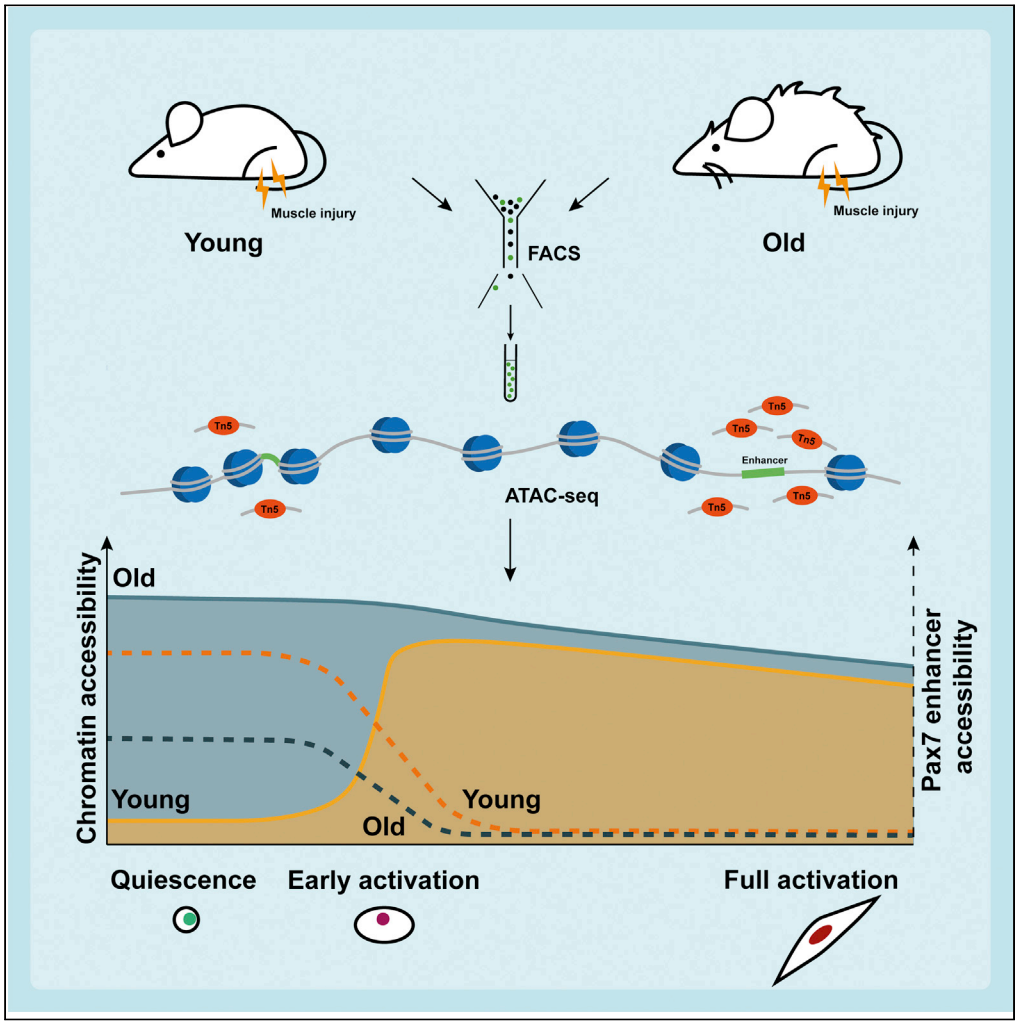
iScience

CellPress
OPEN ACCESS

Article

Global chromatin accessibility profiling analysis reveals a chronic activation state in aged muscle stem cells



Anqi Dong, Jing Liu, Kangning Lin, Wenshu Zeng, Wai-Kin So, Shenyuan Hu, Tom H. Cheung

tcheung@ust.hk

Highlights

Performed time-course profiling of *in situ* fixed muscle stem cells on activation

Showed rapid changes in chromatin accessibility in muscle stem cell early activation

Identified Pax7 enhancers for muscle stem cell quiescence and activation

Aged quiescent muscle stem cells display chronically activated chromatin signature

Dong et al., iScience 25, 104954
September 16, 2022
https://doi.org/10.1016/j.isci.2022.104954

iScience

CellPress
OPEN ACCESS

Article

Global chromatin accessibility profiling analysis reveals a chronic activation state in aged muscle stem cells

