed with 4.0% paraformaldehyde for 10 minutes at room temperature. Coverslips were washed once with PBS, then permeabilized with 0.2% Triton X-100/PBS. Cells were blocked with incubation solution (3% Bovine Serum Albumin (BSA) (Sigma) in 0.05% Triton X-100/PBS) for 1 hour at room temperature. Coverslips were incubated with primary antibody in incubation solution overnight at 4°C, washed with 0.05% Triton X-100/PBS, and incubated in the dark with secondary antibody in incubation solution for 2 hours at room temperature. Coverslips were washed, shielded form light, with 0.05% Triton X-100/PBS, followed by a short PBS

wash and a short wash with ddH_2O. Coverslips were mounted on slides with Vectashield® mounting media containing DAPI (Vector Laboratories). Cells were imaged with a Nikon 90i Eclipse epifluorescence microscope equipped with an Orca ER camera (Hamamatsu) using a 100X oil objective and Volocity software (PerkinElmer). Images in figures were prepared using Fiji software.

RIME—RIME was performed on ~10^8 frozen, crosslinked embryonic stem cells as previously described (Mohammed et al., 2016), with some modifications. Cells were crosslinked as described below for ChIP-Seq. Cells were resuspended in 10mL LB1 (50mM HEPES-KOH pH 7.5, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% Igepal CA-630, 0.25% Triton X-100) and rotated at 4°C for 10 minutes. Nuclei were pelleted at 2000g for 5 minutes at 4°C, then resuspended in 10mL LB2 (10mM Tris-HCl pH 8.0, 200mM NaCl, 1mM EDTA, 0.5mM EGTA). Samples were rotated at 4°C for 5 minutes, then centrifuged at 2000g for 5 minutes at 4°C. Nuclei were resuspended in 300μL LB3 (10mM Tris-HCl pH 8.0, 100mM NaCl, 1mM EDTA, 0.5mM EGTA, 0.1% sodium deoxycholate, 0.5% N-lauroylsarcosine) per 2×10^7 cells and sonicated using a QSonica Q800R3 Sonicator as described below for ChIP-Seq. Triton X-100 was added to a final concentration of 1%, and lysate was clarified by centrifugation at 20,000g for 10min at 4°C. 460μL protein A Dynabeads (Invitrogen) were incubated with 10μg antibody in PBS with 5mg/mL BSA for 1 hour at room temperature with rotation. Beads were washed once in PBS/BSA, then added to cleared cell lysate. Samples were rotated at 4°C overnight, then washed ten times in RIPA buffer (50mM HEPES pH 7.6, 1mM EDTA, 0.7% sodium deoxycholate, 1% NP-40, 0.5M LiCl). After washing, beads were boiled in 2x Laemmli buffer for 10 minutes, and the supernatant was loaded onto a polyacrylamide gel. The protein gel was stained with Coomassie, and lanes were cut into four segments each for mass spectrometry analysis. In-gel digestion, micro-capillary LC/MS/MS analysis, protein database search and data analysis were performed by the Taplin Mass Spectrometry Facility at Harvard Medical School.

RNA-Seq—ESCs were de-MEFed before collection for RNA-Seq. RNA was isolated from pelleted cells according to manufacturer instructions using the PureLink™ RNA Mini Kit (Invitrogen). Samples were subsequently treated with DNase using the TURBO DNA-*free*™ Kit (Invitrogen). Ribosomal RNA was depleted using NEBNext® rRNA Depletion Kit (New England Biolabs), and cDNA was generated using NEBNext® Ultra II Directional RNA First and Second Strand Synthesis Modules (New England Biolabs), and resuspended in 10uL of ddH_2O.

Sequencing libraries were prepared from cDNA as previously described (Bowman et al., 2013). End repair was performed by incubating 10μL of sample with 2μL 10mM dNTPs, 5μL T4 ligase buffer, 2.5μL T4 polymerase, 0.5μL Klenow polymerase, 2.5μL T4 PNK, and 27.5μL ddH_2O (all enzymes from New England Biolabs), in a thermocycler at 20°C for 30 minutes. AMPure SPRI bead (Beckman Coulter) cleanup was performed at a 1.8X bead ratio, and samples were eluted in 16μL ddH_2O. For A-tailing, 2.5μL NEB Buffer #2, 5μL 1mM dATP, and 1.5μL Klenow 3’−5’ exo minus (New England Biolabs) were added to samples, and the reaction was incubated at 37°C for 30 minutes. AMPure SPRI bead cleanup was performed at a 1.8X bead ratio, and samples were eluted in 9μL

ddH_2O. Adapters compatible with Illumina sequencing were ligated by adding 12.5μL 2X rapid T4 Ligase Buffer (Enzymatics), 2.5μL T4 rapid ligase (Enzymatics), 0.2μL 1μM universal adapters, and 0.8μL ddH_2O to samples and incubating the reaction at room temperature for 15 minutes. AMPure SPRI bead cleanup was performed at a 1.6X bead ratio, and samples were eluted in 10μL ddH_2O. Libraries were amplified by a qPCR reaction containing 25μL Phusion master mix (New England Biolabs), 1uL Illumina universal primer (10μM), 1uL Illumina barcoded primer (10μM), 1μL SYBR Green (1:2000 dilution of stock; Thermo Fisher), 10μL sample, and 11.5μL ddH_2O. Reaction was stopped at the end of the logarithmic phase of the reaction, and samples were cleaned with AMPure SPRI beads at 1.2X ratio. Samples were resuspended in 20μL ddH_2O, and further diluted and pooled to equimolar amounts for sequencing.

Chromatin Immunoprecipitation (ChIP)—Ten to twenty million harvested cells were crosslinked in 1% formaldehyde-PBS for 10 minutes at room temperature. Crosslinking was quenched with glycine, and cells were pelleted by centrifugation at 4°C and washed once in cold PBS. Cells were resuspended in cold lysis buffer (50mM Tris-HCl pH 8.1, 10mM EDTA pH 8.0, 1% SDS, PIC, 1mM DTT) and incubated for 10 minutes on ice prior to sonication using a QSonica Q800R3 Sonicator. Insoluble chromatin was pelleted by centrifugation, and supernatant was diluted 1:10 with IP Dilution Buffer (20mM Tris-HCl pH 8.1, 150mM NaCl, 2mM EDTA pH 8.0, 0.1% Triton X-100, PIC, 1mM DTT). Sonicated *Drosophila* chromatin was added at a known ratio to experimental chromatin to serve as a spike-in as previously described (Egan et al., 2016). Triton X-100 was added to a final concentration of 1%, input was collected and diluted to 200μL, and chromatin was added to Protein A Dyanbeads (Invitrogen) that had been previously incubated with experimental and spike-in antibodies. Samples were rotated at 4°C overnight. Beads were washed with IP Dilution Buffer with 1% Triton X-100, followed by Wash Buffer A (50mM HEPES pH 7.9, 500mM NaCl, 1mM EDTA pH 8.0, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS, PIC, DTT), followed by Wash Buffer B (20mM Tris-HCl pH 8.0, 250mM LiCl, 1mM EDTA pH 8.0, 0.1% Na-deoxycholate, 1% Igepal, PIC, DTT), followed by a quick TE wash. Beads were incubated at 37°C in 200μL Extraction Buffer (1% SDS, 0.1M $NaHCO_3$) for 15 minutes with shaking. Supernatant was collected, and 12μL 2M TrisHCl, 2μL RNase A (500μg/mL, Roche), and 12μL 5M NaCl were added to ChIP samples and inputs. Crosslinks were reversed overnight at 65°C, followed by 1 hour incubation at 55°C with Proteinase K. DNA was isolated by phenol:chloroform extraction and ethanol precipitation. For ChIP-Seq, sequencing libraries were prepared as described above for RNA-Seq (Bowman et al., 2013).

