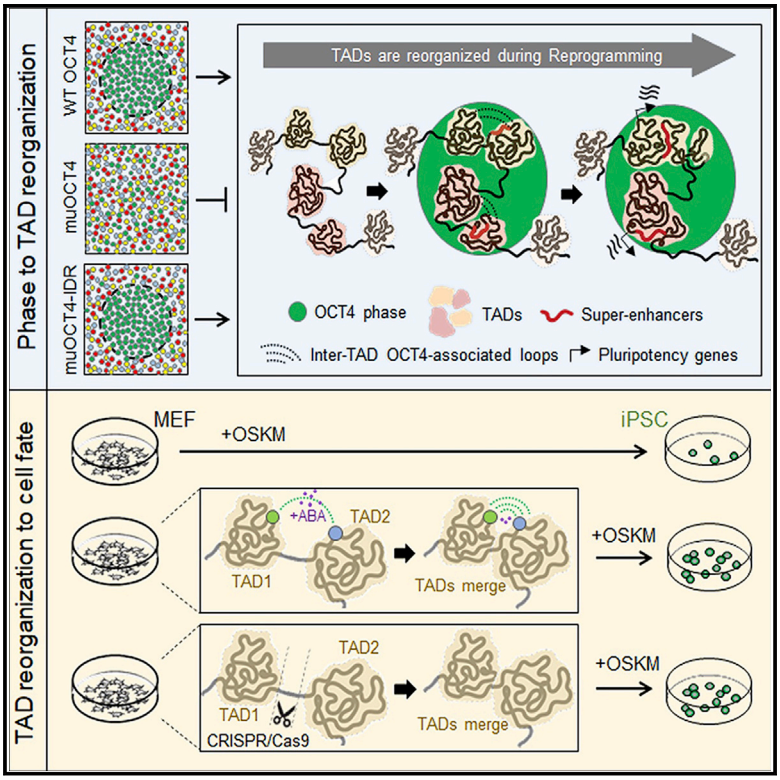
Cell Stem Cell

Article

Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions

Graphical abstract



Highlights

- TADs are reorganized during somatic cell reprogramming
- Manipulation of TAD reorganization promotes reprogramming
- OCT4 phase separation regulates TAD reorganization
- Reprogramming regulators are identified by a TAD-reorganization-based algorithm, TADMAN

Authors

Jia Wang, Haopeng Yu, Qian Ma, ..., Wei Zhang, Yiliang Miao, Junjun Ding

Correspondence

dingjunj@mail.sysu.edu.cn

In brief

Wang et al. provide evidence that TAD reorganization contributes to cell fate transition, and Oct4 phase separation regulates TAD reorganization. This provides insights into how a master transcription factor regulates 3D genome topology, especially TAD structures, for cell fate control.

Wang et al., 2021, Cell Stem Cell *28*, 1–16
October 7, 2021
https://doi.org/10.1016/j.stem.2021.04.023

CellPress

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

Cell Stem Cell

CellPress

Article

Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions

Jia Wang,[1,2,3,12] Haopeng Yu,[4,5,12] Qian Ma,[1,2,3,12] Pengguihang Zeng,[1,2,3] Danya Wu,[6] Yingping Hou,[7] Xinyi Liu,[1,2,3] Lumeng Jia,[8] Jun Sun,[1,2,3] Yilong Chen,[4,5] Diana Guallar,[9] Miguel Fidalgo,[9] Jiahao Chen,[1,2,3] Yangyinhui Yu,[3] Shaoshuai Jiang,[1,2,3] Fenjie Li,[1,2,3] Cai Zhao,[1,2,3] Xianglin Huang,[1,2,3] Jianlong Wang,[11] Cheng Li,[7,8] Yujie Sun,[10] Xiaoxi Zeng,[4,5] Wei Zhang,[4,5] Yiliang Miao,[6] and Junjun Ding[1,2,3,13,*]

[1]RNA Biomedical Institute, Sun Yat-Sen Memorial Hospital, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou 510080, China

[2]Center for Stem Cell Biology and Tissue Engineering, Key Laboratory for Stem Cells and Tissue Engineering, Ministry of Education, Sun Yat-Sen University, Guangzhou 510080, China

[3]Department of Cell Biology, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou 510080, China

[4]West China Biomedical Big Data Center, West China Hospital/West China School of Medicine, Sichuan University, Chengdu 610041, China

[5]Med-X Center for Informatics, Sichuan University, Chengdu 610041, China

[6]Institute of Stem Cell and Regenerative Biology, College of Animal Science and Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, China

[7]Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China

[8]School of Life Sciences, Peking University, Beijing 100871, China

[9]Centro de Investigacionen Medicina Molecular e Enfermidades Cronicas, Universidade de Santiago de Compostela, Santiago de Compostela 15782, Spain

[10]State Key Laboratory of Membrane Biology, Biomedical pioneering innovation center (BIOPIC) School of Life Sciences, Peking University, Beijing 100871, China

[11]Department of Medicine, Columbia Center for Human Development, Columbia Stem Cell Initiative, Columbia University Irving Medical Center, New York, NY 10032, USA

[12]These authors contributed equally

[13]Lead contact

*Correspondence: dingjunj@mail.sysu.edu.cn

https://doi.org/10.1016/j.stem.2021.04.023

SUMMARY

Topological-associated domains (TADs) are thought to be relatively stable across cell types, although some TAD reorganization has been observed during cellular differentiation. However, little is known about the mechanisms through which TAD reorganization affects cell fate or how master transcription factors affect TAD structures during cell fate transitions. Here, we show extensive TAD reorganization during somatic cell reprogramming, which is correlated with gene transcription and changes in cellular identity. Manipulating TAD reorganization promotes reprogramming, and the dynamics of concentrated chromatin loops in OCT4 phase separated condensates contribute to TAD reorganization. Disrupting OCT4 phase separation attenuates TAD reorganization and reprogramming, which can be rescued by fusing an intrinsically disordered region (IDR) to OCT4. We developed an approach termed TAD reorganization-based multiomics analysis (TADMAN), which identified reprogramming regulators. Together, these findings elucidate a role and mechanism of TAD reorganization, regulated by OCT4 phase separation, in cellular reprogramming.

INTRODUCTION

The eukaryotic three-dimensional (3D) genome is organized in a hierarchical order mainly comprising compartments, topological-associated domains (TADs), and loops from large to small scales (Dixon et al., 2012; Lieberman-Aiden et al., 2009; Nora et al., 2012; Rao et al., 2014). Compartments undergo dynamic switch during cell fate transitions, although such a switch weakly correlates with changes in gene expression (Dixon et al., 2015). In contrast, TADs are stable among different cell types and species (Dixon et al., 2012, 2015). However, recent studies have reported the loss of TADs during pluripotent stem cell (PSC) differentiation (Bonev et al., 2017; Zhang et al., 2019), indicating that they are likely to reorganize in these biological processes. Evidence showed that disruption of certain TADs or insulated neighborhood boundaries enables *de novo* enhancer-promoter interactions, which lead to abnormal gene expression and diseases (Hnisz et al., 2016; Lupiáñez et al., 2015). These studies indicate that alteration of TAD or sub-TAD structures may influence transcription and cell fate. Supporting that, TAD-TAD interactions

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

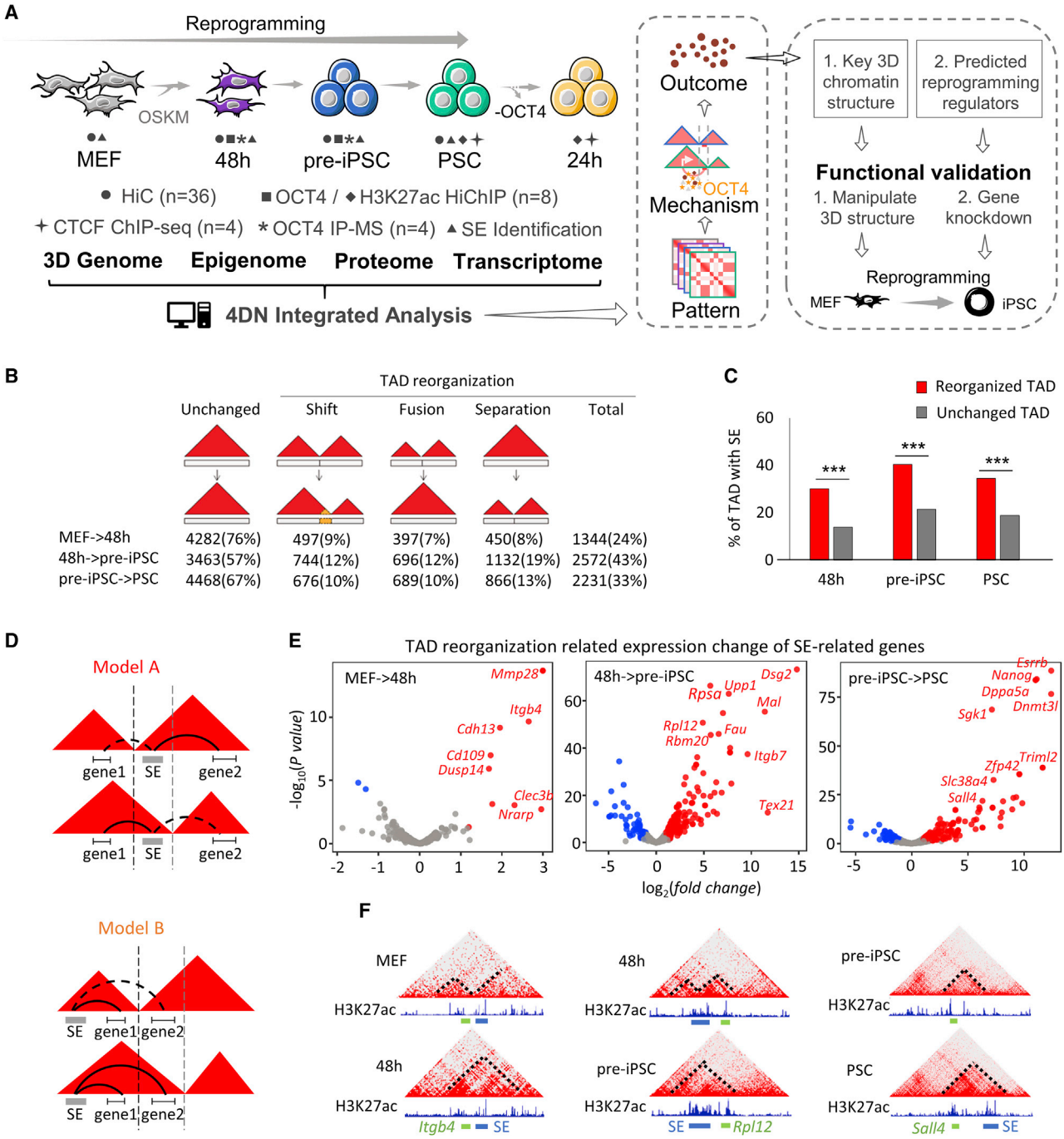


Figure 1. TADs are reorganized during somatic cell reprogramming

(A) Scheme of the research route and methods in this study. Gray, purple, blue, and green cells represent the four stages of reprogramming from MEF to PSC, whereas yellow cells represent the PSC 24 h after OCT4 depletion. The dark circle, square, rhombus, quadrilateral, asterisk, and triangle beside the cells represent the experiments or analyses of HiC, OCT4 HiChIP, H3K27ac HiChIP, CTCF ChIP-seq, OCT4 IP-MS, and SE identification in corresponding cells, respectively. The 4DN pattern, mechanism, and outcome during somatic cell reprogramming were obtained by integrating the multiomics experiments and analyses. The functional significance of key higher-order 3D structures and reprogramming regulators were validated in reprogramming.

(B) Number and percentage of unchanged TADs versus reorganized TADs at each stage during reprogramming. TAD reorganization comprises TAD shift, fusion, and separation.

(C) The distribution percentage of the TADs with SEs. SEs likely localize in reorganized TADs (red), rather than unchanged TADs (gray). Fisher's exact test was used for statistics. ***$p < 0.001$.

(D) Two models depicting how TAD reorganization contributes to SE-related gene expression. TAD reorganization alters SE targets by redistributing SEs (Model A) or SE-related genes (Model B) in TADs. Dotted lines and arcs represent TAD boundaries and inexistent loops, while solid lines represent existent loops.

(legend continued on next page)

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

between neighboring TADs were reported to undergo 20% of reorganization during neural differentiation, which was associated with changes in gene expression (Fraser et al., 2015). In addition, TAD cliques, a cluster of at least three fully interacting TADs, are altered during differentiation, which affects the expression of differentiation-related genes (Paulsen et al., 2019). Furthermore, a TAD fusion event was observed in T cell acute lymphoblastic leukemia, which was correlated with the activation of the *MYC* gene (Kloetgen et al., 2020). All these results suggest that TAD reorganization may contribute to cell fate transitions by regulating key genes that govern the cell identity. However, the mechanism between TAD reorganization and cell fate transitions has not been fully understood.

It is well established that cell-specific 3D architectures are shaped during cell fate transitions to influence transcriptional regulation (Beagan et al., 2016; Bonev et al., 2017; Krijger et al., 2016; Siersbæk et al., 2017; Stadhouders et al., 2018). Because cell fate transitions are usually driven by the master transcription factors (Takahashi and Yamanaka, 2006; Vierbuchen et al., 2010; Wapinski et al., 2013; Zhang et al., 2010), recent studies have begun to reveal the contributions of the master transcription factors to 3D chromatin structures (Dall'Agnese et al., 2019; de Wit et al., 2013; Stadhouders et al., 2018). However, the mechanism by which master transcription factors direct the dynamics of 3D architectures for cell fate control has yet to be clarified. Phase separation capacity of the master transcription factors allows functional chromatin interactions to regulate gene transcription (Boija et al., 2018; Sabari et al., 2018) and help organize and regulate eukaryotic genome (Gibson et al., 2019). These findings indicate the potential link between phase separation of master transcription factors and 3D genome organizations. However, whether and how phase separation of master transcription factors regulates the formation of the higher-order 3D structure, such as TAD reorganization, remains unknown.

OCT4, SOX2, KLF4, and MYC (OSKM) are master transcription factors for somatic cell reprogramming (Takahashi and Yamanaka, 2006). In this study, we explored the dynamic pattern of a 3D genome and the underlying mechanisms during somatic cell reprogramming. We found that TADs are reorganized during reprogramming, which is correlated with gene regulation and cell fate transitions. Manipulation of TAD reorganization affects the efficiency of induced pluripotent stem cell (iPSC) generation. Moreover, OCT4 phase separation contributes to TAD reorganization. Finally, we applied TAD reorganization-based multiomics analysis (TADMAN) to the discovery of reprogramming regulators.

RESULTS

TADs are reorganized during somatic cell reprogramming

To determine the pattern and underlying mechanisms of 3D genome dynamics during somatic cell reprogramming, we integrated multiomics experiments and analyses containing 3D genome, proteome, epigenome, and transcriptome in mouse embryonic fibroblasts (MEFs) (Abad et al., 2013), cells 48 h after reprogramming initiation, pre-induced PSCs (pre-iPSCs) (Sridharan et al., 2013), and PSCs (Figure 1A). The functional higher-order 3D architectures and reprogramming regulators were obtained by integrative analysis, whose functions were validated in reprogramming (Figure 1A). The genome is organized into A and B compartments, related to active and inactive regions, respectively (Lieberman-Aiden et al., 2009). Our compartment analysis showed that around 25% of the genomic regions underwent compartment A/B switch during reprogramming (Figure S1A; Table S1). However, a large portion of the pluripotency genes was not associated with a compartment switch (Figures S1B and S1C), indicating involvement of other mechanisms in activating them. To elucidate that, we analyzed TAD dynamics during reprogramming (Zufferey et al., 2018). Although TADs are relatively conserved in different cell types and species (Dixon et al., 2012, 2015), loss of TADs has been observed during PSC differentiation (Bonev et al., 2017; Zhang et al., 2019). Consistent with that, we found that the number of TADs increased during reprogramming (Figure S1D), indicating TADs may be reorganized during that biological process. Three types of TAD reorganization, including TAD shift, fusion, and separation (Figure 1B), were observed by considering the significant changes in TAD boundaries or size.

To assess the biological significance of TAD reorganization, we identified super-enhancers (SEs) and SE-related genes at each stage of reprogramming (Table S2) because SEs are the key *cis*-regulatory elements that control cell fate (Hnisz et al., 2013; Lovén et al., 2013; Whyte et al., 2013). As expected, the SE-related genes are more stage specific than typical enhancer (TE)-related genes (Figure S1E), indicating the former may be more related to cell fate. Notably, SEs preferentially localize in reorganized TADs as opposed to unchanged TADs (Figure 1C), suggesting a correlation between SE distribution and TAD reorganization. Two models hypothesize the role of TAD reorganization in regulating SE-related genes, in which TAD reorganization redistributes SEs (model A) or SE-related genes (model B) in TADs to change SE targets (Figures 1D and S1F). Furthermore, the activation of SE-related genes, including many critical pluripotency genes, was more correlated to TAD reorganization than it was to compartment switching during reprogramming (Figures 1E, 1F, S1G, and S1H; Table S2), indicating a potential role for TAD reorganization in reprogramming.

