The Company of **Biologists**

Lineage-specific reorganization of nuclear peripheral heterochromatin and H3K9me2 domains

Kelvin See[1,2,3,4], Yemin Lan[2,4], Joshua Rhoades[1,2,3,4,5], Rajan Jain[1,2,3], Cheryl L. Smith[1,2,3,4] and Jonathan A. Epstein[1,2,3,4,*]

ABSTRACT

Dynamic organization of chromatin within the three-dimensional nuclear space has been postulated to regulate gene expression and cell fate. Here, we define the genome-wide distribution of nuclear peripheral heterochromatin as a multipotent P19 cell adopts either a neural or a cardiac fate. We demonstrate that H3K9me2-marked nuclear peripheral heterochromatin undergoes lineage-specific reorganization during cell-fate determination. This is associated with spatial repositioning of genomic loci away from the nuclear periphery as shown by 3D immuno-FISH. Locus repositioning is not always associated with transcriptional changes, but a subset of genes is upregulated. *Mef2c* is specifically repositioned away from the nuclear periphery during early neurogenic differentiation, but not during early cardiogenic differentiation, with associated transcript upregulation. *Myocd* is specifically repositioned during early cardiogenic differentiation, but not during early neurogenic differentiation, and is transcriptionally upregulated at later stages of cardiac differentiation. We provide experimental evidence for lineage-specific regulation of nuclear architecture during cell-fate determination in a mouse cell line.

KEY WORDS: H3K9me2 domains, Nuclear architecture, Nuclear peripheral heterochromatin

INTRODUCTION

A recent focus of research on gene regulation and epigenetics involves understanding the mechanisms by which DNA is packaged within the nucleus, and how this packaging and organization is dynamically regulated over time in specific cell types. This has been the focus of the federally funded '4D Nucleome Project' (Dekker et al., 2017) among many other research programs. One mechanism of chromatin organization involves localization of specific regions of the genome to the nuclear periphery in so-called lamina associated domains (LADs) (Guelen et al., 2008). LADs have been characterized as transcriptionally repressive heterochromatin at the nuclear periphery (Guelen et al., 2008; van Steensel and Belmont, 2017). Recent efforts to identify changes in LADs during differentiation of mouse embryonic stem cells (mESCs) to neural

precursors and astrocytes implicate a reorganization of LADs during neuronal differentiation (Dekker et al., 2017; Guelen et al., 2008; Peric-Hupkes et al., 2010). *Drosophila* neuroblasts undergo developmentally regulated sub-nuclear genome reorganization of the *hunchback* gene locus that correlates with loss of progenitor competence (Guelen et al., 2008; Kohwi et al., 2013). Likewise, differentiation of C2C12 myoblasts into myotubes suggests that there may be repositioning of myogenic genes away from the periphery during myogenic differentiation (Guelen et al., 2008; Robson et al., 2016; van Steensel and Belmont, 2017). Nevertheless, it remains unclear whether lineage-specific spatial reorganization of nuclear peripheral heterochromatin accompanies lineage specification when the multipotent progenitor cell adopts different cell fates in a single system.

We sought to examine lineage-specific chromatin organization and nuclear peripheral heterochromatin changes that occur during the acquisition of different cell fates. Using a reductionist approach, we assessed changes in nuclear peripheral heterochromatin in P19 multipotent embryonal carcinoma cells, which are competent to adopt a neurogenic or a cardiogenic fate in a simple, well-established model system. P19 cells are karyotypically normal cells that are multipotent and able to form cells of all three germ layers when injected into blastocysts (Rossant and McBurney, 1982). They differentiate into neuronal cells upon exposure to retinoic acid (Jones-Villeneuve, 1982) or into cardiomyocytes and skeletal muscle upon exposure to dimethyl sulfoxide (DMSO) (Edwards et al., 1983; Jasmin et al., 2010). Our group and others have recently reported that the histone modification H3K9me2 is specifically associated with LADs and is restricted to nuclear peripheral heterochromatin (Kind et al., 2013; Peric-Hupkes et al., 2010; Poleshko et al., 2017; van Steensel and Belmont, 2017). We used H3K9me2 ChIP-seq to monitor changes in nuclear peripheral heterochromatin in the P19 differentiation system as cells specifically adopted a neuronal fate or cardiac mesodermal fate.

Here, we report that there is dynamic local reorganization of nuclear peripheral heterochromatin during cell-fate acquisition that is specific to the particular cell fate being adopted. In addition, we demonstrate that genomic loci of crucial developmental master regulators are released from the nuclear periphery according to the specific cell fate during this process. This provides evidence that spatial organization of chromatin with respect to the nuclear periphery provides a layer of regulation that is specific to the distinct cell fate acquired.

RESULTS

P19 cells are competent to adopt neuronal or cardiac cell fates

Using well established protocols (Jasmin et al., 2010; Martins et al., 2005), multipotent P19 embryonic carcinoma cells can be induced to differentiate along separate cell lineages. To promote

[1]Department of Medicine, Perelman School of Medicine, University of Pennsylvania, PA 19104, USA. [2]Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania, PA 19104, USA. [3]Institute for Regenerative Medicine, Penn Cardiovascular Institute, Perelman School of Medicine, University of Pennsylvania, PA 19104, USA. [4]Penn Epigenetics Institute, Perelman School of Medicine, University of Pennsylvania, PA 19104, USA. [5]Institute for Biomedical Informatics, Perelman School of Medicine, University of Pennsylvania, PA 19104, USA.

*Author for correspondence (epsteinj@pennmedicine.upenn.edu)

K.S., 0000-0003-0647-6977; J.A.E., 0000-0001-8637-4465

DEVELOPMENT

differentiation along a neuronal lineage, P19 cells were cultured with 1 µM *all-trans* retinoic acid (RA). Alternately, to differentiate cells along a mesodermal lineage, cells were grown in 1% DMSO (Fig. 1A). Undifferentiated P19 cells were cultured at low density to prevent ectopic differentiation and to maintain multipotency. The morphologies of RA-treated and DMSO-treated cells were distinct at later stages of treatment: RA-treated cells displayed long cellular processes that resembled neurite extensions by day 5 (Fig. S1), whereas spontaneous beating foci were observed in DMSO-treated

cultures by day 10 (Movie 1). In addition to monitoring morphology, we compared protein levels of several known markers of differentiation 3 days after each treatment and in the undifferentiated cells. The pluripotency marker Oct4 (also known as Pou5f1) was downregulated at day 3 (D3) in both RA- and DMSO-treated cells, which indicates loss of pluripotency by D3 in both treatment conditions (Fig. 1B). RA treatment resulted in specific upregulation of the neurogenic transcription factor EphA3 at D3 (Fig. 1B). DMSO treatment induced specific upregulation of the
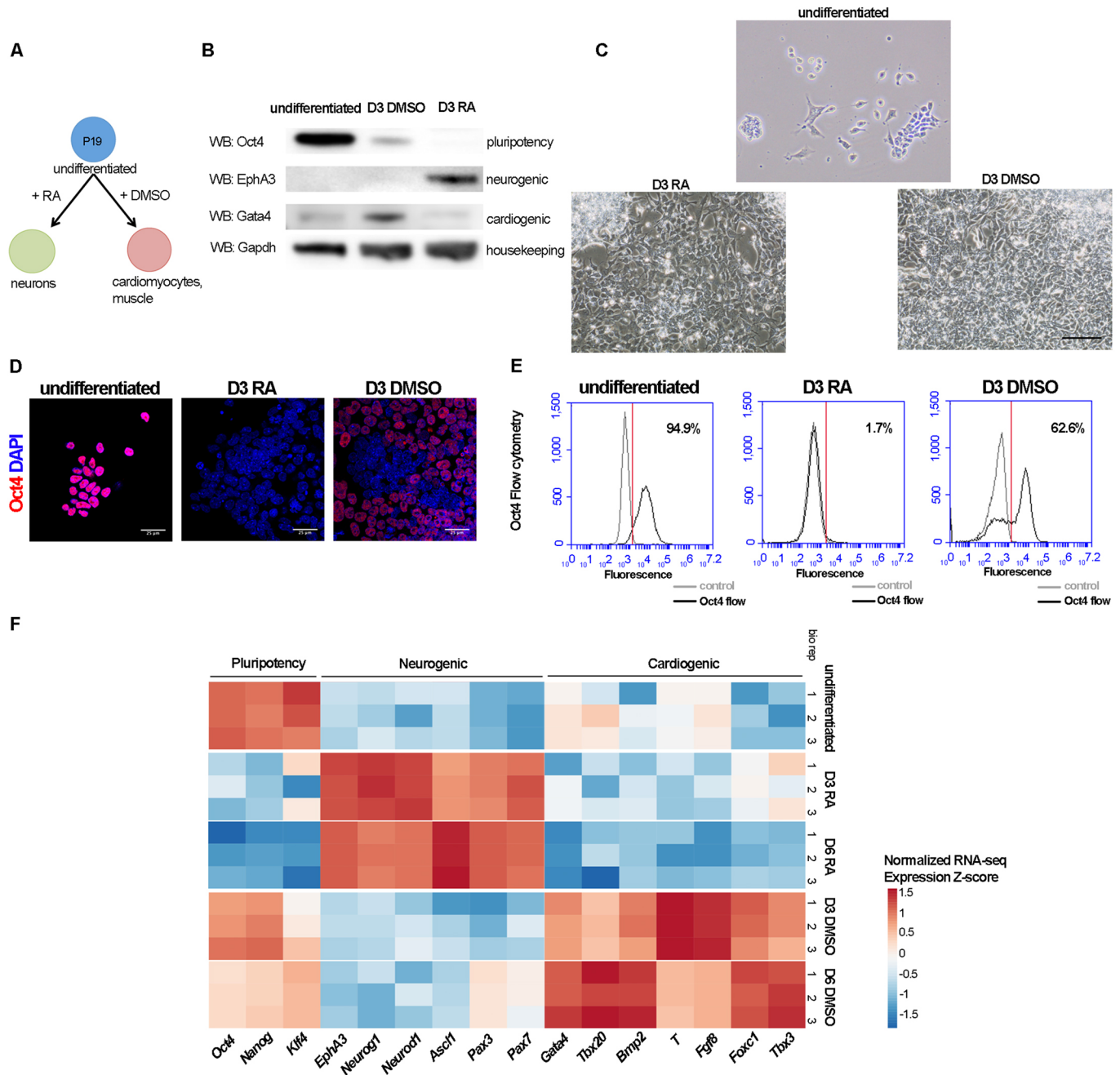
Fig. 1. P19 cells are competent to adopt a neural fate or cardiac fate in a model system of cell-fate choice. (A) Scheme for differentiation of P19 cells. (B) Western Blot (WB) of pluripotency marker (Oct4), neurogenic transcription factor (EphA3), cardiogenic transcription factor (Gata4) and housekeeping marker (Gapdh) during differentiation of P19 cells. (C) Representative images of undifferentiated P19, D3 RA- and D3 DMSO-treated cells. (D) Immunofluorescence of pluripotency factor Oct4 in undifferentiated, D3 RA- and D3 DMSO-treated P19 cells. (E) Flow cytometry of undifferentiated, D3 RA- and D3 DMSO-treated P19 cells stained for Oct4 to determine the percentage of Oct4[+] cells. Gray line represents negative control; black line represents Oct4-stained sample. (F) Heatmap representation of RNA-seq of P19 cells during differentiation. bio rep, biological replicate. Normalized RNA-seq expression z-score provided. Scale bars: 250 µm in C; 25 µm in D.