CUT&RUN—CUT&RUN was performed as previously described, with described modifications for high-calcium/low-salt conditions (Meers et al., 2019; Skene et al., 2018). One million fresh, non-crosslinked cells per IP were washed twice in wash buffer (20mM HEPES pH 7.5, 150mM NaCl, 1x EDTA-free protease inhibitor), then resuspended in 1mL wash buffer with bead slurry (Concanavalin A-coated beads in 20mM HEPES-KOH pH 7.9, 10mM KCl, 1mM $CaCl_2$, 1mM $MnCl_2$). Cells were rotated for 10 minutes at room temperature, then divided into separate tubes and placed on magnetic stands to remove the supernatant. 50μL of antibody buffer (20mM HEPES pH 7.5, 150mM NaCl. 0.1% digitonin, 2mM EDTA, 1x protease inhibitor) containing the appropriate antibody (at a

1:100 dilution) was added, and samples were incubated on ice for 2 hours with regular shaking. Supernatants were removed, and samples were washed once with wash buffer. 50μL of wash buffer was added, followed by pA-MNase to a final concentration of 700ng/mL. Samples were incubated on ice for 1 hour with regular shaking, then supernatant was removed and samples were washed twice with wash buffer. Beads were washed once in rinse buffer (20mM HEPES pH 7.5, 1x protease inhibitor), then 100μL ice-cold incubation buffer (3.5mM HEPES pH 7.5, 10mM CaCl2) was added and samples were incubated on ice for 30 minutes with regular shaking. Supernatant was removed and 200μL stop buffer (170mM NaCl, 20mM EGTA, 20μg/mL glycogen, 25μg/mL RNase A) was added. Beads were incubated at 37°C for 30 minutes, supernatant was collected, and DNA was extracted by phenol:chloroform extracted. Libraries were prepared as described above for RNA-Seq (Bowman et al., 2013).

Quantification and Statistical Analysis

Processing and analysis of RNA-Seq data—To analyze RNA-Seq data, adapter sequences were removed from reads using Cutadapt v1.14 with "--minimum-length 30" (Martin, 2011). Reads were aligned to the mm10 genome using STAR v2.7.0f (Dobin et al., 2013). Files were converted to BAM format using Samtools v1.9 (Li et al., 2009). Tag directories and bigwig files were generated using HOMER v4.10 (Heinz et al., 2010). Reads were counted using featureCounts (Liao et al., 2014). Differentially expressed genes were identified from three replicates using the edgeR package in R Studio, and batch effects from different differentiation time courses were accounted for (Robinson et al., 2010).

Figures were generated in RStudio using data generated from edgeR analysis. Genes were filtered to PRC1 targets using RING1B ChIP-Seq peaks from ESCs (see below). For changes in gene expression over time, genes were considered up- or down-reg*ul*ated if they showed an FDR < .05 and at least 2-fold change in expression over time as determined in EdgeR. Supervised heatmap clusters were generated by using increasingly stringent cutoffs, from 2-fold to 6-fold changes in gene expression over time. Clusters were designed to address whether gene expression changes over time were similar, greater, or reduced in CBX7-CaPS compared to WT cells. To identify genes that were differentially regulated over time, the cutoff for up- or down-regulation was made less stringent for each cluster to minimize artificial differences between cell lines. For example, genes that were 2-fold upregulated in WT were considered upregulated in CaPS if they changed expression at least 1.5-fold, and genes that were 4-fold upregulated in WT were considered similarly upregulated in CaPS if they changed expression at least 3-fold.

Gene set enrichment analysis was performed using the GenePattern platform using MSigDB gene sets (Ashburner et al., 2000; Liberzon et al., 2011; Mootha et al., 2003; Subramanian et al., 2005). Gene ontology analysis was performed on differentially regulated genes using the "GO_Biological_Process_2018" database in enrichR in R Studio (Chen et al., 2013; Kuleshov et al., 2016). For boxplots, statistical significance was determined in R using a Wilcoxon test.

Processing and analysis of ChIP-Seq and CUT&RUN data—To analyze ChIP-Seq data, adapter sequences were removed from reads using Cutadapt v1.14 with "--minimum-length 30" (Martin, 2011). Reads were aligned to the mm10 and dm6 genomes using Bowtie 2 v2.3.4.3 with one mismatch permitted per seed, a maximum insert of 2kb, and the "--no-discordant" option specified (Langmead and Salzberg, 2012). Files were converted to BAM format and PCR duplicates were removed using Samtools v1.9 (Li et al., 2009). To normalize ChIP-Seq data, mm10 reads were randomly downsampled by a factor determined by the formula:

$$Downsampling\ Factor = \frac{Drosophila\ Reads\ in\ Input}{Mouse\ Reads\ in\ Input} * \frac{Mouse\ Reads\ in\ IP}{Drosophila\ Reads\ in\ IP} * \alpha$$

where α represents the number that results in the largest downsampling factor equaling 1. Including the ratio of *Drosophila* to mouse reads in the input was intended to account for minor variations in spike-in mixing. This normalization strategy has been previously described (Fursova et al., 2019; Hu et al., 2015). Reads were downsampled using Samtools, and tag directories and bigwig files were generated using HOMER v4.10 (Heinz et al., 2010). CUT&RUN data processing was performed as in ChIP-Seq, but data was normalized to total read number by random downsampling. Peaks were called and annotated on normalized tag directories using HOMER by comparing to input, with "-style histone" and "-minDist 10000". Confident peaks were identified by filtering to a minimum tags per ten million (RPTM) value based on enrichment of the peak over background in IGV browser view. This cutoff was generally near 10RPTM. Peaks were also filtered to be within 10kb of their associated TSS.

All ChIP-Seq figures were generated from confident peaks as described above, using replicates as described in figure legends. Quantitative comparisons were only made between experiments performed at the same time. Heatmaps were created using deepTools computeMatrix and plotHeatmap (Ramírez et al., 2016). Comparisons of tags at ChIP-Seq peaks were made in R using DiffBind, normalizing for background signal (Stark and Brown, 2011). A P-Value < 0.05 was considered statistically significant.

Gene ontology analysis was performed on genes with 2-fold enriched ChIP-Seq signal in a given cell type using the "GO_Biological_Process_2018" database in enrichR in R Studio (Ashburner et al., 2000; Chen et al., 2013; Kuleshov et al., 2016). For boxplots, statistical significance was determined in R using a Wilcoxon test.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank the Kingston lab, C. Kadoch, J. Lee, and F. Winston for helpful discussions on this project; J. Kim for critical reading of the manuscript; the G. Ruvkin and J. Lee labs for help with and access to microscopes; and R. Tomaino at the Taplin Mass Spectrometry facility for mass spectrometry analysis. This work was supported by a graduate fellowship from the Department of Defense (to E.S.J.) and by NIH grants R01GM043901 and R35GM131743 (to R.E.K).