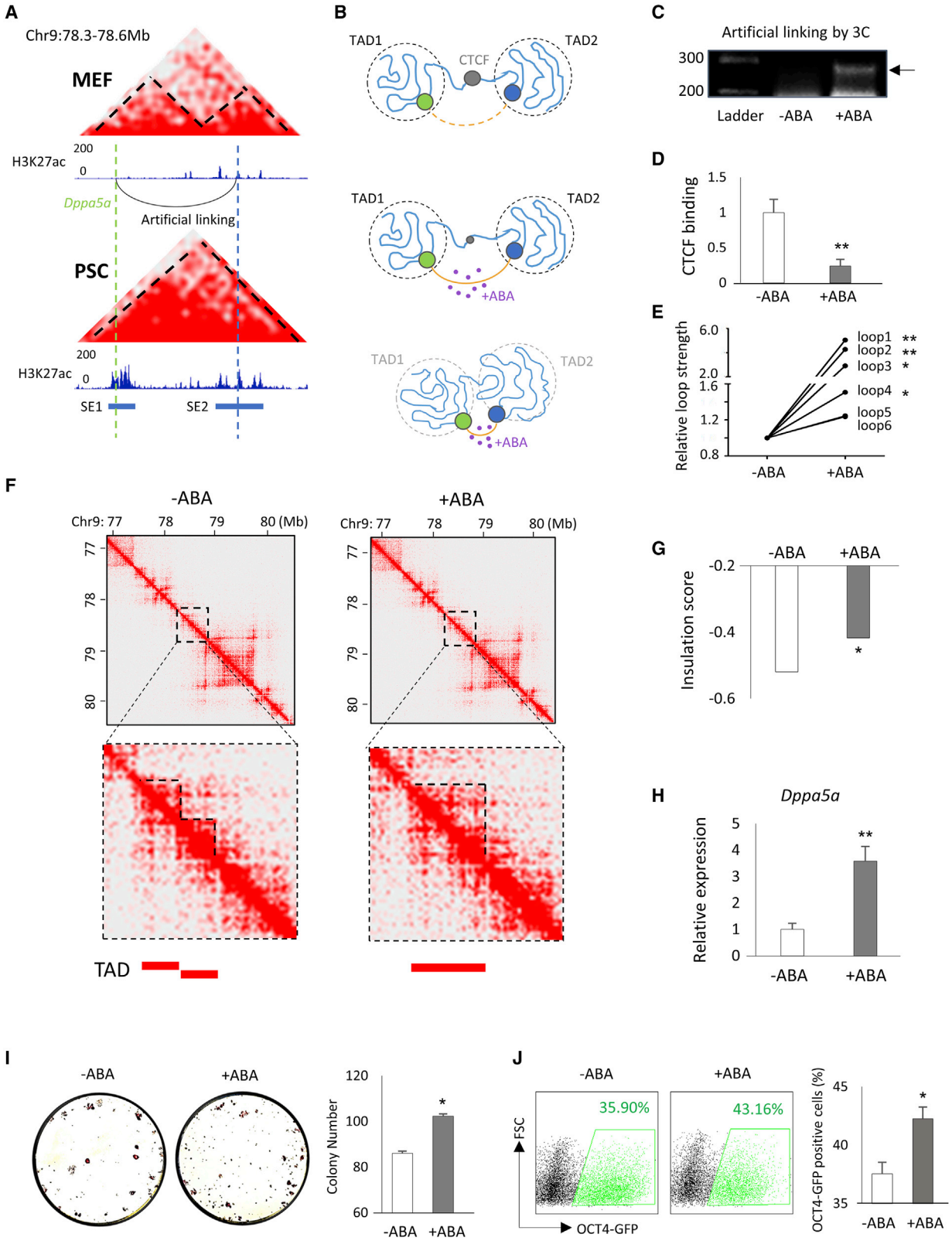
Taken together, TADs are reorganized during reprogramming, which is correlated with the transcriptional regulation of the pluripotency genes.

Chemical-dependent genome linking induces TAD reorganization

To elucidate the role of TAD reorganization in regulating reprogramming, we artificially induced TAD reorganization in MEFs by

(E) Volcano plot showing the SE-related genes with significant expression change in response to TAD reorganization in each stage of reprogramming. Red points stand for the significant upregulated genes, whereas blue points represent the downregulated genes. Most of the SE-related genes were upregulated in response to TAD reorganization. Gene expression was obtained from RNA-seq data (Chronis et al., 2017).
(F) Three examples showing the TAD-reorganization-related activations of the SE-related genes in each stage corresponding to (E). Blue and green bars stand for SEs and SE-related genes, respectively. Dotted lines outline TADs.

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023



(legend on next page)

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

a dCas9-based chemical-inducible linking tool (Morgan et al., 2017). The TAD containing *Dppa5a* and its neighboring TAD were selected for artificial TAD reorganization because the two TADs were merged during reprogramming (Figure 2A), indicating a potential correlation between *Dppa5a* activation and TAD reorganization. Moreover, *Dppa5a* is remarkably highly expressed in PSCs (Figure 1E), indicating the potential importance of *Dppa5a* in the late stage of pluripotency establishment. Supporting that, knockout (KO) of *Dppa5a* in MEFs significantly reduced reprogramming efficiency, whereas overexpression (OE) of *Dppa5a* significantly enhanced reprogramming efficiency (Figures S2A–S2C). The impaired reprogramming of *Dppa5a* KO MEFs may be due to increased apoptosis (Figure S2D).

Our strategy for artificial TAD reorganization is linking two genomic loci belonging to the two neighboring TADs, which provides a driving force for TAD fusion (Figure 2B). The two loci were, respectively, targeted by dCas9-ABI and dCas9-PYL1 and inducibly linked by abscisic acid (ABA) (Figure 2B). The artificial linking was successfully established in MEFs (Figure 2C), which further inhibited CTCF binding at the TAD boundary (Figure 2D) and increased the overall inter-TAD interactions (Figure 2E). Importantly, Hi-C results validated the merging of the two TADs in MEFs with artificial linking compared with the control MEFs (Figure 2F), and the former also showed a higher insulation score (Figure 2G). Moreover, *Dppa5a* expression was upregulated on day 4 during reprogramming in response to artificial linking (Figure 2H). The early activation of *Dppa5a* during reprogramming is likely to correlate with TAD reorganization, because the expressions of most genes in *Dppa5a* TAD are also induced on day 4 of reprogramming (Figures S2E–S2H). Furthermore, the artificial linking enhanced reprogramming efficiency, as evidenced by increased iPSC colonies (Figure 2I), OCT4-GFP-positive cell numbers (Figure 2J), and the expressions of key pluripotency genes (Figure S2I), compared with those of the control MEFs.

Taken together, TAD reorganization can be induced by chemical-dependent genome linking and may contribute to reprogramming.

Boundary deletion-induced TAD reorganization promotes reprogramming

To further validate the role of TAD reorganization in reprogramming, we produced genetically engineered MEFs, in which the common boundary between the *Dppa5a* TAD and the neighboring TAD was deleted using CRISPR-Cas9 genome editing tools (Figure 3A). Boundary-deleted homozygotes and heterozygotes were confirmed by PCR using specific primers (Figures S3A and S3B). Importantly, deletion of the TAD boundary resulted in TAD reorganization, as evidenced by merging of the two TADs bordering the same boundary (Figure 3B) and an increased insulation score (Figure 3C). Moreover, the expression of *Dppa5a* was upregulated on day 4 after reprogramming initiation in response to boundary deletion as compared with that of the control cells (Figure 3D). Furthermore, deletion of the TAD boundary led to enhanced reprogramming efficiency in terms of increased iPSC colonies (Figures 3E and S3C) and increased expression levels of the pluripotency genes (Figures 3F and S3D).

Taken together, these findings demonstrate that manipulation of TAD reorganization by boundary deletion promotes reprogramming.

Concentrated chromatin loops of OCT4 contribute to TAD reorganization

To understand whether and how master transcription factor contributes to TAD reorganization, we focused on the core Yamanaka reprogramming factor OCT4 (Takahashi and Yamanaka, 2006; Theunissen and Jaenisch, 2014) and identified OCT4-associated chromatin loops (OCT4 loops) in 48-h cells, pre-iPSCs, and PSCs by an *in situ* Hi-C library followed by chromatin immunoprecipitation (HiChIP) (Mumbach et al., 2016), given that OCT4 is expressed in these stages of reprogramming (Figures 4A and S4A). Loops analysis revealed that the OCT4 loops were stage specific (Figure 4B; Table S3) and closely associated with promoter-promoter and SE-promoter loops (Figure S4B). Intriguingly, the OCT4 loop count within reorganized boundary regions was significantly higher than that within unchanged boundary regions (Figures 4C and S4C), indicating a correlation between concentrated OCT4 loops and TAD reorganization.

Figure 2. Chemical-dependent genome linking induces TAD reorganization

(A) Diagram showing the activation of *Dppa5a* during reprogramming may be associated with TAD reorganization. Artificial TAD reorganization may be induced by linking two loci belonging to two neighboring TADs (black arc). Two loci are linked in MEFs; one is near the *Dppa5a* gene (green dotted line), and the other is within the neighboring TAD (blue dotted line).

(B) The hypothesis model depicting the possible influence of chemical-inducible construction of the linking on CTCF binding, inter-TAD loop strength, and TAD alteration. The two selected loci belong to different TADs in MEFs. Thus, artificial linking of the two loci in MEFs may drive TAD structure alteration. Gray, green, and blue circles represent CTCF, the dCas9-ABI-*Dppa5a* anchor, and the dCas9-PYL1 anchor, respectively. Purple points represent the chemical abscisic acid (ABA). The yellow arc stands for the artificial linking.

(C) Successful linking of the two loci in MEFs is validated by 3C (arrow). A 3C assay was performed for two biological replicates.

(D) Artificial linking in MEFs decreases CTCF binding at the boundary by ChIP qPCR. Data are presented as means ± SD. $^{**}p < 0.01$. Three biological replicates were assayed for ChIP-qPCR experiment.

(E) Artificial linking increases the strength of inter-TAD interactions between the two neighboring TADs, indicating the distance between the two TADs is decreased. The intra-TAD interactions were not influenced. Loops 1–4 are inter-TAD interactions, whereas loops 5–6 are intra-TAD interactions. The interaction strength was determined by 3C-qPCR. $^{*}p < 0.05$, $^{**}p < 0.01$ compared to control cells. Three biological replicates were assayed for 3C-qPCR experiment.

(F) Surrounding Hi-C heatmaps showing merging of two TADs in response to artificial linking. Red bars stand for TADs.

(G) Insulation score is significantly increased in response to artificial linking. $^{*}p < 0.05$. Bins with a high insulation effect have a low insulation score. p value is calculated by diffHic, which is based on domain-boundary-wise negative binomial generalized-linear models.

(H) Artificial linking in MEFs results in higher *Dppa5a* expression on day 4 after reprogramming initiation as compared with WT MEF. Data are presented as means ± SD. $^{**}p < 0.01$ compared with control cells. *Dppa5a* expression is from three biological replicates.

(I and J) Artificial linking in MEFs enhances reprogramming efficiency by increasing the number of iPSC colonies (I) and OCT4-GFP-positive cells (J). $^{*}p < 0.05$. iPSC colony numbers and the ratios of OCT4-GFP-positive cells at day 12 of reprogramming are from three biological replicates.

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

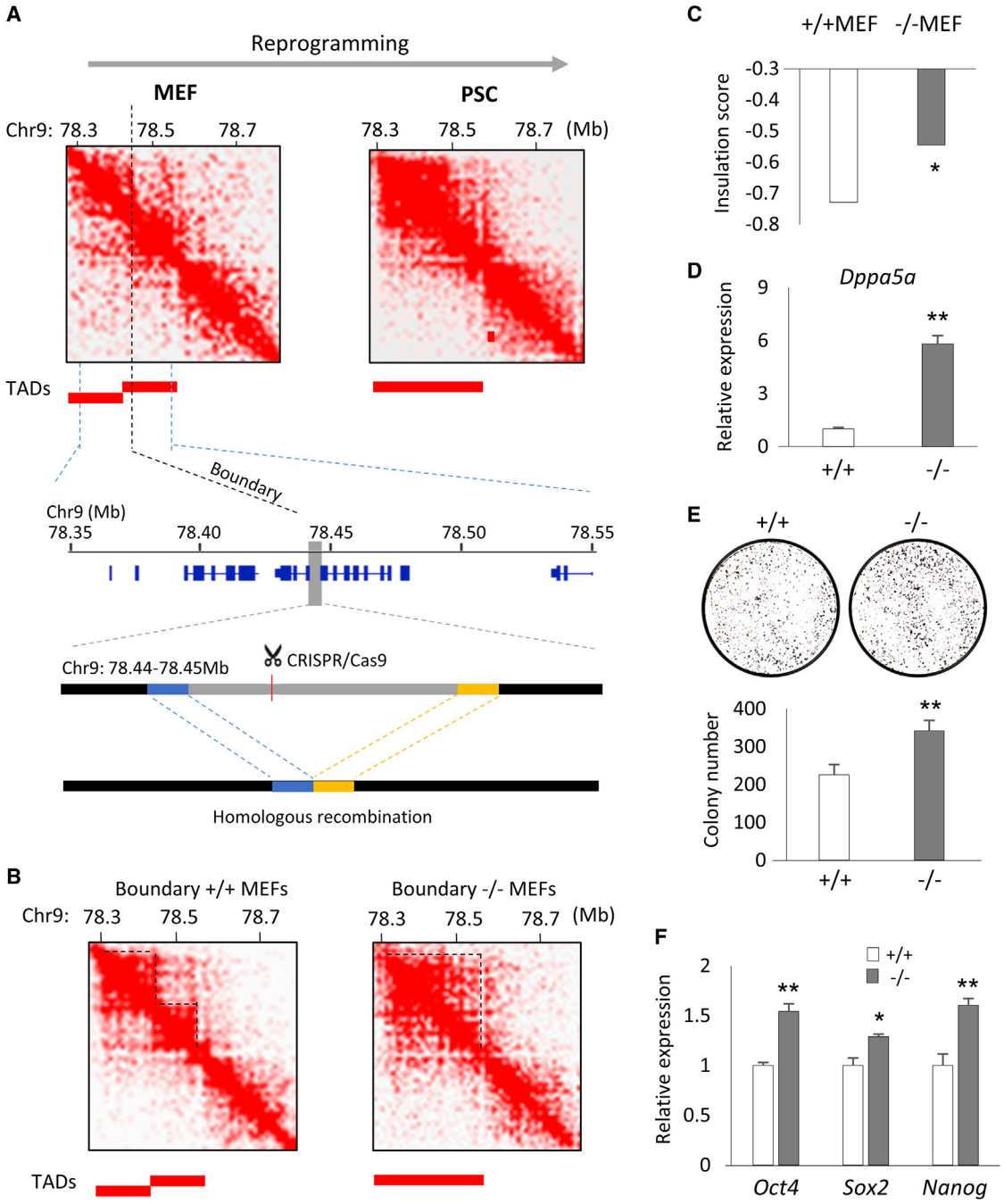


Figure 3. Boundary deletion-induced TAD reorganization promotes reprogramming

(A) TAD boundary between the TAD containing *Dppa5a* and the neighboring TAD is deleted by CRISPR-Cas9 genome-editing tools. The boundary is specific in MEFs and disappears in PSC during reprogramming. The boundary site is provided (dotted line). Red bars stand for TADs. Gray, blue, and yellow bars stand for the deletion region, the left homologous arm, and the right homologous arm, respectively.

(B) Surrounding HiC heatmaps showing merging of TADs in response to boundary deletion.

(C) Insulation score was significantly increased in response to the boundary deletion. *$p < 0.05$.

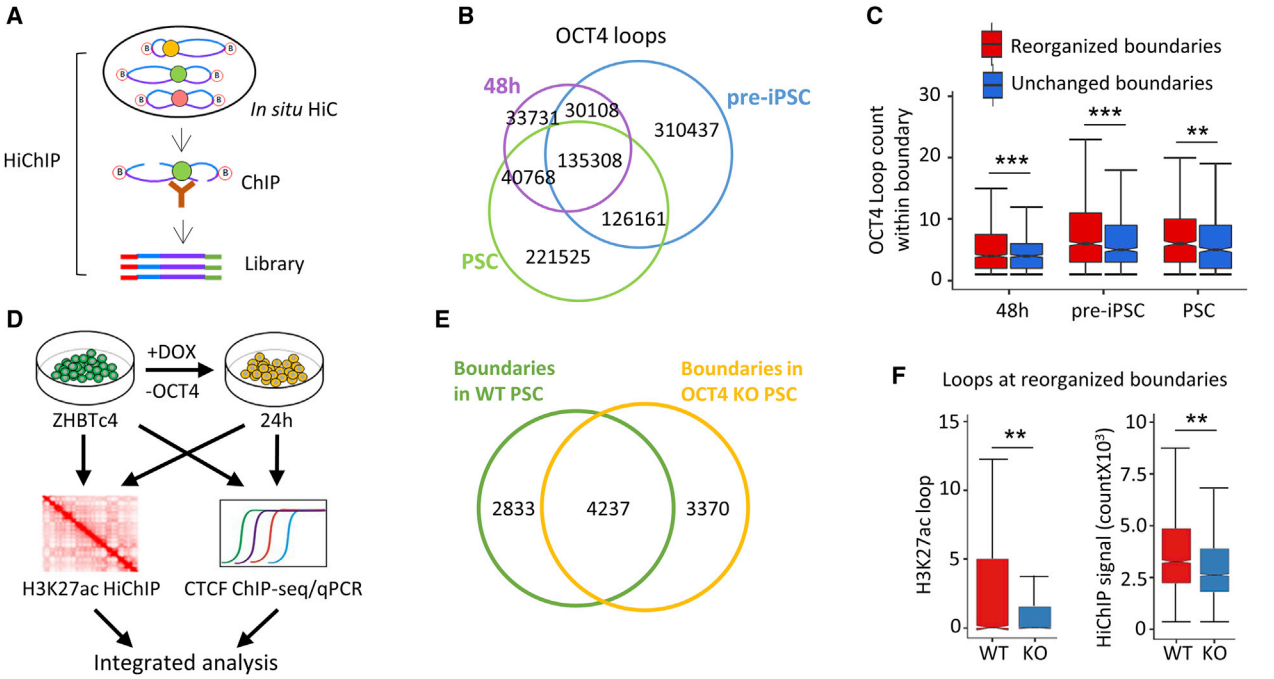
(D) The *Dppa5a* expression of the −/− MEFs is significantly higher than +/+ MEFs at the day 4 during reprogramming. **$p < 0.01$. Results are from three biological replicates.