cardiogenic transcription factor Gata4 at D3 (Fig. 1B). This suggests that the molecular markers of cell-fate specification were expressed at D3, even before the observed morphological changes that were induced by DMSO or RA (Fig. 1C).

Using both immunofluorescence and flow cytometry, we found that pluripotency factor Oct4 was highly expressed in almost all undifferentiated cells (94.9% by flow cytometry; Fig. 1D,E) but was found to be dramatically depleted to ~1.7% Oct4$^+$ cells after 3 days of RA treatment, and reduced to ~62.6% Oct4$^+$ cells after 3 days of DMSO treatment (Fig. 1D,E). This suggests that D3 RA-treated cells were more efficiently differentiated compared with D3 DMSO-treated cells.

To further characterize the differences between RA- and DMSO-treated P19 cells, we assessed global changes in gene expression at early stages of differentiation, before morphological changes were readily apparent. We extracted RNA and performed RNA-seq on three biological replicates each of undifferentiated P19 cells as well as RA- and DMSO-differentiated cells that had been treated for 3 days or 6 days. Relative levels of gene expression for replicates of each condition cluster together, and the overall pattern of changes in transcription reveals clear differences between D3 RA- and D3 DMSO-treated cells (Fig. 1F). Differential gene expression reflects a transition from undifferentiated towards either a neurogenic fate following 3 days of RA treatment (upregulation of genes involved in nervous system development, axon guidance and neurogenesis; Table S1, Fig. S2A) or a cardiac mesodermal fate after 3 days of DMSO treatment (upregulation of genes required for somitogenesis, heart morphogenesis and heart development; Table S2, Fig. S2B). Following differentiation in either condition, a number of genes that encode pluripotency factors (*Oct4*, *Nanog*, *Klf4*) were downregulated (Fig. 1F).

The D3 and day 6 (D6) RNA-seq data demonstrate that specific differences in transcript abundance discriminate between treatment conditions. Several neurogenic transcription factors, including *EphA3*, *Neurog1*, *Neurod1*, *Ascl1*, *Pax3* and *Pax7* were specifically upregulated in D3 RA-treated cells, but not in D3 DMSO-treated cells (Fig. 1F). Cardiogenic transcription factors, including *Gata4*, *Tbx20*, *Bmp2*, *T*, *Fgf8*, *Foxc1* and *Tbx3* were specifically upregulated in D3 DMSO-treated cells, but not in D3 RA-treated cells (Fig. 1F). We validated a subset of genes by real-time quantitative PCR (RT-qPCR) and confirmed the depletion of *Oct4* in D3 RA-treated cells and downregulation of *Oct4* in D3 DMSO-treated cells (Fig. S2C), specific upregulation of neurogenic transcription factors (*Pax3*, *Neurog1*) in D3 RA-treated cells (Fig. S2D,E) and specific upregulation of cardiogenic *Gata4* transcription factor in D3 DMSO-treated cells (Fig. S2F). These data demonstrate that P19 cells can be induced to adopt a treatment-specific cell fate – either neuronal or cardiac mesodermal – and that the differences are reflected in distinctive morphologies and gene expression patterns.

H3K9me2-marked heterochromatin reflects nuclear architecture in P19 cells

To identify the changes in nuclear architecture that accompany lineage specification, we examined the distribution pattern of genome-wide H3K9me2-marked chromatin in undifferentiated P19 cells, and in D3 RA- and D3 DMSO-treated P19 cells. As shown previously, H3K9me2 preferentially marks chromatin at the nuclear periphery and H3K9me2 domains are highly correlated with LADs (Poleshko et al., 2017). We performed H3K9me2 ChIP-seq of undifferentiated D3 RA- and D3 DMSO-treated P19 cells with three biological replicates per condition (Fig. 2A,B; merged biological

replicates for visualization). We chose this time point to assess H3K9me2 domains as molecular markers of lineage specificity were already observed following 3 days of either treatment (Fig. 1F, Tables S1 and S2). Control H3 ChIP-seq was performed in parallel for all samples in order to assess the specific contribution of H3K9me2. The H3K9me2 signal, as assayed in 10 kb bins, was highly correlated within each set of biological replicates [average H3K9me2 ChIP-seq inter-replicate correlation (see Materials and Methods): Spearman's correlation coefficient $r=0.87$ undifferentiated, $r=0.94$ RA, $r=0.9$ DMSO] (Fig. S3A). In comparison, the correlation coefficients across treatment conditions were lower (average of 10 kb bins individual replicates: undifferentiated versus RA, $r=0.73$; undifferentiated versus DMSO, $r=0.85$; RA versus DMSO, $r=0.84$; Fig. S3A). This confirms that there was greater variability across differentiation conditions than within individual biological replicates per condition.

We observed genome-wide coverage of H3K9me2-marked regions in P19 cells to be ~40% of the mouse genome (ranging from 38.5% in undifferentiated to 41.7% in D3 RA and 40.8% in D3 DMSO; Fig. S3B). This is consistent with the genome coverage that is reported for LADs in human fibroblast cells (40%) (Guelen et al., 2008) and various mouse cells (35-43%) (Meuleman et al., 2013; Peric-Hupkes et al., 2010; Poleshko et al., 2017). The H3K9me2 domains that are found in undifferentiated P19 cells were also similar to the LaminB or H3K9me2 domains that were previously observed in mESCs (Poleshko et al., 2017) (Fig. S3C,D).

H3K9me2 domains vary according to cell fate

Using Enriched Domain Detector (EDD; Lund et al., 2014), we defined H3K9me2 domains and asked whether these domains varied by treatment in the P19 differentiation system. We performed principal component analysis (PCA) on the genome-wide H3K9me2 domains and leveraged the replicate data to compare domains across differentiation conditions. We found that the three biological replicates were similar and cluster together within each condition, whereas significant variation separates the samples by treatment condition (Fig. 2C, 95% confidence interval ellipses). The H3K9me2 domains were sufficient to accurately distinguish between undifferentiated, D3 RA-treated and D3 DMSO-treated P19 cells.

To assess the extent of similarity across the various conditions, we also measured the Jaccard indices and found that the H3K9me2 domains in undifferentiated samples showed differences between either D3 RA- or D3 DMSO-treated samples (undifferentiated versus D3 RA=0.76, undifferentiated versus D3 DMSO=0.78; Fig. 2D).