Anqi Dong,[1] Jing Liu,[1] Kangning Lin,[1] Wenshu Zeng,[1] Wai-Kin So,[1] Shenyuan Hu,[1] and Tom H. Cheung[1,2,3,4,*]

SUMMARY

Regulation of chromatin accessibility is critical for cell fate decisions. Chromatin structure responds to extrinsic environments rapidly. The traditional adult stem cell isolation approach requires tissue dissociation, which triggers stem cell activation and leads to alterations in chromatin structure. To preserve the *in vivo* chromatin states, we utilized the PFA-perfusion-based isolation approach and characterized the DNA regulatory landscapes during muscle stem cell quiescence exit and aging. We showed that aged SCs display a chronically activated chromatin signature. Detailed analysis of the chromatin accessibility profiles identified key enhancer elements for SC quiescence. Constant activation of the enhancer elements promotes stemness and prevents SCs from differentiation, whereas genetic deletion causes cell-cycle arrest and leads to defects in activation. Our comprehensive characterization of the chromatin accessibility and transcriptomic landscapes in SC quiescence and aging broadens our understanding of these processes and identifies key distal regulatory elements for SC function.

INTRODUCTION

Adult stem cells are essential for tissue regeneration and homeostasis maintenance (Rumman et al., 2015; Weissman, 2000). Skeletal muscle possesses a remarkable regeneration capacity after acute injury because of its resident stem cells, muscle stem cells, or satellite cells (SCs), located between the basal lamina and the muscle fiber sarcolemma (Mauro, 1961). In a resting skeletal muscle, SCs remain in a reversible G0 state called the quiescent state for a long period (Cheung and Rando, 2013). Upon stimuli such as acute injury, quiescent SCs exit from quiescence to activate and re-enter the cell cycle for proliferation. They will further differentiate to repair the damaged tissue (Relaix and Zammit, 2012). Some activated SCs will return to quiescence through the self-renewal process to replenish the stem cell pool (Collins et al., 2005; Zammit et al., 2004). Pax7 is an indispensable marker for quiescent SCs (Lepper et al., 2011; Von Maltzahn et al., 2013; Oustanina et al., 2004; Sambasivan et al., 2011; Seale et al., 2000). Depleting Pax7 in SCs results in SC pool depletion because of cell-cycle arrest and precocious differentiation (Von Maltzahn et al., 2013), suggesting the fundamental role of Pax7 in maintaining SC quiescence and proper activation. Recently, different approaches have been used to examine the transcriptomic signature of quiescent SCs (Machado et al., 2017; Van Velthoven et al., 2017; Yue et al., 2020). However, measuring gene expression alone is insufficient in identifying the regulome for cell fate determination. Understanding the chromatin dynamics during SC quiescence exit and self-renewal will enable the identification of regulatory elements for quiescence maintenance.

The eukaryotic DNA is highly organized into a nuclear structure called chromatin (Probst et al., 2009; Sitbon et al., 2017). The accessibility to the regulatory DNA elements restrains the gene expression, therefore defining the cell identity. Genome-wide chromatin-based assays are sensitive methods for assaying the activity of regulatory elements. Specifically, Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) (Buenrostro et al., 2013) provides a straightforward measurement of chromatin accessibility and can be used to discover functional regulatory elements. Recent findings using ATAC-seq suggested that distal regulatory elements are more cell-type-specific than promoter regions (Corces et al., 2016), highlighting the importance of distal regulatory elements in cell fate specification. Although much effort has been made in identifying enhancers for SC quiescence, the distal regulatory network of SC remains to be explored.