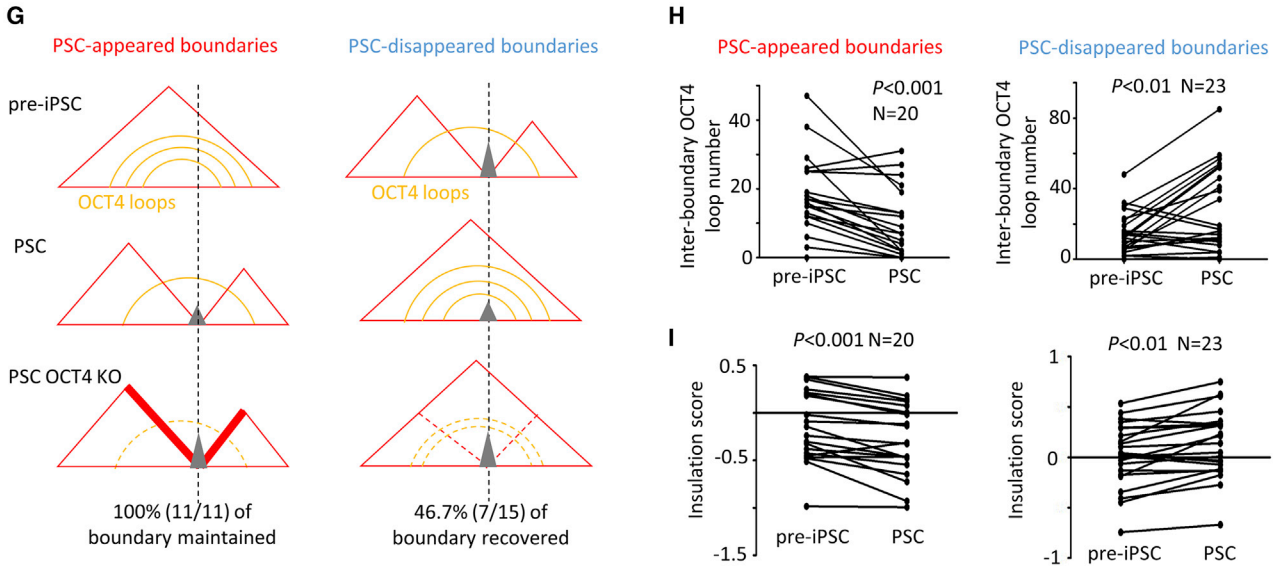
(E) Boundary deletion enhances reprogramming efficiency by increasing the number of iPSC colonies. **$p < 0.01$. Results are from three biological replicates.

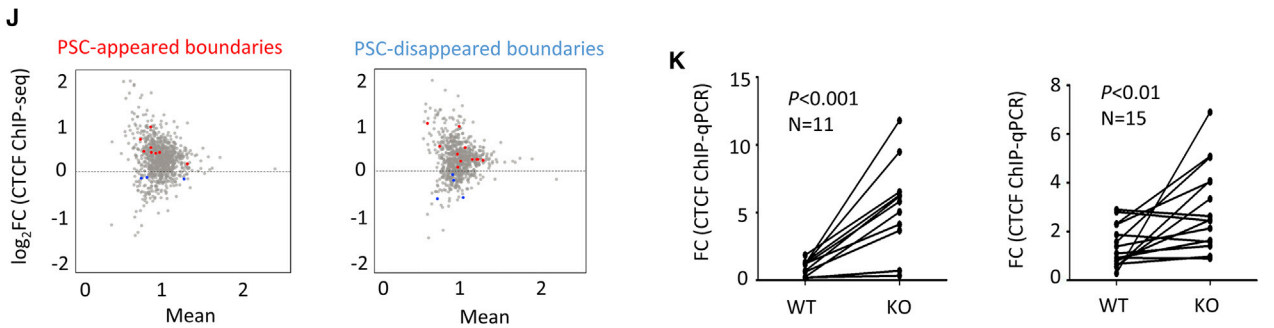
(F) Boundary deletion enhances reprogramming efficiency by increasing the expressions of key pluripotency genes of iPSC colonies. **$p < 0.05$, *$p < 0.01$. Results are from three biological replicates.

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

CellPress







(legend on next page)

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

To further dissect the underlying relationship between OCT4 loops and TAD reorganization, OCT4 was acutely depleted in the ZHBTc4 PSCs by 24 h of doxycycline (Dox) treatment (Niwa et al., 2000), and the changes in 3D chromatin architectures were examined by H3K27ac HiChIP, CTCF ChIP sequencing (ChIP-seq), and qPCR (Figure 4D). Acute depletion of OCT4 did not affect the expression of NANOG, structuring factors, or histone modification (Figure S4D), indicating that the cell status was not changed. Importantly, a large number of boundaries are reorganized upon OCT4 depletion (Figure 4E), indicating that OCT4 may directly regulate TAD reorganization. Moreover, loss of OCT4 significantly reduced the number and strength of H3K27ac loops at the reorganized boundaries (Figure 4F), further supporting the correlation between chromatin loops and TAD reorganization.

The mechanism underlying how OCT4 loops regulate TAD boundaries was analyzed during pre-iPSC to PSC transition (Figure 4G; Table S4). The inter-boundary OCT4 loops and inter-boundary insulation score were significantly reduced on the newly formed PSC boundary sites (Figures 4H and 4I, left), whereas both were significantly increased on the disappeared boundary sites (Figures 4H and 4I, right), suggesting that the dynamics of OCT4 loops may regulate TAD boundaries. Because CTCF has a crucial role in defining TAD boundaries (Dixon et al., 2012; Nora et al., 2017), we also determined the influence of OCT4 loss on CTCF binding strength. Intriguingly, OCT4 depletion increased CTCF binding strength at the newly formed boundaries in PSCs (Figures 4J and 4K, left), and the enhanced CTCF binding maintained those boundaries (11/11, 100%), despite OCT4 depletion (Figure 4G, left). Meanwhile, CTCF binding strength also increased at the boundary sites that disappeared in the PSCs (Figures 4J and 4K, right), resulting in the recovery of nearly half (7/15, 46.7%) of those boundaries in response to OCT4 deletion (Figure 4G, right). Because a previous study reported that DNA methylation influences CTCF binding (Bell and Felsenfeld, 2000), we investigated whether the increased CTCF binding upon OCT4 loss was associated with DNA methylation. However, bisulfite sequencing showed that this effect was not significantly correlated with DNA methylation (Figure S4E).

Taken together, these results demonstrate that the dynamics of concentrated OCT4 loops contributes to TAD reorganization through CTCF binding.

OCT4 is required for structuring factors binding on a subset of loop anchors

To determine whether and how OCT4 contributes to the formation of chromatin loops, we purified and characterized OCT4 protein complexes in the later three stages of reprogramming to construct stage-specific OCT4 interactomes by immunoprecipitation with mass spectrometry (IP-MS) (Figure 5A). We identified 76 and 140 OCT4-interacting proteins at the 48 h and pre-iPSC stages and cited the 198 OCT4-interacting proteins in PSCs published previously (Figure 5B; Table S5) (Ding et al., 2012). The encoding genes of most OCT4-interacting proteins were highly expressed in the pluripotency status (Figure S5A), and a significant proportion had positive RNAi hits in PSCs and reprogramming (Figure S5B), as reported previously (Chia et al., 2010; Ding et al., 2012; Fazzio et al., 2008; Golipour et al., 2012; Hu et al., 2009). This indicates that OCT4-interacting proteins are crucial for establishing and maintaining pluripotency.

Notably, in combinations of IP-MS with co-immunoprecipitation (co-IP) results, we observed that OCT4 physically interacts with chromatin-structuring factors, chromatin-remodeling factors, and the candidate-structuring factors (Figures 5B–5E) previously identified by ChIP-MS (Weintraub et al., 2017). Genome-wide ChIP-seq analysis showed that the binding of the key structuring factors CTCF, YY1, and SMC1/3 were enriched at the anchors of OCT4 loops (Figures 5F, S5C, and S5D), indicating that

Figure 4. Concentrated chromatin loops of OCT4 contribute to TAD reorganization

(A) Scheme of HiChIP method used in this study. The details are in the STAR Methods

(B) Venn diagrams of OCT4 loops in 48 h (purple border), pre-iPSC (blue border), and PSC (green border). OCT4 loops are cell specific.

(C) Boxplot showing the difference of OCT4 loop count between unchanged and reorganized boundary regions. Boundary is the genomic region that two boundary loci extend 10% of the TAD length toward the inside. Welch two-sample t test was used for statistics. $^{**}p < 0.01$. $^{***}p < 0.001$.

(D) Scheme of OCT4 KO study. H3K27ac HiChIP is performed in WT and OCT4 KO ZHBTc4 PSCs to detect the changes of loops and TAD boundaries in response to OCT4 depletion. CTCF ChIP-seq and ChIP-qPCR assays are performed to detect the binding change of CTCF in response to OCT4 depletion.

(E) Venn diagrams of TAD boundaries in WT (green circle) and OCT4 KO (yellow circle) ZHBTc4 PSCs by H3K27ac HiChIP analysis. A portion of 40% (2833/7070) of TAD boundaries in WT ZHBTc4 cells disappear and 44% (3370/7607) of TAD boundaries in OCT4 KO ZHBTc4 cells appear in response to OCT4 depletion.

(F) Boxplot showing the decrease of both H3K27ac loop number (left) and strength (right) at reorganized boundaries in response to OCT4 depletion. Wilcoxon rank-sum and signed-rank tests were used for statistics. $^{**}p < 0.01$.

(G) Two models for boundary change. The left and right models represent boundary appearance and disappearance from pre-iPSC to PSC, respectively. The change of boundaries in response to OCT4 depletion in each model is also displayed. Red triangles are TADs. Red dotted lines mean the boundary may recover. Thick red lines mean the boundary strength is enhanced. Yellow arcs are inter-boundary OCT4 loops. Gray triangle is CTCF. Black dotted lines are the loci with boundary change.

(H) The change of inter-boundary OCT4 loops corresponding to each model from pre-iPSC to PSC. For the 20 randomly selected genome loci corresponding to the model of boundary appearance (left panel), the number of inter-boundary OCT4 loops significantly decreases from pre-iPSC to PSC. In contrast, for the 23 randomly selected genome loci corresponding to the model of the boundary disappearance (right panel), the number of inter-boundary OCT4 loops significantly increases from pre-iPSC to PSC. The inter-boundary OCT4 loops are the OCT4 loops whose anchors belong to different TADs.

(I) The change of insulation score corresponding to each model from pre-iPSC to PSC. For the randomly selected genome loci corresponding to the models of the boundary appearance (left panel) or disappearance (right panel), the insulation score significantly decreases or increases from pre-iPSC to PSC, respectively.

(J) The global fold change (FC) of CTCF binding in response to OCT4 depletion for both models mentioned in (G). For both models, OCT4 KO significantly increases the global binding of CTCF by ChIP-seq. The 11 boundary-appearance loci and 15 boundary-disappearance loci mentioned in (K) are marked by color points, in which red stands for increased CTCF binding, whereas blue stands for decreased CTCF binding. Two biological replicates are assayed for ChIP-seq experiment.

(K) The FC of CTCF binding strength in response to OCT4 depletion for both boundary appearance (left, n = 11) and disappearance (right, n = 15) models. For both models, OCT4 KO significantly enhances the CTCF binding by ChIP-qPCR. Three biological replicates were assayed for ChIP-qPCR experiment.

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

CellPress

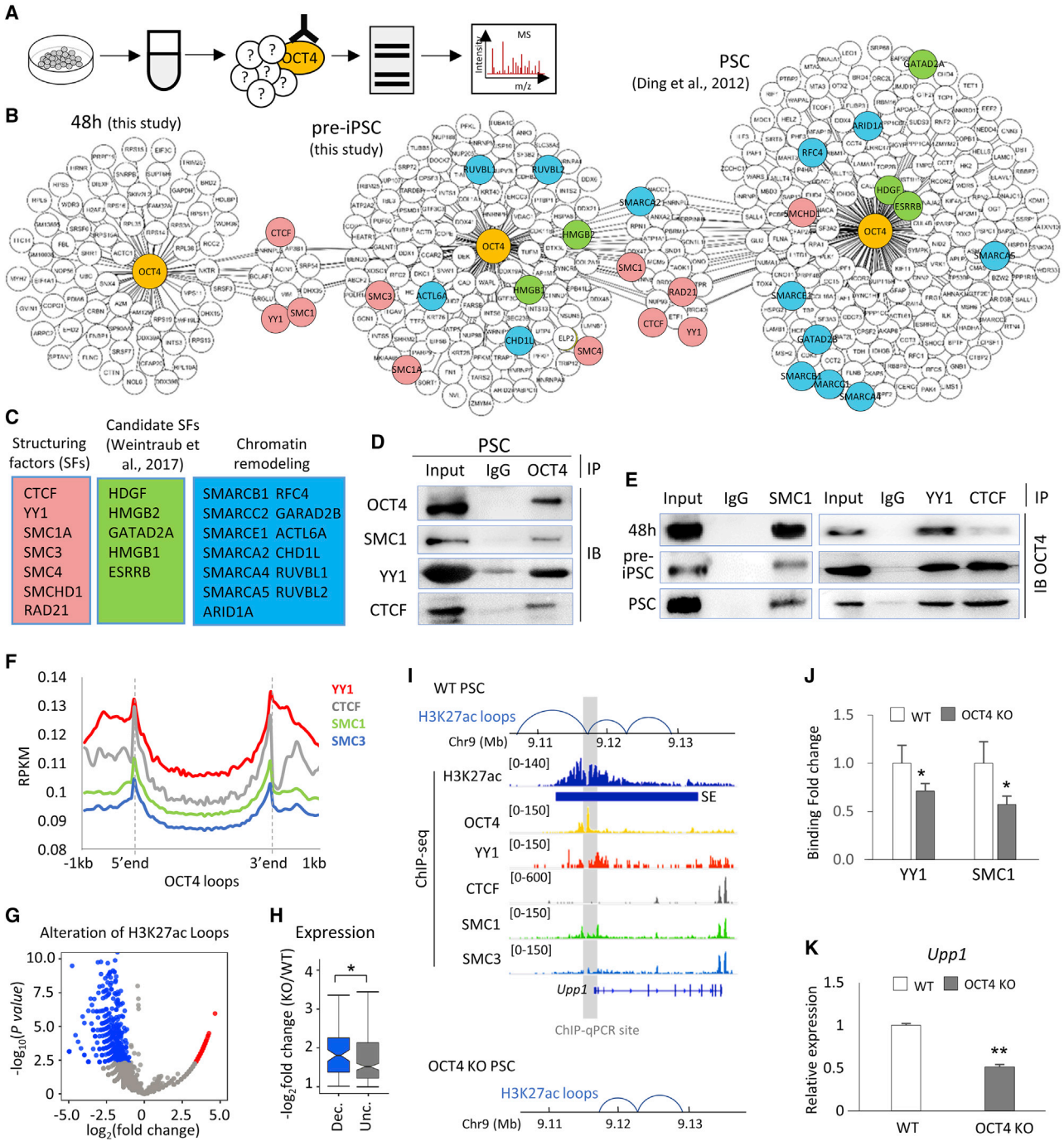


Figure 5. OCT4 is required for structuring factors binding on a subset of loop anchors

(A) Scheme showing the antibody-dependent IP-MS system for identifying OCT4 interactomes during reprogramming. The details are in the STAR Methods.

(B) Diagram depicting the OCT4 interacting proteins in each stage of reprogramming. SAINTexpress version 3.3 was used as a statistical tool to calculate the probability value of each potential protein-protein interaction compared with the background contaminants using default parameters (Teo et al., 2014). OCT4 is yellow, the well-known structuring factors are pink, several candidate structuring factors identified by ChIP-MS are green (Weintraub et al., 2017), and chromatin remodeling proteins are blue.

(C) The OCT4 interacting proteins with genome-shaping functions are listed in boxes with different colors. Pink box lists the well-known structuring factors, green box lists several candidate structuring factors identified by ChIP-MS (Weintraub et al., 2017), and blue box lists the chromatin remodeling proteins.

(D) Co-IP shows the interactions between OCT4 and the structuring factors. The nuclear extract (NE) of PSCs is immunoprecipitated by the OCT4 antibody, followed by immunoblotted with the antibodies of the structuring factors SMC1, YY1, and CTCF. Two biological replicates are assayed for Co-IP experiment.

(legend continued on next page)

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

the chromatin structuring factors cooperated with OCT4 to mediate the formation of OCT4 loops. This was further supported by a significant overlapping of OCT4 and SMC1 loops (Figures S5E–S5G). Moreover, different structuring factors preferred specific loop types. For instance, cohesin proteins were enriched at all kinds of loops, whereas OCT4, YY1, and CTCF were, respectively, enriched at SE, promoter, and boundary loops (Figure S5H). Furthermore, acute OCT4 depletion led to the loss of considerable H3K27ac-associated loops (Figure 5G), which, in turn, significantly affected gene expression (Figure 5H). The strength and number of loops at SE regions were also significantly reduced upon OCT4 loss (Figure S5I). For instance, OCT4 depletion results in loss of SE-loop in the nearby *Upp1* locus, reduces binding strength of YY1 and SMC1, and downregulates *Upp1* expression (Figures 5I–5K).

Taken together, these results demonstrate that OCT4 is required for the binding of structuring factors on a subset of loop anchors to form chromatin loops.

OCT4 phase separation regulates TAD reorganization

To further elucidate the relationship between the concentrated OCT4 loops and TAD reorganization, we defined OCT4 loop clusters as the genomic regions with a high density of OCT4 loop anchors (Table S3). Importantly, we found that TAD reorganization is likely to happen at the chromatin regions occupied by OCT4 loop clusters (Figure 6A). Because OCT4 loop clusters are the regions with aggregations of OCT4 protein, and OCT4 phase-separated condensates are enriched by OCT4 protein and its targeting chromatin loops (Boija et al., 2018; Hnisz et al., 2017), we hypothesized that the OCT4-phase-separated condensates may be enriched at OCT4 loop cluster regions and may contribute to TAD reorganization. To prove it, first, we found that OCT4 was phase separated *in vivo* during reprogramming, which was sensitive to 1,6-hexanediol (Figure S6A). Moreover, OCT4 can form droplets *in vitro* with MED1 intrinsically disordered regions (IDRs) (Figure S6B), which were rapidly recovered from photobleaching (Figures S6C and S6D) and fused (Figure S6E). These results indicate that the OCT4 phase separation has liquid-like behavior (Lin et al., 2015). Second, a combination analysis of DNA fluorescence *in situ* hybridization (FISH) and immunofluorescence (IF) showed a higher co-localization rate between OCT4-phase-separated condensates and OCT4-loop-cluster regions, rather than non-loop-cluster regions (Figures 6B and 6C). Given that OCT4 loop clusters are highly correlated with TAD reorganization, the latter is very likely to happen within OCT4 phase separation.