Although a large subset of H3K9me2 domains (68.1% of H3K9me2 domain genome coverage in base pairs) was constitutive across all conditions, there were distinct regions that varied between undifferentiated and differentiated P19 cells (31.9%) (Fig. 2A,B,E). These variable regions include a subset (3.0%) that are found in undifferentiated cells, but absent in both D3 RA- and D3 DMSO-treated samples (shared loss; Fig. 2E). Similarly, 8.3% of H3K9me2-marked domains were not observed in undifferentiated cells, but were present in both RA- and DMSO-treated cells (shared gain; Fig. 2E). Notably, there were twice as many regions (20.6%) that were specific to either the neuronal or myogenic lineage within the set of H3K9me2 domains that changed as a result of differentiation (specific; Fig. 2E). In particular, regions that lost H3K9me2 specifically in RA or DMSO (9.8%) had three times more genome coverage than regions that were lost in both conditions (3.0%) (specific loss versus shared loss; Fig. 2E). There were also
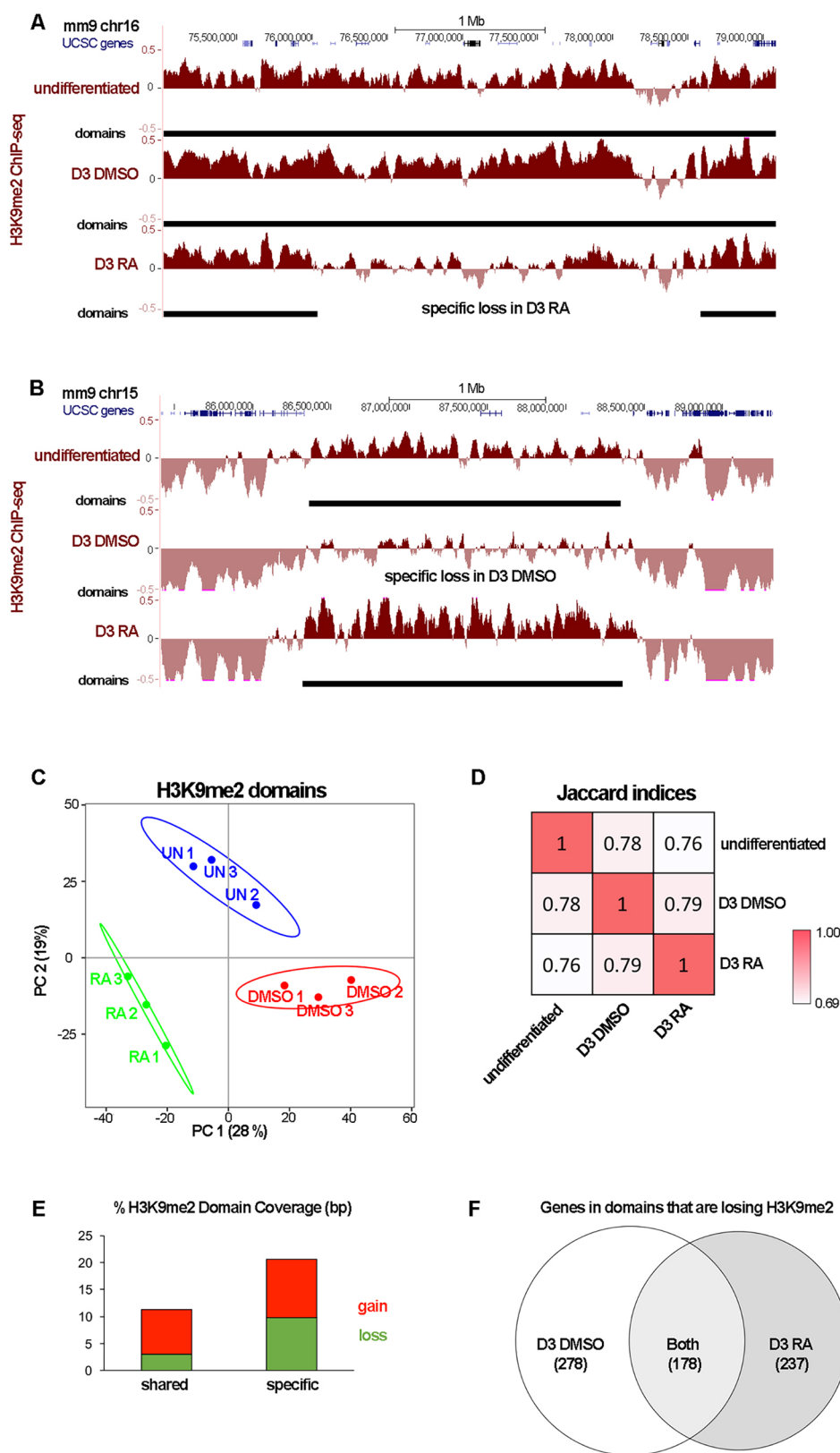
Fig. 2. H3K9me2-marked nuclear peripheral heterochromatin displays lineage-specific local reorganization during cell-fate choice. (A,B) Representative merged H3K9me2 ChIP-seq tracks that show regions of specific loss in D3 RA- or in D3 DMSO-treated cells. Black bars indicate H3K9me2 domains called by EDD. (C) PCA plot of H3K9me2 domains accurately stratifies individual biological replicates by treatment condition. Ellipses show 95% confidence interval. UN, undifferentiated; DMSO, D3 DMSO-treated; RA, D3 RA-treated. (D) Jaccard indices of H3K9me2 domains found in undifferentiated, D3 DMSO- or D3 RA-treated P19 cells. (E) Comparison of % H3K9me2 domain coverage (base pairs) in shared or specific regions (D3 DMSO only or D3 RA only). Gain, regions that gained residency within H3K9me2 domains; loss, regions that lost residency within H3K9me2 domains. (F) Number of genes found in H3K9me2 domains that were specifically lost in either D3 DMSO or D3 RA.

more genes found in regions of specific loss in D3 RA-treated or D3 DMSO-treated compared with regions of shared loss (Fig. 2F; Tables S5-S7).

We defined genes to be within the H3K9me2 domains if the gene has at least 75% overlap with H3K9me2 domain (see Materials and

Methods). We then asked whether genes found within regions that lost the identifying H3K9me2 mark during lineage specification had any function relevant to the cell fate the P19 cells would adopt. We found that there were many neuronal genes (*Mef2c*, *Sv2b*, *Nrp1*, *Sox3*, *Dpysl3*, *App*, *Mecom*, *Top2b*, *Ptn*, *Rarb*, *Camk4*, *Gap43*,

Rai2, *St7*, *Syt10*, *Tcf4*) with functions that are important for neuronal development that lost the H3K9me2 mark specifically in D3 RA-treated samples (Table S5), which suggests that the genes were highly relevant and specific to the neuronal cell fate being adopted. Conversely, we found that there were many cardiac genes [*Myocd*, *Gfra2*, *Ghr*, *Map2k4*, *Wisp1* (*Ccn4*), *Nr3c1*, *Plxna4*, *Adcy2*, *Odz2* (*Tenm2*), *Bzw2*, *Tnik*, *Pld1*, *Ptger3*, *Ndrg1*, *Prune2*, *Esrrg*, *Lmcd1*, *Lmo7*, *Palmd*, *Hmgxb4*, *Tbc1d4*, *Sox11*, *Syne1*, *Dock1*, *Zfp161* (*Zbtb14*), *Adamts1*] that are important for cardiac muscle development that lost the H3K9me2 mark specifically in D3 DMSO-treated samples (Table S6). Importantly, these neuronal or cardiac genes were found in regions that lost the H3K9me2 mark in a lineage-specific fashion. This suggests that genes losing lineage-specific H3K9me2-marked domains and nuclear peripheral localization upon differentiation have functions that are important and specific to the cell fate to be adopted.

Genomic loci reposition from the nuclear periphery with lineage specificity

To investigate global cellular changes of the H3K9me2-marked nuclear peripheral heterochromatin during P19 differentiation, we examined H3K9me2 by immunofluorescence in undifferentiated, D3 RA-treated and D3 DMSO-treated cells. Using high-resolution confocal microscopy to visualize H3K9me2, we observed H3K9me2-marked heterochromatin at the nuclear periphery, adjacent to the nuclear lamina, in both undifferentiated and differentiated P19 cells (Fig. 3A). There was no apparent alteration in the nuclear peripheral heterochromatin layer as assessed by the thickness of the H3K9me2 immunofluorescent signal following DMSO or RA treatment compared with untreated cells (average thickness 0.34 µm; Fig. 3B). We also examined the total amount of the H3K9me2 modification in nuclear lysates by western blot (Fig. 3C). We observed very similar levels of H3K9me2 (normalized to total H3) in nuclear lysates from untreated and differentiated P19 cells (Fig. 3D).

We hypothesized that, in undifferentiated cells, genes that encode lineage-specific factors would be found in nuclear peripheral heterochromatin and would be released from the nuclear periphery upon differentiation. We examined the nuclear localization of individual loci that are differentially marked by H3K9me2 according to the differentiation conditions. Myocyte enhancer factor 2c (*Mef2c*) is a transcription factor that is involved in neurogenesis (Leifer et al., 1993) as well as cardiogenesis (Lin et al., 1997). H3K9me2 ChIP-seq data indicate that the *Mef2c* locus is H3K9me2-marked in undifferentiated P19 cells and in D3 DMSO-differentiated cells, but H3K9me2 is reduced at the *Mef2c* locus in D3 RA-differentiated cells (average H3-normalized H3K9me2 counts per kb: undifferentiated=19.6, D3 RA-treated= −16.2, D3 DMSO-treated=16.2). We performed high-resolution 3D immunofluorescent *in situ* hybridization (immuno-FISH) (Poleshko et al., 2017) to assess the position of the *Mef2c* locus and to quantify the 3D distance between the locus and the nuclear periphery. *Mef2c*-specific DNA FISH probes were used to visualize the locus relative to the nuclear lamina (LaminB immunofluorescence). Given that DMSO-treatment results in efficient differentiation of only a subset of cells at D3 (Fig. 1D,E), we included immunofluorescence of Oct4 in the immuno-FISH protocol and focused on Oct4$^-$ cells in the D3 DMSO-treated analysis. Through high-resolution 3D immuno-FISH, we verified that the *Mef2c* locus was positioned at the nuclear periphery in both undifferentiated P19 cells (Oct4$^+$) and D3 DMSO-treated Oct4$^-$ cells, but was repositioned away from the nuclear periphery in D3 RA-treated cells (Oct4$^-$) (Fig. 3E,F). We quantified

the 3D distance to the nuclear lamina for each allele in each condition and observed a statistically significant difference between undifferentiated and D3 RA-treated cells as well as D3 DMSO-treated Oct4$^-$ cells compared with D3 RA-treated cells (Fig. 3F). For each condition, we also scored the number of alleles (0, 1, or 2) at the nuclear periphery in each cell and found that only D3 RA-treated samples had a majority of cells with at least one allele repositioned away from the nuclear periphery (Fig. 3G). These data are consistent with our H3K9me2 ChIP-seq findings and suggest that the release of the *Mef2c* locus from the nuclear periphery is specific to the neuronal cell-fate choice at D3 of RA-induced differentiation.

Another gene locus that displayed lineage-specific changes in H3K9me2 is myocardin (*Myocd*), which is a master regulator of cardiomyocytes and cardiac smooth muscle (Li et al., 2003; Wang et al., 2001). Comparison of the H3K9me2 ChIP-seq signal at the *Myocd* locus in undifferentiated P19 cells with that found at *Myocd* in D3 DMSO-treated cells revealed specific loss of H3K9me2 in D3 DMSO-treated, but not in D3 RA-treated samples (average H3-normalized H3K9me2 counts per kb: undifferentiated=15.3, D3 RA-treated=21.3, D3 DMSO-treated=−4.3). Consistent with the H3K9me2 ChIP-seq data, observations of the *Myocd* genomic locus by 3D immuno-FISH revealed that this locus was positioned at the nuclear periphery in undifferentiated P19 cells but localized to the nuclear interior in D3 DMSO-treated Oct4$^-$ cells (Fig. 3H-J). In D3 RA-treated samples, the majority of the *Myocd* locus remained at the nuclear periphery (Fig. 3I,J). These results indicate that release of the *Myocd* locus from the nuclear periphery is specific to the D3 DMSO-induced cardiomyocyte lineage.