[1]Division of Life Science, Center for Stem Cell Research, HKUST-Nan Fung Life Sciences Joint Laboratory, State Key Laboratory of Molecular Neuroscience, Molecular Neuroscience Center, The Hong Kong University of Science and Technology, Hong Kong, China

[2]Hong Kong Center for Neurodegenerative Diseases, Hong Kong, China

[3]Guangdong Provincial Key Laboratory of Brain Science, Disease and Drug Development, Shenzhen-Hong Kong Institute of Brain Science, HKUST Shenzhen Research Institute, Shenzhen, China

[4]Lead contact

*Correspondence: tcheung@ust.hk

https://doi.org/10.1016/j.isci.2022.104954

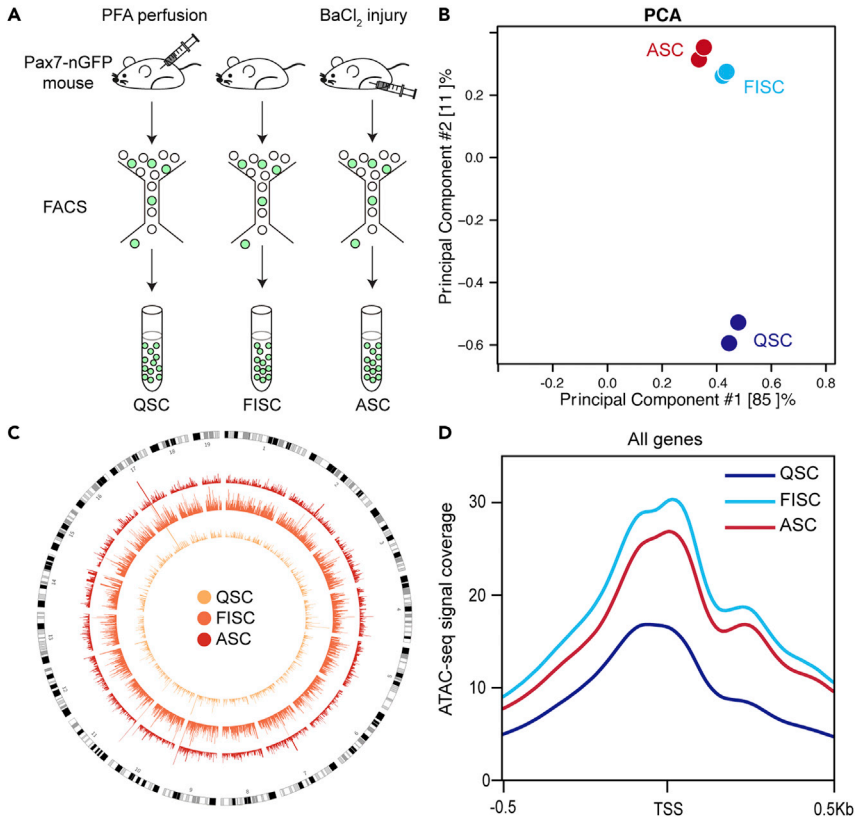


Figure 1. Quiescent satellite cells have a highly compact chromatin environment (See also Figure S1)

(A) Illustration of the isolation scheme for quiescent satellite cells (QSCs), freshly isolated satellite cells (FISCs), and activated satellite cells (ASCs).

(B) Principal component analysis (PCA) plot of ATAC-seqdata for QSCs, FISCs, and ASCs.

(C) The Circos plot of ATAC-seq signal coverage across mouse genome for QSCs, FISCs, and ASCs.

(D) ATAC-seq signal coverage across transcription start sites (TSS) in QSCs, FISCs, and ASCs.

In this study, we examined the chromatin accessibility changes of SCs, from quiescence exit, early activation, and regeneration, showing the trajectory of chromatin environment changes for SC activation in young and aged conditions. We found that the old SCs exhibit a much more open chromatin environment, suggesting that the old SCs exhibit a chronically activated chromatin state. Further analysis of the chromatin accessibility profiles identified two Pax7 enhancers that facilitate SC function. Deleting these enhancers reproduces the phenotype of Pax7-deficient SCs, suggesting the essential role of these enhancers in regulating Pax7 expression. Taken together, our study provided a database for the chromatin accessibility and corresponding transcriptome landscapes covering the progression of SC quiescence exit, activation, regeneration, and aging process, which can be further utilized to identify functional distal regulatory elements for SC function.