REFERENCES

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet 25, 25–29. [PubMed: 10802651]

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal Al., Feil R, Schreiber SL, and Lander ES (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125, 315–326. [PubMed: 16630819]

Bernstein E, Duncan EM, Masui O, Gil J, Heard E, and Allis CD (2006). Mouse Polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. Mol. Cell. Biol. 26, 2560–2569. [PubMed: 16537902]

Bibel M, Richter J, Lacroix E, and Barde YA (2007). Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. Nat. Protoc. 2, 1034–1043. [PubMed: 17546008]

Blackledge NP, Farcas AM, Kondo T, King HW, McGouran JF, Hanssen LL, Ito S, Cooper S, Kondo K, Koseki Y, Ishikura T, Long HK, Sheahan TW, Brockdorff N, Kessler BM, Koseki H, Klose RJ (2014). Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and Polycomb domain formation. Cell 157(6):1445–1459. [PubMed: 24856970]

Blackledge NP, Fursova NA, Kelley JR, Huseyin MK, Feldmann A, and Klose RJ (2020). PRC1 catalytic activity is central to Polycomb system function. Mol. Cell 77, 857–874. [PubMed: 31883950]

Bowman SK, Simon MD, Deaton AM, Tolstorukov M, Borowsky ML, and Kingston RE (2013). Multiplexed Illumina sequencing libraries from picogram quantities of DNA. BMC Genomics 14, 466. [PubMed: 23837789]

Boyle S, Flyamer IM, Williamson I, Sengupta D, Bickmore WA, Illingworth RS (2020). A central role for canonical PRC1 in shaping the 3D nuclear landscape. Genes Dev. 34, 931–949. [PubMed: 32439634]

Brookes E, de Santiago I, Hebenstreit D, Morris KJ, Carroll T, Xie SQ, Stock JK, Heidemann M, Eick D, Nozaki N, Kimura H, Ragoussis J, Techmann SA, and Pombo A (2012). Polycomb associates genome-wide with a specific RNA polymerase II variant, and regulates metabolic genes in ESCs. Cell Stem Cell 10, 57–170.

Chan HL, and Morey L (2019). Emerging roles for Polycomb-group proteins in stem cells and cancer. Trends Biochem. Sci. 44, 788–700.

Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, and Ma'ayan A (2013). EnrichR: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics 14, 128. [PubMed: 23586463]

Cohen I, Zhao D, Bar C, Valdes VJ, Dauber-Decker KL, Nguyen MB, Nakayama M, Rendl M, Bickmore WA, Koseki H, Zheng D, Ezhkova E (2018). PRC1 fine-tunes gene repression and activation to safeguard skin development and stem cell specification. Cell Stem Cell. 22, 726–739. [PubMed: 29727681]

Cooper S, Dienstbier M, Hassan R, Schermelleh L, Sharif J, Blackledge NP, De Marco V, Elderkin S, Koseki H, Klose R, Heger A, Brockdorff N (2014). Targeting Polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119 in PRC2 recruitment. Cell Rep. 7, 1456–1470. [PubMed: 24857660]

Dignam JD, Lebovitz RM, Roeder RG (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. 11, 1475–1489. [PubMed: 6828386]

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. [PubMed: 23104886]

Du Z, Zheng H, Kawamura YK, Zhang K, Gassler J, Powell S, Xu Q, Lin Z, Xu K, Zhou Q, Ozonov EA, Véron N, Huang B, Li L, Yu G, Liu L, Yeung WKA, Wang P, Chang L, Wang Q, He A, Sun

Y, Na J, Sun Q, Sasaki H, Tachibana K, Peters AHFM, Xie W (2020). Polycomb group proteins regulate chromatin architecture in mouse oocytes and early embryos. Mol. Cell 77, 825–839.e7. [PubMed: 31837995]

Dyer PN, Edayathumangalam RS, White CL, Bao Y, Chakravarthy S, Muthurahan UM, Luger K (2004). Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzymol. 375, 23–44. [PubMed: 14870657]

Egan B, Yuan CC, Craske ML, Labhart P, Guler GD, Arnott D, Maile TM, Busby J, Henry C, Kelly TK, Tindell CA, Jhunjhunwala S, Zhao F, Hatton C, Bryant BM, Classon M, and Trojer P (2016). An alternative approach to ChIP-Seq normalization enables detection of genome-wide changes in histone H3 lysine 27 trimethylation upon EZH2 inhibition. PLoS One 11, e0166438. [PubMed: 27875550]

Forzati F, Federico A, Pallante P, Abbate A, Esposito F, Malapelle U, Sepe R, Palma G, Troncone G, Scarfò M, Arra C, Fedele M, and Fusco A (2012). CBX7 is a tumor suppressor in mice and humans. J. Clin. Invest. 122, 612–623. [PubMed: 22214847]

Francis NJ, Kingston RE, and Woodcock CL (2004). Chromatin compaction by a Polycomb group protein complex. Science 306, 1574–1577. [PubMed: 15567868]

Fursova NA, Blackledge NP, Nakayama M, Ito S, Koseki Y, Farcas AM, King HW, Koseki H, and Klose RJ (2019). Synergy between variant PRC1 complexes defines Polycomb-mediated gene repression. Mol. Cell 74, 1020–1036. [PubMed: 31029541]

Gibson BA, Doolittle LK, Schneider MWG, Jensen LE, Gamarra N, Henry L, Gerlich DW, Redding S, and Rosen MK (2019). Organization of chromatin by intrinsic and regulated phase separation. Cell 179, 470–484. [PubMed: 31543265]

Grau DJ, Chapman BA, Garlick JD, Borowski M, Francis NJ, and Kingston RE (2011). Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. Genes Dev. 25, 2210–2221. [PubMed: 22012622]

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, and Glass CK (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophase and B cell identities. Mol. Cell 38, 576–589. [PubMed: 20513432]

Hu B, Petela N, Kurze A, Chan KL, Chapard C, and Nasmyth K (2015). Biological chromodynamics: a general method for measuring protein occupancy across the genome by calibrating ChIP-Seq. Nucleic Acids Res. 43, e132. [PubMed: 26130708]

Isono K, Endo TA, Ku M, Yamada D, Suzuki R, Sharif J, Ishikura T, Toyoda T, Bernstein BE, and Koseki H (2013). SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. Dev. Cell 26, 565–577. [PubMed: 24091011]

Keenan MM, Brown D, Brennan LD, Renger R, Khoo H, Carlson CR, Huang B, Grill SW, Narlikar GJ, Redding S (2021). HP1 proteins compact DNA into mechanically and positionally stable phase separated domains. Elife 10, e64563. [PubMed: 33661100]

Kent S, Brown K, Yang C, Alsaihati N, Tian C, Wang H, Ren X (2020). Phase-separated transcriptional condensates accelerate target-search process revealed by live-cell single-molecule imaging. Cell Rep. 33, 108248. [PubMed: 33053359]

Kim S, Kim D, Cho SW, Kim J, and Kim JS (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res. 24, 1012–1019. [PubMed: 24696461]

King IFG, Emmons RB, Francis NJ, Wild B, Müller J, Kingston RE, and Wu CT (2005). Analysis of a Polycomb group protein defines regions that link repressive activity on nucleosomal templates to in vivo function. Mol. Cell Biol. 25, 6578–6591. [PubMed: 16024794]

Klauke K, Radulović V, Broekhuis M, Weersing E, Zwart E, Olthof S, Ritsema M, Burggeman S, Wu X, Helin K, Bystrykh L, de Haan G (2013). Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. Nat. Cell Biol. 15, 353–362. [PubMed: 23502315]

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, and Ma'ayan A

(2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 44, W90–97. [PubMed: 27141961]

Kuroda MI, Kang H, De S, Kassis JA (2020). Dynamic competition of Polycomb and Trithorax in transcriptional programming. Annu. Rev. Biochem. Advance online publication; 10.1146/annurev-biochem-120219-103641.

Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, and Reinberg D (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev. 16, 2893–2905. [PubMed: 12435631]

Kundu S, Ji F, Sunwoo H, Jain G, Lee JT, Sadreyev RI, Dekker J, and Kingston RE (2017). Polycomb Repressive Complex 1 generates discrete domains that change during differentiation. Mol. Cell 65, 432–446. [PubMed: 28157505]

Lau MS, Schwartz MG, Kundu S, Savol AJ, Wang PI, Marr SK, Grau DJ, Schorderet P, Sadreyev RI, Tabin CJ, and Kingston RE (2017). Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning. Science 355, 1081–1084. [PubMed: 28280206]

Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. [PubMed: 22388286]

Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, Agard DA, Redding S, Narlikar GJ (2017). Liquid droplet formation by HP1a suggests a role for phase separation in heterochromatin. Nature 547, 236–240. [PubMed: 28636604]

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079. [PubMed: 19505943]

Liao Y, Smyth GK, and Shi W (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930. [PubMed: 24227677]

Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics 27, 1739–1740. [PubMed: 21546393]

Loubiere V, Papadopoulos GL, Szabo Q, Martinez A-M, Cavalli G (2020). Widespread activation of developmental gene expression characterized by PRC1-dependent chromatin looping. Sci Adv. 6, eaax4001. [PubMed: 31950077]

Ma RG, Zhang Y, Sun TT, and Cheng B (2014). Epigenetic regulation by Polycomb group complexes: focus on roles of CBX proteins. J. Zhejiang Univ. Sci. B 15, 12–428.

Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ 3rd, Voigt P, Martin SR, Taylor WR, De Marco V, Pirrotta V, Reinberg D, Gamblin SJ (2009). Role of the Polycomb protein EED in the propagation of repressive histone marks. Nature 461,762–767. [PubMed: 19767730]

Martin M (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10–12.

Meers MP, Bryson TD, Henikoff JG, and Henikoff S (2019). Improved CUT&RUN chromatin profiling tools. Elife 8, e46314. [PubMed: 31232687]

Min IM, Waterfall JJ, Core LJ, Munroe RJ, Schimenti J, and Lis JT (2011). Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. Genes Dev. 25, 42–754.

Mohammed H, Taylor C, Brown GD, Papachristou EK, Carroll JS, and D'Santos CS (2016). Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. Nat. Protoc. 11, 316–326. [PubMed: 26797456]

Mootha VK, Lindgren CM, Eriksson K-F, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstråle M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, Groop LC (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. 34, 267–273. [PubMed: 12808457]

Morey L, Pascual G, Cozzuto L, Roma G, Wutz A, Benitah SA, and Di Croce L (2012). Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. Cell Stem Cell 10, 47–62. [PubMed: 22226355]

O'Loghlen A, Muñoz-Cabello AM, Gaspar-Maia A, Wu HA, Banito A, Kunowska N, Racek T, Pemberton HN, Beolchi P, Lavial F, Masui O, Vermeulen M, Carroll T, Graumann J, Heard E,

Dillon N, Azuara V, Snijders AP, Peters G, Bernstein E, and Gil J (2012). MicroRNA regulation of Cbx7 mediates a switch of Polycomb orthologs during ESC differentiation. Cell Stem Cell 10, 33–46. [PubMed: 22226354]

Pasini D, and Di Croce L (2016). Emerging roles for Polycomb proteins in cancer. Curr. Opin. Genet. Dev. 36, 50–58. [PubMed: 27151431]

Plys AJ, Davis CP, Kim J, Rizki G, Keenan MM, Marr SK, and Kingston RE (2019). Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. Genes Dev. 33, 1–15. [PubMed: 30602436]

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, and Zhang F (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 8, 2281–2308. [PubMed: 24157548]

Robinson MD, McCarthy DJ, and Smyth GK (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. [PubMed: 19910308]

Riback JA, Zhu L, Ferrolino MC, Tolbert M, Mitrea DM, Sanders DW, Wei M-T, Kriwacki RW, Brangwynne CP (2020). Composition-dependent thermodynamics of intracellular phase separation. Nature 581, 209–214. [PubMed: 32405004]

Schuettengruber B, and Cavalli G (2009). Recruitment of Polycomb group complexes and their role in the dynamic regulation of cell fate choice. Development 136, 3531–3542. [PubMed: 19820181]

Seif E, Kang JJ, Sasseville C, Senkovich O, Kaltashov A, Boulier EL, Kapur I, Kim CA, Francis NJ (2020). Phase separation by the polyhomeotic sterile alpha motif compartmentalizes Polycomb Group proteins and enhances their activity. Nat. Commun. 11, 5609. [PubMed: 33154383]

Shin Y and Brangwynne CP (2017). Liquid phase condensation in cell physiology and disease. Science. 357, eaaf4382. [PubMed: 28935776]

Skene PJ, Henikoff JG, and Henikoff S (2018). Targeted in situ genome-wide profiling with high efficiency for low cell numbers. Nat. Protoc. 13, 1006–1019. [PubMed: 29651053]

Stark R, and Brown G (2011). DiffBind: differential binding analysis of ChIP-Seq peak data. http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf

Statt A, Casademunt H, Brangwynne CP, Panagiotopoulos AZ (2020). Model for disordered proteins with strongly sequence-dependent liquid phase behavior. J. Chem. Phys. 152, 075101. [PubMed: 32087632]

Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH (2017). Phase separation drives heterochromatin domain formation. Nature 547, 241–245. [PubMed: 28636597]

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U.S.A. 102, 15545–15550. [PubMed: 16199517]

Tamburri S, Lavarone E, Fernández-Pérez D, Conway E, Zanottie M, Manganaro D, and Pasini D (2020). Histone H2AK119 mono-ubiquitination is essential for Polycomb-mediated transcriptional repression. Mol. Cell 77, 840–956. [PubMed: 31883952]

Tatavosian R, Kent S, Brown K, Yao T, Duc HN, Huynh TN, Zhen CY, Ma B, Wang H, and Ren X (2019). Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation. J. Biol. Chem. 294, 1451–1463. [PubMed: 30514760]

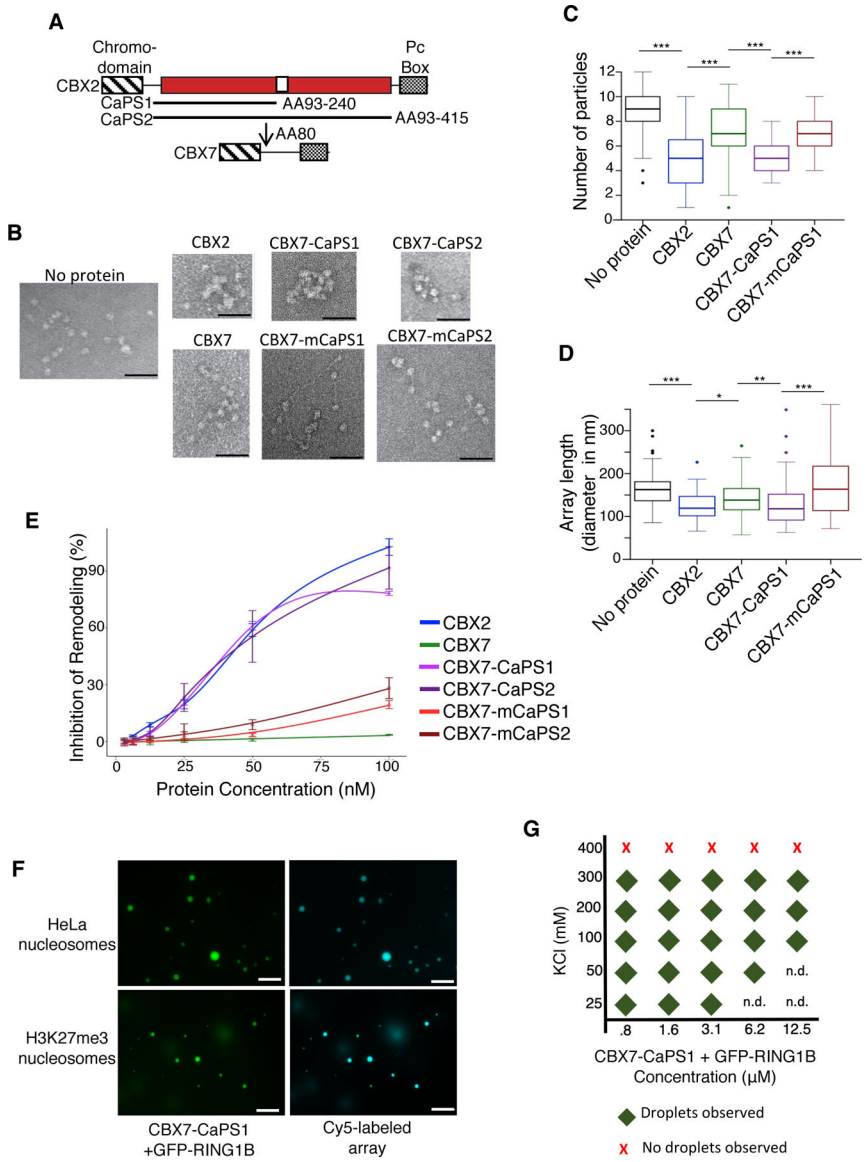

Figure 1. The CBX2 CaPS domain confers *in vitro* chromatin compaction and phase separation activities on CBX7.