Loss of locus-specific H3K9me2 mark correlates with transcriptional activation for a subset of genes

Several studies have shown that positioning of genomic loci at the nuclear periphery is correlated with transcriptional repression (Peric-Hupkes et al., 2010; Robson et al., 2016) and inhibition of G9a, the histone methyl-transferase that is responsible for H3K9 dimethylation, which results in upregulated expression of lamin-associated genes (Yokochi et al., 2009). We therefore asked whether those genes that have changes in H3K9me2 and nuclear localization also vary in transcriptional activity during lineage specification.

We examined genome-wide changes in gene expression during lineage-specific H3K9me2 reorganization through RNA-seq analysis of undifferentiated, RA- and DMSO-treated P19 cells. Across all treatment conditions, genes that are found within H3K9me2 domains have significantly lower gene expression compared with those outside of H3K9me2 domains (Fig. S4A, $P<2.2\times10^{16}$ Mann–Whitney *U*-test). We assessed the connection between transcriptional activity and the changes in nuclear peripheral positioning that occurred during our differentiation protocol by examining the expression level of genes that have reduced H3K9me2 specifically in D3 RA-treated or D3 DMSO-treated samples relative to undifferentiated P19 cells. Although most genes that show an RA treatment-specific reduction in H3K9me2 domain residency had no corresponding change in gene expression (69.9%), a subset of genes (27.6%), including *Mef2c*, was transcriptionally upregulated in D3 RA (Fig. 4A, dots sized by quartile rank in undifferentiated RNA-seq; Table S8). This suggests that the release of a significant number of genes from the repressive environment of the nuclear periphery, which occurred during RA differentiation, was associated with upregulation of gene expression in D3 RA-treated P19 cells. We did not observe any significant number of genes that show DMSO-specific reduction of H3K9me2 to be transcriptionally upregulated in D3 DMSO (Fig. 4B; Table S9).
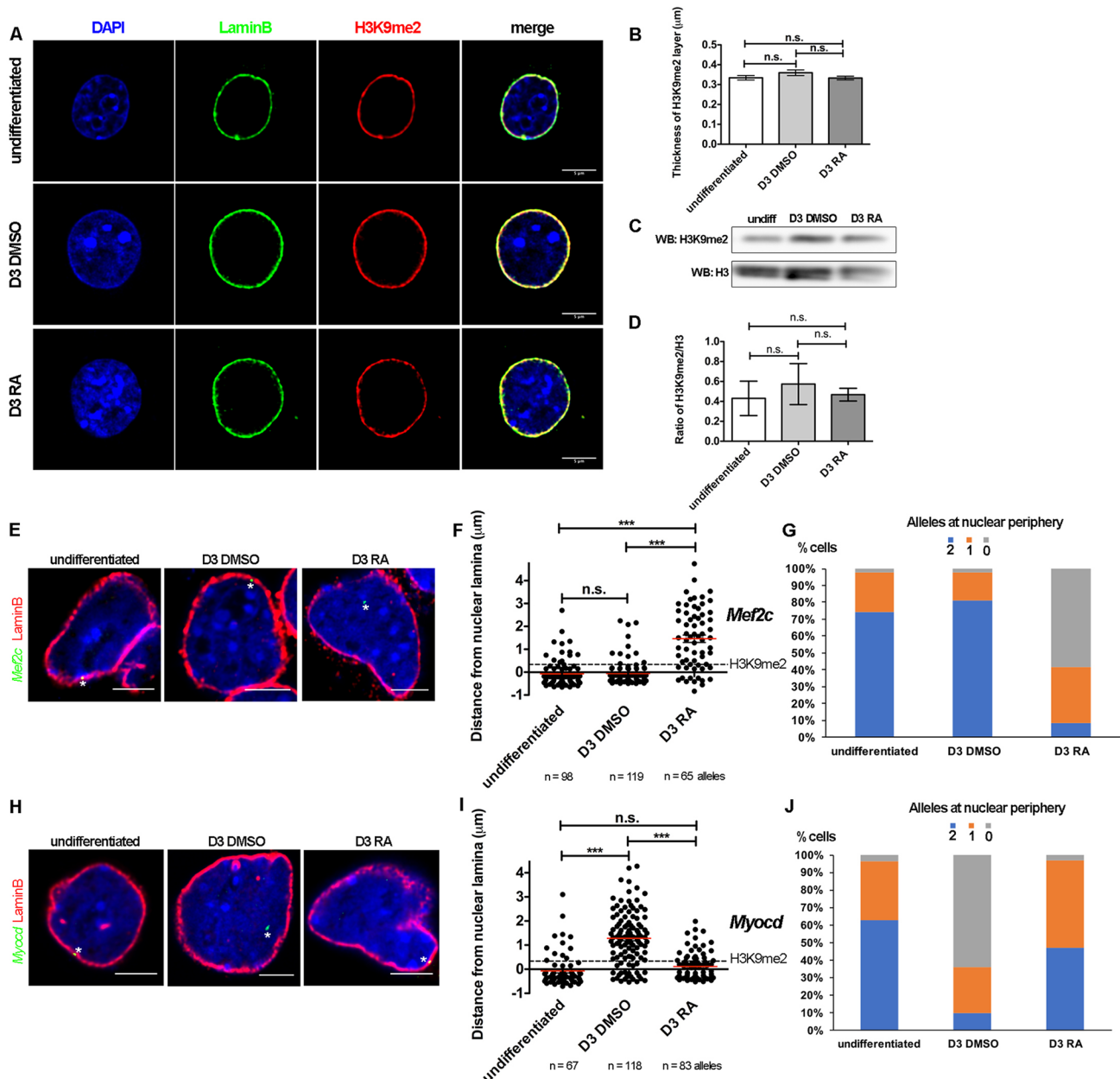
Fig. 3. Lineage-specific master regulators *Mef2c* and *Myocd* are specifically repositioned from nuclear peripheral heterochromatin during cell-fate choice. (A) Immunofluorescence of H3K9me2 and LaminB reveals nuclear peripheral heterochromatin in undifferentiated, D3 DMSO- and D3 RA-treated P19 cells. (B) Quantification of H3K9me2 layer thickness. One-way ANOVA test. *n*=30 measurements per condition. Error bars indicate s.e.m. (C) Representative western blot of H3K9me2 and H3 in nuclear lysates of undifferentiated, D3 DMSO- and D3 RA-treated P19 cells. (D) Quantification of H3K9me2 (normalized by H3) western blot levels. One-way ANOVA test. *n* =3 biological replicates. (E) High-resolution 3D immuno-FISH reveals that the *Mef2c* genomic locus is specifically released from the nuclear periphery in D3 RA-treated, but not in undifferentiated or D3 DMSO-treated cells. Representative single confocal *z*-slice of immuno-FISH of indicated loci in each cell type immunostained with LaminB. (F) Quantification of 3D distance between *Mef2c* locus and the nuclear periphery. One-way ANOVA test. ***$P<0.001$. Dotted line represents the thickness of H3K9me2; red line indicates mean; *n*, number of alleles counted per condition. For the D3 DMSO-treated condition, Oct4⁻ cells were quantified. (G) Percentage of cells with 0, 1 or 2 alleles of *Mef2c* at the nuclear periphery. Average number of cells is 31 cells per condition. (H) High-resolution 3D immuno-FISH reveals that the *Myocd* genomic locus is specifically released from the nuclear periphery in D3 DMSO-treated, but not in undifferentiated or D3 RA-treated cells. Representative single confocal *z*-slice of immuno-FISH of indicated loci in each cell type immunostained with LaminB. (I) Quantification of 3D distance between *Myocd* locus and the nuclear periphery. One-way ANOVA test. ***$P<0.001$. Dotted line represents the thickness of H3K9me2; red line indicates mean; *n*, number of alleles counted per condition. For the D3 DMSO-treated condition, Oct4⁻ cells were quantified. (J) Percentage of cells with 0, 1 or 2 alleles of *Myocd* at the nuclear periphery. Average number of cells is 36 cells per condition. n.s., not significant. Asterisks (E,H) highlight the position of the genomic locus. Scale bars: 5 µm in A; 5 µm in E,H.