RESULTS

Dissecting the chromatin accessibility of quiescent satellite cells

Previous transcriptomic examination of SC quiescence and activation has revealed that the transcriptome landscape changes drastically during the isolation process (Machado et al., 2017; Van Velthoven et al., 2017; Yue et al., 2020). Previous studies have examined chromatin accessibility in freshly isolated SCs, activated SCs (FISCs and ASCs), and cultured myoblasts (Chow et al., 2021; Dong et al., 2020). Using our perfusion-based sorting technique (Yue and Cheung, 2020), we dissect the chromatin accessibility landscape in the quiescent SCs (QSCs) (Figures 1A, S1A, and S1B). We have tested and confirmed our fixation approach does not alter the chromatin accessibility landscape (Figure S1C). Principal component analysis (PCA) shows that the chromatin accessibility landscape of QSCs is significantly different from that of FISCs and ASCs (Figure 1B). The QSCs have the lowest ATAC-seq signal coverage across the whole genome as

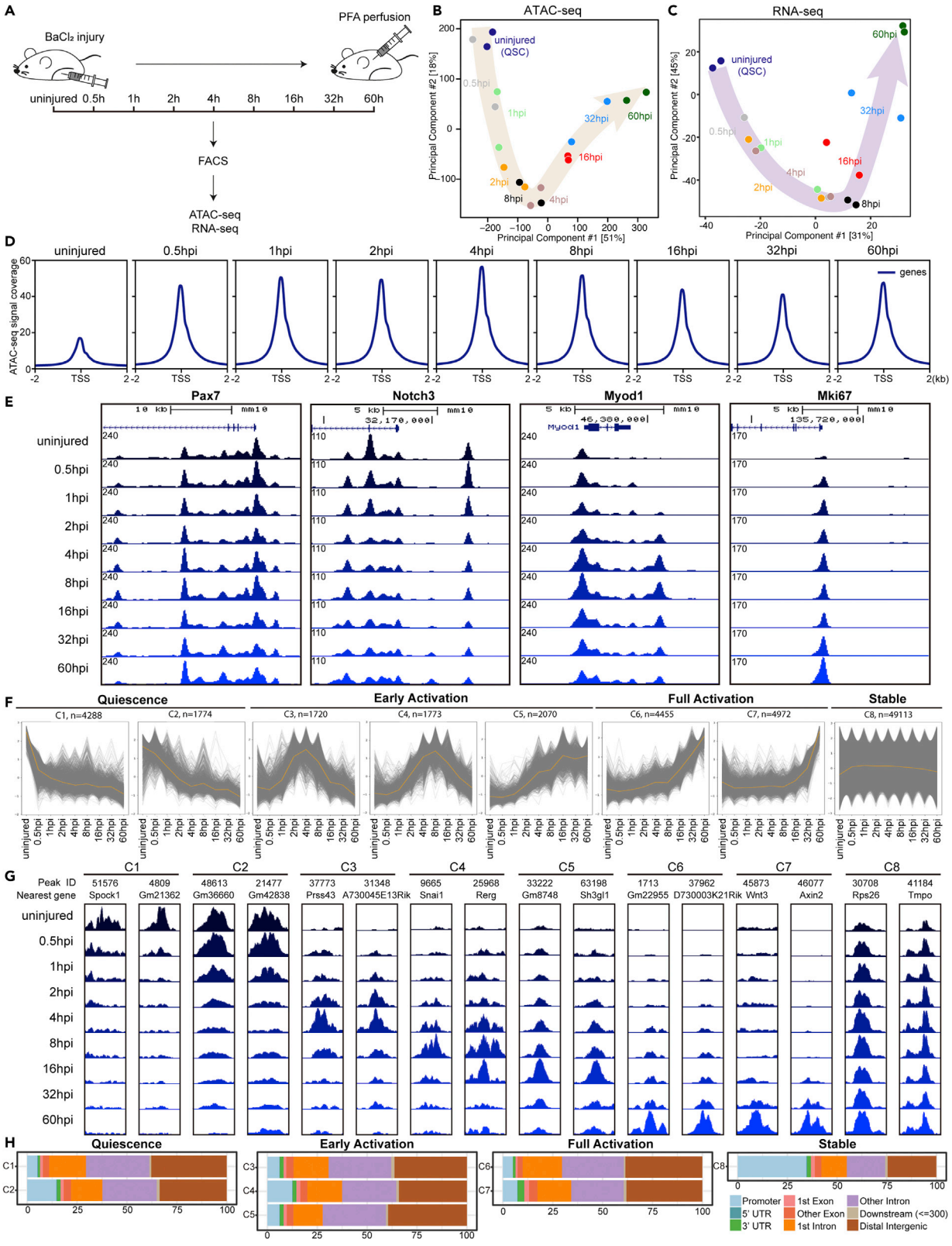

A
BaCl₂ injury
PFA perfusion
uninjured 0.5h 1h 2h 4h 8h 16h 32h 60h
FACS
ATAC-seq
RNA-seq
B
ATAC-seq
uninjured (QSC)
0.5hpi
1hpi
2hpi
4hpi
8hpi
16hpi
32hpi
60hpi
Principal Component #2 [18%]
Principal Component #1 [51%]
C
RNA-seq
uninjured (QSC)
0.5hpi
1hpi
2hpi
4hpi
8hpi
16hpi
32hpi
60hpi
Principal Component #2 [45%]
Principal Component #1 [31%]
D
uninjured
0.5hpi
1hpi
2hpi
4hpi
8hpi
16hpi
32hpi
60hpi
genes
ATAC-seq signal coverage
TSS
(kb)
E
Pax7
Notch3
Myod1
Mki67
uninjured
0.5hpi
1hpi
2hpi
4hpi
8hpi
16hpi
32hpi
60hpi
F
Quiescence
Early Activation
Full Activation
Stable
C1, n=4288
C2, n=1774
C3, n=1720
C4, n=1773
C5, n=2070
C6, n=4455
C7, n=4972
C8, n=49113
G
C1
C2
C3
C4
C5
C6
C7
C8
Peak ID
Nearest gene
51576 Spock1
4809 Gm21362
48613 Gm36660
21477 Gm42838
37773 Prss43
31348 A730045E13Rik
9665 Snai1
25968 Rerg
33222 Gm8748
63198 Sh3gl1
1713 Gm22955
37962 D730003K21Rik
45873 Wnt3
46077 Axin2
30708 Rps26
41184 Tmpo
uninjured
0.5hpi
1hpi
2hpi
4hpi
8hpi
16hpi
32hpi
60hpi
H
Quiescence
Early Activation
Full Activation
Stable
Promoter
5' UTR
3' UTR
1st Exon
1st Intron
Other Exon
Other Intron
Downstream (<=300)
Distal Intergenic