To clarify the role of OCT4 phase separation on TAD reorganization, we generated two types of OCT4 mutants to disrupt OCT4 phase (Figure 6D). Mutant 1 was an acidic mutant, wherein all acidic residues in the IDRs were replaced with alanine (Boija et al., 2018). Because acidic amino acids in OCT4 IDRs may facilitate the interaction with MED IDR, this mutation type may be defective in its ability to phase separate (Boija et al., 2018). To further verify whether TAD reorganization is correlated with the OCT4 phase, we generated mutant 2, which was a deletion mutant lacking three predicted phase-separation-related amino acids in the C-terminal IDR (amino acid 324–326) (Figures 6D and S6F). We observed that both mutants displayed a disabled capability to phase separate *in vitro* and *in vivo* (Figures 6E, 6F, S6G, and S6H). Importantly, both mutants attenuated TAD reorganization induced by wild-type (WT) OCT4 (Figures 6G–6I and S6I–S6M). As a result, the reprogramming efficiency was reduced upon these mutations (Figures 6J and 6K).

To further validate the inhibitory role of OCT4 mutants on TAD reorganization and reprogramming resulted from their disabled phase-separated ability, we generated the rescued OCT4 mutants by fusing the IDR of FUS, which is known to drive phase separation (Figure 6D) (Liu et al., 2020; Shin et al., 2017). Intriguingly, IDR fusion rescued the phase-separation capability of the OCT4 mutants (Figures 6E, 6F, S6G, and S6H). Importantly, the attenuated TAD reorganization of both OCT4 mutants was restored by phase rescue via IDR fusion (Figures 6G–6I). As a result, the reduced reprogramming efficiency of the OCT4 mutants was recovered in response to IDR fusion (Figures 6J and 6K). These results illustrate that disruption of OCT4 phase separation suppresses TAD reorganization and reprogramming, which can be restored by phase rescue via IDR fusion.

Taken together, these results demonstrate that OCT4 phase separation has a key role in TAD reorganization for efficient reprogramming.

Identification of reprogramming regulators by TADMAN

Because TAD reorganization is important for reprogramming, we established a new algorithm that integrates TAD reorganization-based multiomics analysis (TADMAN) to identify the reprogramming regulators. The genes encoding the regulators should meet the following criteria via different techniques: (1) they are regulated by SEs in reorganized TAD (HiC); (2) the genes and SEs are connected by OCT4 loops (HiChIP); (3) both the genes and SEs are occupied by OCT4 peaks (ChIP-seq); and (4) the genes are the top 10% highly expressed genes at any stage of reprogramming and have at least a 10-fold change in expression compared with the previous or next stage (by RNA sequencing [RNA-seq]) (Figure 7A). Accordingly, we obtained 59 candidate regulators (Figure 7B). The candidates contain several well-known

(E) Co-IP shows the interactions between the structuring factors and OCT4. The NEs of 48 h, pre-iPSC, and PSC are immunoprecipitated by the antibodies of the structuring factors SMC1, YY1, and CTCF, respectively, followed by immunoblotted with OCT4 antibody. Two biological replicates are assayed for Co-IP experiment.

(F) ChIP-seq RPKM (reads per kilo base per million mapped reads) of the structuring factors YY1, CTCF, SMC1, and SMC3 around the center of OCT4 loops in PSC.

(G) Volcano plot showing the fold change of H3K27ac loops strength between ZHBTc4 WT and OCT4 KO cells. Blue, red, or gray scatter points stand for the loops with decreased, increased, or unchanged strength in response to OCT4 depletion, respectively.

(H) The decreased loops (blue) significantly affect gene expression more than the unchanged loops (gray) by RNA-seq analysis (King and Klose, 2017). *$p < 0.05$.

(I) One example is given near the *Upp1* locus, which shows loss of OCT4 induced by the loss of the SE loop.

(J) The binding fold change of YY1 and SMC1 at the disappeared loop anchors the nearby *Upp1* promoter in response to OCT4 depletion by ChIP-qPCR. Data are presented as means ± SD. *$p < 0.05$. Three biological replicates are assayed for ChIP-qPCR experiment.

(K) OCT4 depletion results in downregulation of *Upp1* expression by RT-qPCR. **$p < 0.01$.

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

CellPress

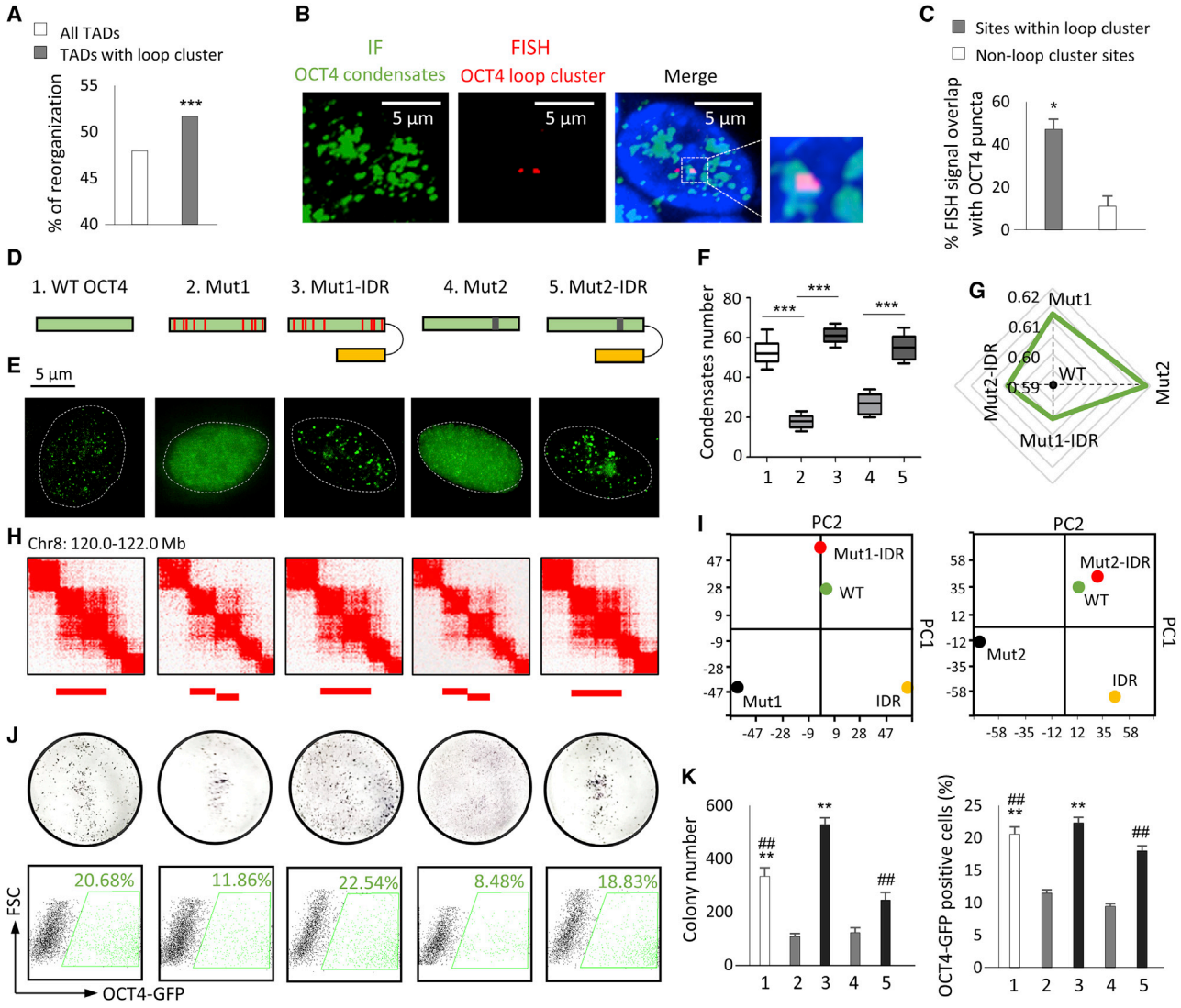


Figure 6. OCT4 phase separation regulates TAD reorganization

(A) Comparison of TAD reorganization incidence between all TADs (white bar) and the TADs occupied by the OCT4 loop cluster (gray bar). A loop anchor cluster with a high OCT4 HiChIP signal was considered as the OCT4 loop cluster. The OCT4 HiChIP signal was not evenly distributed across the loop anchor regions. We found the x axis point for which a line with a slope of 1.0 was tangent to the curve. We define loop anchors above that point to be the OCT4 loop cluster. A hypergeometric test was used for statistics. ***$p < 0.001$.

(B) Representative images showing the co-localization of OCT4 phase-separated condensates (IF) with OCT4 loop clusters (DNA FISH).

(C) Column showing the quantification of co-localization between OCT4 phase-separated condensates and OCT4 loop cluster sites or non-OCT4 loop cluster sites. *$p < 0.05$.

(D) The pictures showing different perturbations of OCT4 protein. (1) WT OCT4; (2) acidic mutant OCT4:glutamic acid (E) and aspartic acid (D) residues to alanine (A) residues, and the red bars stand for A; (3) fusing acidic mutant OCT4 with FUS IDR (yellow bar); (4) del324-326 OCT4, and gray bar stands for the deletion site; and (5) fusing del324-326 OCT4 with FUS IDR.

(E) IF imaging showing the *in vivo* phase-separated ability of the WT, mutants, and IDR-fused mutants. The areas enclosed by white dotted lines are nuclei.

(F) Comparison of *in vivo* phase-separated condensate numbers among WT, mutants, and IDR-fused mutants. ***$p < 0.001$.

(G) Spidergram depicting the Jaccard distance to the TAD pattern of WT OCT4 between mutants and IDR-fused mutants. The core of the spidergram stands for the WT OCT4. Higher distance to WT OCT4 means higher dissimilarity.

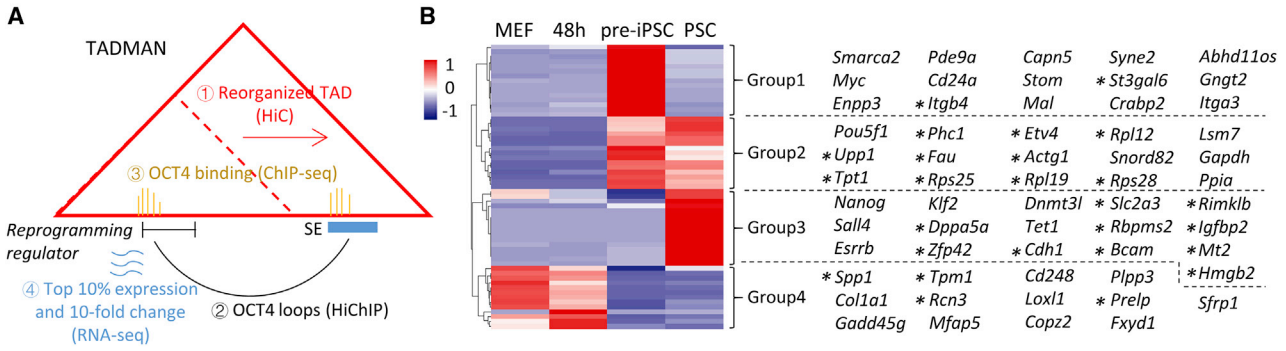
(H) Hi-C heatmaps showing the difference of TAD patterns among WT OCT4, mutants, IDR-fused mutants, or FUS IDR overexpressing MEFs. The difference between WT and mutants is compensated by IDR fusion. Red bars stand for TADs.

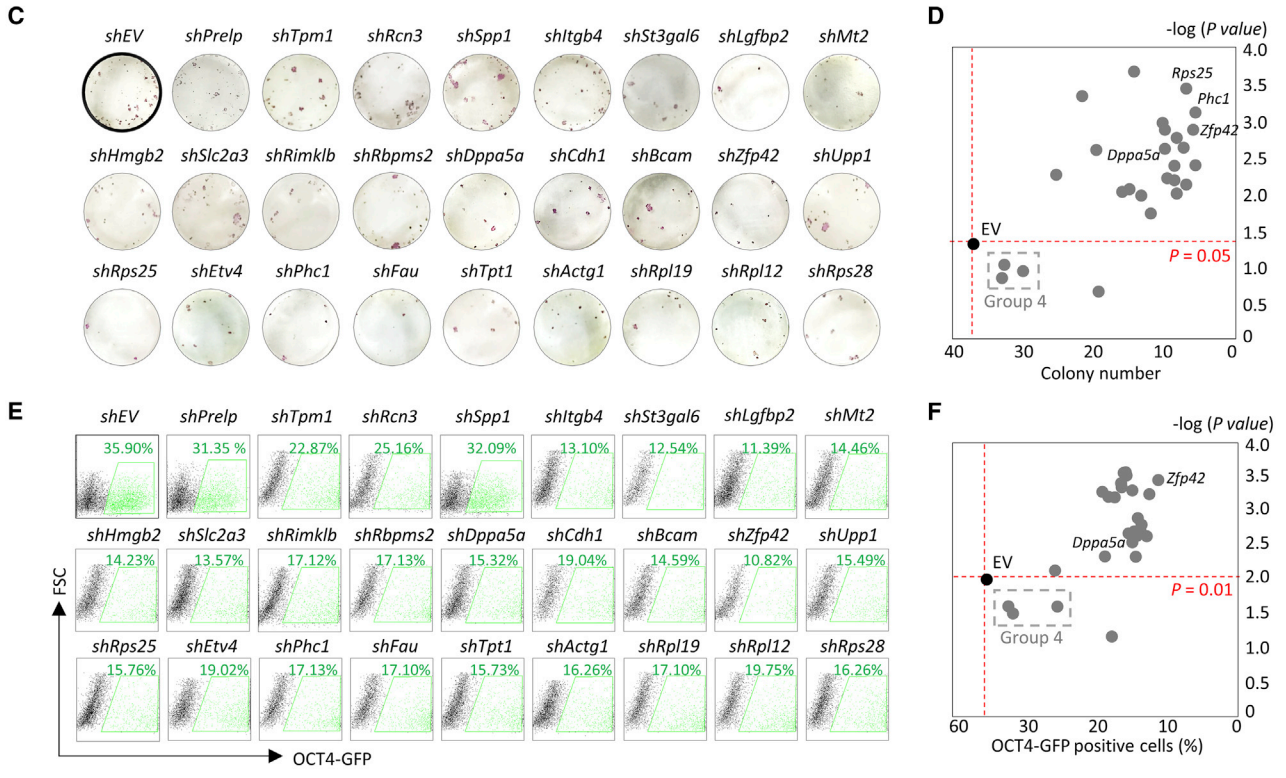
(I) Principal component analysis (PCA) of TAD boundaries of WT, mutants, IDR-fused mutants, or IDR overexpressing MEFs. Top panel is for acidic mutant, whereas the bottom panel is for deletion mutant.

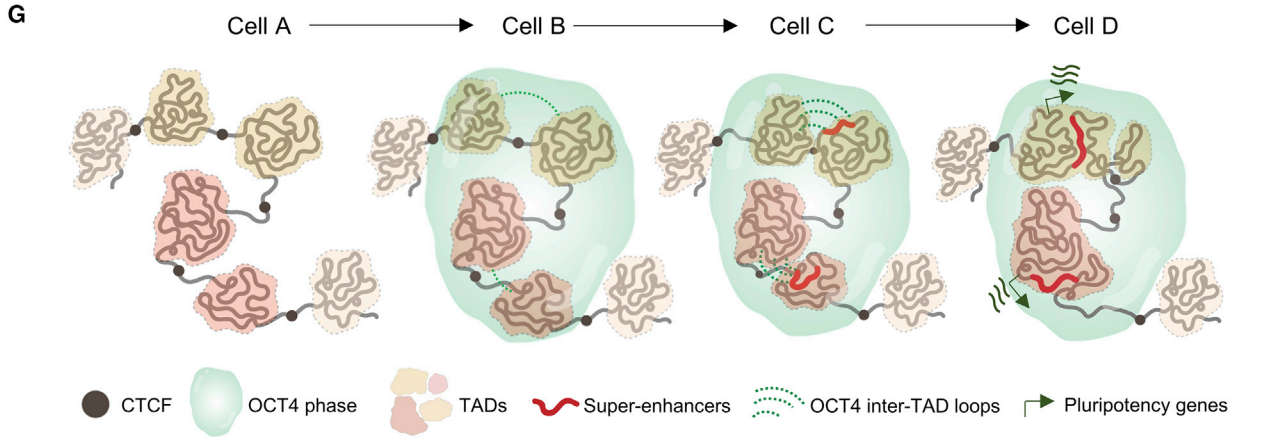
(J) iPSC colony formation assay (top panel) and OCT4-GFP-positive cell flow cytometry (bottom panel) are performed to compare the influence of WT, mutants, or IDR-fused mutants on reprogramming efficiency. Mutations of OCT4 reduce the reprogramming efficiency, which is recovered by IDR fusion.

(K) Quantification of iPSC colonies (left panel) and OCT4-GFP-positive cells (right panel). Data are presented as means ± SD. **$p < 0.01$ compared with Mut1; ##$p < 0.01$ compared with Mut2. iPSC colony numbers and OCT4-GFP-positive cells are from three biological replicates.

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023







(legend on next page)

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

pluripotency genes, such as *Oct4*, *Sall4*, *Nanog*, *Zfp42*, and *Esrrb*, indicating the potential reliability of TADMAN. All candidates were divided into the upregulated (group 1–3) and downregulated (group 4) groups according to their expression profiles during reprogramming (Figure 7B).

We randomly selected 26 of the candidates for functional validation in reprogramming. Importantly, knockdown (KD) of most candidates in group 1–3 significantly decreased reprogramming efficiency by reducing the number of iPSC colonies (Figures 7C and 7D), OCT4-GFP-positive cells (Figures 7E and 7F), and the expression levels of the key pluripotency genes (Figure S7A). In contrast, KD of the candidates in group 4 had little influence on reprogramming efficiency (Figures 7C–7F and S7A). These results demonstrate that TADMAN is an effective and reliable method to identify reprogramming regulators with increased expression during reprogramming. Therefore, we removed group 4 from the candidate list and positively validated 20 reprogramming regulators from the 22 (91%) candidates (Figure S7B). TADMAN is also applicable for predicting neural-cell-fate regulators during PSC to neural progenitor cell (NPC) differentiation (Figures S7C–S7E) (Bonev et al., 2017), whose effects were validated by functional studies (Figures S7F–S7H).