To determine whether genes that were released from the nuclear periphery might be transcriptionally activated at a later stage of differentiation, we analyzed RNA-seq data from D6 RA-treated samples. We found no significant difference in gene expression profiles for the same set of genes between D3 RA- and D6 RA-treated ($P=0.295$, 2×3 Chi² test, Fig. 4C; Table S10). Similarly, we found that most loci with DMSO-specific reduction in H3K9me2 according to ChIP-seq
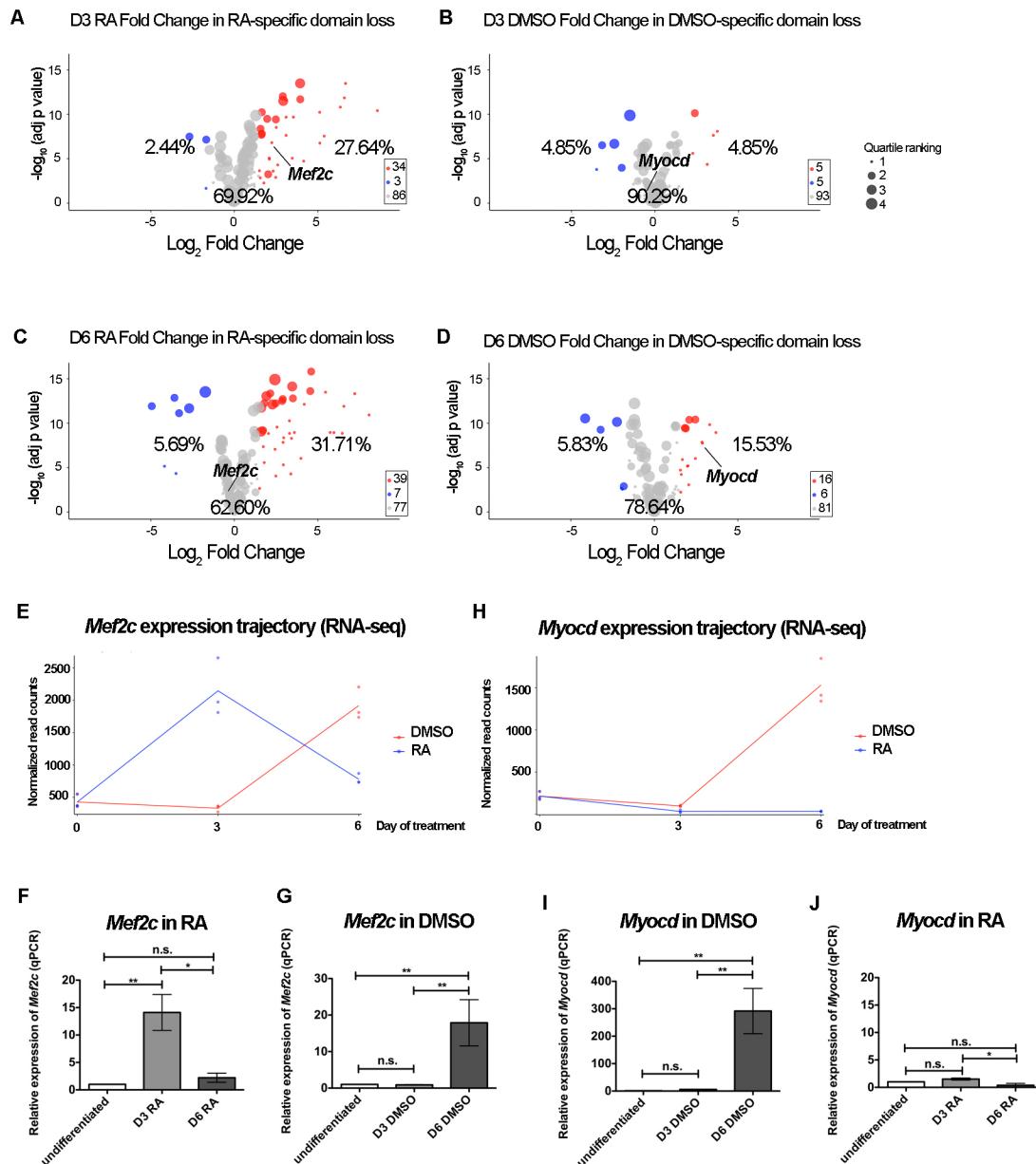
Fig. 4. Loss of residency within H3K9me2 domains results in transcriptional activation for subset of genes. (A) Volcano plot showing D3 RA RNA-seq expression profile of genes that lost residency within H3K9me2 domains specifically in D3 RA-treated cells. A substantial proportion of genes (23.93%) are transcriptionally upregulated when spatially repositioned from the nuclear periphery. Dots are sized by quartile ranking of expression in undifferentiated P19 cells. Red dots indicate significantly upregulated genes ($\log_2$ fold change$\geq$1.5, $P\leq$0.05), blue dots indicate significantly downregulated genes ($\log_2$ fold change$\leq-$1.5, $P\leq$0.05), gray dots are non-significantly changed. Fold changes calculated for D3 RA-treated cells with respect to undifferentiated cells. (B) Volcano plot showing D3 DMSO RNA-seq expression profile of genes that lost residency within H3K9me2 domains specifically in D3 DMSO-treated cells. Most genes that repositioned away from the nuclear periphery in D3 DMSO-treated cells did not have associated transcriptional changes. Fold changes calculated for D3 DMSO-treated cells with respect to undifferentiated cells. (C) Volcano plot showing D6 RA RNA-seq expression profile of genes that lost residency within H3K9me2 domains specifically in D3 RA-treated cells. Fold changes calculated for D6 RA-treated cells with respect to undifferentiated cells. (D) Volcano plot showing D6 DMSO RNA-seq expression profile of genes that lost residency within H3K9me2 domains specifically in D3 DMSO-treated cells. Fold changes calculated for D6 DMSO-treated cells with respect to undifferentiated cells. (E) *Mef2c* expression trajectory across day 3 (D3) and day 6 (D6) of treatment based on RNA-seq data. (F,G) Relative gene expression levels of *Mef2c* across RA or DMSO differentiation by RT-qPCR. One-way ANOVA. *P<0.05, **P<0.01; n=3-5 biological replicates. (H) *Myocd* expression trajectory across day 3 (D3) and day 6 (D6) of treatment based on RNA-seq data. (I,J) Relative gene expression levels of *Myocd* across DMSO or RA differentiation by RT-qPCR. One-way ANOVA. *P<0.05, **P<0.01. n=3-5 biological replicates. n.s. not significant.

analysis showed little significant difference in gene expression when assayed at D3- or D6- DMSO treatment (Fig. 4D, P=0.0354, 2×3 Chi^2 test; Table S11). This suggests that most genes that were released at D3 were not transcriptionally upregulated by D6 of differentiation, though we cannot rule out that they would be upregulated at later time points.

According to our RNA-seq analysis, expression of *Mef2c* is ~fourfold higher in D3 RA-treated cells than in undifferentiated P19 cells (Fig. 4E), which is consistent with both its reduction in H3K9me2 signal and localization change from the nuclear periphery towards the nuclear interior (Fig. 3F). By day 6 of RA treatment, however, *Mef2c* expression is relatively diminished (Fig. 4C,E). In

addition to neurogenesis, *Mef2c* is also known to play a role in cardiac and skeletal muscle development (Lin et al., 1997) and has been previously shown to be transcriptionally upregulated in P19 cells by day 6 of DMSO treatment (Grépin et al., 1997; Skerjanc et al., 1998). Our RNA-seq results confirmed that *Mef2c* is expressed in D6 DMSO (Fig. 4E). We further validated these results with RT-qPCR (Fig. 4F,G). *Mef2c* expression was upregulated in D3 RA, but not in D6 RA samples (Fig. 4F), whereas DMSO treatment led to *Mef2c* expression being unchanged in D3 DMSO samples, but dramatically upregulated at D6 DMSO (Fig. 4G). The expression data for *Mef2c* correlates with nuclear localization and demonstrates a link between transcriptional activity and spatial positioning of a genomic locus in the nucleus. Notably, this illustrates lineage-specific modification of the H3K9me2 mark and its association with gene expression during cell-fate specification.

Given the less efficient differentiation induced by DMSO at D3, the lack of transcriptional upregulation of most genes in D3 DMSO might be because of the heterogeneity of this population of cells. *Myocd* is an example of a locus that was released from the nuclear periphery specifically in the D3 DMSO-treated cells. *Myocd* was transcriptionally unchanged at D3 DMSO treatment, but was dramatically upregulated in D6 DMSO cells (Fig. 4B,D,H). This was verified through RT-qPCR analysis (Fig. 4I). Notably, this upregulation is specific to D6 DMSO and is not found in RA differentiation (Fig. 4H,J). Thus, *Myocd* is an example of a gene that is first released from the nuclear periphery at D3 DMSO, before the observed transcriptional upregulation at D6 DMSO. This provides evidence for genomic loci to be repositioned away from the nuclear periphery during early stages of differentiation before transcriptional activation at later stages of differentiation. Taken together, this demonstrates that the local reorganization of H3K9me2-marked nuclear peripheral heterochromatin is lineage-specific, but is not always associated with changes in transcriptional activity.

DISCUSSION
We report that H3K9me2-marked nuclear peripheral heterochromatin undergoes dynamic local reorganization during differentiation and that the loss of H3K9me2-marked domains displays lineage specificity. These lineage-specific regions contain genes that are functionally important for the future fate of the cell. Using high-resolution 3D immuno-FISH, we confirm that the changes in H3K9me2 domains as determined by ChIP-seq are consistent with spatial repositioning of genomic loci (*Mef2c*, *Myocd*) with respect to the nuclear periphery. Most genes that are released from the nuclear periphery do not have transcriptional changes, but a subset of genes has associated transcriptional upregulation. We provide evidence that the dynamic local reorganization of H3K9me2-marked nuclear peripheral heterochromatin is lineage specific.

Understanding the mechanisms that drive cell-fate specification is at the heart of developmental biology and a number of studies have described the transcriptional changes that accompany cell fate choice. In addition, the three-dimensional nuclear architecture and its impact on the localization of relevant gene loci within different compartments of the nucleus should be recognized and examined in the context of cell-fate specification. In this study, by demonstrating that there is lineage-specific reorganization of H3K9me2-marked nuclear peripheral heterochromatin during cell-fate choice, we provide evidence that the same multipotent progenitor cell reorganizes its peripheral heterochromatin specifically according to the distinct cell fate it is adopting.

Spatial regulation of transcriptional activation
We found that loss of the H3K9me2 mark in specific domains corresponds to repositioning of genomic loci away from the nuclear periphery, suggesting that H3K9me2 ChIP-seq accurately predicts spatial positioning of genomic regions within nuclear peripheral heterochromatin. In *Caenorhabditis elegans*, MET-2 and SET-25 (orthologs of G9a and SetDB1), enzymes that promote mono-/di-methylation or tri-methylation of lysine 9 on histone 3 (H3K9), respectively, are necessary for anchoring chromatin to the nuclear lamina and for transcriptional repression of heterochromatin (Gonzalez-Sandoval and Gasser, 2016; Towbin et al., 2012). Loss of SET-25 and MET-2 leads to loss of H3K9me1/2/3, transcriptional de-repression, and subsequent detachment of chromosomal regions from the nuclear lamina, which is generally consistent with our observations.

