MyoD is a 3D genome structure organizer for muscle cell identity

Ruiting Wang\textsuperscript{1,12}, Fengling Chen\textsuperscript{2,12}, Qian Chen\textsuperscript{1,12}, Xin Wan\textsuperscript{1,12}, Minglei Shi\textsuperscript{3}, Antony K. Chen\textsuperscript{4}, Zhao Ma\textsuperscript{4,5}, Guohong Li\textsuperscript{6,7}, Min Wang\textsuperscript{6,7}, Yachen Ying\textsuperscript{4,5}, Qinyao Liu\textsuperscript{1}, Hu Li\textsuperscript{1,8}, Xu Zhang\textsuperscript{9}, Jinbiao Ma\textsuperscript{10}, Jiayun Zhong\textsuperscript{10}, Meihong Chen\textsuperscript{1}, Michael Q. Zhang\textsuperscript{2,3,11}, Yong Zhang\textsuperscript{1,8}*, Yang Chen\textsuperscript{1,2,3}*, & Dahai Zhu\textsuperscript{1,8}*

The genome exists as an organized, three-dimensional (3D) dynamic architecture, and each cell type has a unique 3D genome organization that determines its cell identity. An unresolved question is how cell type-specific 3D genome structures are established during development. Here, we analyzed 3D genome structures in muscle cells from mice lacking the muscle lineage transcription factor (TF), MyoD, versus wild-type mice. We show that MyoD functions as a “genome organizer” that specifies 3D genome architecture unique to muscle cell development, and that H3K27ac is insufficient for the establishment of MyoD-induced chromatin loops in muscle cells. Moreover, we present evidence that other cell lineage-specific TFs might also exert functional roles in orchestrating lineage-specific 3D genome organization during development.
Recent Hi-C analysis of comprehensive interaction maps over large regions or whole genomes has indicated that the genome is hierarchically organized into chromosome territories, A/B compartments, topologically associated domains (TADs), and chromatin loops. Many studies have shown that the 3D structure of the genome differs widely among different cell types. However, the molecular mechanisms underlying the establishment of cell-type-specific 3D genome organization are largely unknown. It has been proposed that “master transcription factors” (TFs) act as anchor proteins to orchestrate cell-type-specific 3D genome architecture. For example, studies have shown that KLF4 is involved in the organization and regulation of pluripotency-associated three-dimensional enhancer networks, and LHX2- and LDB1-mediated trans-interactions regulate olfactory receptor choice, and Pax3 cooperates with Ldb1 to direct local chromosome architecture during myogenic-lineage specification.

MyoD (myogenic differentiation 1) and Myf5 are members of the myogenic regulatory factor (MRF) family and are expressed in somites. These proteins function as a basic helix-loop-helix transcription factors and are required for myogenic determination during early embryogenesis. This was initially evidenced by the complete blockade of skeletal muscle formation in MyoD and Myf5 double-mutant mice. Later studies demonstrated that the phenotype observed in Myod-Myf5 double-knockout (KO) mice was caused by impaired expression of Mrf4, Myf5, and Myf5 synergistically regulate the expression levels of myogenic genes genome-wide by binding their consensus E-box sequences.

Previous work and our present data indicate that Myf5 expression is significantly upregulated in MyoD-null cells. The single inactivation of either MyoD or Myf5 in mice results in apparently normal muscle development but delayed myogenic differentiation. Moreover, we detected a clear A/B compartment shift in proliferating myocytes (Fig. 1c). Moreover, we also performed ChIP-seq for CTCF (CCCTC-binding factor) and H3K27ac (histone H3 lysine 27 acetylation) (Fig. 1b).

We then examined the impacts of MyoD depletion on chromatin A/B compartments in muscle cells by performing the exploratory principal component analysis (PCA) based on the compartment values for the four muscle sample types examined. We observed clear distinctions between the WT and MKO genotypes for both proliferating and differentiating muscle cells (Fig. 1c). Moreover, we detected a clear A/B compartment shift in proliferating MKO cells vs. WT cells (1.6% B→A, 0.84% A→B) (Fig. 1d left and 1e left). For differentiating cells, there was a clear compartment switch for MKO cells compared to differentiating WT cells (1.38% B→A, 1.28% A→B) (Fig. 1d left and 1e right).

We also analyzed the dynamics of chromatin contact domain boundaries (CDBs) in muscle cells in response to MyoD knockout (Fig. 1c right). A total of 931 CDBs exhibited differential extents of insulation in MKO cells compared to WT cells in the proliferating and differentiating states (Fig. 1d right and f), implying that MyoD may somehow facilitate CDB insulation. Pursuing this, we compared insulation scores for CDBs bound by MyoD or CTCF. In general, the insulation scores of MyoD-bound CDBs were similar to those of CTCF-bound CDBs (Fig. 1g). However, the CDBs bound solely by MyoD exhibited significantly weaker insulation compared to the CTCF-bound CDBs. This indicates that MyoD may work in concert with CTCF to regulate the insulation of CDBs in muscle cells.

MyoD regulates A/B compartment switching and the formation of contact domain boundaries (CDBs) in muscle cells. In light of the aforementioned MyoD data, we first performed RNA-seq analysis of muscle cells that were isolated from hind-limb skeletal muscles of wild-type (WT) and MyoD-null (MKO) mice and cultured in growth medium (GM, Pax7+, and MyoD+) or differentiation medium (DM, MyoG+) (Supplementary Fig. 1b). We then combined our results with existing MyoD ChIP-seq data to examine potential functional link(s) between MyoD occupancy and transcriptional activation. We extracted MyoD-binding peaks localized at promoter regions and identified their cognate genes within ±3 kb (Fig. 1a). Intriguingly, we found that knockout of MyoD did not affect the expression of a majority of these genes (2691/3510 (76.7%) in primary proliferating myoblasts, 1987/3509 (56.6%) in differentiating myocytes) (Fig. 1a and Supplementary Data 1), suggesting that the presence of MyoD at these genes (even in their promoter regions) frequently does not directly impact their transcriptional activation. This result further supports the hypothesis that, beyond its canonical function as an activator of myogenic gene expression during development, MyoD may exert broad functional impacts, potentially on higher-order genome structures.

To directly investigate the potential genome architectural roles of MyoD in muscle cells, we examined muscle samples from four groups (WT-GM, WT-DM, MKO-GM, and MKO-DM) using bridge-linker Hi-C (BL-Hi-C), which is a modified in situ Hi-C with improved sensitivity and specificity for structural and regulatory chromatin loop detection. In the BL-Hi-C experiments, 28 libraries were sequenced to a total depth of over 12 billion reads. The high-quality deep-sequencing data were validated with a high cis-interaction rate (Supplementary Fig. 1c) and reached 5 kb resolution (Supplementary Fig. 1d), which ensured the rigor of our subsequent computational analyses for each hierarchical level of chromatin structure in muscle cells. As an integrative analysis of the function of MyoD in orchestrating 3D genome organization, we also performed ChIP-seq for CTCF (CCCTC-binding factor) and H3K27ac (histone H3 lysine 27 acetylation) (Fig. 1b).

In this study, we explored the cell-lineage-specific master transcription factor-mediated specification of developmental-context-specific 3D genome organization, with a focus on MyoD. We chose MyoD for the following reasons: First, MyoD has been well-established as a master TF in myogenic cell-lineage specification during development and trans-differentiation. It is known to regulate the expression of myogenesis-specific genes through binding to consensus E-box (CAGNTG)-containing cis-regulatory elements, which number more than 14 million across the genome. Second, over 40,000 MyoD-binding peaks have been identified by independent ChIP-seq studies in muscle cells; of them, only 15% are located in promoter regions (Supplementary Fig. 1a). Finally, in addition to its functions in binding and trans-activating genes during differentiation, MyoD also constitutively binds at tens of thousands of additional sites throughout the genome in proliferating muscle stem cells (Supplementary Fig. 1a). Together, these data emphasize the attractively large empirical scope represented by genome-wide functional analysis of the potential roles of MyoD in muscle-cell-lineage-specific 3D genome organization. In this report, we provide computational and experimental evidence that uncovers a previously unappreciated role of MyoD as a genome organzer that contributes to establishing the unique 3D genome architecture in muscle cells, beyond its well-known functions as a TF in activating myogenic gene expression during development.