DISCUSSION

In this study, we found that the master transcription factor OCT4 forms phase-separated condensates, which contribute to TAD reorganization and cell fate transitions (Figure 7G). Phase separation has been proven to cluster chromatin loops in forming special regulatory elements that compartmentalize and concentrate the transcription apparatus for robust gene expressions (Boija et al., 2018; Hnisz et al., 2017; Sabari et al., 2018). However, no direct evidence shows that phase separation of the master transcription factors regulates the organization of higher-order 3D structures, such as TADs. Here, we demonstrate that OCT4 phase separation has a key role in TAD reorganization. This observation extends the traditional view of the master transcription factors in regulating gene expression and cell fate. For example, previous studies showed that OCT4 functions as both the pioneer factor to open heterochromatin and the transcription factor to recruit stage-specific enhancers for reprogramming (Chronis et al., 2017; Cirillo et al., 2002; Soufi et al., 2012, 2015). The newly identified function of OCT4 in directly regulating TAD reorganization via phase separation establishes a bridge between the master transcription factor and the 3D genome. Furthermore, OCT4 phase separation concentrates chromatin loops to regulate TAD reorganization, which is correlated to SE redistribution in TADs to change their targets. These findings provide several new insights into the 3D genome and cell fate transitions. First, TAD reorganization may directly contribute to transcriptional regulation and cell fate transitions. Second, lower-order 3D structures, such as chromatin loops, may determine the organization of higher-order 3D-structure TADs. Third, TAD reorganization can be manipulated to promote cell fate transitions.

Disruption of TAD boundaries could lead to abnormal gene transcriptions and diseases (Lupiáñez et al., 2015), indicating manipulation of TAD structures may influence cell fate transitions. The 4D nucleome (4DN)-related studies provide platforms to identify key TAD architectures for cell fate (Dekker et al., 2017), and it is supposed that pre-establishment of the TAD structures of the late stage cells in the early stage cells may promote cell fate transitions. Deletion of TAD boundary is one of the ideal strategies to manipulate TAD structure, which has proven that deletion of HERV-H sequences in TAD boundaries lead to the merging of TADs in human PSCs, resulting in gene expression change and increased efficiency of PSC to cardiomyocyte differentiation (Zhang et al., 2019). Another strategy is to shorten the distance between two neighboring TADs by artificial linking, which may result in TAD merging. A chemically inducible dCas9-based tool has been validated to effectively establish chromatin loops and influence gene expression (Morgan et al., 2017). That tool has two advantages: first, it does not change the genome, and second, the artificial looping is reversible by withdrawing the chemical (Morgan et al., 2017). In this study, we assessed the feasibility of

Figure 7. Identification of reprogramming regulators by TADMAN

(A) Scheme for TADMAN, which was used to identify reprogramming regulators. The model comprises four conditions: (1) *in situ* Hi-C, the encoding genes of the regulators are regulated by SEs in reorganized TADs; (2) HiChIP, OCT4 loops link the encoding genes of the regulators with SEs; (3) ChIP-seq, the SEs and the encoding genes are co-occupied by OCT4 peaks; and (4) RNA-seq: the encoding genes are the top 10% of the highly expressed genes at any stage of reprogramming and are at least 10-fold change of expression from the previous or next stage.

(B) Heatmap showing the dynamic expression of the 59 candidate reprogramming regulators by TADMAN during reprogramming (left). According to the expression profile, the candidates are divided into four groups, in which groups 1–3 are composed of the candidates with upregulated expressions during reprogramming, whereas group 4 is composed of the candidates with downregulated expressions (right). Asterisk stands for the 26 randomly selected candidates for functional validation in this study.

(C and D) The reprogramming efficiency is evaluated by iPSC colony formation assay. After 12 days of reprogramming, the iPSC colonies of each knockdown (KD) test are stained by alkaline phosphatase (C). Both colony number (compared with *shEV*) and p value ($p < 0.05$) were considered for significance analysis (D). The candidates in group 4 (marked with a dotted line box) were not effectively validated. The area on the right of the vertical red dotted line means that the iPSC colony number is more than that of the control. The area above the horizontal red dotted line is $p < 0.05$ as compared with that of the control. iPSC colony numbers at day 12 of reprogramming are from three biological replicates.

(E and F) The reprogramming efficiency is evaluated by the ratio of OCT4-GFP-positive cells. After 12 days of reprogramming, the cells in each KD test were harvested for flow cytometry analysis to detect the ratio of OCT4-GFP-positive cells (E). Both the ratio of OCT4-GFP-positive cells (compared with *shEV*) and the p value ($p < 0.01$) were considered for significance analysis (F). The candidates in group 4 (marked by dotted line box) were not effectively validated. The area on the right of the vertical red dotted line means that the OCT4-GFP-positive cells were more than that of the control. The area above the horizontal red dotted line is $p < 0.01$ as compared with that of the control. The ratios of OCT4-GFP-positive cells at day 12 of reprogramming are from three biological replicates.

(G) Model showing how OCT4 condensates contribute to TAD reorganization for reprogramming. In brief, OCT4 forms condensates at the place at which TAD reorganization would intensively take place. In the condensates, the inter-TAD OCT4 loops close the distance with the neighboring TADs. As the numbers of inter-TAD loops increase, CTCF binding at the TAD boundaries decreases, which results in TAD reorganization. TAD reorganization is correlated to the SE redistribution in the TADs, which activates the expressions of genes with cell identity.

both strategies and observed that both boundary deletion and chemical induction could effectively manipulate TAD structures.

Identification of cell fate regulators in the light of the dynamics of 3D genome is also a promising way for cell fate guidance. However, very few studies so far have provided the related models. Our TADMAN approach, which can precisely predict cell fate regulators fills the gap. One unique feature of TADMAN is that it integrates both chromatin loops and TAD reorganization. Therefore, according to a TADMAN prediction, artificial pre-establishment of pluripotent chromatin loops in somatic cells not only links two loci on the genome but also pre-shapes the pluripotent TAD structures, which markedly improves cell fate transitions. However, TADMAN can only effectively identify cell fate drivers, rather than repressors, although both are cell fate regulators. This may be because the key features of TADMAN, such as TAD reorganization and SEs, are mainly for transcriptional activation rather than repression. Nonetheless, the reliability of TADMAN has been validated in different cell fate transition systems, which makes it a promising tool for predicting biomarkers of diseases, such as carcinogenesis and aging.

Limitations of study

One of the main findings of the study is that the artificial organization of a higher-order 3D structure, such as TAD, can regulate gene expression and promote cell fate transitions. However, whether these artificial 3D structures would replace the master transcription factors for cell fate control remains unknown. For example, could representative artificial chromatin loops or TADs replace OCT4 for reprogramming? It will be of interest to examine whether the essential 3D structures would take the place of the master transcription factors for cell fate decision. Moreover, we concluded that phase separation contributes to TAD reorganization during cell fate transitions. Whether phase separation contributes to other higher-order 3D structures, such as compartment switch and the formation of long-range chromatin loops, is still unknown. Furthermore, we established a TAD-reorganization-based algorithm for the prediction of cell fate regulators, whose effectiveness and accuracy have been proven in reprogramming and NPC differentiation. Whether that algorithm could be widely used to identify biomarkers in pathological cell fate transitions is unclear. It would be of interest to test the potential application of this algorithm in carcinogenesis, aging, or other pathological processes. Therefore, further studies should provide important insights into these issues and describe the integral network among phase separation, 3D genome, and cell fate.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.stem.2021.04.023.

ACKNOWLEDGMENTS

This research was funded by grants from the National Key Research and Development Program (2016YFA0101700 and 2017YFA0102800), the National Natural Science Foundation of China (grant nos. 31970811 and 31771639), the Frontier Research Program of Bioland Laboratory (Guangzhou Regenerative Medicine and Health Guangdong Laboratory) (2018GZR110105007), the Guangdong Innovative and Entrepreneurial Research Team Program (2016ZT06S029) to J.D., and the National Natural Science Foundation of China (grant no. 81703086) to J.W. We thank Professor Kathrin Plath for the gift of pre-iPSCs, Professor Jiekai Chen for the gift of OG2 MEFs and OD14 MEFs, and Professor Richard A. Young for the gift of the pETEC-OCT4-acidicmutant-GFP vector.

AUTHOR CONTRIBUTIONS

Conceptualization, J.D., Jia Wang, H.Y., and Q.M.; methodology, Jia Wang, H.Y., Q.M., P.Z., X.L. J.S., and J.D.; software, H.Y., P.Z., and X.L.; formal analysis, H.Y., P.Z., Jian Wang, Q.M., X.L., J.S., and Y.C.; investigation, Q.M., Jia Wang, D.W., J.C., Y.Y., Q.T., C.Z., and X.H.; resources, C.L., Y.S., and Y.M.;

writing – original draft, Jian Wang, Q.M., and H.Y.; writing – review & editing, all authors; visualization, Jia Wang, H.Y., Q.M., P.Z., X.L., and Y.Y.; supervision, J.D.; funding acquisition, J.D. and Jia Wang.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 2, 2019
Revised: August 6, 2020
Accepted: April 19, 2021
Published: May 25, 2021

REFERENCES

Abad, M., Mosteiro, L., Pantoja, C., Cañamero, M., Rayon, T., Ors, I., Graña, O., Megías, D., Domínguez, O., Martínez, D., et al. (2013). Reprogramming *in vivo* produces teratomas and iPS cells with totipotency features. Nature *502*, 340–345.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Beagan, J.A., Gilgenast, T.G., Kim, J., Plona, Z., Norton, H.K., Hu, G., Hsu, S.C., Shields, E.J., Lyu, X., Apostolou, E., et al. (2016). Local genome topology can exhibit an incompletely rewired 3D-folding state during somatic cell reprogramming. Cell Stem Cell *18*, 611–624.

Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482–485.

Bergsland, M., Ramsköld, D., Zaouter, C., Klum, S., Sandberg, R., and Muhr, J. (2011). Sequentially acting Sox transcription factors in neural lineage development. Genes Dev. *25*, 2453–2464.

Boija, A., Klein, I.A., Sabari, B.R., Dall'Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K., Manteiga, J.C., Hannett, N.M., et al. (2018). Transcription factors activate genes through the phase-separation capacity of their activation domains. Cell *175*, 1842–1855.e1816.

Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.P., Tanay, A., and Cavalli, G. (2017). Multiscale 3D genome rewiring during mouse neural development. Cell *171*, 557–572.e24.

Chia, N.Y., Chan, Y.S., Feng, B., Lu, X., Orlov, Y.L., Moreau, D., Kumar, P., Yang, L., Jiang, J., Lau, M.S., et al. (2010). A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. Nature *468*, 316–320.

Choi, H., Larsen, B., Lin, Z.Y., Breitkreutz, A., Mellacheruvu, D., Fermin, D., Qin, Z.S., Tyers, M., Gingras, A.C., and Nesvizhskii, A.I. (2011). SAINT: probabilistic scoring of affinity purification-mass spectrometry data. Nat. Methods *8*, 70–73.

Chronis, C., Fiziev, P., Papp, B., Butz, S., Bonora, G., Sabri, S., Ernst, J., and Plath, K. (2017). Cooperative binding of transcription factors orchestrates reprogramming. Cell *168*, 442–459.e20.

Cirillo, L.A., Lin, F.R., Cuesta, I., Friedman, D., Jarnik, M., and Zaret, K.S. (2002). Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. Mol. Cell *9*, 279–289.

Dall'Agnese, A., Caputo, L., Nicoletti, C., di Iulio, J., Schmitt, A., Gatto, S., Diao, Y., Ye, Z., Forcato, M., Perera, R., et al. (2019). Transcription factor-directed re-wiring of chromatin architecture for somatic cell nuclear reprogramming toward trans-differentiation. Mol. Cell *76*, 453–472.e8.

de Wit, E., Bouwman, B.A., Zhu, Y., Klous, P., Splinter, E., Verstegen, M.J., Krijger, P.H., Festuccia, N., Nora, E.P., Welling, M., et al. (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature *501*, 227–231.

Dekker, J., Belmont, A.S., Guttman, M., Leshyk, V.O., Lis, J.T., Lomvardas, S., Mirny, L.A., O'Shea, C.C., Park, P.J., Ren, B., et al.; 4D Nucleome Network (2017). Corrigendum: The 4D nucleome project. Nature *552*, 278.

Ding, J., Xu, H., Faiola, F., Ma'ayan, A., and Wang, J. (2012). Oct4 links multiple epigenetic pathways to the pluripotency network. Cell Res. *22*, 155–167.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331–336.

Dray, S., and Dufour, A. (2007). The ade4 Package: Implementing the Duality Diagram for Ecologists. Journal of Statistical Software *22*, 1–20. https://doi.org/10.18637/jss.v022.i04.

Fazzio, T.G., Huff, J.T., and Panning, B. (2008). An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. Cell *134*, 162–174.

Fraser, J., Ferrai, C., Chiariello, A.M., Schueler, M., Rito, T., Laudanno, G., Barbieri, M., Moore, B.L., Kraemer, D.C., Aitken, S., et al.; FANTOM Consortium (2015). Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. Mol. Syst. Biol. *11*, 852.

Gibson, B.A., Doolittle, L.K., Schneider, M.W.G., Jensen, L.E., Gamarra, N., Henry, L., Gerlich, D.W., Redding, S., and Rosen, M.K. (2019). Organization of chromatin by intrinsic and regulated phase separation. Cell *179*, 470–484.e21.

Golipour, A., David, L., Liu, Y., Jayakumaran, G., Hirsch, C.L., Trcka, D., and Wrana, J.L. (2012). A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. Cell Stem Cell *11*, 769–782.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell *155*, 934–947.

Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.L., Bak, R.O., Li, C.H., Goldmann, J., Lajoie, B.R., Fan, Z.P., Sigova, A.A., et al. (2016). Activation of proto-oncogenes by disruption of chromosome neighborhoods. Science *351*, 1454–1458.

Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., and Sharp, P.A. (2017). A phase separation model for transcriptional control. Cell *169*, 13–23.

Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S.H., and Elledge, S.J. (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes Dev. *23*, 837–848.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430–435.

King, H.W., and Klose, R.J. (2017). The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. eLife *6*, e22631.

Kloetgen, A., Thandapani, P., Ntziachristos, P., Ghebrechristos, Y., Nomikou, S., Lazaris, C., Chen, X., Hu, H., Bakogianni, S., Wang, J., et al. (2020). Three-dimensional chromatin landscapes in T cell acute lymphoblastic leukemia. Nat. Genet. *52*, 388–400.

Krijger, P.H., Di Stefano, B., de Wit, E., Limone, F., van Oevelen, C., de Laat, W., and Graf, T. (2016). Cell-of-origin-specific 3D genome structure acquired during somatic cell reprogramming. Cell Stem Cell *18*, 597–610.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lareau, C.A., and Aryee, M.J. (2018). hichipper: a preprocessing pipeline for calling DNA loops from HiChIP data. Nat. Methods *15*, 155–156.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Li, G., Chen, Y., Snyder, M.P. and Zhang, M.Q. (2016) ChIA-PET2: a versatile and flexible pipeline for ChIA-PET data analysis. Nucleic acids research. https://doi.org/10.1093/nar/gkw809.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al.

(2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289–293.

Lin, Y., Protter, D.S., Rosen, M.K., and Parker, R. (2015). Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. Mol. Cell *60*, 208–219.

Liu, X., Shen, J., Xie, L., Wei, Z., Wong, C., Li, Y., Zheng, X., Li, P., and Song, Y. (2020). Mitotic implantation of the transcription factor prospero via phase separation drives terminal neuronal differentiation. Dev. Cell *52*, 277–293.e8.

Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell *153*, 320–334.

Lun, A.T., and Smyth, G.K. (2015). diffHic: a Bioconductor package to detect differential genomic interactions in Hi-C data. BMC Bioinformatics *16*, 258.

Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell *161*, 1012–1025.

Morgan, S.L., Mariano, N.C., Bermudez, A., Arruda, N.L., Wu, F., Luo, Y., Shankar, G., Jia, L., Chen, H., Hu, J.F., et al. (2017). Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. Nat. Commun. *8*, 15993.

Mumbach, M.R., Rubin, A.J., Flynn, R.A., Dai, C., Khavari, P.A., Greenleaf, W.J., and Chang, H.Y. (2016). HiChIP: efficient and sensitive analysis of protein-directed genome architecture. Nat. Methods *13*, 919–922.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat. Genet. *24*, 372–376.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381–385.

Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. Cell *169*, 930–944.e22.

Paulsen, J., Liyakat Ali, T.M., Nekrasov, M., Delbarre, E., Baudement, M.O., Kurscheid, S., Tremethick, D., and Collas, P. (2019). Long-range interactions between topologically associating domains shape the four-dimensional genome during differentiation. Nat. Genet. *51*, 835–843.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Sabari, B.R., Dall'Agnese, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C., et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. Science *361*, eaar3958.

Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.J., Vert, J.P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. *16*, 259.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., and Ren, B. (2012). A map of the *cis*-regulatory sequences in the mouse genome. Nature *488*, 116–120.

Shin, H., Shi, Y., Dai, C., Tjong, H., Gong, K., Alber, F., and Zhou, X.J. (2016). TopDom: an efficient and deterministic method for identifying topological domains in genomes. Nucleic Acids Res. *44*, e70.

Shin, Y., Berry, J., Pannucci, N., Haataja, M.P., Toettcher, J.E., and Brangwynne, C.P. (2017). Spatiotemporal control of intracellular phase transitions using light-activated optoDroplets. Cell *168*, 159–171.e14.

Siersbæk, R., Madsen, J.G.S., Javierre, B.M., Nielsen, R., Bagge, E.K., Cairns, J., Wingett, S.W., Traynor, S., Spivakov, M., Fraser, P., and Mandrup, S. (2017). Dynamic rewiring of promoter-anchored chromatin loops during adipocyte differentiation. Mol. Cell *66*, 420–435.e5.

Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell *151*, 994–1004.

Soufi, A., Garcia, M.F., Jaroszewicz, A., Osman, N., Pellegrini, M., and Zaret, K.S. (2015). Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. Cell *161*, 555–568.

Sridharan, R., Gonzales-Cope, M., Chronis, C., Bonora, G., McKee, R., Huang, C., Patel, S., Lopez, D., Mishra, N., Pellegrini, M., et al. (2013). Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1γ in reprogramming to pluripotency. Nat. Cell Biol. *15*, 872–882.

Stadhouders, R., Vidal, E., Serra, F., Di Stefano, B., Le Dily, F., Quilez, J., Gomez, A., Collombet, S., Berenguer, C., Cuartero, Y., et al. (2018). Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. Nat. Genet. *50*, 238–249.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Teo, G., Liu, G., Zhang, J., Nesvizhskii, A.I., Gingras, A.C., and Choi, H. (2014). SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. J. Proteomics *100*, 37–43.

Theunissen, T.W., and Jaenisch, R. (2014). Molecular control of induced pluripotency. Cell Stem Cell *14*, 720–734.

Vella, P., Barozzi, I., Cuomo, A., Bonaldi, T., and Pasini, D. (2012). Yin Yang 1 extends the Myc-related transcription factors network in embryonic stem cells. Nucleic Acids Res. *40*, 3403–3418.

Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Südhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. Nature *463*, 1035–1041.

Wang, L., Feng, Z., Wang, X., Wang, X., and Zhang, X. (2010). DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics *26*, 136–138.

Wapinski, O.L., Vierbuchen, T., Qu, K., Lee, Q.Y., Chanda, S., Fuentes, D.R., Giresi, P.G., Ng, Y.H., Marro, S., Neff, N.F., et al. (2013). Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. Cell *155*, 621–635.

Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S., Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L., et al. (2017). YY1 is a structural regulator of enhancer-promoter loops. Cell *171*, 1573–1588.e1528.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307–319.

Yang, T., Zhang, F., Yardımcı, G.G., Song, F., Hardison, R.C., Noble, W.S., Yue, F., and Li, Q. (2017). HiCRep: assessing the reproducibility of Hi-C data using a stratum-adjusted correlation coefficient. Genome Res. *27*, 1939–1949.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhang, X., Huang, C.T., Chen, J., Pankratz, M.T., Xi, J., Li, J., Yang, Y., Lavaute, T.M., Li, X.J., Ayala, M., et al. (2010). Pax6 is a human neuroectoderm cell fate determinant. Cell Stem Cell *7*, 90–100.

Zhang, Y., Li, T., Preissl, S., Amaral, M.L., Grinstein, J.D., Farah, E.N., Destici, E., Qiu, Y., Hu, R., Lee, A.Y., et al. (2019). Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. Nat. Genet. *51*, 1380–1388.

Zufferey, M., Tavernari, D., Oricchio, E., and Ciriello, G. (2018). Comparison of computational methods for the identification of topologically associating domains. Genome Biol. *19*, 217.

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

CellPress

STAR★METHODS

KEY RESOURCES TABLE

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
Antibodies		
OCT4	Santa Cruz	Cat# sc-5279; RRID: AB_628051
NANOG	Laboratories	Cat# A300-397A; RID: AB_386108
CTCF	Santa Cruz	Cat# sc-398149
CTCF	Millipore	Cat# 07-729; RRID: AB_441965
SMC1	Bethyl	Cat# A300-055A; RRID: AB_2192467
YY1	Santa Cruz	Cat# sc-7341; RRID: AB_2257497
YY1	Abcam	Cat# ab109237; RRID: AB_10890662
TUBULIN	Abcam	Cat# ab6046; RRID: AB_2210370
β-actin	Sigma	Cat# A5441; D: AB_476744
H3	Santa Cruz	Cat# sc-517576; RRID: AB_2848194
H3K27ac	Abcam	Cat# ab4729; RRID: AB_2118291
Normal Mouse IgG	Millipore	Cat# 12-371; RRID: AB_145840
Normal Rabbit IgG	Millipore	Cat# 12-370; RRID: AB_145841
Anti-Mouse IgG HRP-linked	Cell signaling	Cat# 7076S; RRID: AB_330924
Anti-Rabbit IgG HRP-linked	Cell signaling	Cat# 7074S; RRID: AB_2099233
Trueblot Anti-Mouse IgG HRP	Rockland	Cat# RK-18-8817-31
Trueblot Anti-Rabbit IgG HRP	Rockland	Cat# RK-18-8816-31
Chemicals, peptides, and recombinant proteins		
DMEM Medium	Hyclone	Cat# SH30022.01
DMEM/F12 1:1 Medium	Gbico	Cat# C11330500BT
Fetal Bovine Serum	VISTECH	Cat# SE100-B
Trypsin/EDTA	Coring	Cat# 25-051
KnockOut™ Serum Replacement	Gbico	Cat# 10828028
PD0325901	SELLECK	Cat# S1036
CHIR99021	SELLECK	Cat# S1263
N2 Supplement	Thermo Fisher	Cat# 17502-048
Puromycin	Sigma-Aldrich	Cat# 540222
B27 Supplement	Gbico	Cat# 17504044
β-mercaptoethanol	Sigma	Cat# M6250
Neurobasal	Thermo Fisher	Cat# 21103-049
NEAA	Thermo Fisher	Cat# 11140050
Glutamax	Thermo Fisher	Cat# 35050061
MboI	New England Biolabs	Cat# R0147
AluI	New England Biolabs	Cat# R0137
Biotin-14-dATP	Life Technologies	Cat# 19524016
dCTP	Invitrogen	Cat# 18253013
dGTP	Invitrogen	Cat# 18254011
dTTP	Invitrogen	Cat# 18255018
DNA Polymerase I, Large (Klenow) Fragment	New England Biolabs	Cat# M0210
T4 DNA Ligase	New England Biolabs	Cat# M0202
Proteinase K	New England Biolabs	Cat# P8102
UltraPure™ 10% SDS	Life Technologies	Cat# 15553-035

(*Continued on next page*)

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

Continued

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
Triton X-100	Sigma	Cat# 93443
T4 PNK	New England Biolabs	Cat# M0201
T4 DNA Polymerase I	New England Biolabs	Cat# M0203
Lambda Exonuclease	New England Biolabs	Cat# M0262
Exonuclease I	New England Biolabs	Cat# M0293
Klenow (3′-5′exo-)	New England Biolabs	Cat# M0212
Quick Ligase	New England Biolabs	Cat# M2200
User Enzyme	New England Biolabs	Cat# E7338A
Proteinase Inhibitor	Roche	Cat# 1169749800
Bovine Serum Albumin	Sigma	Cat# A7906
RNaseA	Thermo	Cat# EN0531
Benzonase	Sigma	Cat# E1014
GelCode™ Blue Safe Protein Stain Buffer	Thermo Scientific	Cat# 24594
5M NaCl Solution	Sigma	Cat# S5150
Abscisic Acid	Sigma	Cat# A1049
Propidium Iodide	eBioscience	Cat# 88-8007-72
DAPI	Sigma	Cat# D9542
IPTG	Solarbio	Cat# I1020-5
PEG-8000	Sigma	Cat# 89510-250G-F
Critical commercial assays		
RNAzol	MRC	Cat# RN190-500
EndoFree Plasmid Midi Kit	Cwbio	Cat# CW2105S
FuGENE	Promega	Cat# E2311
TruePrep DNA Library Prep Kit V2 for Illumina	Vazyme	Cat# TD501
TruePrep Index Kit V2 for Illumina	Vazyme	Cat# TD202
CLOuD9	System Biosciences	Cat# CASCL9-100A
ANNEXIN V APOP DETECT KIT APC	eBioscience	Cat# 88-8007-72
E.Z.N.A. Gel Extraction Kit	Omega	Cat# D2500-02
TIANquick Midi Purification Kit	TIANGEN	Cat# DP204-03
DNA Bisulfite Conversion Kit	TIANGEN	Cat# DP215-02
TIANamp Genomic DNA Kit	TIANGEN	Cat# DP304-02
His-tag Protein Purification Kit	Beyotime	Cat# P2226
SYBR qPCR Master Mix	Vazyme	Cat# Q711-00
Leukocyte Alkaline Phosphatase Kit	Sigma	Cat# 86R
Alkaline Phosphatase Stain Kit	YEASEN	Cat# 40749ES60
PHENOL:CHLOROFORM pH 6.7/8.0	Amresco	Cat# 0883-400ML
AMPure XP Beads	Beckman Coulter	Cat# A63881
Dynabeads MyOne Streptavidin T1 Beads	Life Technologies	Cat# 65602
Protein G Agarose Beads	Thermo Scientific	Cat# 10004D
NEBNext Multiplex Oligos for Illumine	New England Biolabs	Cat# E7335
Qubit™ 1X dsDNA HS Assay Kit	Invitrogen	Cat# Q33230
Cytobuster	Millipore	Cat# 134044
SDS-PAGE Gel Kit	Beyotime	Cat# P0012A
NuPAGE™ Sample Reducing Agent	Thermo Scientific	Cat# NP0009
Clarity™ Western ECL Substrate	Bio-Rad	Cat# 1705061
Centrifugal Filter	Millipore	Cat# UFC803096

(*Continued on next page*)

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

Continued

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
ZymoTaq™ Premix	Zymo Research	Cat# E2003
DreamTaq Green PCR Master Mix	Thermo Fisher	Cat# K1081
Primesoript RT Master Mix	Takara	Cat# RR036A
Cell lines		
rtTA-OG2-OSKM transgenic MEFs	Laboratory of Shaorong Gao	N/A
Pre-iPSC line 1 (12.1)	Laboratory of K. Plath	N/A
Mouse embryonic stem cell line V6.5	Laboratory of R. Jaenisch	RRID: CVCL_C865
ZHBTc4 PSCs	Laboratory of Shaorong Gao	N/A
OG2 MEFs	Laboratory of Jiekai Chen	N/A
OD14 MEFs	Laboratory of Jiekai Chen	N/A
Plat-E	Laboratory of Jiekai Chen	N/A
293T cells	ATCC	Cat# CRL-3216; RID: CVCL_0063
Oligonucleotides		
See Table S6	N/A	N/A
Deposited data		
In situ HiC in MEF, 48h, pre-iPSC	This study	SRA:PRJNA558517
In situ HiC in boundary-deleted MEFs and control MEFs	This study	SRA:PRJNA721515
In situ HiC in MEFs overexpressing WT OCT4, mut1, mut2, mut1-IDR, mut2-IDR and IDR only	This study	SRA:PRJNA721515
BL-HiC in artificial linking MEFs and control MEFs	This study	SRA:PRJNA650173
BL-HiC in only WT OCT4, acidic mutant OCT4 or empty vector overexpressed MEFs	This study	SRA:PRJNA650173
BL-HiC in WT OCT4-SKM and acidic mutant OCT4-SKM reprogrammed cells	This study	SRA:PRJNA650173
OCT4 HiChIP in 48h, pre-iPSC	This study	SRA:PRJNA558518
H3K27ac HiChIP in WT and OCT4 KO ZHBTc4 PSC	This study	SRA:PRJNA558518
CTCF ChIP-seq in WT and OCT4 KO ZHBTc4 PSC	This study	SRA:PRJNA584854
OCT4 IP-MS in 48h, pre-iPSC	This study	Table S5
In situ HiC in NPC and PSC	(Bonev et al., 2017)	SRA:SRP101791
RNA-seq in NPC and PSC	(Bonev et al., 2017; Mumbach et al., 2016)	SRA:SRP101791
H3K27ac ChIP-seq in NPC	(Bonev et al., 2017)	SRA:SRP101791
OCT4 HiChIP in PSC	(Mumbach et al., 2016)	SRA:SRP074188
SMC1 HiChIP in PSC	(Mumbach et al., 2016)	SRA:SRP074188
OCT4 IP-MS in PSC	(Ding et al., 2012)	Figure 5B
ATAC-Seq in reprogramming	(Chronis et al., 2017)	SRA:SRP094576
RNA-Seq in reprogramming	(Chronis et al., 2017)	SRA:SRP094578
OCT4, H3K27ac ChIP-Seq in PSC	(Chronis et al., 2017)	SRA:SRP094580
YY1 ChIP-Seq in PSC	(Vella et al., 2012)	SRA:SRP008009
SMC1, SMC3 ChIP-Seq in PSC	(Kagey et al., 2010)	SRA:SRP002778
CTCF ChIP-Seq in PSC	(Shen et al., 2012)	SRA:SRP006786
SOX1 ChIP-Seq in NPC	(Bergsland et al., 2011)	SRA:SRP108564
RNA-seq in WT and OCT4 KO ZHBTc4 PSC	(King and Klose, 2017)	SRA:SRP091444
Software and algorithms		
TADMAN	This study	https://github.com/corephi/TADMAN

(*Continued on next page*)

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

Continued

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
mm10	ENSEMBL release 90	https://asia.ensembl.org/Mus_musculus/Info/Index
Trim Galore 0.4.4_dev	https://www.bioinformatics.babraham.ac.uk/projects/trim_galore	https://github.com/FelixKrueger/TrimGalore
bowtie2 v2.3.0	(Langmead and Salzberg, 2012)	https://github.com/BenLangmead/bowtie2
samtools v1.3.1	(Li et al., 2009)	https://github.com/samtools/samtools
Bedtools v2.26.0	(Quinlan and Hall, 2010)	https://github.com/arq5x/bedtools
macs2 v2.1.2	(Zhang et al., 2008)	https://github.com/macs3-project/MACS
HTSeq-count v0.11.2	(Anders et al., 2015)	https://github.com/simon-anders/htseq
edgeR v3.26.5	(Robinson et al., 2010)	https://bioconductor.org/packages/release/bioc/html/edgeR.html
HiC-Pro v2.10.0	(Servant et al., 2015)	https://github.com/nservant/HiC-Pro
HiCRep v1.8.0	(Yang et al., 2017)	https://bioconductor.org/packages/release/bioc/html/hicrep.html
ChIA-PET2 v0.9.3	(Li et al., 2016)	https://github.com/GuipengLi/ChIA-PET2
TopDom v0.0.2	(Shin et al., 2016)	https://github.com/HenrikBengtsson/TopDom
diffHic v1.16.0	(Lun and Smyth, 2015)	https://bioconductor.org/packages/release/bioc/html/diffHic.html
Hichipper v0.7.5	(Lareau and Aryee, 2018)	https://github.com/aryeelab/hichipper
DEGSeq v1.38.0	(Wang et al., 2010)	https://www.bioconductor.org/packages//2.10/bioc/html/DESeq.html
ade4 1.7-15	(Dray and Dufour, 2007)	https://cran.r-project.org/web/packages/ade4/index.html
SAINTexpress v3.3	(Choi et al., 2011; Teo et al., 2014)	http://saint-apms.sourceforge.net/Main.html
ROSE v0.1	(Lovén et al., 2013)	https://bitbucket.org/young_computation/rose

Other		
PVDF membrane	Bio-Rad	Cat# 1620177
BL21 E.coli	TransGen Biotech	Cat# CD201-02
Nikon Eclipse Ts2R-FL	Nikon	N/A
Covaris sonicator E220	Covaris	N/A
Qubit™ 4 Fluorometer	Invitrogen	Cat# Q33226
Nikon A1+ Confocal Microscope	Nikon	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Junjun Ding (stemcellding@163.com).

Materials availability

DNA constructs and other research reagents generated by the authors will be distributed upon request to other researchers.

Data and code availability

All datasets used are summarized in the key resources table. The accession number for the Sequence Read Archive (SRA) BioProject where the HiC, HiChIP, ChIP-seq data are available is: SRA: PRJNA558517, SRA: PRJNA558518, SRA: PRJNA584854, SRA: PRJNA650173, and SRA: PRJNA721515. The algorithm of TADMAN generated in this study are available at https://github.com/corephi/TADMAN. Oligonucleotides used can be found in Table S6.

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

CellPress

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines

Preparation of rtTA-OG2-OSKM transgenic mouse embryonic fibroblasts (MEFs)

rtTA-OG2-OSKM mice were generously provided by Professor Shaorong Gao at the Tongji University. The embryos of 12.5-13.5 days were obtained and washed in PBS. Successively remove the fetal membrane and cut off the placenta, wash with PBS every time. After removal of the heads, the visceral, the limbs and the tails, the retained embryo was transferred into a centrifuge tube and cut 2-5 min. Trypsinize the tissue 5 min with 0.25% trypsin/EDTA (Coring) in the 37°C water bath. Suspend cells with cell culture medium containing DMEM (Hyclone) supplemented with 10% Fetal Bovine Serum (FBS) (VISTECH), in the 100 mm Petri dish culture in the 37°C, 5% CO_2 incubator.

OG2 MEF and OD14 MEF

OG2 MEF (control) and OD14 MEF (*Dppa5a* KO) were generously provided by Professor Jiekai Chen at the Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences.

48h cells

For preparation of the 48h time point cells after reprogramming initiation, 2°MEFs were cultured in PSC media containing 2 ug/ml doxycycline for 48h to induce the expression of OSKM.

Pre-iPSC line 12.1

Pre-iPSC line 12.1 is a gift from Professor Kathrin Plath at University of California Los Angeles.

PSC lines R1, V6.5 and ZHBTc4

PSC lines R1, V6.5 and ZHBTc4 were used in this study.

Boundary deletion MEFs

Cas9/sgRNA vectors and homologous arm vectors were co-transfected by FuGENE (Promega) to delete certain TAD boundary (chr9:78443250-78447870). After 24h of transfection, 1ug/ml puromycin (Sigma-Aldrich) was added to select positive transfected cells. More than 200 monoclones were collected under microscope after puromycin selection. Boundary deletion efficiency was validated by PCR using specific primers, and the homozygotes (−/−) and heterozygotes (-/+) were collected respectively, which were used to evaluate TAD reorganization by HiC, as well as reprogramming efficiency by iPSC colony formation.

Cell culture conditions

PSC lines culture. All PSC lines and the pre-iPSC line 12.1 were grown on 0.1% gelatinized tissue culture dishes in PSC media: DMEM supplemented with 15% FBS, 0.1 mM β-mercaptoethanol (Sigma), NEAA (Thermo Fisher), 2 mM Glutamax (Thermo Fisher), Nucleoside MIX, LIF. For V6.5 and ZHBTc4 PSC lines, two inhibitors, 1uM PD0325901 (SELLECK) and 3uM CHIR99021 (SELLECK), were added into the PSC media.