We found that the release of genomic loci from the nuclear periphery is not always concomitant with transcriptional activation, but a subset of released genes is transcriptionally upregulated. Consistent with this, there was no strong genome-wide correlation of gene detachment and upregulation in MET-2/SET-25 mutants, though a subset of genes that detach from the nuclear lamina was strongly upregulated (Gonzalez-Sandoval et al., 2015; Towbin et al., 2012). We interpret these findings to suggest that detachment from the nuclear periphery equates to competence for gene activation, but that additional transcriptional activation events are necessary for gene transcription subsequent to or simultaneous with detachment.

A recent study identified CEC-4 in *C. elegans* as a protein that specifically recognizes H3K9me2 and mediates the perinuclear localization of heterochromatin independent of lamin (Gonzalez-Sandoval et al., 2015). Loss of CEC-4 resulted in repositioning of loci away from the nuclear periphery, but these genes remained transcriptionally repressed and there was no global derepression observed (Gonzalez-Sandoval et al., 2015). This suggests that the spatial positioning of a locus can be uncoupled from transcriptional control (Gonzalez-Sandoval and Gasser, 2016). Indeed, it has been reported that chromatin decondensation is sufficient to reposition genomic loci away from nuclear periphery in mESCs without transcriptional activation (Therizols et al., 2014). This underscores the importance of defining the genome-wide H3K9me2 distribution to inform both the spatial positioning and repressive potential of genomic loci. In our study, we found that the loss of H3K9me2-marked domains accurately predicts spatial repositioning away from the nuclear periphery. We propose that this spatial repositioning increases the accessibility of the genomic loci for subsequent activation by enhancers or transcription factors during appropriate stages of lineage specification.

Spatial and temporal regulation of master regulators *Mef2c* and *Myocd*
Mef2c encodes a master regulator transcription factor that has been reported to be involved in neurogenesis (Leifer et al., 1993) and cardiogenesis (Lin et al., 1997). However, it remains unclear how the same transcription factor can be involved in two distinct lineage commitments. Interestingly, we observed that the timing of repositioning of the *Mef2c* locus from nuclear peripheral heterochromatin is highly specific to the fate that a cell adopts. We found that *Mef2c* is repositioned away from the nuclear periphery in D3 RA-induced early neurogenic differentiation and associated with transcriptional upregulation for neurogenesis. Conversely, we found that *Mef2c* remains at the nuclear periphery in D3 DMSO-induced early cardiogenic differentiation, but is

DEVELOPMENT

transcriptionally upregulated only later at D6 after DMSO-induced cardiac differentiation. Consistent with this, a previous study that described cardiogenesis in P19 cells reported that *Mef2c* is upregulated at D6 DMSO and, together with Nkx2-5, initiates cardiogenesis (Skerjanc et al., 1998). This suggests that the same genomic locus (*Mef2c*) can be repositioned away from the nuclear periphery at lineage-specific time points during cell-fate choice. This would provide an additional layer of spatial regulation of accessibility for the same transcription factor that is involved in two distinct cell fates.

Myocd encodes a master regulator for cardiac and smooth muscle (Li et al., 2003; Wang et al., 2001). We found that *Myocd* was released at D3 DMSO-induced early cardiogenic differentiation, but was not transcriptionally upregulated until D6 DMSO-treatment. This was not unexpected as *Myocd* has previously been reported to be expressed at D6 DMSO together with *Mef2c* in P19 cells, and to be required for cardiogenesis (Ueyama et al., 2003). Indeed, Mef2c has been shown to directly activate a distal *Myocd* enhancer during cardiovascular development (Creemers et al., 2006). Consistent with this result, we found that the *Myocd* locus is released from the periphery at D3 DMSO, but the expression of both *Myocd* and *Mef2c* are highly upregulated only at D6 DMSO. This suggests that the genomic locus (*Myocd*) can first be repositioned away from the nuclear periphery before transcriptional activation at later stages of development when other appropriate transcriptional activators (e.g. *Mef2c*) are expressed.

Taken together, we find that the spatial repositioning of genomic loci, such as master regulators *Mef2c* and *Myocd*, away from the nuclear periphery is tightly controlled during cell-fate decisions and the timing of release is tightly regulated in the process of lineage specification. Our data are consistent with a model in which lineage-specific master regulators positioned at the nuclear periphery in undifferentiated cells are specifically released and accessible for transcriptional activation during cell-fate determination. The dynamic reorganization of nuclear peripheral heterochromatin and H3K9me2 domains provides an additional layer of spatial regulation of transcription during lineage specification. Future studies will aim to uncover the molecular mechanisms that regulate lineage-specific reorganization of nuclear peripheral heterochromatin.

MATERIALS AND METHODS
Cell culture
P19 cells were obtained from ATCC (CRL-1825) and maintained at low confluency in α-Minimum Essential Media (α-MEM, 12571048, Invitrogen) supplemented with 10% fetal bovine serum (charcoal stripped FBS, F6765 Sigma-Aldrich) and 1% penicillin/streptavidin. P19 cells were differentiated in 1 μM RA (R2625, Sigma-Aldrich) or in 1% DMSO (D2650, Sigma-Aldrich) by forming embryoid bodies (EBs) for 2 days in bacteriological petri dishes, before transferring to T75 cell culture flasks and replacing with maintenance media.

Western blots
Protein lysates were run on 4%-12% NuPAGE Bis-Tris protein gels (NP0335, Invitrogen) and blots were probed with anti-H3K9me2 (1:2500, ab1220, Abcam), anti-H3 (1:2000, ab1791, Abcam), anti-Gata4 (1 μg/ml, ab84593, Abcam), anti-EphA3 (1:1000, sc-514209, Santa Cruz Biotechnology), anti-Pou5f1 (1:1000, ab19857, Abcam), anti-Gapdh (1:1000, 2118S, Cell Signaling Technology) according to manufacturer's instructions. Secondary antibodies used were: anti-rabbit IgG conjugated HRP (1:2000, 7074S, Cell Signaling Technology) or anti-mouse IgG conjugated HRP (1:5000, 7076S, Cell Signaling Technology). ECL Prime Western Blotting System (RPN2232, GE Healthcare) and Amersham Imager 600 (GE Healthcare) were used for visualization and imaging.

ChIP-seq library preparation
ChIP was performed similar to previously described (Poleshko et al., 2017). Briefly, P19 cells were resuspended and crosslinked in media with 1% methanol-free formaldehyde (Thermo Fisher Scientific) for 10 min at room temperature (rtp) with gentle rotation. Crosslinking was quenched with addition of glycine (125 mM final concentration) and incubated at rtp for 5 min with gentle rotation. Cells were pelleted by centrifugation (300 *g* for 3.5 min rtp) and resuspended in PBS, transferred to Eppendorf tubes and pellets were obtained by centrifugation (300 *g* for 3.5 min). Supernatant was discarded and pellets were flash frozen on dry ice and stored at −80°C.

For each ChIP sample, 50 μl of Dynabeads Protein G for Immunoprecipitation (10004D, Invitrogen) was washed in block buffer [0.5% bovine serum albumin (BSA) in PBS] 3× and beads were resuspended in 500 μl block buffer with 5 μg antibody (anti-H3K9me2, ab1220, Abcam; anti-H3, ab1791, Abcam) and incubated at 4°C with rotation for 6 h. Another 50 μl of Dynabeads Protein G per ChIP sample was washed in block buffer 3×, resuspended in 500 μl block buffer and set aside for pre-clearing of sonicated chromatin.

To obtain nuclei from the frozen cross-linked pellets, cells were resuspended in cold Cell Lysis Buffer 1 (50 mM HEPES-KOH, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100, fresh protease inhibitors) and rotated at 4°C for 10 min, followed by centrifugation (300 *g* for 4 min). Supernatant was discarded and pellet was resuspended in cold Cell Lysis Buffer 2 (10 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, fresh protease inhibitors) and rotated for 10 min at rtp. Pellet was obtained by centrifugation (300 *g* for 4 min) and resuspended in 1 ml cold Lysis Buffer 3 (10 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-deoxycholate, 0.5% N-laurylsarcosine, fresh protease inhibitors) and transferred to a pre-chilled 1 ml Covaris AFA tube. Samples (10 million cells per tube) were sonicated for 10 min at 4°C in a sonicator (Covaris S220). Sonicated lysate was transferred to Eppendorf LoBind tubes and Triton X-100 added (to a final concentration of 1%). Samples were centrifuged at maximum speed for 10 min at 4°C and supernatant was transferred to 500 μl pre-clearing beads and rotated for 2 h at 4°C. Pre-cleared sonicated chromatin was collected as supernatant by centrifugation at maximum speed for 10 min at 4°C. Quantification of DNA was performed using Qubit dsDNA HS assay kit (Q32854, Invitrogen). Antibody-conjugated beads were washed thrice in block buffer and resuspended in 100 μl of block buffer. Each ChIP sample had 15 μg of pre-cleared sonicated chromatin added to antibody-conjugated beads and incubated with rotation at 4°C overnight. We set aside 1.5 μg of sonicated chromatin per sample for input.

On the second day, beads were collected with magnetic holder (Dynamag-2 magnet, Thermo Fisher Scientific) and washed 5× in 1 ml RIPA wash buffer (50 mM HEPES-KOH, 500 mM LiCl, 1 mM EDTA, 1% NP-40, 0.7% Na-deoxycholate) with 5 min rotation at rtp per wash. Beads were washed in 1 ml final ChIP wash buffer [50 mM NaCl, 1× Tris-EDTA buffer (TE)] and 10% aliquot set aside for ChIP-western blotting before final resuspension in ChIP elution buffer (10 mM EDTA, 50 mM Tris-HCl, 1% SDS) and eluted at 65°C for 30 min. Around 10% of beads/input were set aside for ChIP-western blotting with anti-H3K9me2 antibody/anti-H3 antibody to ensure that there was effective immunoprecipitation of target antibody. Eluted ChIP DNA or thawed input DNA was transferred to a new LoBind tube and incubated at 65°C overnight to reverse crosslinks.