Figure 2. SCs exhibit different phases of chromatin accessibility changes during quiescence exit (See also Figures S2 and S3, Table S1 and S2)
(A) Illustration of the experimental design for isolation of injured SCs.
(B and C) Principal component analysis (PCA) plots of ATAC-seq (B) and RNA-seq (C) datasets of uninjured and injured SCs.
(D) ATAC-seq signal coverage across transcription start sites (TSS) in injured SCs.
(E) Representative genome tracks of ATAC-seq signal coverage across the promoter regions of quiescence-specific genes (*Pax7* and *Notch3*) and proliferation genes (*Myod1* and *Mki67*).
(F) Fuzzy cluster analysis of ATAC-seq signal of injured SCs. Line plots show standardized ATAC-seq signal, with individual gray lines representing individual loci, and the orange line representing the values for the cluster center.
(G) Representative genome tracks of ATAC-seq signal coverage across peaks from different clusters.
(H) The proportion of different genomic annotations of ATAC-seq peaks from different categories.

well as across transcription start sites (TSS) compared with FISCs and ASCs (Figures 1C and 1D), suggesting that QSCs have a relatively compact chromatin environment. The relative genomic distribution of ATAC peaks was similar between different cell states, with ~50% within the promoter regions, ~25% in the intronic regions, and ~20% in the intergenic regions (Figure S1D). However, QSCs have the least ATAC-seq peaks identified compared with FISCs and ASCs, especially in the distal region (Figure S1E). When examining the correlation between chromatin accessibility and gene expression during the activation process, we found a large number of genes whose chromatin environments become open in FISCs or ASCs, but do not have their gene expression increased (Figures S1F–S1G). This result indicates that the chromatin environment starts to open up on activation, whereas the actual transcription has not yet caught up with the pace. Altogether, our results suggest that the QSCs undergo rapid alterations in their chromatin environment during the isolation process. Therefore, it is necessary to use the perfusion-based sorting scheme to capture the chromatin landscape during SC quiescence exit *in vivo*.