Results
MyoD regulates A/B compartment switching and the formation of contact domain boundaries (CDBs) in muscle cells. MyoD is an anchor protein for chromatin loop formation in muscle cells. The most well-characterized anchor protein in vertebrates is CTCF, which mediates chromatin looping in concert with cohesin, probably through a loop-extrusion mechanism. Considering our finding that MyoD apparently functions together with CTCF to regulate the insulation of CDBs, we questioned whether MyoD might control the formation of chromatin loops in...
Fig. 1 MyoD regulates A/B compartment switching and contact domain boundary (CDB) formation in muscle cells. a Association of MyoD-binding peaks at promoter regions and related gene expression. Heatmaps show the MyoD-binding peaks at promoters (centralized at the TSS within 3 kb distance) and related gene expression trends in WT and MKO myoblasts (left panel). Corresponding data for WT and MKO myocytes is shown in the right panel.

b Strategy for genetic investigation of the chromatin architectural roles of MyoD.

c Principal Component Analysis (PCA) of the values of compartments or the relative insulation scores of contact domain boundaries (CDBs) among the four indicated sample types. Each dot represents an individual biological replicate.

d Bar plots showing the percentage of A/B compartment shifts between WT and MKO myoblasts and myocytes (left panel). Bar plots showing the number of relative insulation-enhanced or -reduced CDBs between WT and MKO in proliferating and differentiating muscle stem cells (right panel).

e Heatmaps of the A/B compartment-shifted regions between WT and MKO myoblasts and myocytes.

f Hi-C aggregation plots centered at differential CDBs between WT and MKO myoblasts and myocytes.

g Probability density of insulation score at genetic regions bound by CTCF or MyoD. Four classes were assessed: CTCF-bound, MyoD-bound; MyoD-bound; and CTCF-unbound within a 20 kb distance; and MyoD-bound and CTCF-unbound within a 40 kb distance.
Fig. 2 MyoD is an anchor protein for chromatin loop formation in muscle cells. a Aggregate Peak Analysis (APA) plots showing the aggregated Hi-C contacts around chromatin loops identified in WT-GM and WT-DM cells. n represents the total number of chromatin loops called by HiCCUPS (see Methods). The APA score at the upper left corner of each aggregation plot was calculated as the ratio of Hi-C contacts at the central pixel to the mean Hi-C contacts of the lower-left pixels. The Hi-C contacts at the central pixel are shown at upper right corner of each aggregation plot. b Enriched motifs in accessible regions at the anchors of the chromatin loops identified in panel a. P-values of motif enrichment were determined using cumulative binomial distributions, as applied by Homer. c Percentage of chromatin loops with anchors bound by MyoD in WT-GM and WT-DM cells. MyoD is indicated as being bound at both anchors (Both, red), one of the two anchors (Xor, blue), or neither of the two anchors (Neither, green). Pseudo-peaks of MyoD binding positions were cataloged all of the identified chromatin loops into four types (see Supplementary Fig. 2c and Supplementary Data 3). We found that 44.18% of chromatin loops in GM cells were MyoD-bound (10.87% with both anchors bound by MyoD, 33.31% with one anchor bound by MyoD), while 43.93% of chromatin loops in DM cells were MyoD-bound (10.75% with both anchors bound by MyoD, 33.31% with one anchor bound by MyoD). Furthermore, the MyoD-binding peaks at loop anchors were concordant with the binding signals of CTCF and the cohesin complex subunit, SMC3 (Supplementary Fig. 2d). We cataloged all of the identified chromatin loops into four types based on the binding of MyoD and/or CTCF at loop anchors in muscle cells: MyoD-MyoD(noCTCF), MyoD-MyoD(CTCF), MyoD-CTCF, and CTCF-CTCF (Fig. 2d, e). This classification was further supported by our ability to observe both colocalization of MyoD/CTCF and non-overlaid MyoD or CTCF signals under direct stochastic optical reconstruction microscopy.
MyoD orchestrates chromatin loop formation in muscle cells. Next, we examined whether MyoD directly mediates the formation of DNA looping in vitro and in cells. First, we performed an in vitro DNA circularization assay using a purified recombinant MyoD protein that was verified to have DNA-binding activity (Supplementary Fig. 3a) and a 3.8 kb linear DNA fragment containing five MyoD-binding sites at each end (Fig. 3a). The in vitro MyoD-mediated DNA looping was visualized by TEM. We found that MyoD did indeed facilitate linear DNA circularization in vitro in a MyoD-binding-motif-dependent manner (Fig. 3a and Supplementary Fig. 3b), indicating that MyoD can directly mediate DNA looping in vitro.

MyoD-instructed chromatin looping was further corroborated by deleting the MyoD-binding peaks from a MyoD-bound loop anchor. As shown in Fig. 3b, the MyoG-Mybph loop containing four E-boxes in the MyoG promoter and seven E-boxes in the Mybph promoter was a MyoD-mediated loop (Fig. 3b). We took this MyoG-Mybph locus as an example to test whether the E-boxes at the MyoD-loop anchor were required for MyoD-mediated looping in cells. To this end, we deleted about 410 bp containing four E-boxes of the MyoG promoter from the genome using the CRISPR strategy (Fig. 3c and Supplementary Fig. 3c). We then examined the formation of the MyoD-instructed loop at this mutant locus by Tn5-mediated fluorescence in situ hybridization (Tn5-FISH) (Fig. 3d). Consistent with the in vitro data, the MyoD-instructed loop was not detected at the E-box-deleted locus during muscle cell differentiation (Fig. 3d). Moreover, the expression of the cognate gene, Mybph, was significantly decreased in these cells (Supplementary Fig. 3d).

Finally, we took advantage of our observation that the MyoD-mediated MyoG-Mybph loop was lost in MyoD-null cells (Fig. 3b), and tested whether this lost MyoD-instructed loop could be re-established by reintroducing MyoD into MKO cells (Supplementary Fig. 3e). We visualized chromatin looping at the MyoG and Mybph locus by dCas9-mediated DNA imaging (Fig. 3e) and found that re-expression of MyoD (Ad-MyoD) was sufficient to re-establish the diminished loop at the MyoG and Mybph locus in MKO cells (Fig. 3e). This rescued loop formation efficiently induced transcription of MyoG and Mybph, as evidenced by the detection of nascent RNA products (Supplementary Fig. 3f).

Collectively, these genetic, biochemical, and cellular imaging results demonstrate that MyoD functions as an anchor protein that orchestrates chromatin loop formation in muscle cells.

MyoD anchors the formation of myogenesis-specific chromatin loops. Based on the above findings, we next asked whether MyoD can mediate the organization of muscle-cell-specific chromatin looping. To test this, Hi-C data from embryonic stem (ES) and cortical neuron (CN) cells were collected and examined for the four types of chromatin loops identified in muscle cells. Relatively few MyoD-MyoD(noCTCF) loops were identified in ES cells (8 loops, 1.8%) and CN cells (1 loop, 0.2%) (Fig. 4a), revealing that the MyoD-MyoD(noCTCF) loops detected in muscle cells represent myogenic-lineage-specific chromatin loops. When we compared the expression tendencies of genes associated with the four loop types in WT-DM cells versus other cell types collected in the ENCODE database (see Methods), we found that the genes harboring MyoD-MyoD (noCTCF) loops exhibited the strongest extent of muscle-lineage-specific expression (Fig. 4b). Accordantly, the genes associated with MyoD-MyoD (noCTCF) chromatin loops were the most highly enriched for GO terms related to muscle cell differentiation (Supplementary Fig. 4).

To seek further molecular insights into MyoD-bound muscle-cell-specific chromatin looping, we examined ChIP-seq data for both MyoD and NeuroD2, which is a neuron-specific TF responsible for neuronal cell-lineage commitment and differentiation, in both muscle and neuron cells. Only 18% of MyoD-binding peaks (34,881 pseudo-peaks) in muscle cells overlapped with NeuroD2-binding peaks (51,803 peaks) in cortical neuron cells, suggesting that MyoD and NeuroD2 might mediate the formation of separate sets of chromatin loops that can subsequently regulate distinct sets of target genes to properly specify lineage-specific cell identities. In support of this, we observed that the Myh gene cluster, which is located within a MyoD-bound muscle-cell-specific chromatin loop, was not detected in neuronal cells, which also lacked NeuroD2 binding at the Foxg1 locus were found to be neuron-specific, and were not observed in muscle cells (Fig. 4d). Taken together, these results illustrate that cell-lineage-specific TFs may play genome-wide architectural roles in modulating 3D genome organization, and thus have a considerably larger functional impact than that held by the traditional concept that TFs affect lineage specification primarily by activating or regulating gene transcription.

MyoD regulates primed architectural chromatin loops in proliferating muscle cells. Given previously published studies showing that MyoD constitutively binds to tens of thousands of sites genome-wide in proliferating cells, together with our present findings that such binding only infrequently results in transcriptional activation at this stage and MyoD functions as an anchor protein for lineage-specific 3D genome organization, we sought to further explore the nature of MyoD’s genome-level architectural functions by reassessing our BL-Hi-C data. We were particularly interested in examining how the presence of MyoD influences internal interactions occurring within chromatin loops.