(A) Schematic of the additions of the CBX2 CaPS domain to CBX7. The red regions within CBX2 represent the region necessary for compaction as described by the lysine/arginine-to-alanine point mutations in Grau et al., 2011. The lines represent the segments of CBX2 that were inserted into CBX7, and the arrow indicates the location of the insertions (AA = amino acid).

(B) Representative negative stain electron microscopy images of nucleosomal array compaction by CBX proteins and chimeras. Scale bar = 50nm.

(C) Quantification of nucleosome compaction as shown in **(B)** by counting the number of individually distinguishable particles per array (n=50 images per sample; **p < .01; ***p < .001).

(D) Diameters of nucleosomal arrays, from electron microscopy images as in **(B)** (n=50 images per sample; * p < .05; **p<.01; *** p < .001).

(E) Inhibition of SWI/SNF nucleosome remodeling by CBX proteins and chimeras as measured by restriction enzyme accessibility. Data represent the mean ± SEM (n=3).
(F) Micrographs of 6.25μM CBX7-CaPS1, purified with GFP-RING1B, with 100nM Cy5-labeled nucleosome array substrate in 100mM KCl solution. Scale bar = 10μm.
(G) Phase diagram representing the salt and protein concentrations at which CBX7-CaPS1 forms phase separated droplets as observed by fluorescence microscopy (n.d.: no data).
See also Figure S1.

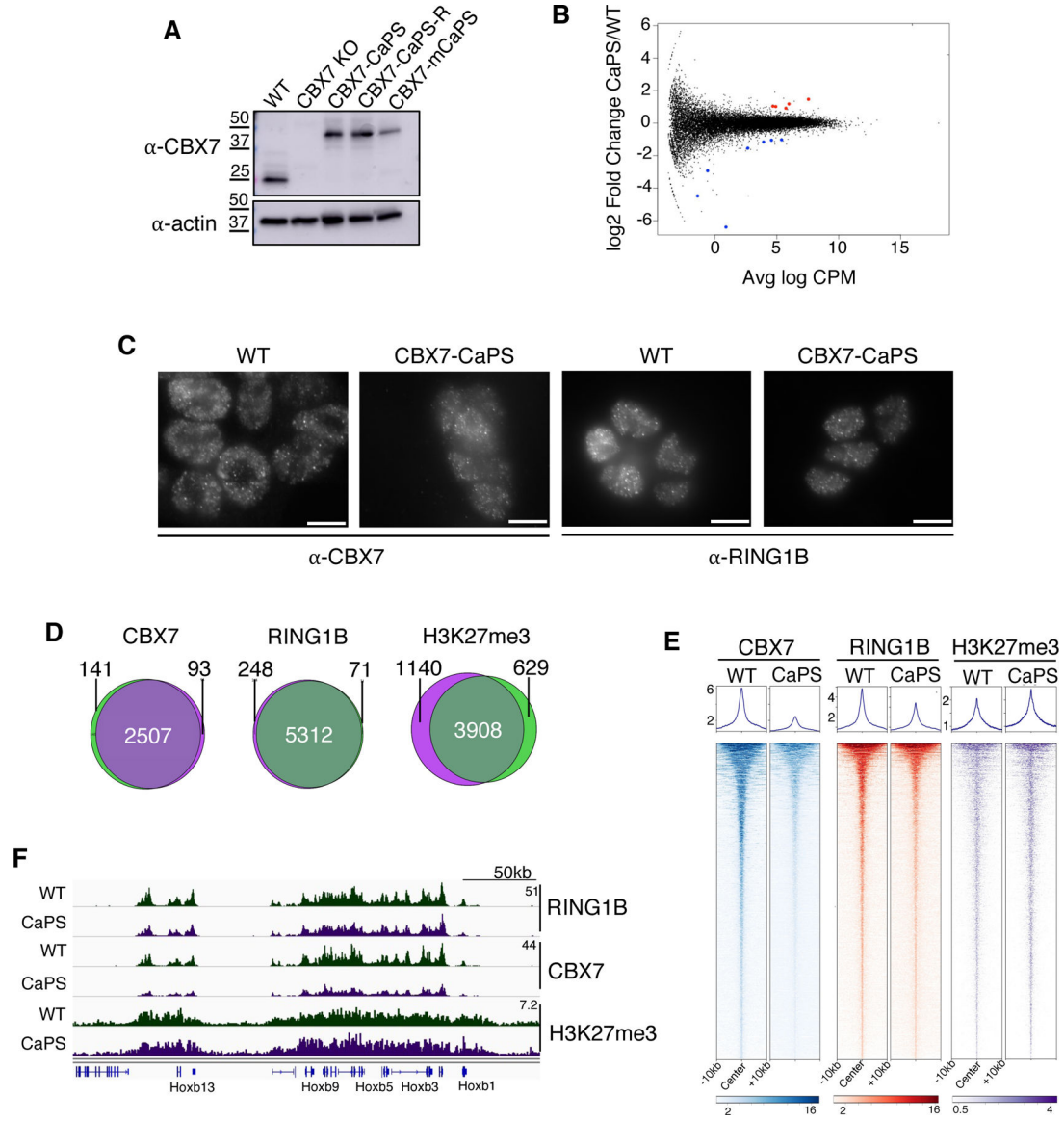

Figure 2. CBX7-CaPS ESCs are phenotypically similar to WT.

(A) Western blot against CBX7 in indicated ESC nuclear extracts, with actin loading control.

(B) Log2(fold change) in gene expression between CBX7-CaPS and WT ESCs from 3 replicates of RNA-Seq. Significantly changed genes in CBX7-CaPS cells (FDR < 0.05) are shown in red (upregulated) or blue (downregulated).

(C) Immunofluorescence against CBX7 and RING1B in WT and CBX7-CaPS ESCs. Scale bar = 10μm.

(D) Overlap of confident CBX7, RING1B, and H3K27me3 ChIP-Seq target genes in WT (green) and CBX7-CaPS (purple) ESCs. CBX7 and RING1B ChIP-Seq data represent the average of three replicates, and H3K27me3 data represent the average of two replicates.

(E) Heatmaps of confident CBX7, RING1B, and H3K27me3 ChIP-Seq peaks in WT and CBX7-CaPS ESCs.

(F) IGV genome browser view of RING1B, CBX7, and H3K27me3 ChIP-Seq signal at the HoxB locus in WT (green) and CBX7-CaPS (purple) ESCs.
See also Figure S2.