The dynamics of chromatin accessibility during satellite cell quiescence exit

We performed time-course profiling of chromatin accessibility (ATAC-seq) and gene expression (RNA-seq) using a perfusion-based sorting scheme to capture the process of chromatin accessibility and transcriptome alterations during SC quiescence exit *in vivo* (Figures 2A and S2A; Table S1). The replicates of different injury time-points showed a highly concordant profile (Figures S2B and S2C). Both the PCA plots of ATAC-seq and RNA-seq datasets show the trajectory of the *in vivo* activation process (Figures 2B and 2C). We also observed that the FISCs profile resembles that of 4–8 hours post-injury (hpi) samples (Figures S2D). This suggests that the isolation process similarly triggers SC activation to that of an acute injury. The ATAC-seq signal reaches the highest at 4 hpi and gradually decreases in later time points until full activation (Figure 2D). Representative genome tracks also showed that the promoter regions of quiescence-specific genes (e.g., *Pax7* and *Notch3*) are highly accessible in QSCs and early activated SCs, whereas genes associated with proliferation (e.g., *Myod1* and *Mki67*) only have their promoter regions accessible on activation (Figure 2E). When examining the relative genomic distribution of ATAC peaks, we found that uninjured QSCs have a relatively high enrichment in peaks located within the promoter regions (Figure S2E). The number of ATAC-seq peaks identified within distal regions increases drastically during activation compared with ATAC-seq peaks identified within promoter regions, implying the important role of distal elements during activation (Figure S2F). Further examination of ATAC-seq data has identified multiple peaks that are differentially accessible during quiescence exit (Figure S3A). This result indicates that activation is a dynamic process that takes place in different phases.

To investigate the dynamics of chromatin accessibility during SC quiescence exit, we performed time-course fuzzy clustering on normalized ATAC-seq read counts in peaks. We have identified eight clusters that exhibit significant changes across different time points and classified them into four distinct categories: Quiescence (C1 and C2), early activation (C3, C4, and C5), full activation (C6 and C7), and stable cluster where chromatin accessibility is constantly stable (C8) (Figures 2F and 2G; Table S2). Examination of the genomic locations shows that the majority of the peaks that exhibit changes in their accessibility during activation (i.e., clusters from quiescence, early activation, and full activation categories) are located in the distal regions. In contrast, a large portion of peaks maintains their accessibility (C8) located within the promoter regions (Figure 2H). Further gene ontology analysis of different categories shows that quiescence enriched peaks (C1 and C2) are associated with extracellular matrix communications; peaks enriched in early activation (C3, C4, and C5) are associated with cell-cell communication and stimuli response; peaks enriched in full activation (C6 and C7) are associated with system development and cell differentiation (Figure S3B). Altogether, these results uncovered the characteristics of chromatin changes during different phases of quiescence exit and highlighted the importance of characterizing the distal regulatory elements during SC quiescence exit.