On the third day, 200 μl of 1× TE was added to the reverse crosslinked ChIP DNA and RNase treated (final concentration 0.2 mg/ml, 1010969001, Roche) at 37°C for 2 h, and Proteinase K treated (final concentration 0.2 mg/ml, CB3210-5, Denville Scientific) at 55°C for 2 h. ChIP and input DNA were purified through phenol:chloroform extraction and resuspended in 10 mM Tris-HCl (pH 8). Quantification of ChIP and input DNA was performed using Qubit dsDNA HS assay kit. ChIP-seq library preparation used 250 ng of ChIP or input DNA for the Next Ultra II DNA Library preparation kit for Illumina (E7645S, New England Biolabs) and NEBNext multiplex oligos for Illumina (dual index primers set I, E7600S, New England Biolabs) according to the manufacturer's instructions. The quality of individual ChIP-seq libraries was assessed using Bioanalyser (High Sensitivity DNA kit, 5067-4626, Agilent) and quantified by qPCR using Kapa Library Quantification Kit (KK4835, Kapa Biosystems). Libraries

were normalized and pooled for multiplex sequencing, re-quantified and sequenced on Nextseq500 platform (75 bp, single end, Illumina).

ChIP-seq analysis

Sequencing reads were aligned to genome assembly (NCBI37/mm9) using Bowtie2 (version 2.2.9) with default parameters and '–local' to allow soft clipping of ends. Alignments were further processed by SAMtools (version 0.1.19) to remove low quality alignments ('-q 10') and read duplicates ('rmdup -s'). Individual biological replicates had high correlation (Fig. S3A) and were merged using SAMtools. Each individual sample had an average of ~37.1 million reads (Table S3).

H3K9me2 domains for individual and pooled libraries were called using EDD (version 1.1.18) (Lund et al., 2014), which is ideal for the discovery of broad enrichment domains from ChIP-seq data. We used the following parameters '–bin-size auto –gap-penalty 3 –fdr 0.05' with H3K9me2 ChIP-seq as 'ChIP' and H3 ChIP-seq as 'input' to define H3K9me2 domains. Blacklist regions from ENCODE mm9 and default EDD mm9 blacklist regions were used and the required fraction of uninformative bins is 0.97.

The visualization track of the individual library was generated using bedtools, for which the genome coverage was normalized to 1 million reads per library. Tracks of log2-fold change between H3K9me2 and H3 ChIP were generated using deepTools (version 3.0.1, 'bigwigCompare –pseudocount=0.00001 –binSize=100'). The log2-fold change tracks were further quantile normalized using bwtool (version 1.0, 'summary -binSize=100') and R package preprocessCore (Bioconductor). A similar track generation approach was performed for merged replicates.

Spearman correlations between ChIP-seq libraries over the whole genome were computed using 10 kb bins and plotted using deepTools ('multiBigwigSummary' and 'plotCorrelation'). Metaplots of ChIP signal flanking differentially bound sites were also generated using deepTools ('computeMatrix' and 'plotHeatmap').

EDD called H3K9me2 domains were used as input into bedtools multiinter, which allowed identification of shared and unique domains of D0, DMSO and RA treatments. ENSEMBL mm9 gtf file was parsed with GTFtools (Li, 2018preprint) to create a gene start and stop coordinate file. Bedtools (version v2.27.1) intersect was used to find overlaps between gene coordinates and H3K9me2 domains. For a gene to be considered overlapping, 75% of the total gene length must lie within an H3K9me2 domain(s). Gene intersection output was used to extract lists of genes (Tables S5-S7) that are associated with shared and unique H3K9me2 domain subsets.

To obtain H3-normalized H3K9me2 counts per kb, we subtracted H3 ChIP-seq reads from H3K9me2 ChIP-seq reads across the region that covered the gene body with +5 kb extensions on both ends. The H3-normalized counts were then normalized by gene length per kb.

RNA isolation and preparation

Total RNA was isolated in Trizol (15596026, Ambion) and purified using RNeasy mini columns (74104, Qiagen) according to the manufacturer's instructions. RNA was digested with DNAse I (AM2238, Ambion) at 37°C for 30 min and heat inactivated at 70°C for 5 min before a second round of RNeasy cleanup. cDNA was synthesized using oligo dT primers with SuperScript III reverse transcriptase (18080044, Invitrogen). RT-qPCR were performed in triplicate with PowerSYBR Green PCR Master Mix (4367659, Applied Biosystems) on StepOnePlus Real-Time PCR System (43776600, Applied Biosystems). *Gapdh* was used as a housekeeping normalization control. The sequence of primers used for RT-qPCR are found in Table S4.

RNA-seq analysis

RNA-seq libraries were prepared using Truseq Stranded Total RNA Library Prep Kit with Ribo-Zero Human/Mouse/Rat Set A (RS-122-2201, Illumina) according to the manufacturer's instructions and sequenced using NextSeq 500 platform (75 bp, single-end, Illumina). Fastq files were assessed for quality using the FastQC (version 0.11.2) program and aligned to the mouse genome assembly NCBI37/mm9 using STAR (version 2.6.0c) (Dobin et al., 2012). Unmapped reads, non-primary alignments and alignments with a MAPQ score of <10 were filtered with SAMtools.

Gene ontology

DAVID Gene Ontology (GO) (Huang et al., 2009) v6.8 was used for Gene Ontology (Biological Process) analysis and only terms with $P<0.05$ and false discovery rate (FDR)<0.05 were considered for enrichment.

Differential expression analysis

Rsubread (Liao et al., 2013) package's featureCounts function was used to count the number of uniquely mapping reads per gene using Ensembl v.67 mm9 gene annotation file. For differential gene expression analysis, genes with <1 count per million (CPM) in less than 25% of samples were removed from differential expression analysis. Limma (Ritchie et al., 2015) voom function was used to log transform and quantile normalize the CPM matrix. Limma was used with a linear model to perform differential gene expression analysis with adjusted *P*-values for multiple comparisons using the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995). Cutoffs of FDR≤0.05 and fold change ≥1.5 or ≤−1.5 were used for significantly upregulated or downregulated genes. Genes that had at least 75% overlap with H3K9me2 domains were considered to be found within domains and were used for downstream analysis. The Mann–Whitney *U*-test was performed in R. Heatmaps were generated using R.

Volcano plots

Volcano plots of protein coding genes that are associated with RA-specific loss or DMSO-specific loss of H3K9me2 domains were visualized using R's ggplot2 (version 2.2.1) package and show upregulated (red), downregulated (blue) and not differentially expressed (gray) genes, based on adjusted *P*-value ≤0.05 and log2-fold change thresholds of ±1.5, respectively. Fold changes of genes used in Fig. 4A-D are found in Tables S8-S11.

Boxplots

To analyze gene expression of all genes associated with H3K9me2 domains, undifferentiated and D3 RNA-seq count data were normalized similar to the undifferentiated, D3 and D6 RNA-seq differential expression analysis except lowly expressed genes were not removed. All non-protein coding genes were removed. Gene lengths were calculated using GTFtools (Li, 2018 preprint), outputting the maximum transcript length for each gene. Reads per kilobase of transcript per million mapped reads values were then generated by normalizing CPM values to gene length. Using the unique and shared H3K9me2 domain gene lists for each treatment, R (version 3.4.3) and ggplot2 (version 2.2.1) were used to construct boxplots of gene expression package.

Immunofluorescence and DNA FISH

P19 cells were grown on glass coverslips (#1, 22×22 mm, 102222, Thermo Fisher Scientific) and fixed in 2% paraformaldehyde (PFA) for 10 min at rtp and washed thrice in PBS. Cells were permeabilized in 0.5% Triton X-100 for 10 min at rtp, washed thrice in PBS and incubated in block solution (1% BSA, PBS, 0.1% Triton X-100, 0.05% Tween 20) for 1 h at rtp. Cells were incubated with primary antibody for 1 h, washed thrice in 1× PBST (PBS, 0.05% Tween 20) and incubated with secondary antibody for 1 h at rtp. Cells were washed in twice in PBS for 5 min each and counterstained with DAPI (D9542, Sigma-Aldrich) for 10 min at rtp. Cells were washed in PBS twice before mounting on glass slides with mounting medium (Vectashield). Primary antibodies used were anti-LaminB1 (1:500, sc-374015, Santa Cruz Biotechnology), anti-LaminB1 (1:500, sc-6216, Santa Cruz Biotechnology), anti-H3K9me2 (1:500, 39239, Active Motif), anti-Oct4 (1:500, ab19857, Abcam) and secondary antibodies used include donkey anti-rabbit Alexa 568 (1:500, A10042, Invitrogen), donkey anti-mouse Alexa 568 (1:500, A10037, Invitrogen), donkey anti-goat Alexa 488 (1:500, A11055, Invitrogen) or goat anti-rabbit Alexa 568 (1:500, A11011, Invitrogen). Cells were either imaged at this stage or processed further for DNA FISH, in which case cells were post-fixed in 2% PFA for 10 min at rtp and permeabilized with ice cold 0.7% Triton X-100/0.1 M HCl for 10 min. Cells were denatured at 75°C for 30 min in 50% formamide/2× saline-sodium citrate buffer (SSC) and transferred to ice cold 2× SSC, before hybridization with labeled probe. Probes were labeled with green dUTP using Nick Translation Kit (32-801300, Abbott Laboratories) and denatured for 30 min at 37°C and 10 min at 75°C immediately before hybridization.