Previous studies indicated that CTCF and cohesin function together to establish architectural loops and that the extent of enhancer-promoter (E-P) interactions is elevated within such loops. As our results revealed a notable colocalization of MyoD-binding peaks with CTCF (Supplementary Fig. 2d), we speculated that MyoD-bound loops might also feature increased intraloop interactions. To test this, we measured the internal interactions for each loop by calculating the “domain score” (D-score) in
WT and MKO proliferating cells. MyoD binding was found to be enriched on the anchors of chromatin loops with reduced internal interactions (65%) in MKO-GM cells compared with WT-GM cells (Fig. 5a, b), indicating that MyoD regulates the internal interactions within chromatin loops by its direct binding on loop anchors, as also seen for CTCF and cohesin. Further analysis showed that interactions changes within chromatin loops were positively correlated with expression changes of the genes within those loops (Fig. 5c).

We next focused on analyzing the internal interaction changes of MyoD-bound loops and the expression changes of their cognate genes (Fig. 5d). We specifically examined MyoD-bound loops that displayed reduced interactions in MKO proliferating cells, because we hypothesized that these loops might be directly regulated by MyoD binding. We found that nearly 50% of the examined loops enclosed downregulated genes, including those exclusively regulated by MyoD in myoblasts and commonly regulated by MyoD in myoblasts and myocytes (Fig. 5d and...
Supplementary Fig. 5a–c). These genes were enriched for known myogenesis-related functions, indicating that interactions within these MyoD-bound loops are necessary for the regulation of gene expression related to muscle identity (Fig. 5d). For example, Rbm24 is known to regulate myogenesis, and its differential expression in WT-GM and MKO-GM cells offers an excellent illustration of a gene that is enclosed within this type of MyoD-bound loop in muscle cells (Fig. 5e and Supplementary Fig. 5c, d).

Intriguingly, most (89%) of the remaining MyoD-bound loops with decreased interactions did not contain any gene that showed differential expression upon MyoD knockout (Fig. 5d). This finding underscores the apparent centrality of the architectural role (rather than the transcriptional activation role) of MyoD for these MyoD-bound loops in proliferating muscle cells. Although reduced interactions within these chromatin loops in MKO cells did not alter gene expression in proliferating cells, we found that genes within these chromatin loops were not turned on or off properly in MKO-DM cells compared to WT-DM cells (Fig. 5f), and this genetic dysregulation was clearly manifest in the defective differentiation phenotype of the MKO cells (Supplementary Fig. 1b)20–22. As the loops were still maintained in differentiated cells (Supplementary Fig. 5e), it appears that the interactions constrained within the loops in proliferating cells are effectively “primed” for muscle cell differentiation by the presence of MyoD,

**Fig. 4** MyoD anchors the formation of myogenesis-specific chromatin loops. a Proportions of muscle chromatin loop subsets present in embryonic stem (ES) or cortical neurons (CN) cells. The numbers of muscle chromatin loops found in ES or CN cells are indicated for each bar. b Boxplots represent the expression Z-scores obtained in WT-DM cells for genes associated with each class of loops. The Z-score of each gene in WT-DM cells was calculated versus those in other ENCODE-collected cell types (ES, spleen, B, T, megakaryocyte, neural progenitor, and CN cells). The loop-related gene numbers are marked in red. P-values were calculated by the one-sided Wilcoxon rank-sum test. Source data are provided as a Source Data file. c Representative region showing muscle-cell-specific loops anchored on MyoD-binding regions at the Myh gene cluster. d Representative region showing neuronal cell-specific loops anchored on NeuroD2-binding regions at the Foxg1 locus.
raising the possibility that some very early signal can somehow direct MyoD to occupy sites across the genome that subsequently permit both muscle-cell-appropriate loop architecture and later rewiring in response to differentiation signals. Supporting this idea, we found that \textit{Lrtm1}, a myogenesis-related gene\textsuperscript{49} that is expressed only in differentiated muscle cells (Supplementary Fig. 5f), contains chromatin loops that have already been defined by MyoD in proliferating cells (Fig. 5g). Thus, mechanistically, our analyses of the MyoD-bound loops and their internal interactions reveal molecular insights into how MyoD exerts its architectural function, even in proliferating cells, to ensure the correct trajectory towards eventual muscle cell identity.

The regulatory loops specified by MyoD are functionally required for muscle cell differentiation. The genome is not only structurally organized within the nucleus but is also dynamically orchestrated in response to various cellular signals\textsuperscript{5,47,50–53}. The
biological function of MyoD in regulating muscle cell differentiation has been well documented, and very recent work showed that MyoD’s transcriptional activity is regulated by its ability to mediate chromatin interactions during trans-differentiation. However, it remains largely unknown how MyoD impacts chromatin conformation dynamics during muscle cell differentiation. Using a modified differential loop detection method (see Methods), we identified 6242 differential chromatin loops (~25%) between proliferating and differentiating muscle cells. Among these loops, 5754 were significantly enhanced whereas 488 were reduced in differentiating cells (Fig. 6a). We detected significant enrichment for muscle cell differentiation-related genes among those associated with the enhanced chromatin loops (Supplementary Fig. 6a). These data also revealed that nearly 25% of chromatin loops that undergo dynamic changes during muscle cell differentiation contribute to 3D genome rewiring in a manner that functionally regulates myogenic differentiation.

A similar analysis performed on MKO cells identified 585 differential chromatin loops between WT-DM and MKO-DM cells. Consistent with the proposed regulatory role of these loops, 81.5% (477/585) were significantly reduced in the differentiated MKO cells (Supplementary Fig. 6b). Moreover, 71.57% (341/477) of the loops reduced in differentiated MKO cells had MyoD-binding peaks, supporting them as bona fide MyoD-mediated chromatin loops in differentiated WT cells (Supplementary Fig. 6c). Among the loops found to be enhanced under differentiation, 285 loops were bound by MyoD and disappeared when MyoD was absent from differentiating muscle cells (Supplementary Data 4), suggesting that they might be directly induced by MyoD. The genes associated with these MyoD-induced loops were significantly upregulated during differentiation and downregulated when MyoD was absent, verifying the regulatory role of MyoD-induced loops (Fig. 6b). In addition, acute depletion of MyoD with a Cre-mediated knockout system (MyoDΔfl, Cre) yielded similar results: MyoD depletion significantly decreased the strength of these loops and the expression levels of their cognate genes during muscle cell differentiation (Supplementary Fig. 6d–l). Together, these findings show that MyoD-mediated loops are required for muscle cell differentiation.

Next, we sought to understand what drives the enhancement of MyoD-induced loops during muscle differentiation. We speculated that H3K27ac modification (Fig. 6c). Based on this observation, we speculated that H3K27ac modification might somehow actively direct the deposition of these modifications on the MyoD-induced loops.

Taking the Me2c locus as a representative MyoD-induced loop, we sought to demonstrate its regulatory role in differentiation. We found that enhancer-promoter loops linked to Me2c emerged during differentiation in a MyoD-dependent manner (Fig. 6d, Supplementary Fig. 6k). Concordantly, H3K27ac modification was enhanced in WT-DM cells but diminished in MKO-DM cells at this locus (Fig. 6d), and Me2c gene expression was not induced in MKO-DM cells (Fig. 6f upper panel and Supplementary Fig. 6k). Similar results were obtained when we examined MyoD-induced loops within the Dyrk2 locus (Fig. 6e and Supplementary Fig. 6l), from which the mRNA transcription product was only abundant in WT-DM cells (Fig. 6f lower panel and Supplementary Fig. 6l).

Collectively, these results demonstrate that MyoD orchestrates the formation of myogenic-lineage-specific chromatin loops that are required for the transcriptional regulation of myogenic genes during muscle cell differentiation.