2°MEF reprogramming. For *in vitro* 2°MEF reprogramming, MEFs were plated at a density of 1.5×10^4 cells per well in 12-well gelatin-coated plates, cultured in iPSC medium (DMEM supplemented with 10% FBS, 10% Knockout serum replacement (GIBCO), 0.1 mM β-mercaptoethanol, NEAA, 2 mM Glutamax, Nucleoside MIX, leukemia inhibitory factor (LIF)) with doxycycline (2 mg/ml). Medium was changed every day for 12 days. Reprogramming efficiency was evaluated by counting the number of iPSC colonies stained with alkaline phosphatase (AP) (Sigma-Aldrich). Simultaneously, the reprogramming cells were harvested at 12th day, wash twice with PBS, and the OCT4-GFP positive cell number was also considered for reprogramming efficiency by flow cytometry. All colony formation assays and OCT4-GFP flow cytometry assays were performed for three biological replicates.

WT MEF reprogramming. WT MEFs were plated at a density of 1.5×10^4 cells per well in 12-well gelatin-coated plates, and reprogramming was initiated by viral transfection of pMXs-*Sox2*, pMXs-*Klf4*, pMXs-*Myc* with pMXs-*Oct4*. The cells were cultured in iPSC medium (DMEM supplemented with 10% FBS, 10% Knockout serum replacement, 0.1 mM β-mercaptoethanol, NEAA, 2 mM Glutamax, Nucleoside MIX, LIF) for 35 days to get iPSC clones. Cells were collected at day 21 for BL-HiC and day 35 for AP staining.

OG2 and OD14 MEF reprogramming. The MEFs were plated at a density of 1.5×10^4 cells per well in 12-well gelatin-coated plates. After transfection with stemcca virus, the cells were cultured in iPSC medium (DMEM supplemented with 10% FBS, 10% Knockout serum replacement, 0.1 mM β-mercaptoethanol, NEAA, 2 mM Glutamax, Nucleoside MIX, LIF) for 12 days to get iPSC clones.

NPC differentiation. The PSC 46C line was plated at a density of 3×10^4 cells/cm2 in 12-well gelatin-coated plates, cultured in N2B27 medium. N2B27 medium consists of a 1:1 mixture of DMEM/F12 (Gbico) supplemented with 1 × N2 (Thermo Fisher), NEAA, 1 mM L-glutamine and 0.1 mM β-mercaptoethanol with Neurobasal (Thermo Fisher) supplemented with B27 (Gbico).

METHOD DETAILS

In situ HiC

In brief, 5×10^7 cells were cross-linked with 1% formaldehyde for 10 min at room temperature. Cells were incubated in 500ul of lysis buffer (10mM Tris-HCl pH8.0, 10mM NaCl, 0.2% Igepal CA630, protease inhibitors cocktail) for 15 min at 4°C. Cells were washed

twice in lysis buffer and incubated in 50 ul of 0.5% SDS at 65°C for 8 min. The SDS was quenched by incubated in a solution of 25ul of 10% Triton-X and 145ul H2O at 37°C for 15 min. 25 μL of 10 × NEBuffer2 and 20ul of MboI (NEB) were added and the chromatin was digested at 37°C overnight with rotation. To inactivate MboI, the sample was incubated at 65°C for 20 min. Restriction fragments were biotinylated by supplementing the reaction with 37.5ul biotin-14-dATP (Life Technologies), 1.5ul of 10mM dCTP (Invitrogen), 1.5ul of 10mM dGTP (Invitrogen), 1.5ul of 10mM dTTP (Invitrogen), 8ul of DNA polymerase I, large (Klenow) fragment (NEB) and incubated at 37°C for 4h. The end-repaired chromatin was added into 663ul H2O, 120ul T4 ligase buffer, 100ul 10% Triton X-100, 12ul 10mg/ml BSA, 5ul T4 DNA ligase (NEB). The ligation was at least for 4h at room temperature. Crosslinks were reversed by adding 50ul of 20 mg/ml proteinase K (NEB), 120ul 10% SDS (Life technologies) at 55°C for 30min, and another addition of 130ul 5M NaCl at 68°C overnight. The DNA was precipitated with ethanol and resuspended in 130 μL of Tris-buffer (10 mM Tris-HCl, pH 8.0) for sonication (Covaris). The samples were recombined and the volume was adjusted to 300ul with Tris-buffer. Fragments were selected using AMPure XP beads (Beckman Coulter). The DNA fragments were eluted off the beads in 300ul of Tris-buffer. To isolate biotin-labeled ligation junctions, 150ul of 10 mg/ml Dynabeads MyOne Streptavidin T1 beads (Life technologies) were washed with 400ul of 1 × Tween washing buffer (5 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween 20), collected with a magnet, resuspended in 300ul of 2 × binding buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 2 M NaCl) and added to the sample. Biotinylated DNA was bound to the beads by incubating the sample for 15 min at room temperature. For end-repair and biotin removal from unligated ends, 88ul 1 × NEB T4 DNA ligase buffer, 2ul dNTPs, 5ul T4 PNK (NEB), 4ul T4 DNA polymerase I (NEB), 1ul DNA polymerase I, large (Klenow) fragment were add to incubate for 30 min at room temperature. The fragment was washed twice in 1 × Tween Wash Buffer 2min at 55°C. A-tailing was performed by incubating in 90ul 1 × NEB buffer 2, 5ul dATP, 5ul Klenow exo minus at 37°C for 30min. The adaptor was ligated by incubating in a mixture of 50ul 1 × Quick Ligation Buffer, 2ul Quick Ligase (NEB) and 3ul Illumina indexed adaptor (NEB) for 15min at room temperature. 2.5ul User Enzyme (NEB) was added for further incubation 15min at 37°C. The fragment was washed by 1 × Tween Wash Buffer twice at 55°C 2 min. Library preparation was performed with an NEBNext DNA Library Prep Kit (NEB), libraries were amplified for 10-12 cycles and were size-selected with AMPure XP beads. Two biological replicates were performed for each cell line.

BL-HiC

In brief, 5 × 10^7 cells were cross-linked with 1% formaldehyde for 10 min at room temperature. Cells were incubated in 500ul of lysis buffer ((0.1%SDS; 50mM HEPES-KOH, pH7.5; 150mM NaCl; 1mM EDTA; 1% Triton X-100; 0.1% Sodium Deoxycholate) containing protease inhibitors, and incubate on ice for 20 min. Cells were washed twice in lysis buffer and incubated in 50 ul of 0.5% SDS at 62°C for 10 min. The SDS was quenched by incubated in a solution of 25ul of 10% Triton-X and 145ul H_2O at 37°C for 15 min. 25 μL of 10 × NEBuffer2 and 100U of AluI (NEB) were added and the chromatin was digested at 37°C overnight with rotation. To inactivate AluI, the sample was incubated at 62°C for 20 min, then cool to room temperature. Nuclei were collected by centrifuging 3500 g at 4°C for 5min. Resuspended in 400 μL of Klenow (3′-5′exo-) solution (40uL NEB buffer2, 8ul 10 mM dATP, 40ul 10% Triton X-100, 304ul H_2O, and 8ul Klenow (3′-5′exo-)), and incubated at 37°C for 1h. Wash nuclei twice with 1 × T4 DNA ligase buffer. Resuspended in 1200ul of proximity ligation solution (120uL T4 DNA ligase buffer, 120ul 10% Triton X-100, 939ul H_2O, 6ul T4 DNA ligase, 12ul 10 mg/ml BSA, 3ul Bridge linker (200ng/ul)), and rotate at room temperature for 4h. Centrifuge 3500 g for 5 min at 4°C to remove the supernatant, and then add 70ul lambda exonuclease buffer, 4ul lambda exonuclease (NEB), 6 ul exonuclease I (NEB), 618uL H_2O, rotate at 37°C for 1h. To reverse crosslink DNA, resuspended chromatin with 1370ul digestion buffer (10mM Tris-HCl, pH = 8, 25mM EDTA, 1% SDS, 1 mg/ml proteinase K), incubate the sample for 1h at 55°C. Add 130ul of 5M NaCl and incubate overnight at 68°C. Cool tubes to room temperature. The DNA was precipitated with ethanol and resuspended in 130ul of Tris-buffer (10 mM Tris-HCl, pH 8.0) for sonication. The DNA fragments were eluted off the beads in 300ul of Tris-buffer. To isolate biotin-labeled ligation junctions, 30ul of 10 mg/ml Dynabeads MyOne Streptavidin T1 beads were washed with 400ul of 1 × Tween washing buffer (5 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween 20), collected with a magnet, resuspended in 300ul of 2 × binding buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 2 M NaCl) and added to the sample. Biotinylated DNA was bound to the beads by incubating the sample for 15 min at room temperature. For end-repair, resuspended the DNAs-on-Dynabeads in 100ul mix: 10ul NEB T4 DNA ligase buffer, 5ul dNTP mix, 5ul T4 PNK, 4ul T4 DNA polymerase I, 1ul DNA polymerase I, large (Klenow) fragment, 75ul water were added to incubate for 30 min at 37°C. Resuspended the DNAs-on-Dynabeads in 100 μL of master mix: 10 μL NEB buffer 2, 5 μL 10mM dATP, 5 μL of 10U/μl NEB Klenow exo- (NEB), 80 μL water for A-tailing. Shake at 900 rpm for 30 minutes at 37°C on the thermomixer. The adaptor was ligated by incubating in a mixture of 25ul Quick Ligation Buffer, 2ul Quick Ligase, 3ul Illumina indexed adaptor and 20ul water for 15min at room temperature. 2.5ul User Enzyme was added for further incubation 15min at 37°C. Library preparation was performed with an NEBNext DNA Library Prep Kit, libraries were amplified for 10-12 cycles and were size-selected with AMPure XP beads. Two biological replicates were performed for each cell line.

HiChIP

In brief, 10^7 cells were cross-linked for 10 min at room temperature with 1% formaldehyde in growth media and quenched in 0.125 M glycine. Cross-linked cell pellets were resuspended in 500 ul of ice-cold HiC lysis buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, and 0.2% IGEPAL CA-630) with protease inhibitor (Roche), and incubated at 4°C for 30 min. Nuclei were pelleted and washed once with 500ul of ice-cold HiC lysis buffer. After removing supernatant, nuclei were incubated in 100ul of 0.5% SDS at 62°C for 10 min. SDS was quenched by adding 285ul water and 50ul 10% Triton X-100, and nuclei were incubated for 15 min at 37°C. After addition of 50ul of 10 × NEB Buffer 2 and 375 U of MboI, chromatin was digested at 37°C overnight. MboI enzyme was then inactivated by incubating

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

at 62°C for 20 min. To fill in the restriction fragment overhangs and mark the DNA ends with biotin, 52ul of fill-in master mix containing 37.5ul of 0.4 mM biotin-dATP, 1.5ul of 10 mM dCTP, 1.5ul of 10 mM dGTP, 1.5ul of 10 mM dTTP, and 10ul of 5 U/ul DNA Polymerase I, Large (Klenow) Fragment was added and incubated at 37°C for 1h with rotation. Proximity ligation was performed by incubating in a solution of 947ul of ligation master mix containing 150ul of 10X NEB T4 DNA ligase buffer, 125ul of 10% Triton X-100, 7.5ul of 20 mg/mL BSA (Sigma), 10ul of 400 U/ul T4 DNA ligase and 655.5ul of water at room temperature for 4h with rotation. After proximity ligation, nuclei were pelleted and resuspended in 1ml of ChIP sonication buffer (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8.0, 1 mM EGTA pH 8.0, 1% Triton X-100, 0.1% sodium deoxycholate, and 0.1% SDS) with protease inhibitor. Nuclei were sonicated using a Covaris S220. 60ul of protein G magnetic beads were washed three times with sonication buffer, resuspended in 50ul of sonication buffer, and added to the sonicated chromatin and incubated for 1h at 4°C with rotation. Beads were then separated on a magnetic stand and the supernatant was mixed with 7.5 ug of OCT4 or H3K27ac antibody, and incubated overnight at 4°C with rotation. 60ul of protein G magnetic beads were washed and resuspended in 100ul of sonication buffer. Samples were incubated with the beads for 2h at 4°C with rotation. Beads were then separated on a magnetic stand and washed three times with 1ml of high salt sonication buffer (50 mM HEPES-KOH pH 7.5, 500 mM NaCl, 1 mM EDTA pH 8.0, 1 mM EGTA pH 8.0, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS) followed by three times with 1ml of LiCl wash buffer (20 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 250 mM LiCl, 0.5% IGEPAL CA-630, 0.5% sodium deoxycholate, 0.1% SDS) and once with 1ml of TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 50 mM NaCl). Beads were then resuspended in 200ul of elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0, 1% SDS) and incubated at 65°C for 15 min to elute. To purify eluted DNA, RNA was degraded by incubation with 2.5ul of 33 mg/mL RNase A (Thermo) at 37°C for 2h. Protein was degraded by incubation of 10ul of 20 mg/mL proteinase K at 55°C for 45 min. Samples were then incubated at 65°C for 5h to reverse cross-links. Phenol: chloroform: isoamyl alcohol extraction was performed followed by an ethanol precipitation. The DNA was then resuspended in TE buffer. Tagmentation of ChIP DNA was performed using the Tn5 (Vazyme). First, 5ul of streptavidin T1 magnetic beads (Invitrogen) was washed with 1ml of tween wash buffer (5 mM Tris-HCl pH 7.5, 0.5 mM EDTA pH 8.0, 1 M NaCl, 0.05% Tween-20) and resuspended in 10ul of 2 × biotin binding buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0, 2 M NaCl). 125 ng purified DNA was added in a total volume of 10ul of water to the beads and incubated at room temperature for 15 min with agitation every 5 min. After capture, beads were separated with a magnet and the supernatant was discarded. Beads were then washed twice with 500ul of tween wash buffer at 55°C for 2 min. After washes, beads were resuspended in 1 × TTBL. Tn5 amount was adjusted linearly for different amounts of post-ChIP DNA, 125 ng of DNA was transposed with 4ul of Tn5. Samples were incubated at 55°C with interval shaking for 10 min. Beads were then placed on a magnet, and supernatant was removed. 50 mM EDTA was added to samples and incubated with interval shaking at 50°C for 30 min. Samples were washed in 10 mM Tris, pelleted and resuspended for PCR amplification (Vazyme). Two biological replicates were performed for each cell line.

3C

Cells were crosslinked with 1% formaldehyde for 10 min at room temperature. Cells were resuspended in 500ul of lysis buffer (10mM Tris-HCl pH8.0, 10mM NaCl, 0.2%Igepal CA630, protease inhibitors cocktail) and incubated for 15 min at 4°C. Cells were washed twice in lysis buffer and incubated in 50ul of 0.5% SDS at 65°C for 8 min. SDS was quenched by incubating in a solution of 25ul of 10% Triton-X and 145ul H_2O at 37°C for 15 min. 25ul of 10 × NEBuffer2 and 20ul of MboI were added and the chromatin was digested at 37°C overnight with rotation. To inactivate MboI, the sample was incubated at 65°C for 20 min. Restriction fragments were biotinylated by supplementing the reaction with 1.5ul dNTP, 8ul of DNA polymerase I, large (Klenow) fragment, 40.5ul H2O and incubated at 37°C for 4h. The end-repaired chromatin was added into a solution of 663ul H_2O, 120ul NEB T4 ligase buffer, 100ul 10% Triton X-100, 12ul 10mg/ml BSA, 5ul T4 DNA ligase. The ligation was performed for 4h at room temperature with rotation. Crosslinks were reversed by adding 50ul of 20 mg/ml proteinase K, 120ul 10% SDS at 55°C for 30min, and another addition of 130ul 5M NaCl at 68°C overnight. The DNA was precipitated with ethanol and resuspended in Tris-buffer for about 100ng/ul DNA. 100ng DNA were amplified in PCR for 35 cycles. Primers used are listed in Table S6.

IP-MS

To identify stage-specific OCT4 partners during reprogramming, we pulled down OCT4 protein complexes from nuclear extraction in 48h and pre-iPSC by OCT4 antibody (Santa Cruz). The 48h and pre-iPSC were expanded to 60 of 150 mm diameter dishes respectively for the preparation of nuclear extracts. Nuclear extracts were pre-cleared with 0.5 mL of Protein G agarose beads (Thermo Scientific) in 16 mL IP DNP buffer in presence of 750 units of Benzonase (Sigma) overnight at 4°C. At the same time, 25 ug OCT4 antibody or IgG conjugated Protein G agarose beads by incubating in IP DNP buffer overnight at 4°C. The pre-cleared nuclear extracts were combined with the antibody-conjugated beads, and rotated for 5h at 4°C. After five washes in Buffer D (20 mM HEPES pH = 7.6, 0.2 mM EDTA, 1.5 mM MgCl2, 100 mM KCl, 20% glycerol) supplemented with 0.02% NP-40, the bound material was eluted by boiling for 5 min in Laemmli sample buffer. Put the solution into the concentration columns, centrifuge 14000 g 30 min at 20°C. Centrifuge until the solution near 40ul, Samples were then fractionated on a 10% NuPAGE™ 4%–12% Bis-Tris Gel (Thermo Scientific) and stained with GelCodeTM Blue Safe Protein Stain buffer (Thermo Scientific). The products from a single purification were subjected to whole lane LC-MS/MS sequencing and data analysis. Two biological replicates were performed for each cell line.

Co-IP

Nuclear extracts were prepared from 48h, pre-iPSC and V6.5 PSCs. Endogenous OCT4 was immunoprecipitated with 5 μg of OCT4 antibody pre-bound to Protein G agarose beads (Roche Diagnostics), and co-immunoprecipitated structuring factors including

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

CTCF, SMC1 and YY1 were identified by western blot with the antibodies of CTCF, SMC1 and YY1. TUBULIN and β-ACTIN were used as the loading control in this study.