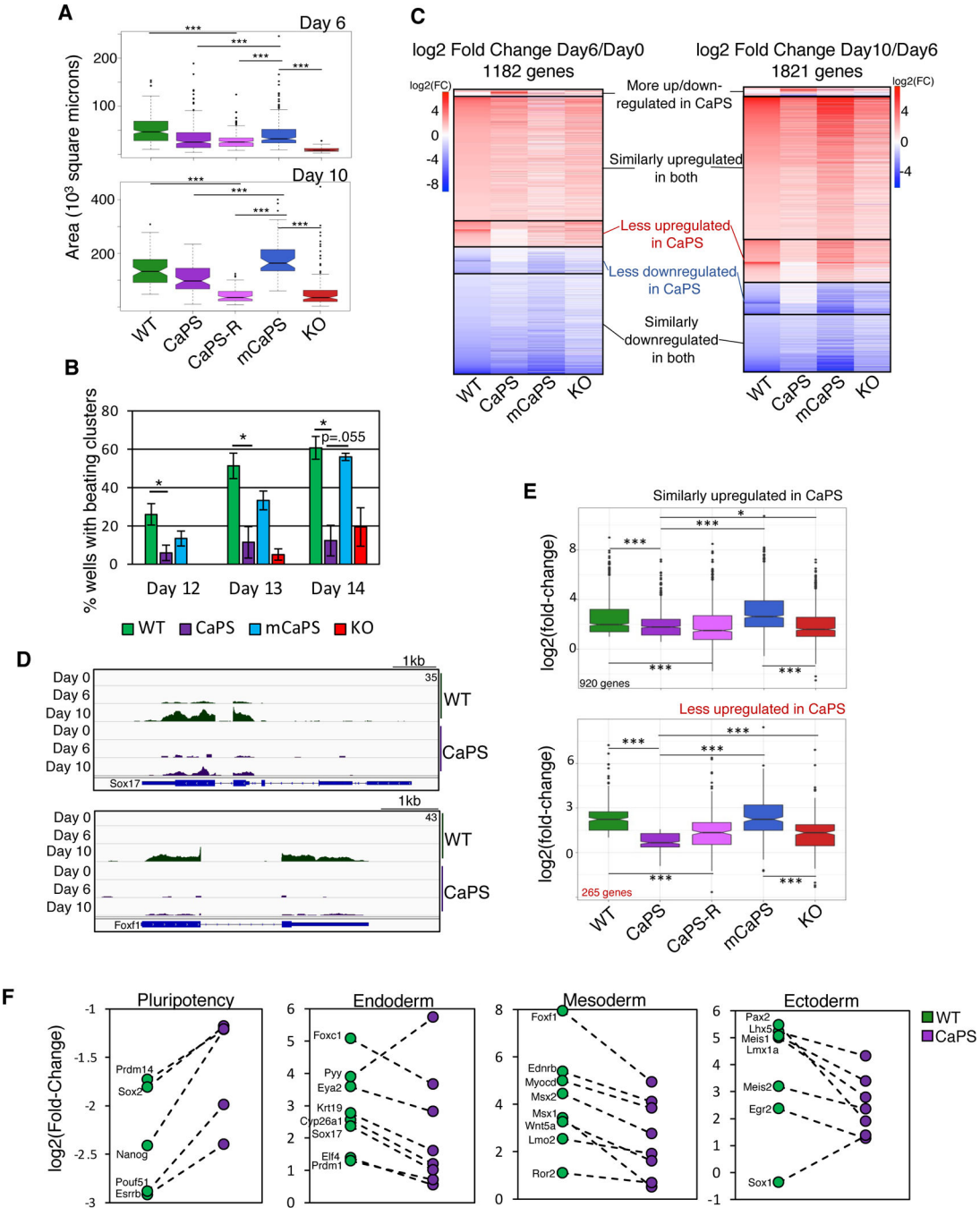

Figure 3. Embryoid bodies expressing CBX7-CaPS show defects in differentiation and gene expression.

(A) Size of embryoid bodies generated from indicated ES cell lines at day 6 and day 10 of differentiation, as measured in ImageJ in 2D from images as in Figure S3A (*** = $p < .001$).

(B) Quantification of beating clusters observed in EB formation (* = $p < .05$). Data represent the mean ± SEM (n=3).

(C) Heatmaps showing log2(fold-change) of PRC1 target gene expression over the course of EB formation. Only PRC1 targets that changed expression more than 2-fold with FDR < .05

in WT or CBX7-CaPS cells are shown. Supervised clusters were generated by comparing log2(fold-change) for a given gene in CBX7-CaPS differentiation to its change in WT differentiation.

(D) IGV genome browser views of RNA-Seq over EB formation for *Sox17* and *Foxf1*.

(E) Comparison of log2(fold-change) in expression of genes from indicated clusters from panel (**C**), from day 6 to day 10 of EB formation (* = $p < 0.05$; *** = $p < 0.001$).

(F) Log2(fold-change) of specific lineage markers from day 6 to day 10 of embryoid body formation in WT (green) and CaPS (purple) differentiation.

See also Figure S3.

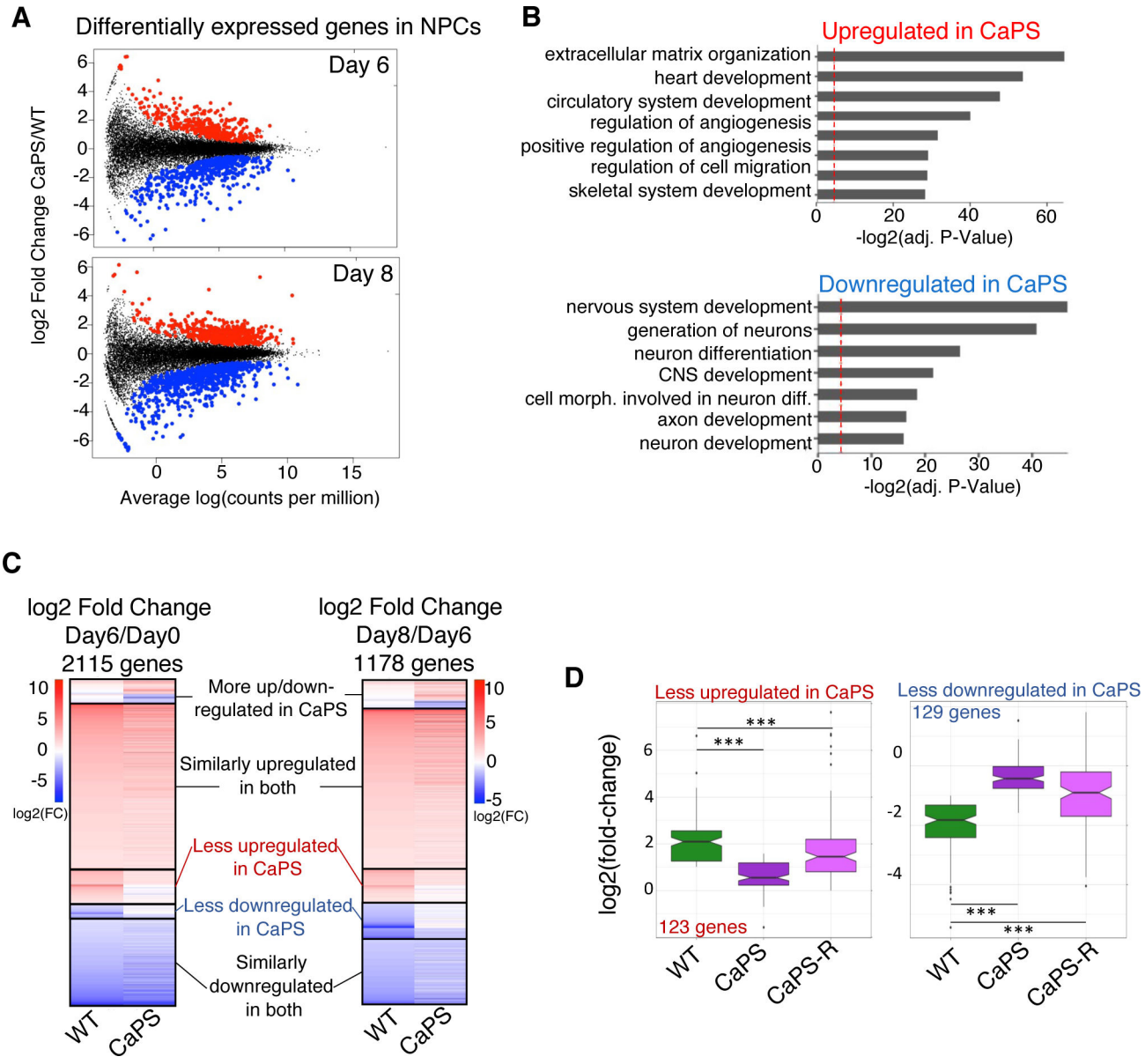

Figure 4. Gene expression is perturbed during neural differentiation in CBX7-CaPS cells.