**In MyoD-knockout cells, H3K27ac is insufficient to drive the formation of MyoD-bound chromatin loops.** The dynamic organization of the 3D genome is known to facilitate the sophisticated interplay between accessible versus inaccessible chromatin states and TF occupancy, but the interdependent impacts (and likely feedback) of chromatin status on 3D genome organization have remained largely elusive. In addition to our observation of MyoD-dependent H3K27ac modification on MyoD-induced loops, we found that during muscle cell differentiation, the loop anchors with increased H3K27ac signals showed overall enrichment for consensus E-box sequences (the known target motif of the MyoD protein) (Fig. 7a). When MyoD was knocked out, H3K27ac modification was remarkably decreased on MyoD-binding regions (Fig. 7b), further supporting our contention that overall H3K27ac modification depends on MyoD during differentiation. As MyoD itself stably binds around many myogenic genes in proliferating muscle cells, MyoD-dependent H3K27ac modification correlates even more specifically with loop formation or disruption. These results and the previously reported positive correlation between H3K27ac and loops prompted us to examine whether the H3K27ac-mediated active chromatin state

---

**Fig. 5 MyoD mediates primed architectural chromatin loops in proliferating muscle cells.** a Aggregated Hi-C maps generated around chromatin loops with enhanced, reduced, or static internal interactions (measure by domain score) between WT and MKO myoblasts (MKO-GM vs. WT-GM). b Percentage of chromatin loops with differential internal interactions (MKO-GM vs. WT-GM) that are occupied by MyoD at anchors. MyoD bound at both anchors (both, red), one of the two anchors (Xor, blue), or neither of the two anchors (Neither, green). Pseudo-peaks of MyoD binding identified from combined ChIP-seq data obtained from public datasets and our study were used for this analysis (see Methods). c Boxplot showing the corresponding expression changes of genes within chromatin loops having enhanced, reduced, or static internal interactions (domain score) identified in (Supplementary Data 4), suggesting that they might be directly associated with gene expression changes of MyoD-bound chromatin loops whose associated genes did not show significantly differential expression in MKO-GM cells compared with WT-GM cells are shown as blue dots. In contrast, internal interaction-reduced MyoD-bound chromatin loops whose associated genes showed significant down-regulation in MKO-GM cells compared with WT-GM cells are shown as red dots. d The Rbm24 locus as a representative internal interaction-reduced MyoD-bound chromatin loop associated with genes that were significantly downregulated in MKO-GM cells compared with WT-GM cells. e Heatmap showing expression (FPKM) Z-scores among the four indicated sample types for genes within interaction-reduced MyoD-bound chromatin loops whose associated genes did not show differential expression in MKO-GM cells compared with WT-GM cells. f The Lmr1 locus as a representative internal interaction-reduced MyoD-bound chromatin loop that is primed in proliferating muscle cells, comparing MKO-GM cells with WT-GM cells.

---

**Fig. 6a** MyoD mediates primed architectural chromatin loops in proliferating muscle cells. **a** Aggregated Hi-C maps generated around chromatin loops with enhanced, reduced, or static internal interactions (measure by domain score) between WT and MKO myoblasts (MKO-GM vs. WT-GM). **b** Percentage of chromatin loops with differential internal interactions (MKO-GM vs. WT-GM) that are occupied by MyoD at anchors. MyoD bound at both anchors (both, red), one of the two anchors (Xor, blue), or neither of the two anchors (Neither, green). Pseudo-peaks of MyoD binding identified from combined ChIP-seq data obtained from public datasets and our study were used for this analysis (see Methods). **c** Boxplot showing the corresponding expression changes of genes within chromatin loops having enhanced, reduced, or static internal interactions (domain score) identified in (Supplementary Data 4), suggesting that they might be directly associated with gene expression changes of MyoD-bound chromatin loops whose associated genes did not show significantly differential expression in MKO-GM cells compared with WT-GM cells are shown as blue dots. In contrast, internal interaction-reduced MyoD-bound chromatin loops whose associated genes showed significant down-regulation in MKO-GM cells compared with WT-GM cells are shown as red dots. **d** The Rbm24 locus as a representative internal interaction-reduced MyoD-bound chromatin loop associated with genes that were significantly downregulated in MKO-GM cells compared with WT-GM cells. **e** Heatmap showing expression (FPKM) Z-scores among the four indicated sample types for genes within interaction-reduced MyoD-bound chromatin loops whose associated genes did not show differential expression in MKO-GM cells compared with WT-GM cells. **f** The Lmr1 locus as a representative internal interaction-reduced MyoD-bound chromatin loop that is primed in proliferating muscle cells, comparing MKO-GM cells with WT-GM cells.

---

**Fig. 7a** MyoD binding during the differentiation (Fig. 6c). However, we found that 88% of MyoD-induced loops were decorated with MyoD-dependent H3K27ac modification (Fig. 6c). Based on this observation, we speculated that H3K27ac might contribute to loop formation during muscle cell differentiation, and that MyoD might somehow actively direct the deposition of these modifications on the MyoD-induced loops.

Taking the Me2c locus as a representative MyoD-induced loop, we sought to demonstrate its regulatory role in differentiation. We found that enhancer-promoter loops linked to Me2c emerged during differentiation in a MyoD-dependent manner (Fig. 6d, Supplementary Fig. 6k). Concordantly, H3K27ac modification was enhanced in WT-DM cells but diminished in MKO-DM cells at this locus (Fig. 6d), and Me2c gene expression was not induced in MKO-DM cells (Fig. 6f upper panel and Supplementary Fig. 6k). Similar results were obtained when we examined MyoD-induced loops within the Dyrk2 locus (Fig. 6e and Supplementary Fig. 6l), from which the mRNA transcription product was only abundant in WT-DM cells (Fig. 6f lower panel and Supplementary Fig. 6l).

Collectively, these results demonstrate that MyoD orchestrates the formation of myogenic-lineage-specific chromatin loops that are required for the transcriptional regulation of myogenic genes during muscle cell differentiation.

**In MyoD-knockout cells, H3K27ac is insufficient to drive the formation of MyoD-bound chromatin loops.** The dynamic organization of the 3D genome is known to facilitate the sophisticated interplay between accessible versus inaccessible chromatin states and TF occupancy, but the interdependent impacts (and likely feedback) of chromatin status on 3D genome organization have remained largely elusive. In addition to our observation of MyoD-dependent H3K27ac modification on MyoD-induced loops, we found that during muscle cell differentiation, the loop anchors with increased H3K27ac signals showed overall enrichment for consensus E-box sequences (the known target motif of the MyoD protein) (Fig. 7a). When MyoD was knocked out, H3K27ac modification was remarkably decreased on MyoD-binding regions (Fig. 7b), further supporting our contention that overall H3K27ac modification depends on MyoD during differentiation. As MyoD itself stably binds around many myogenic genes in proliferating muscle cells, MyoD-dependent H3K27ac modification correlates even more specifically with loop formation or disruption. These results and the previously reported positive correlation between H3K27ac and loops prompted us to examine whether the H3K27ac-mediated active chromatin state
directly triggers the formation of MyoD-bound chromatin loops during muscle cell differentiation.

We tested this possibility by treating MKO cells with the histone deacetylase inhibitor, Tacedinaline (Tac), to enrich their H3K27ac signal throughout muscle cell differentiation. We found that Tac treatment led to the accumulation of H3K27ac modification in 18,110 genomic regions, including the anchors of MyoD-induced loops (96/285) (Fig.7c). Furthermore, analysis of BL-Hi-C data from the Tac-treated MKO and DMSO-treated control cells showed that, in the absence of MyoD, the accumulation of H3K27ac per se was not sufficient to trigger the formation of new loops or enhance the signals for MyoD-bound or CTCF-CTCF chromatin loops (Fig.7d, e). Consistently, Tac treatment could not rescue the differentiation defect of MyoD-null cells (Supplementary Fig. 7a, b).

To further validate the above results at a specific MyoD-mediated loop, we specifically increased the H3K27Ac level at the Myog-Mybph locus in MyoD-null cells using the dCas9-p300 system. Toward this end, we applied sgRNAs to guide the dCas9-p300-core to each anchor of the MyoG-Mybph loop in MyoD-null cells (Supplementary Fig. 7c). The p300-facilitated acetylation of H3K27 at anchors of MyoG-Mybph loop was confirmed by ChIP-qPCR (Supplementary Fig. 7d), and chromatin loop formation at the MyoG-Mybph locus was examined by dCas9-mediated DNA imaging (Fig. 7f). We observed that the increased levels of H3K27ac mediated by dCas9-p300-core were not able to rescue MyoG-Mybph loop formation or activate transcription of MyoG or Mybph in the MyoD-null cells (Fig. 7g, h). Cells isolated from WT mice were used as a control (Fig. 7f–h).
**Fig. 7 H3K27ac enhancement in MKO cells is not sufficient to drive the formation of MyoD-bound chromatin loops.**

*a* Motif enrichment at loop-anchor H3K27ac peaks gained or lost in WT-DM cells compared with WT-GM cells. *p*-values of motif enrichment were determined using cumulative binomial distributions, as applied by Homer.