BAC DNA clones (Thermo Fisher Scientific) used were *Mef2c* (RP23-187H18) and *Myocd* (RP23-379F14). The immuno-FISH protocol contains a denaturation step of 75°C for 30 min, which is essential to allow the cellular DNA to be denatured before hybridization with the DNA FISH probes. This time-critical step is carefully administered and we do not observe fragmented nuclei or compromised nuclear integrity after this step. The immunostaining in some cells may be slightly disrupted after the denaturation step. If the cells displayed compromised nuclear integrity, they were excluded from analysis.

Flow cytometry
Cells were counted and aliquoted at $5×10^5$ per fluorescence-activated cell sorting (FACS) tube. Cells were washed once in FACS wash buffer (2% FBS in PBS), collected as a pellet by centrifugation at $500\,\boldsymbol{g}$ for 10 min and resuspended in 250 µl of Fixation/Permeabilization solution (BD Biosciences) to incubate for 20 min at 4°C. Cells were washed once in 1× Perm/Wash Buffer (BD Biosciences) and collected as a pellet by centrifugation ($500\,\boldsymbol{g}$ for 10 min). Anti-Oct3/4 antibody conjugated to PE (560186, BD Pharmingen) was prepared in 20 µl of 1× Perm/Wash buffer and the cell pellet was resuspended in antibody solution to incubate for 30 min at 4°C in the dark. Cells were washed twice in 1× Perm/Wash buffer and resuspended in 200 µl 1× Perm/Wash buffer for flow cytometry (BD Accuri). Negative controls included non-stained cells and mouse isotype IgG1$_k$ conjugated to PE (555749, BD Pharmingen)-stained cells.

Image analysis and quantification
All confocal images were taken on laser scanning confocal microscope Leica 3X STED from the University of Pennsylvania Cell and Developmental Biology (CDB) Microscopy Core. Confocal 3D images were taken using a *z*-step size of 0.04 µm with a range of 80-200 *z*-slices per cell. Images were deconvoluted using Huygens Professional (Scientific Volume Imaging) and 3D reconstructions were performed on Imaris software 7.4 (Bitplane). Measurements of 3D distance from the center of each DNA FISH spot to the nuclear lamina surface are as previously described (Poleshko et al., 2017). In brief, the DNA FISH dots are generated using the spots tool at its intensity center of mass with a 300 nm diameter. The lamin surface and DAPI surface are generated using the surface tool and generate a 3D spheroid structure based on the entirety of confocal *z*-stacks that have a resolution of 0.04 µm (*x*,*y*,*z*). As the cells are allowed to attach to the glass coverslip, the nucleus is not a perfect sphere but rather more spheroid in shape. Spots that are embedded within the nuclear lamina return negative distances. For calculating the percentages of cells with 0, 1 or 2 alleles in Fig. 3G,J, only cells in which both alleles were visible by immuno-FISH were included for quantification. All statistical testing for one-way analysis of variance (ANOVA) with post hoc Tukey tests were performed using Prism 5 (Graphpad).

References
Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* **57**, 289-300.

Creemers, E. E., Sutherland, L. B., McAnally, J., Richardson, J. A. and Olson, E. N. (2006). Myocardin is a direct transcriptional target of Mef2, Tead and Foxo proteins during cardiovascular development. *Development* **133**, 4245-4256.

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. (2017). The 4D nucleome project. *Nature* **549**, 219-226.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2012). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21.

Edwards, M. K., Harris, J. F. and McBurney, M. W. (1983). Induced muscle differentiation in an embryonal carcinoma cell line. *Mol. Cell. Biol.* **3**, 2280-2286.

Gonzalez-Sandoval, A. and Gasser, S. M. (2016). On TADs and LADs: spatial control over gene expression. *Trends Genet.* **32**, 485-495.

Gonzalez-Sandoval, A., Towbin, B. D., Kalck, V., Cabianca, D. S., Gaidatzis, D., Hauer, M. H., Geng, L., Wang, L., Yang, T., Wang, X. et al. (2015). Perinuclear anchoring of H3K9-methylated chromatin stabilizes induced cell fate in C. elegans embryos. *Cell* **163**, 1333-1347.

Grépin, C., Nemer, G. and Nemer, M. (1997). Enhanced cardiogenesis in embryonic stem cells overexpressing the GATA-4 transcription factor. *Development* **124**, 2387-2395.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M. B., Talhout, W., Eussen, B. H., de Klein, A., Wessels, L., de Laat, W. et al. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**, 948-951.

Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44-57.

Jasmin, R. A., Spray, D. C., Campos de Carvalho, A. C. and Mendez-Otero, R. (2010). Chemical induction of cardiac differentiation in p19 embryonal carcinoma stem cells. *Stem Cells Dev.* **19**, 403-412.

Jones-Villeneuve, E. M. (1982). Retinoic acid induces embryonal carcinoma cells to differentiate into neurons and glial cells. *J. Cell Biol.* **94**, 253-262.

Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S. S., Janssen, H., Amendola, M., Nolen, L. D., Bickmore, W. A. and van Steensel, B. (2013). Single-cell dynamics of genome-nuclear lamina interactions. *Cell* **153**, 178-192.

Kohwi, M., Lupton, J. R., Lai, S.-L., Miller, M. R. and Doe, C. Q. (2013). Developmentally regulated subnuclear genome reorganization restricts neural progenitor competence in Drosophila. *Cell* **152**, 97-108.

Leifer, D., Krainc, D., Yu, Y. T., McDermott, J., Breitbart, R. E., Heng, J., Neve, R. L., Kosofsky, B., Nadal-Ginard, B. and Lipton, S. A. (1993). MEF2C, a MADS/MEF2-family transcription factor expressed in a laminar distribution in cerebral cortex. *Proc. Natl. Acad. Sci. USA* **90**, 1546-1550.

Li, H. (2018). GTFtools: a Python package for analyzing various modes of gene models. *bioRxiv*.

Li, S., Wang, D.-Z., Wang, Z., Richardson, J. A. and Olson, E. N. (2003). The serum response factor coactivator myocardin is required for vascular smooth muscle development. *Proc. Natl. Acad. Sci. USA* **100**, 9366-9370.

Liao, Y., Smyth, G. K. and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **41**, e108-e108.

Lin, Q., Schwarz, J., Bucana, C. and Olson, E. N. (1997). Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. *Science* **276**, 1404-1407.

Lund, E., Oldenburg, A. R. and Collas, P. (2014). Enriched domain detector: a program for detection of wide genomic enrichment domains robust against local variations. *Nucleic Acids Res.* **42**, e92.

Martins, A. H. B., Resende, R. R., Majumder, P., Faria, M., Casarini, D. E., Tárnok, A., Colli, W., Pesquero, J. B. and Ulrich, H. (2005). Neuronal differentiation of P19 embryonal carcinoma cells modulates kinin B2 receptor gene expression and function. *J. Biol. Chem.* **280**, 19576-19586.

Meuleman, W., Peric-Hupkes, D., Kind, J., Beaudry, J.-B., Pagie, L., Kellis, M., Reinders, M., Wessels, L. and van Steensel, B. (2013). Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res.* **23**, 270-280.

Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S. W. M., Solovei, I., Brugman, W., Gräf, S., Flicek, P., Kerkhoven, R. M., van Lohuizen, M. et al. (2010). Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* **38**, 603-613.

DEVELOPMENT

Poleshko, A., Shah, P. P., Gupta, M., Babu, A., Morley, M. P., Manderfield, L. J., Ifkovits, J. L., Calderon, D., Aghajanian, H., Sierra-Pagán, J. E. et al. (2017). Genome-nuclear lamina interactions regulate cardiac stem cell lineage restriction. *Cell* **171**, 573-580.e14.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W. and Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47-e47.

Robson, M. I., de Las Heras, J. I., Czapiewski, R., Lê Thành, P., Booth, D. G., Kelly, D. A., Webb, S., Kerr, A. R. W. and Schirmer, E. C. (2016). Tissue-specific gene repositioning by muscle nuclear membrane proteins enhances repression of critical developmental genes during myogenesis. *Mol. Cell* **62**, 834-847.

Rossant, J. and McBurney, M. W. (1982). The developmental potential of a euploid male teratocarcinoma cell line after blastocyst injection. *J. Embryol. Exp. Morphol.* **70**, 99-112.

Skerjanc, I. S., Petropoulos, H., Ridgeway, A. G. and Wilton, S. (1998). Myocyte enhancer factor 2C and Nkx2-5 up-regulate each other's expression and initiate cardiomyogenesis in P19 cells. *J. Biol. Chem.* **273**, 34904-34910.

Therizols, P., Illingworth, R. S., Courilleau, C., Boyle, S., Wood, A. J. and Bickmore, W. A. (2014). Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. *Science* **346**, 1238-1242.

Towbin, B. D., González-Aguilera, C., Sack, R., Gaidatzis, D., Kalck, V., Meister, P., Askjaer, P. and Gasser, S. M. (2012). Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell* **150**, 934-947.

Ueyama, T., Kasahara, H., Ishiwata, T., Nie, Q. and Izumo, S. (2003). Myocardin expression is regulated by Nkx2.5, and its function is required for cardiomyogenesis. *Mol. Cell. Biol.* **23**, 9222-9232.

van Steensel, B. and Belmont, A. S. (2017). Lamina-associated domains: links with chromosome architecture, heterochromatin, and gene repression. *Cell* **169**, 780-791.

Wang, D.-Z., Chang, P. S., Wang, Z., Sutherland, L., Richardson, J. A., Small, E., Krieg, P. A. and Olson, E. N. (2001). Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor. *Cell* **105**, 851-862.

Yokochi, T., Poduch, K., Ryba, T., Lu, J., Hiratani, I., Tachibana, M., Shinkai, Y. and Gilbert, D. M. (2009). G9a selectively represses a class of late-replicating genes at the nuclear periphery. *Proc. Natl Acad. Sci. USA* **106**, 19363-19368.

DEVELOPMENT