(A) Log2(fold change) in gene expression between CBX7-CaPS and WT day 6 and day 8 NPCs, from 3 replicates of RNA-Seq. Significantly changed genes in CBX7-CaPS cells (FDR < 0.05) are shown in red (upregulated) or blue (downregulated).

(B) Biological processes enriched in significantly up- and down-regulated genes in day 8 NPCs, from **(A)**. Red line indicates adjusted pValue = .05.

(C) Heatmaps showing log2(fold-change) of PRC1 target gene expression over the course of NPC formation. Only PRC1 targets that changed expression more than 2-fold with FDR < .05 in WT or CBX7-CaPS cell lines are shown. Supervised clusters were generated by comparing log2(fold-change) for a given gene in CBX7-CaPS differentiation to its change in WT differentiation.

(D) Comparison of log2(fold-change) in expression of genes from indicated clusters from **(C)**, from day 6 to day 8 of NPC differentiation (*** = $p < 0.001$).
See also Figure S4.

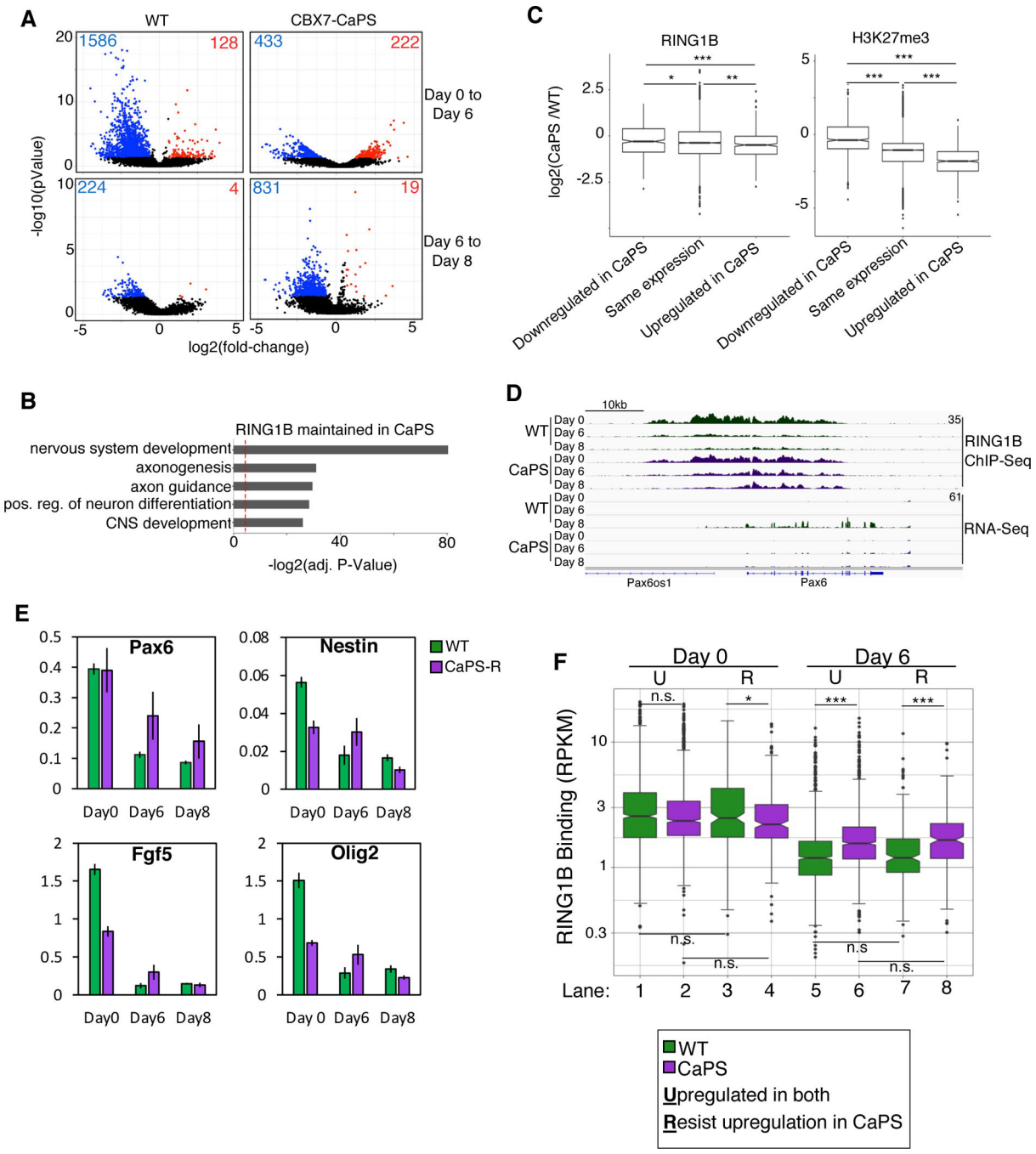

Figure 5. PRC1 redistribution is impaired in CBX7-CaPS NPC differentiation.

(A) Volcano plots comparing RING1B ChIP-Seq signal from Day 0 to Day 6 and from Day 6 to Day 8 in WT and CBX7-CaPS NPC differentiation. Colored points represent at least 1.5-fold change in signal over time with a p-value < 0.05, with blue indicating a decrease and red indicating an increase in signal over time. Data represents average of 2 replicates. **(B)** Biological processes enriched in genes that showed a significant decrease in RING1B binding in WT from Day 0 to Day 6 of NPC formation but did not show a significant change in binding in CBX7-CaPS differentiation over the same time period (related to panel **A**).

(C) Difference in RING1B and H3K27me3 binding intensity at misregulated genes in CBX7-CaPS day 8 NPCs, as identified in Figure 4A. Boxplots were generated based on log2(CBX7-CaPS/WT) binding intensity at genes belonging to each group (*p < .05, **p < .01, ***p < .001).

(D) IGV ChIP-Seq and RNA-Seq tracks at the *Pax6* locus for WT (green) and CBX7-CaPS (purple) NPC differentiation time courses.

(E) ChIP-qPCR of RING1B at indicated genes across NPC differentiation in WT (green) and CBX7-CaPS-R (purple) cells. Data represent the mean of two biological replicates ± SEM.

(F) RING1B ChIP-Seq signal (RPKM) at the indicated time points for genes that show similar upregulation patterns or a resistance to upregulation from day 0 to day 6 of CBX7-CaPS NPC differentiation when compared to WT. See also Figure S5B (n.s. = not significant, *p < .05,**p < .01, ***p < .001).

See also Figure S5.