*b* Heatmaps showing H3K27ac peaks found to be enhanced and reduced on ChIP-seq analysis of MKO-DM cells versus WT-DM cells. MyoD-binding peaks at those differential H3K27ac peaks are also shown. *c* Heatmaps showing H3K27ac peaks enhanced and reduced on ChIP-seq analysis of MKO cells treated with the selective class I HDAC inhibitor, Tacedinaline (Tac), in DM for 12 h. DMSO-treated MKO cells served as a control.

*d* Loop-strength fold changes of the loops with gained, lost, static or no H3K27ac marks on their anchors between MKO cells treated with Tac or DMSO. Loop strength was defined as relative Hi-C contacts found at loop pixels compared with those at surrounding donut-shaped background pixels (see Methods). The loop numbers are marked in red. *p*-values were calculated by the one-sided Wilcoxon rank-sum test. Source data are provided as a Source Data file.

*e* APA plots showing aggregated Hi-C contacts around the four types of chromatin loops (classified in 2d) in MKO cells treated with Tac or DMSO (control). Representative dCas9-mediated DNA images visualizing the chromatin loops linking MyoG (red) and Mybph (green) in WT and MKO cells transfected with dCas9-p300core and cultured in DM for 24 h. dCas9-EGFP served as a control.

**f** The distance between the two anchors visualized in e. The scale bar represents 5 µm. Data are presented as mean ± SEM. *p*-values were determined by ANOVA followed by Tukey’s multiple comparison test, ***p < 0.001.** Source data are provided as a Source Data file.

**g** Relative expression of MyoG and Mybph in the cells described in e, as determined by qRT-PCR. Data are presented as mean ± SEM. *p*-values were determined by ANOVA followed by Tukey’s multiple comparison test. **n = 3 biologically independent samples.** Source data are provided as a Source Data file.
Together, our data support the notion that H3K27ac accumulation in MyoD-null cells is insufficient to reconstruct MyoD-mediated chromatin loops, and further indicate that the master TF, MyoD, is a critical genome organizer in establishing the unique 3D genome architecture of muscle cells.

Discussion

Each cell type acquires a unique stage-specific organization of its 3D genome during development. The cell-type-specific organization of the 3D genome can be regarded as an emergent property that is mediated by the interplay between TFs and chromatin-associated proteins. Based on this model, the components involved in regulating the emergent property of chromatin should be coordinately orchestrated by a genome organizer that instructs the specific organization of the 3D genome in each cell type at a specific developmental stage. In this report, we provide compelling evidence indicating that MyoD exerts a previously unappreciated function as a genome organizer in muscle stem cells, and thereby has broad functional impacts on genome structure beyond its known functions in activating gene expression during development. Mechanically, we reveal internal interaction changes in MyoD-bound loops that suggest insights into the potential mechanism through which cell-lineage pioneer TFs ultimately specify the positions and nature of chromatin loops representing the cell-type-specific internal interactions that enable differentiation to yield the correct lineage-specific cell identity. Furthermore, we biochemically demonstrate that interaction between MyoD and the structural protein, CTCF, drives the formation of different types of loops, ultimately supporting a lineage-specific genetic program. Finally, our group and others demonstrate that MyoD-null myoblasts exhibit decreased levels of H3K27ac and H3K4me1 at MyoD-bound enhancers, raising the possibility that H3K27ac might be implicated in chromatin looping. By directing the CRISPR-mediated site-specific accumulation of H3K27ac at MyoD-bound loop anchors, we show that MyoD-instructed chromatin looping is independent of H3K27ac levels in differentiating myoblasts, thus supporting a model that MyoD might also mediate chromatin looping via a distinct mechanism.

It will be interesting to determine which signals ultimately initiate MyoD’s loop specification layer of genetic regulation at the earliest possible stage of myogenic-lineage-specific cell fate determination. It was recently reported that quiescent muscle stem cells (also known as satellite cells, SCs) have relatively dense heterochromatin, and depletion of the paired-box transcription factor, Pax7, significantly reduces heterochromatin condensation in rare surviving SCs, indicating that Pax7 regulates heterochromatin structure in quiescent muscle cells. Pax7 also contributes to the generation of committed myogenic progenitors by directly regulating Myf5 and MyoD. It was speculated that Pax7 might act as an upstream regulator of MyoD-instructed chromatin looping by upregulating MyoD expression in SCs during myogenic development.

To extend our insights beyond the MyoD-mediated architectural regulation of muscle cell fate, we performed similar preliminary experiments with the well-studied neuron-specific TF, NeuroD2, in neuronal cells. Our results suggest that NeuroD2 has a similar loop specification function that apparently underlies the neuron lineage-specific 3D genome architecture and impacts (downstream) the genetic programming of future cell fate. We are eager to see if other lineage-specific TFs in different developmental systems and/or organisms play a general role, similar to that reported herein for MyoD, as genome organizers that orchestrate the lineage-specific chromatin and 3D genome structures required for cell fate determination during development.

Methods

C2C12 cell culture and differentiation. Mouse C2C12 cells (ATCC, CRL-1772) were cultured in growth medium (GM) consisting of Dulbecco’s modified Eagle’s medium (Gibco, Cat.N: C11995-01) supplemented with 4.5 g/L glucose, 10% fetal bovine serum (Ausbion, Cat.N: V5000T), and 1% penicillin and streptomycin at 37 °C in a 5% CO2 atmosphere. For the differentiation of C2C12 myoblasts, cells were transferred to Dulbecco’s modified Eagle’s medium containing 2% horse serum and 1% antibiotics, and then cultured for 24 h. C2C12 cells were grown at 80–90% confluence before the induction of differentiation.

Mouse lines and animal care. MyoD-knockout (MKO) mice (#002523) were obtained from the Jackson Laboratory. The floxed-MyoD mice in the C57BL/6j background were generated by the Model Animal Research Center of Nanjing University. Mice were housed in a pathogen-free facility and had free access to water and standard rodent chow under the following conditions: 21 °C ambient temperature, 50–60% humidity, 12 h dark/light cycle. All animal procedures were approved by the Animal Ethics Committee of Peking Union Medical College, Beijing, China (ACUC-A01-2016-003).
Primary myoblast isolation, culture, and differentiation. Primary myoblasts were isolated from hind-limb skeletal muscles of MKO, WT littermates, or Flx/MyoD mice at 2–3 weeks old, minced, and digested in a mixture of Trypsin II collagenase and dispase. Cells were filtered from debris, centrifuged, and purified to eliminate fibroblasts by differential attachment for 2 × 10 min. The obtained cells were cultured in a growth medium (F-10 Ham’s medium supplemented with 20% foetal bovine serum, 10 ng/ml basic fibroblast growth factor, 1% antibiotics) on collagen-coated cell culture plates at 37°C in 5% CO2. For the differentiation of primary myoblasts, cells were transferred to Dulbecco’s modified Eagle’s medium ( Gibco, Cat.N: C11995500RT) containing 2% horse serum ( Hyclone, Cat.N: 300740.03) and 1% penicillin and streptomycin, and then cultured for 24 h. All cells were reseeded with cells of similar cell density by 70–80% confluence before induction of differentiation. For H3K27ac accumulation experiments, MKO myoblasts were treated with 2.5 µM Tacedinaline (Selleck, Cat.N: CI994) in DM for 12 h to enrich H3K27ac. DMSO- treated MKO myoblasts grown in DM for 12 h served as a control. The cells were washed twice with cold low-salt wash buffer (20 mM Tris pH 8.1, 2 mM EDTA, 150 mM NaCl, 1% Triton X-100, 0.1% SDS), twice with high-salt wash buffer (20 mM Tris pH 8.1, 2 mM EDTA, 500 mM NaCl, 1% Triton X-100, 0.1% SDS), twice with cold LiCl wash buffer (10 mM Tris pH 8.1, 1 mM EDTA, 250 mM LiCl, 1% NP-40, 1% sodium deoxycholate), and once with cold 1 × TE buffer (10 mM Tris pH 8.0, 1 mM EDTA). The DNA/protein complexes were eluted twice for 10 min at 70°C in 150 µl elution buffer (100 mM NaHCO3, 1 mM EDTA, 1% SDS). The complexes were combined with 5 M NaCl, RNase A, and proteinase K and reverse crosslinked at 65°C for 6 h or overnight. The beads were washed using QiaQuick spin columns (Qiagen) and resuspended in 4 × PCR primer mix (Promega). 

Cell preparation for BL-Hi-C and library construction. BL-Hi-C library construction was performed as previously described39 with some modifications. In brief, primary myoblasts were treated with 2% formaldehyde (Sigma, Cat.N: F8775) for 15 min (room temperature, RT) to fix the DNA and protein, and 0.1% Triton X-100, and the reaction was quenched by adding glycine solution (F.C. 0.2 M) for 5 min (R.T.). The cells were scraped, transferred, centrifuged, and the collected cells were split into aliquots of 3–5 × 10^6 cells. Cell pellets were frozen in liquid nitrogen and stored at −80°C until use.