ChIP

1% formaldehyde in PBS was used to crosslink the cells for 10 min, followed by quenched with 125mM glycine on ice. Cells were collected and flash frozen in liquid nitrogen, then stored at −80°C for use. Frozen crosslinked cells were thawed on ice and then re-suspended in lysis buffer I (50 mM HEPES-KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100, protease inhibitors). After rotated for 10 min at 4°C, the cells were collected, and resuspended in lysis buffer II (10 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, protease inhibitors). After rotated for 10 min at 4°C, the cells were collected, and resuspended in sonication buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA pH 8.0, 0.1% SDS, and 1% Triton X-100, protease inhibitors) for sonication. Sonicated lysates were cleared once by centrifugation at 16,000 g for 10 min at 4°C. Input material was reserved. The remainder was incubated with magnetic beads bound with antibody to enrich for DNA fragments overnight at 4°C. Beads were washed with wash buffer (50 mM HEPES-KOH pH 7.5, 500 mM LiCl, 1mM EDTA pH 8.0, 0.7% Na-Deoxycholate, 1% NP-40) and TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 50 mM NaCl) in order. Beads were removed by incubation at 65°C for 30 min in elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS). Cross-links were reversed overnight at 65°C. To purify eluted DNA, 200 mL TE was added and then RNA was degraded by incubation in 8ul 10 mg/ml RNase A at 37°C for 2h. Protein was degraded by addition of 4ul 20 mg/ml proteinase K and incubation at 55°C for 2h. Phenol: chloroform: isoamyl alcohol extraction was performed followed by an ethanol precipitation. The DNA was then resuspended in 50ml TE and used for qPCR. For ChIP-qPCR experiments, qPCR was performed using SYBR qPCR Master mix (Vazyme). For ChIP-qPCR, the primers used are listed in Table S6.

DNA extraction and PCR

Cells were lysed in 50ul of DNA extraction buffer (10mM Tris-HCl pH8.0, 0.05% SDS, 25 ug/ml Proteinase K) and incubated at 37°C for 1h. Then the cell lysis was incubated at 80°C for 30min and centrifuged at max speed. The supernatant was collect for PCR. To detect the boundary deletion efficiency, PCR was performed using DreamTaq Green PCR Master Mix (Thermo Scientific).

RNA isolation and qRT-PCR

Total RNA was extracted from cell pellets using RNAzol reagent (MRC) and cDNA was synthesized using Primesoript RT Master Mix (Takara). Real time qPCR was performed using SYBR qPCR Master Mix on LightCycler 480 II system. The exhibited data represents the fold change (FC) of experimental group versus control group. In brief, $\triangle$Ct was calculated as $\triangle$Ct = Ct (test gene) - Ct (Ref. gene). $\triangle\triangle$Ct was calculated as $\triangle\triangle$Ct = $\triangle$Ct (experimental group) - $\triangle$Ct (control group). The FC of a test gene in experimental group versus control group was calculated as FC = 2^(−ΔΔCt). Each gene tested in triplicates in every independent experiment, and all experiments were triplicated. Primers used are listed in Table S6.

Protein extraction and western blots

Cells were lysed in Cytobuster (Millipore) at room temperature for 10 min. Lysate was separated on 10% Bis-Tris gel, and then wet transferred to a 0.45 um PVDF membrane (Bio-Rad) in ice-cold transfer buffer for 6h. After blocked with 5% non-fat milk in TBS for 1h at room temperature, the membrane was incubated with the primary antibody overnight at 4°C. After washed three times with TBST, the membrane was incubated with 1:2000 secondary antibodies for 1h at room temperature. After washed three times with TBST. The membrane was developed with ECL substrate and imaged using a CCD camera or exposed using film or with high sensitivity ECL (Bio-Rad).

IF

Cells were plated on the 14 mm slides, washed twice with 2ml warm PBS. The cells were fixed using 4% paraformaldehyde for 15 min at room temperature. Then paraformaldehyde was removed and the cells were permeabilized by 0.5% Triton X-100 for 5 min. Cells were blocked with 200ul blocking buffer (consists of 5% BSA, 0.5% Triton X-100, 94.5% PBS; Triton X-100 was added freshly just before use) for 30 minutes at room temperature. After aspirating blocking buffer, primary antibody was added at a certain concentration in blocking buffer at 4°C overnight or 1h at room temperature. Cells were washed with PBS five times followed by incubation with labeled secondary antibody at a certain concentration of 1:2000 in blocking buffer for 1h at room temperature without light exposure. After two washes of PBS, cells were stained with 3 uM DAPI (Sigma) in PBS for 10 min at room temperature. Finally, images were acquired at a Nikon A1R super-resolution microscope with N-STORM software.

FISH

DNA-FISH Cells were grown on glass coverslip and fixed with 4% PFA in PBS for 12 min, washed 5min in PBS for three times and stored at 4°C. After permeating cells with 1% Triton X-100 for 10min at RT, cells were washed 5 min in PBST for three times, followed by incubated at 37°C in 100 μg/mL RNaseA in 1x PBST 45 min to remove all RNA, and wash three times in 2 × SSCT. Then incubation at 4°C overnight in 50% formamide in 2x SSCT to loosen the chromatin. Then pre-hybridized with 50% formamide in 2x SSCT at 78°C for 10 min and turn to 86°C for 3 min. Keep the dish on the ice. Incubated for 1 min in 70% ethanol, for 1 min in 85% ethanol and for 1 min in 100% ethanol. After air-drying the coverslips, cells were incubated in 70%, 85% and 100% ethanol in order for 1 min at RT

respectively. 5ul primary probe and the secondary probe 1ul 100uM were mixed with 35ul 100% formamide, shake at 37°C for 15 min. Add 35ul 20% dextron sulfate with 4x SSC, 1ul Triton, shake at 37°C for 15 min. Heat the dish to 86°C for 3 min and put it on the ice immediately. 37°C for 16-20h for hybridization. Wash the dish with 2 × SSCT for 5 min. Wash cell with 60°C preheat 2x SSCT for 5 min. Change to 2x SSCT at room temperature. The resultant glass coverslip was then used for IF treatment.

In Vitro Droplet Formation

cDNA encoding the WT OCT4, OCT4 mutants, rescued OCT4 mutants or MED1 IDR were cloned into pET28a expression vector. The base vector was engineered to include a 5′-6 x HIS followed by either GFP or mCherry and a 6 amino acid linker sequence ''GGATCC.'' Vectors expressing GFP or mCherry alone contain the linker sequence followed by a STOP codon. The interest sequence was cloned into the vector at the position between the linker sequence and the STOP codon. AD-acidicmutant-Oct4 sequences were kindly supplied by Professor Richard Young at the Massachusetts Institute of Technology, Cambridge, USA. All expression constructs were sequenced to ensure sequence identity. For protein expression, plasmids were transformed into BL21 *E. coli* (TransGen Biotech) and grown as follows. A fresh bacterial colony was inoculated into LB media containing kanamycin and grown overnight at 37°C. Then the cells were diluted 1:30 in 300ml LB with freshly added kanamycin and grown 2h at 37°C to make sure OD600 up to 0.6-0.8. Then temperature was decreased to 16°C and IPTG (Solarbio) was added to 0.3 mM and growth continued overnight. After centrifugation, cell pellets were resuspended in 15ml of nondenaturing lysis buffer (Beyotime) containing protease inhibitors and sonicated (30 cycles of 30 s on, 15 s off). The lysates were cleared by centrifugation at 3,000 g for 30 minutes and added to 1ml of BeyoGold™ His-tag Purification Resin (Beyotime) that had been pre-equilibrated with 5 volumes of the same buffer. Tubes containing this agarose slurry were rotated for 1.5h. The slurry was poured into a column, washed with 10 volumes of the washing buffer and eluted 6 X with elution buffer containing 250mM imidazole. Each fraction was run on a 12% gel and proteins of the correct size were dialyzed by dialysis buffer (50 mM Tris pH 7.5, 125 mM NaCl, 1 mM DTT and 10% Glycerol. The Recombinant GFP or mCherry fusion proteins were concentrated in centrifugal filters (Millipore) for use. The proteins with an appropriate concentration mixed with 10% PEG-8000 (Sigma), and solve in 125 mM NaCl (Sigma). The protein solution was immediately loaded onto a homemade chamber comprising a glass slide, and imaged with a fluorescence microscopy (Nikon Eclipse Ts2R-FL).

FRAP

mESCs stably expressing OCT4-GFP were cultured in gelatin-coated glass bottom dish for 24 hours. Fluorescence images of GFP were acquired on a Nikon A1+ confocal microscope using a 100x oil-immersion objective lens (HP Apo TIRF 100xH, 1.49 NA, Nikon). The fluorescence intensity of bleached cell at each time point was normalized by fluorescence intensity at background region and fluorescence intensity of adjacent unbleached cell. The images were analyzed using NIS-Elements software. The inner fluidity of *in vitro* droplets were also evaluated by FRAP.

CLOuD9

CLOuD9 (System Biosciences) was performed to link two loci belonging to two neighboring TADs. For respective targeting the selected loci, sp and sa gRNAs were designed online by https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design. In Brief, the primary MEFs were infected by the virus containing sp(SE)dCas9-PLY1-puro and sa(promoter)dCas9-ABI-hygro respectively. The infected MEFs were screened by 2ug/ml puromycin and 300ug/ml hygromycin for 5 days. After selection, the second passage cells were used for loop validation, TAD reorganization and reprogramming. 3C was performed to validate artificial looping. The influence of the artificial looping on TAD reorganization was tested by BL-HiC. In brief, cells containing the chemical inducible CRISPR-GO systems and sgRNAs were treated with or without 3mM abscisic acid (ABA, Sigma) for 48h before harvest. For reprogramming, the cells were cultured in iPSC medium with or without 3mM ABA for 12 days. and reprogramming efficiency was investigated by iPSC colony formation assay and flow cytometry counting the ratio of OCT4-GFP positive cells. The sample without ABA treatment was considered as a control.

Apoptosis

Cell apoptosis induced by *Dppa5a* KO during reprogramming was determined by double-staining cells with annexin V and propidium iodide (PI) (eBioscience) with subsequent flow cytometry analysis. Reprogramming was initiated by transfection of stemcca virus in OG2 MEF (control) and OD14 MEF (*Dppa5a* KO), and day 2, 6 and 10 cells were collected for apoptosis analysis. Cell populations were counted as viable (annexin V-negative, PI-negative), early apoptotic (annexin V-positive, PI-negative), late apoptotic (annexin V-positive, PI-positive) or necrotic (annexin V-negative, PI-positive).

Bisulfite sequencing

DNA methylation changes at TAD boundaries in response to OCT4 acute depletion were evaluated by bisulfite sequencing. In brief, DNA was extracted from ZHBTc4 WT and OCT4 KO cells, followed by bisulfite treatment for 60 min. PCR amplification was performed using ZymoTaq™ premix (Zymo Research). Direct Sanger sequencing of the PCR products was used to evaluate methylation levels at multiple CpG sites. The proportion of methylation was calculated as the peak ratio of cytosine to the sum of cytosine and thymine at each site. Primers used are listed in Table S6.

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

QUANTIFICATION AND STATISTICAL ANALYSIS

Quality control of sequencing reads

All the Illumina sequencing reads used in the study were first quality controlled by Trim Galore. In details, we removed the bases with quality below 20 and the adaptor sequences from the 3′ end, and filtered the reads with length less than 50nt.

ChIP-seq and ATAC-seq data analysis

We aligned the ChIP-Seq and ATAC-Seq data to the mm10 reference genome by bowtie2 with default parameter, followed by removing the multiple aligned reads, PCR duplications with samtools. To eliminate the impact of "Problematic genomic regions," we downloaded the ENCODE blacklist, and discarded the reads aligned this region through bedtools. Specially, for ATAC-Seq data, we additionally removed the aligned mitochondria. Finally, we used macs2 to calling peaks with control, setting a q value cutoff of 0.05.

RNA-seq data analysis

For read alignment and expression quantification, we first removed low quality reads, and trimmed the adaptor sequence with Trim Galore. Then we mapped the remaining pair-end reads to the reference genome using STAR with ENCODE option bundles. Using HTSeq-count, we counted the uniquely mapped reads, and normalized the read count by trimmed mean of M values (TMM), and transformed to reads per kilobases per million reads (RPKM) by edgeR. With an expression cutoff of RPKM $\geq$ 1 in at least one sample group, we removed low abundant genes, and detected the differentially expressed genes using edgeR. Genes were considered differentially expressed when the overall false discovery rate (FDR) < 0.01 and fold change is above 2.0.

In situ HiC data analysis

To call the valid HiC interactions, we pooled the quality controlled biological replicate reads together, and aligned them to mm10 reference genome by HiC-Pro. Using HiC-Pro, we removed the duplicate reads, assigned the reads to MboI restriction fragments, and filter the invalid interactions. Then we binned the interaction to 100k resolution for compartmentation, and 40k for TAD identification.

Compartmentalization

To eliminate the ambiguous compartment switch, we first smoothed 100k matrix by smoothMat implemented in HiCRep in bin level. Then we calculated the dominant eigenvector of the smoothed contact matrices using matrix2compartment.pl available at (https://github.com/dekkerlab/cworld-dekker). Considering a TAD is always contained in one compartment, we finally smoothed the compartmentation in TAD level. We assigned TADs to either the A- or the B- compartments, by calculating the average dominant eigenvector of each TAD.

TAD reorganization identification

We choose TopDom to identify TAD boundaries, because TopDom not only calls high quality and reproducible TADs, but also is insensitive to sequencing depth, normalization strategies and resolutions. To gain a better accuracy, we set the windows size gradient from 2 to 7, and found that a widow size of 3 have the higher Pearson's correlation coefficient among all stages. Thus, we run TopDom with a window size of 3 from the 40kb ICE-normalized matrices in this study. After getting the raw boundaries, we detected differential domain boundaries using diffHiC. In details, we first converted the valid HiC interaction reads to HD5 file by each replicate, and then calculated the Direction Index using domainDirections function setting width as 40k and span as 7. Finally, we performed the significant differences test with replicates by glmQLFTest function, which integrated edgeR for dispersion estimation and GLM fitting. Only the raw TopDom TAD boundaries with FDR < 0.05 were treated as reorganization, and their corresponding TAD were classified as shift, fusion and separation by homemade Perl script. For the definition of TAD reorganization cluster, in brief, a genomic region with at least 80% of TAD reorganized, and the reorganization frequency is more than 2.0 per Mb length. One TAD reorganization cluster is comprised of at least five TADs.

BL-HiC data analysis

After quality control of the BL-HiC reads, we trimmed the linker using trimLinker, which was embedded in ChIA-PET2, setting parameter as '-A ACGCGATATCTTATC -B AGTCAGATAAGATAT -m 1 -k 1 -l 16'. We called valid chromatin interactions and identified TAD reorganization same as *in situ* HiC data. To assess the TAD alteration level between different samples, we calculated Jaccard distance from TAD boundary locus by R package ade4.

HiChIP data analysis

The raw HiChIP interactions were called by the same procedure as described above *in situ* HiC. All HiChIP valid reads were further processed to call loops using hichipper with default parameter, only loops with FDR < 0.05 were kept for further analysis. To systematically assess OCT4-associated loops across reprogramming, we used diffHiC to call the consensus loops and read count, after removing self-circles and low abundance loops. Then the loops read counts were normalized by TMM implemented in edgeR, and filtered by a cutoff of normalized reads count < 3 for each stage. Since H3K27ac HiChIP data has not biological replicate, we

Please cite this article in press as: Wang et al., Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions, Cell Stem Cell (2021), https://doi.org/10.1016/j.stem.2021.04.023

adopted DEGSeq for differential loops calling, loops were considered as changing if their interaction were significantly different (q value < 0.01) and with above 2-fold change. For the definition of OCT4 loop cluster, in brief, a loop anchor cluster with high OCT4 HiChIP signal was considered as OCT4 loop cluster. OCT4 HiChIP signal was not evenly distributed across the loop anchor regions. We found the x axis point for which a line with a slope of 1.0 was tangent to the curve. We define loop anchors above this point to be OCT4 loop cluster.

IP-MS data analysis

SAINTexpress version 3.3 was used as a statistical tool to calculate the probability value of each potential protein-protein interaction compared to background contaminants using default parameters. A main probability score (AvgP) of 0.7 was used as a cutoff for significant protein-protein interaction.

Definition of super-enhancers

To identify typical enhancers, we initially stitched the H3K27ac ChIP-seq peaks along 500bp away, but re-split them if the gaps coverage is above 0.4. Then the stitched raw enhancer together with aligned H3K27ac and input reads were used to run the ROSE algorithm. The parameters were set as "-s 8667 -g mm10." We classified the genes as SE-related genes if they meet the following criteria: 1) promoter region was within TAD containing SE, 2) top ten percent highly expressed in that stage, 3) and annotated as protein-coding genes.

Prediction of key amino acid residues for phase separation

Key amino acid residues of OCT4 for phase separation was predicted. The prediction strategy is that each three consecutive amino acids are considered as a unit. The phase-separated probability of each unit-deleted OCT4 is calculated and compared with WT OCT4 to generate the probability score. The lower the score is, The higher influence of the unit on phase separation. We selected the key residues by considering three aspects: 1) lower probability score, 2) lower DNA binding, and 3) at the disordered regions. In the light of this standard, deletions of amino acid 324-326 was considered as a potential perturbation to disrupt OCT4 phase separation.

TADMAN algorithm

The reprogramming regulators could be identified by integrating TAD reorganization-based Multiomics Analysis (TADMAN). TADMAN filters SE-related genes with muiltiomics analysis to get the reprogramming regulator candidates, including HiC, OCT4 HiChIP, OCT4 ChIP-Seq and RNA-Seq. TADMAN stipulates that the genes encoding the regulators should meet the following criteria: 1) they are regulated by SEs in reorganized TAD (HiC); 2) the NRR genes and SEs are connected by OCT4 loops (OCT4 HiChIP); 3) both SEs and their target genes are occupied by OCT4 peaks (OCT4 ChIP-seq); and 4) the genes are the top 10% highly expressed genes at any stage of reprogramming and showed at least 10-fold change in expression compared to the previous or next stage (RNA-seq). Taken together, if a SE-related gene would be considered as a regulator candidate by TADMAN, it should locate in reorganized TAD, interact with SE by OCT4 loops, be occupied by OCT4, and highly expressed.