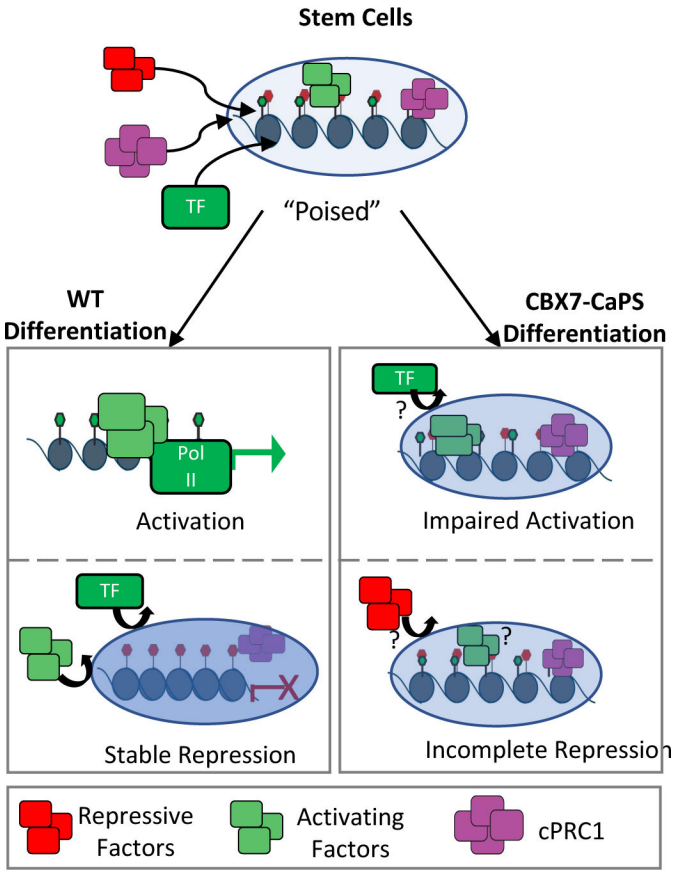

Figure 6: Model for the prevention of Polycomb gene activation and repression by CBX7-CaPS.

In normal stem cell differentiation, CBX7 and activating factors contribute to a poised chromatin environment that allows for perturbation by other repressive or activating factors, leading to the efficient resolution of a bivalent state towards either activation or more stable repression. As CBX7-CaPS ESCs differentiate, target genes resist both up- and down-regulation. This may occur through exclusion of both repressive and activating factors (TFs) from target genes, or by the prolonged maintenance of both active and repressive factors during differentiation.

Table 1:
Peptide counts from immunoprecipitation following mass spectrometry in ESCs.

Table showing unique and total peptide counts from crosslinked-IP followed by mass spectrometry (RIME) analysis for the indicated target proteins in ESCs. Shown are interactions with Polycomb group proteins; complete RIME dataset is in Table S1.

IP:	IgG		CBX7 Rep 1		CBX7 Rep 2		CBX7-CaPS		CBX2	
	Unique	Total	Unique	Total	Unique	Total	Unique	Total	Unique	Total
PHC1	0	0	6	8	11	39	8	11	4	7
RING1B	0	0	4	5	2	2	2	4	3	4
JARID2	0	0	7	9	18	35	9	11	4	5
SUZ12	0	0	4	8	12	12	6	8	0	0
EZH2	0	0	3	3	5	8	4	4	0	0
EED	0	0	3	4	1	1	2	4	0	0
MTF2	0	0	2	4	5	7	2	3	0	0

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-Cbx7 (WB)	Millipore	Cat No. 07–981 (RRID:AB_10807034)
Mouse monoclonal anti-β-Actin	Sigma-Aldrich	Cat No. A2228 (RRID:AB_476697)
Rabbit polyclonal anti-Cbx7 (ChIP, CUT&RUN)	Abcam	Cat No. ab21873 (RRID:AB_726005)
Rabbit polyclonal anti-Cbx2	Bethyl	Cat No. A302–524A (RRID:AB_1998943)
Rabbit polyclonal IgG	Abcam	Cat No. ab37415 (RRID:AB_2631996)
Rabbit monoclonal anti-Ring1B	Cell Signaling Technology	Cat No. 5694 (RRID:AB_10705604)
Rabbit polyclonal anti-H2Av	Active Motif	Cat No. 39715 (RRID:AB_2793318)
Bacterial and virus strains		
Biological samples		
Chemicals, peptides, and recombinant proteins		
Critical commercial assays		
Cellfectin II Reagent	Thermo Fisher Scientific	Cat No. 10362100
PureLink RNA Mini Kit	Thermo Fisher Scientific	Cat No. 12183018A
NEBNext rRNA Depletion Kit	New England BioLabs	Cat No. E6310L
NEBNext UltraII RNA First Strand Synthesis Module	New England Biolabs	Cat No. E7771S
NEBNext UltraII Non-Directional RNA Second Strand Synthessis Module	New England Biolabs	Cat No. E6111S
Deposited data		
All NGS Data, GEO SuperSeries	This Study	GEO: GSE151901
RNA-Seq	This Study	GEO: GSE151900
ChIP-Seq & CUT&RUN	This Study	GEO: GSE151899
Experimental models: cell lines		
SF9	ATCC	CRL-1711
Mouse ESC: J1	ATCC	SCRC-1010
Mouse ESC: CBX7-CaPS	This Study	N/A
Mouse ESC: CBX7-CaPS-R	This Study	N/A
Mouse ESC: CBX7-mCaPS	This Study	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse ESC: CBX7 −/−	This Study	N/A
Experimental models: organisms/strains		
Oligonucleotides		
CBX7-CaPS sgRNA	ACCTGCAGCAGAAGCCGCCT	N/A
CBX7 KO sgRNA #1	TCGTAGGCCATGACAAGGCG	N/A
CBX7 KO sgRNA #2	TGTTTGCGGTGGAGAGCATC	N/A
Pax6 ChIP-qPCR Forward	GCCTAGCAGGTTTACCCAAG	N/A
Pax6 ChIP-qPCR Reverse	GCGACAAAGTTTACCCTCCA	N/A
Nestin ChIP-qPCR Forward	GCCGCGTAACTTCTTCACTA	N/A
Nestin ChIP-qPCR Reverse	GACGTCCTACCCTACCCTAC	N/A
Olig2 ChIP-qPCR Forward	GCCTGAGGTTTGTTAGGGAC	N/A
Olig2 ChIP-qPCR Reverse	TGAGGATGACCCATAACCGT	N/A
Fgf5 ChIP-qPCR Forward	CAAGATGCACTTAGGACCCC	N/A
Fgf5 ChIP-qPCR Reverse	GAAGACGAAGAAAACGTCGC	N/A
Recombinant DNA		
pFastBac-Cbx7	Plys et al., 2019	N/A
pFastBac-Cbx2	Grau et al., 2011	N/A
pFastBac-Cbx2KRA	Grau et al., 2011	N/A
pFastBac-Cbx7-CaPS1	This Study	N/A
pFastBac-Cbx7-mCaPS1	This Study	N/A
pFastBac-Cbx7-CaPS2	This Study	N/A
pFastBac-Cbx7-mCaPS2	This Study	N/A
pMax-CaPSinsertion	This Study	N/A
pMax-mCaPSinsertion	This Study	N/A
Software and algorithms		
Cutadapt v1.14	Martin, 2011	https://cutadapt.readthedocs.io/en/stable/
STAR v2.7.0f	Dobin et al., 2013	https://github.com/alexdobin/STAR
Samtools v1.9	Li et al., 2009	http://www.htslib.org/
HOMER v4.10	Heinz et al., 2010	http://homer.ucsd.edu/homer/index.html
edgeR	Robinson et al., 2010	https://bioconductor.org/packages/release/bioc/html/edgeR.html
MSigDB	Subramanian et al., 2005	https://www.gsea-msigdb.org/gsea/msigdb/
enrichR	Chen et al., 2013; Kuleshov et al., 2016	https://maayanlab.cloud/Enrichr/
Bowtie2 v2.3.4.3	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
deepTools	Ramírez et al., 2016	https://deeptools.readthedocs.io/en/develop/
DiffBind	Stark and Brown, 2011	http://bioconductor.org/packages/release/bioc/html/DiffBind.html

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Other		