For primary contamination control, the collected cells were resuspended in 1 ml of 1% SDS lysis buffer (50 mM HEPES-KOH, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 1% SDS) at R.T. for 15 min and centrifuged at 500 × g for 2 min. The pellet was washed with 1 ml of 0.1% SDS lysis buffer (50 mM HEPES-KOH, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% SDS) and centrifuged at 500 × g for 2 min. The supernatant was discarded, and the cells were resuspended with 50 µl of NP-40 lysis buffer. The mixture was incubated for 5 min at 37°C. The SDS was quenched by the addition of 145 µl H2O and 12.5 µl 20% Triton X-100, and the genome was digested into blunt-end fragments with 5 µl HaeIII (10 K), and loaded onto a pre-equilibrated HiLoad Superdex 75 16/60 column (GE Healthcare). The MyoD-FL protein present in the MyoD-FL fraction was eluted from the column pure by Phenol:Chloroform extraction and ethanol precipitation. 

Library quality control and sequencing. Before sequencing, the libraries were quantified by qPCR and the size distribution was assessed using an Agilent 2100 Bioanalyzer. BL-Hi-C libraries were sequenced on an Illumina NovaSeq 6000 platform using an Illumina NextSeq550 NovaSeq S2 platform. ChiP-seq libraries were sequenced at × 50 single-end on the Illumina NovaSeq550 NovaSeq S2 platforms. RNA-seq libraries were sequenced in 2 × 150 bp paired-end on the Illumina NovaSeq S2 HiSeq X10 platforms. All experiments were performed by ANNONROAD and Novagen.

Recombinant MyoD purification. The PCR-amplified DNA fragment encoding full-length MyoD (residues 1–318) was cloned into a modified pET28a vector with an N-terminal His6-SUMO tag and Up1 protein expression plasmate. The construction expression vector, named pET28a-SUMO-MyoD-FL, was transformed into Escherichia coli strain BL21 (DE3) (Agilent Technologies, Cat.N: 200131). The cells were grown in LB medium supplemented with 50 mg/ml kanamycin at 37°C until the OD600 = 0.6–0.8 was reached, and then 0.2 mg/ml of isopropyl-β-D-1-thiogalactopyranoside (IPTG, Cat.N: 367-93-1) was added to induce protein expression and the cells were cultured overnight at 18°C. The cells were harvested by centrifugation at 3000 rpm (Thermo Scientific Sigma) for 15 min. The pellet was resuspended in buffer containing 20 mM Tris-HCl pH 8.0, 500 mM NaCl, and 25 mM imidazole pH 8.0. Lysed under high pressure with a cell crusher (NIBIO N-02C), and further clarified by centrifugation at 17,000 rpm for 60 min at 4°C (Beckman Coulter). The supernatant was loaded onto 5 ml HisTrap FF columns (GE Healthcare) pre-equilibrated with 1 mM EDTA, 500 mM NaCl, 1% Triton X-100, 0.1% SDS) and 10 mM imidazole pH 8.0 for 20 min. The imidazole was eluted from the column with 100 mM imidazole pH 8.0, and 500 mM NaCl, 1% Triton X-100, 0.1% SDS. The eluted protein was concentrated by centrifugal ultrafiltration, and excess nucleic acids were removed using a HiTrap SP FF column (GE Healthcare). The MyoD-FL protein present in the MyoD-FL fraction was eluted by Phenol:Chloroform extraction and ethanol precipitation.
plasmid was digested with EcoRI (NEB, R0101I) and gel purified. For the ligation assay, the DNA and proteins were incubated in binding buffer (20 mM Tris 7.9, 50 mM NaCl, 1 mM EDTA) at 25°C for 20 min; the DNA:protein ratio was 1:100, wherein the final concentration of DNA was 80 nM and the final concentration of protein was 8 μM. The DNA–protein complexes were purified by gravity-flow gel filtration (4 ml of 2% agarose, ABT E-01508S-2B) using TE buffer (10 mM Tris 7.9, 1 mM EDTA). Peak fractions of DNA–protein complexes were analyzed by TEM. Briefly, samples were fixed with 0.4% glutaraldehyde in TE buffer on ice for 30 min. Then, DNA–protein complexes were mixed with a buffer containing spermidine (final concentration, 2 mM) to enhance the absorption of chromatin to the grids. Samples were loaded to glow-discharged carbon-coated EM grids, which were incubated for 2 min and then blotted. The grids were washed 20 ml baths of 0%, 25%, 50%, 75%, and 100% ethanol solution for 4 min each (at R.T.), air-dried, and then shadowed with tungsten at an angle of 10° with rotation. Finally, the samples were examined using an FEI Tecnai G2 Spirit 120-kV TEM. Micrographs are shown in reverse contrast.

E-box deletion. E-box deletion at the MyoG promotor was performed using a CRISPR/Cas9-mediated DNA editing system. Briefly, two sgRNAs (L and R) targeting the MyoG promotor were designed (Supplementary Data 5) and cloned to a backbone vector carrying a codon-optimized version of Cas9 with EGFP (Cas9- pcx58). The resulting plasmid, which harbored the two sgRNAs, was named Cas9- sgRNA-L&R. For E-box deletion in muscle cells, C2C12 cells were transfected with the backbone vector carrying a codon-optimized version of Cas9 with EGFP (Cas9- pcx58). The EGFP-positive cells were FACS-sorted and plated to a 12-well plate. When the cells reached 80% confluence, they were changed to DM for 24 h. Finally, the cells were genotyped by PCR using four primers (Supplementary Data 5) and subjected to measurement of chromatin loop formation by Tn5-mediated 3D-FISH. The EGFP-negative cells served as negative controls.

DNA FISH. DNA FISH probes were constructed according to a previously published paper (Tn5-FISH)44. Briefly, the probe library was generated by PCR amplification and recovered using a DNA Cleanup kit (Zymo Research, Cat.N: D4014). The primer sequences for probe amplification are listed in Supplementary Data 5. The Tn5-FISH probes were amplified by a second PCR with fluorescence-tagged primers. In the in situ hybridization procedure of Tn5-FISH was similar to that of traditional FISH, as previously described44. Briefly, cells were seeded on coverslips and fixed with 4% paraformaldehyde for 15 min, washed with PBS for 5 min, and permeabilized with 0.1% Triton X-100/saponin solution for 10 min. Next, the cells were washed with PBS for 5 min, incubated in 20% glycerol for 20 min, snap freeze-thawed three times with liquid nitrogen, air-dried, and washed with PBS. The cells were further treated with 0.1 M HCl, permeabilized with 0.5% Triton X-100/saponin, digested with 100 μg/ml RNase A for 30 min at 37°C, and then balanced in 50% deionized formamide/2 × SSC solution for 30 min. Each Tn5-FISH probe (10 ng) was mixed with hybridization buffer and applied to the slide, then the coverslips were placed on top and the slide was incubated at 37°C for 5 min and then overnight at 37°C. The coverslips were washed with wash buffer, stained with DAPI, and then mounted on slides with AntiFade mounting medium (Solarbio, Cat.N: S2100). Microscopic imaging was performed on a Leica TCS SP8 STED equipped with the spectral flexibility of white-light laser (WLL) for excitation and an HC PL APO ×100/1.4 oil objective.

Immunofluorescent staining. Cells were fixed with 4% paraformaldehyde in PBS, blocked in 3% BSA, and stained according to standard protocols using primary antibodies against MyoD (Santa Cruz, sc37578, 1:500), CTCF (Millipore, 07-229, 1:200), PAX7 (Dilish, 1:15), Myog (DSHB, F5D, 1:500), and MHC (DSHB, MF-20, 1:500). As secondary antibodies (for all but the dSTORM experiments), goat anti-mouse Alexa Fluor 488 was used for MyoD and goat anti-rabbit Alexa Fluor 647 (Life Technologies, Cat.N: A-11017, 1:1000) and goat anti-rabbit Alexa Fluor 488 (Life Technologies, Cat.N: A-21246, 1:1000).

Western blot analysis. The cells were lysed in a buffer containing 50 mM Tris pH 7.5, 150 mM NaCl, 0.5% Nonidet P40, and protease and phosphatase inhibitors. Proteins in lysates were resolved by SDS-PAGE and transferred to a polyvinylidenedifluoride membrane. Immunoblotting was performed using primary antibodies against MyoD (Santa Cruz, sc37578, 1:500), PAX7 (Dilish, 1:15), Myog (DSHB, F5D, 1:500), and MHC (DSHB, MF-20, 1:500). As secondary antibodies (for all but the dSTORM experiments), goat anti-mouse Alexa Fluor 488 (Life Technologies, Cat.N: A-11017, 1:1000) and goat antibody rabbit Alexa Fluor 647 (Life Technologies, Cat.N: A-21246, 1:1000).

Direct stochastic optical reconstruction microscopy (dSTORM). Coverslips (Warner Instruments, 25 mm, #1.5 thickness) were used. C2C12 cells were cultured on 12-well plates coated with 2 μg/ml fibronectin (2 mg/ml in PBS, pH 7.4-7.5) and supplemented with 10% FBS. After cells were immunostained for MyoD and CTCF (see Immunofluorescent staining section above), they were blocked twice with 1 × PBS solution containing 50 mg glycine for 10 min, washed once in 1 × PBS, and then subjected to dual-color imaging in dSTORM imaging buffer (1 × PBS, 100 mM β-mercaptoethanol (Sigma, Cat.N: 39070), pH 7.4) containing TetraSpeck beads (Life Technologies, Cat.N: T7279), which were used as fiducial markers for x–y drift correction and for overlaying the two-color images. dSTORM images were acquired on an Olympus IX83 motorized inverted fluorescence microscope equipped with the CellTIRF-Live system. The full-length tracrRNA pool plus dCas9-p300core plasmid were annealed using a custom-written MATLAB code. The percentage of MyoD signals that colocalized with CTCF signals (%MyoD colocalizing with CTCF) was calculated by dividing the number of pixels containing both MyoD and CTCF signals by the total number of pixels containing MyoD signals. The percentage of CTCF signals that colocalized with MyoD signals (%CTCF colocalizing with MyoD) was calculated by dividing the number of pixels containing both MyoD and CTCF signals by the total number of pixels containing CTCF signals.

CRISPR-mediated DNA imaging. To visualize chromatin loop anchors, we took advantage of the CRISPR-mediated DNA labeling system described20 and designed modifications to achieve non-repetitive DNA imaging. Briefly, 15 guide RNAs (gRNA) targeting each anchor of the examined chromatin loop (MyoG intron fly loop) were designed. The backbone sequences of the crRNA and tracrRNA were as previously described43. The sequences of crRNAs and tracrRNAs used in this study are presented in Supplementary Data 5. The crRNAs were synthesized with fluorescent labeling at the 5'-end (Sango Biotech). The full-length tracrRNA was non-fluorescently synthesized by Integrated DNA Technologies (IDT). The crRNAs and tracrRNAs were annealed and incubated with dCas9 protein (BDT) to form fluorescent RNA protein complexes (rRNPs) as previously described20. Primary myoblasts or C2C12 cells were pretreated with 100 ng/ml nocodazole (Sigma, M1494) for 16 h and transfected with the pre-assembled fRNPs or preannealed crRNA pool plus dCas9-p300core plasmid (elts from Bo Huang) by electroporation using an SE Cell line 4D-Nucleofector™ X Kit (Lonza, Catalog#: V4XC10-2D4). The electroporated cells were plated in collagen-coated Nunc Glass Bottom Dishes (Thermo Fisher, 150680) and cultured for 12 h before imaging. Microscopic imaging was performed on a Leica TCS SP8 STED equipped with the spectral flexibility of WLL for excitation and an HC PL APO ×100/1.4 oil objective. Imaging was performed using maximum intensity projection with Z stacks from 0.2 to 3 μm. The optical image processing was performed with STED-imaging. Some images were cropped to obtain key regions needed to measure the distance between DNA loci, with maximum values identified by Fiji (Image J)26 after maximum intensity projection. Nuclei were visualized using DAPI for fixed cells and Hoechst 33342 (Thermo Fisher, 62249) for living cells.
Hi-C data processing, Hi-C loop calling, and differential loop detection. BL-Hi-C data were processed for WT and MKO primary myoblasts at proliferation (GM) and early differentiation (DM) stages, while MyoD ChIP-seq analyses were performed in WT myoblasts and myotubes. H3K27ac ChIP-seq analyses were also performed in MKO cells with or without Tacedinaline treatment. Hi-C seq data were mapped to the mm9 genome with Bowtie2 v2.3.59 using default parameters. Aligned reads were filtered with a minimum MAPQ of 20, and duplicates were removed using SAMtools v1.9. Peaks were called using MACS2 v2.3.59 with default parameters for CTCF and MyoD. H3K27ac peaks were called by MACS2 using the broad peak mode. Signal tracks were generated using the -SPMR option in MACS2. Then, the UCSC Genome Browser utility89, bedGraphToBigWig, was used to transform the bedgraph files to bigwig files. Differential peaks of ChIP-seq experiments were called with the R package, deepTools v3.0.11. ChIP-seq peak annotation was done by ChIPseeker v1.20.095. The public ChIP-seq data used in this study were processed in the same manner. More specifically, we merged our detected MyoD peaks with those called from public MyoD ChIP-seq data (GSE56131) and determined the differential peaks by examining the MyoD-binding strength at merged peaks, comparing our WT-DM and WT-GM MyoD ChIP-seq data on DiffBind v2.10.0. WT-GM pseudosequences were combined while WT-DM pseudosequences were combined often in WT-GM, while WT-DM pseudosequences were combined. The MyoD peaks with enhanced signal in WT-DM cells.

Motif enrichment analysis. The Homer script findMotifsGenome was used with default parameters to enrich for the motif of ATAC-seq peaks overlapped with loop anchors in WT primary myoblasts. The Homer script findPeaks was used with option --mot to find nucleosome-free regions in the H3K27ac peaks. The nucleosome-free regions of the broad H3K27ac peaks were applied to call motifs, using findMotifsGenome with the size set to 200 bp.

Hi-C data processing, Hi-C loop calling, and differential loop detection. BL-Hi-C data were processed for WT and MKO primary myoblasts at proliferation (GM) and early differentiation (DM) stages. MyoD, Ctcf, and Myod peaks were called using Homer with default parameters.读后性contacts were determined if their promoters (±3 kb around the TSS) overlapped with loops. Dynamic loops to obtain explicit gene-loop pairs. In other sections, genes associated with loops were determined if their promoters (+3 kb around the TSS) overlapped loops. To compare the muscle-lineage specificity of genes associated with MyoD-bound and CTCF-bound loops, we calculated the Z-score of gene expression in WT-DM cells compared with those at other ENCODE-collected cell types (embryonic stem cells, spleen cells, B cells, T cells, megakaryocytes, neural progenitors, and cortical neuron cells). ClusterProfiler v3.12.096 was used to enrich and compare GO terms for different sets of loops bound by CTCF or MyoD.

Visualization. Tracks of Hi-C maps and ChIP-seq data were generated by pyGenomeTracks v3.1.2 and UCSC, respectively100. Hi-C maps of each condition were normalized by cis-interaction pairs.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability. The data that support this study are available from the corresponding authors upon reasonable request. The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive (GSA)101 under accession number CRA002490. The raw sequence data of this paper have been deposited in the Gene Expression Omnibus (GEO) database under accession number GSE157339. Accession codes for the published data in GEO used in this study are as follows: MyoD ChIP-seq of WT primary myoblasts, GSE56131; MyoD ChIP-seq of WT primary myotubes, SRR001761; ATAC-seq of WT primary myoblasts, GSE53573; NeuroD1 ChIP-seq of embryonic cortical neuron cells, GSE67339; Hi-C data of mouse neural development as well as CTCF and H3K27ac ChIP-seq of embryonic cortical neuron cells, GSE96107; 5M6C ChIP-seq of C2C12 myoblasts and myotubes, GSE113248. Data were aligned to the mm9 genome. Source data are provided with this paper.

Code availability. Custom scripts described in the Methods will be made available upon request.

Received: 30 December 2020; Accepted: 15 December 2021; Published online: 11 January 2022.

References

1. Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293 (2009).
2. Smit, A. , S. A. et al. A 3D map of the human genome at base-pair resolution reveals principles of chromatin looping. Cell 159, 1665–1680 (2014).
3. Dixon, J. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–380 (2012).
4. Zheng, H. & Xie, W. The role of 3D genome organization in development and cell differentiation. Nat. Rev. Mol. Cell Biol. 20, 355–350 (2019).
5. Bonev, B. et al. Multiscale 3D genome reorganization during mouse neural development. Cell 171, 557–572 e524 (2017).
6. Grubert, F. et al. Landscape of cohesin-mediated chromatin loops in the human genome. Nature 583, 737–743 (2020).
Rudnicki, M. A., Braun, T., Hinuma, S. & Jaenisch, R. Inactivation of MyoD in skeletal muscle precursor cells. *Cell, 75*, 1351–1359 (1993).

Kassar-Duchossoy, L. et al. Mrf4 determines skeletal muscle identity in mice. *Development, 111*, 1097–1107 (1991).

Sassoon, D. et al. Expression of two myogenic regulatory factors myogenin and MyoD1 during mouse embryogenesis. *Nature, 341*, 303–307 (1989).

Tajbakhsh, S. et al. Gene targeting the myf5 locus with nlaCZ reveals expression of this myogenic factor in mature skeletal muscle fibres as well as early embryonic muscle. *Dev. Dyn., 206*, 291–300 (1996).

Tzahov, M. A. et al. MyoD or Myf5 is required for the formation of skeletal muscle. *Cell, 75*, 1351–1359 (1993).

Kassar-Duchossoy, L. et al. Mrf4 determines skeletal muscle identity in Myf5/Myod double-mutant mice. *Nature, 431*, 466–471 (2004).

Conley, M. L., Yao, Z., Zong, J. W., Groudie, M. & Tappott, S. J. Distinct activities of Myf5 and MyoD indicate separate roles in skeletal muscle lineage specification and differentiation. *Dev. Cell, 35*, 375–385 (2016).

Rudnicki, M. A., Braun, T., Hinuma, S. & Jaenisch, R. Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. *Cell, 71*, 383–390 (1992).

Megency, L. A., Kablar, B., Garrett, K., Anderson, J. E. & Rudnicki, M. A. MyoD is required for myogenic stem cell function in adult skeletal muscle. *Genes Dev., 10*, 1117–1123 (1996).

Montarras, D., Lindon, C., Pinset, C. & Doyenne, P. Cultured myf5 null and myoD null muscle precursor cells display distinct growth defects. *Biol. Cell, 92*, 565–572 (2000).

Sabbourin, L. A., Girgis-Gabardo, A., Seale, P., Asakura, A. & Rudnicki, M. A. Reduced differentiation potential of primary MyoD-/- myogenic cells derived from adult skeletal muscle. *J. Cell. Biol., 144*, 631–643 (1999).

White, J. D. et al. Myotube formation is delayed but not prevented in MyoD-deficient skeletal muscle: studies in regenerating whole muscle grafts of adult mice. *J. Histochem. Cytochem., 48*, 1531–1544 (2000).

Davis, R. L., Weintraub, H. & Lassar, A. B. Expression of a single transfected cDNA converts C2C12 to myoblasts. *Cell, 81*, 987–1000 (1995).

Tapscott, S. J. et al. MyoD1: a nuclear phosphoprotein requiring a Myc homology region to convert fibroblasts to myoblasts. *Science, 242*, 405–411 (1988).

Weintraub, H. et al. Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc. Natl Acad. Sci. USA, 86*, 5434–5438 (1989).

Lassar, A. B., Paterson, B. M. & Weintraub, H. Transfection of a DNA locus that mediates the conversion of 10T1/2 fibroblasts to myoblasts. *Cell, 67*, 469–476 (1986).

Lassar, A. B. et al. MyoD is a sequence-specific DNA binding protein requiring a region of myc homology to bind to the muscle specific kinase enhancer. *Cell, 58*, 823–831 (1989).

Cao, Y. et al. Genome-wide MyoD binding in skeletal muscle cells: a potential for broad cellular reprogramming. *Dev. Cell, 18*, 662–674 (2010).

Mousavi, K. et al. eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. *Mol. Cell, 51*, 606–617 (2013).

Umanovsky, K. B. et al. Lats1 transcription factor is required for myoblasts proliferation during muscle regeneration. *PLoS Genet., 11*, e1005457 (2015).

Fong, A. P. et al. Genetic and epigenetic determinants of neurogenesis and myogenesis. *Dev. Cell, 22*, 721–735 (2012).

Liang, Z. et al. Bl-Hi-C is an efficient and sensitive approach for capturing structural and regulatory chromatin interactions. *Nat. Commun., 8*, 1622 (2017).

Chen, F. Li, G., Zhang, M. Q. & Chen, Y. HiCDB: a sensitive and robust method for detecting contact domain boundaries. *Nucleic Acids Res., 46*, 11239–11250 (2018).

Busslinger, G. A. et al. Cohesin is positioned in mammalian genomes by transcription. *Nature, 544*, 503–507 (2017).

Haahr, J. H. I. et al. The cohesin release factor WAPL restricts chromatin loop extension. *Cell, 169*, 693–707,6614 (2017).
72. Seale, P. et al. Pax7 is required for the specification of myogenic satellite cells. J. Cell Biol. 172, 91–102 (2006).

73. Gauthier, S. & et al. Myf-5 positive satellite cells contribute to Pax7-dependent long-term maintenance of adult muscle stem cells. Cell Stem Cell 13, 590–601 (2013).

74. Berberoglu, M. A. et al. Satellite-like cells contribute to pax7-dependent skeletal muscle repair in adult zebrafish. Dev. Biol. 424, 162–180 (2017).

75. Malatesta, M., Costanzo, M., Cisterna, B. & Zancanaro, C. Satellite cells in skeletal muscle of the hibernating dormouse, a natural model of quiescence and re-activation: focus on the cell nucleus. Cells 9, 1050, https://doi.org/10.3390/cells9041050 (2020).

76. McKinnell, I. W. et al. Pax7 activates myogenic genes by recruitment of a histone methyltransferase complex. Nat. Cell Biol. 10, 77–84 (2008).

77. Relax, F. et al. Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. J. Cell Biol. 172, 91–102 (2006).

78. Zammit, P. S. et al. Pax7 and myogenic progression in skeletal muscle satellite cells. J. Cell Sci. 119, 1824–1832 (2006).

79. Bayani, J. & Squire, J. A. Fluorescence in situ Hybridization (FISH). Curr. Protoc. Cell Biol. Chapter 22, Unit 22.24 (2004).

80. Chen, B. et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell 155, 1479–1491 (2013).

81. Wang, H. et al. CRISPR-mediated live imaging of genome editing and transcription. Science 365, 1301–1305 (2019).

82. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682 (2012).

83. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotye. Nat. Biotechnol. 37, 907–915 (2019).

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

85. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014).

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

87. Li, H. et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079 (2009).

88. Zhang, Y. et al. Model-based analysis of ChIP-seq (MACS). Genome Biol. 9, R137 (2008).

89. Kuhn, R. M., Haussler, D. & Kent, W. J. The UCSC genome browser and regulatory or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/.

Acknowledgements

This work was supported by grants from the National Key Research and Development Program of China (2016YFA0100703: PI D.Z.; 2018YFA0801402: PI Y.C.), the National Natural Science Foundation of China (91949106 and 31971080: PI Y.Z.; 31871343: PI Y.C.), the National Science Foundation in Beijing (1792125; PI Y.Z.), and the CAMS Innovation Fund for Medical Sciences (2021-1-21M-019: PI Y.Z.; 2020-RC310-009: PI Y.C.; 2016-12M-1-017: PI D.Z.). The funding agencies were not involved in any of the research activities in the present study. We thank Professor Cheng Li and Professor Bing Zhu for useful discussions and critical reading of the manuscript.

Author contributions

R.W. carried out BL-Hi-C, RNA-seq, T5-FISH analysis, and colocalization of MyoD and CTCF, and helped to write the manuscript. F.C. performed all computational and bioinformatics analysis, generated the figures and tables as well as wrote the manuscript. Q.C. conducted BL-Hi-C, ChIP-seq, RNA-seq, and helped to write the manuscript. X.W. performed dCas9-mediated DNA imaging, site-specific rescue of H3K27ac with dCas9/p300-core, site-specific rescue loop in MKO cells with Ad-MyoD. M.S. helped to construct BL-Hi-C libraries. A.K.C., Z.M. and Y.Y. performed dSTORM analysis for colocalization of MyoD and CTCF. G.L. and M.W. conducted in vitro DNA circularization and transmission electron microscopy. Q.L. performed differentiation assays of MKO cells treated with Tac. H.L. and M.C. helped with data analysis. X.Z. helped to conduct 3D-FISH. J.M.L and J.Z. purified recombinant MyoD protein. Y.C. and M.Q.Z. supervised computational and bioinformatics analyses. Y.Z. helped to design and supervised the experiments and wrote the manuscript. D.Z. conceived and supervised the project and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-27865-6.

Correspondence and requests for materials should be addressed to Michael Q. Zhang, Yong Zhang, Yan Chen or Dahai Zhu.

Peer review information Nature Communications thanks the anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information are available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.