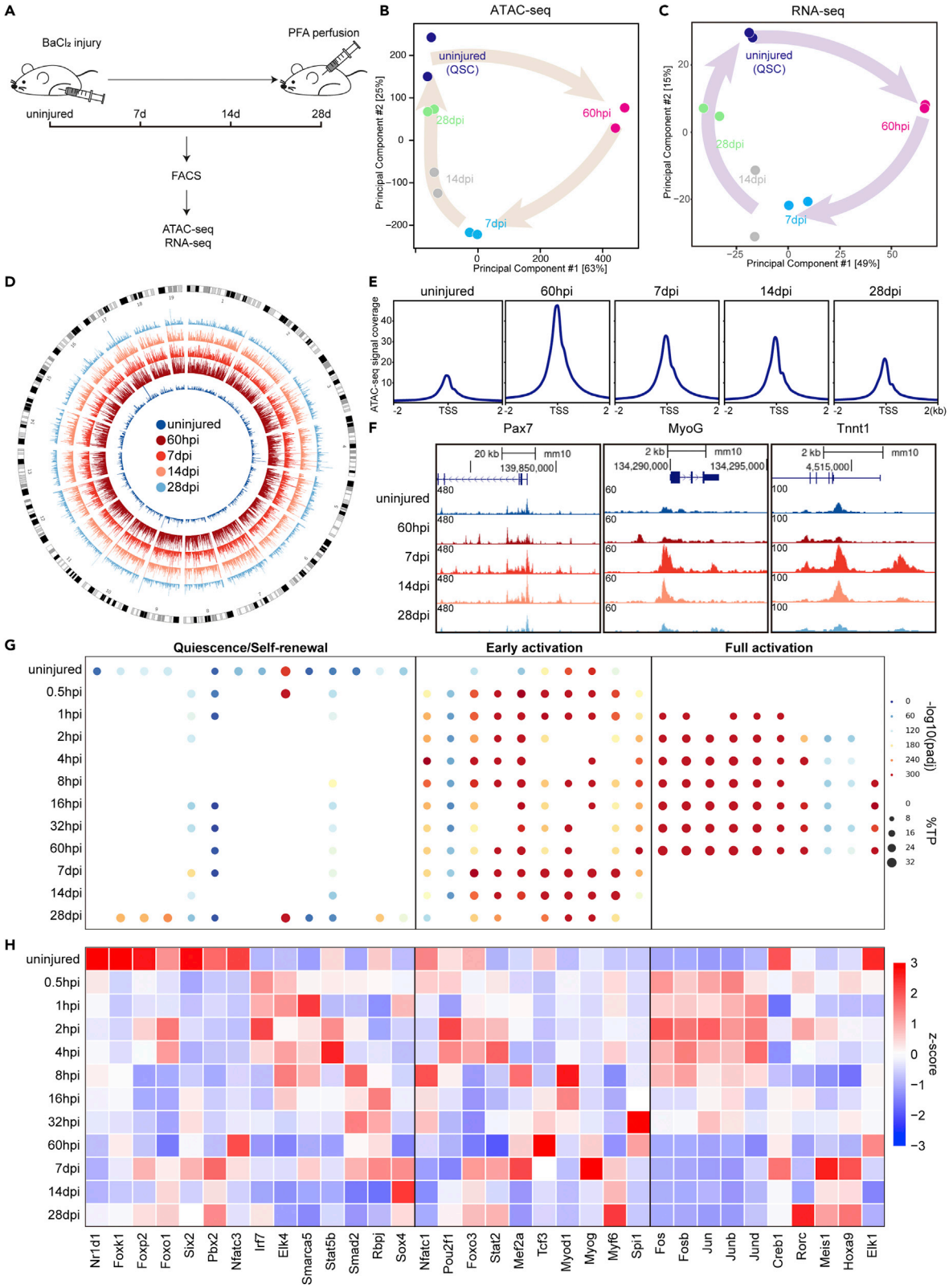

A
BaCl₂ injury
PFA perfusion
uninjured
7d
14d
28d
FACS
ATAC-seq
RNA-seq
B
ATAC-seq
uninjured (QSC)
28dpi
60hpi
14dpi
7dpi
Principal Component #2 [25%]
Principal Component #1 [63%]
C
RNA-seq
uninjured (QSC)
28dpi
60hpi
14dpi
7dpi
Principal Component #2 [15%]
Principal Component #1 [49%]
D
uninjured
60hpi
7dpi
14dpi
28dpi
E
uninjured
60hpi
7dpi
14dpi
28dpi
ATAC-seq signal coverage
TSS
2(kb)
F
Pax7
MyoG
Tnnt1
mm10
uninjured
60hpi
7dpi
14dpi
28dpi
G
Quiescence/Self-renewal
Early activation
Full activation
uninjured
0.5hpi
1hpi
2hpi
4hpi
8hpi
16hpi
32hpi
60hpi
7dpi
14dpi
28dpi
-log10(padj)
%TP
H
uninjured
0.5hpi
1hpi
2hpi
4hpi
8hpi
16hpi
32hpi
60hpi
7dpi
14dpi
28dpi
z-score
Nr1d1
Foxk1
Foxp2
Foxo1
Six2
Pbx2
Nfatc3
Irf7
Elk4
Smarca5
Stat5b
Smad2
Rbpj
Sox4
Nfatc1
Pou2f1
Foxo3
Stat2
Mef2a
Tcf3
Myod1
Myog
Myf6
Spi1
Fos
Fosb
Jun
Junb
Jund
Creb1
Rorc
Meis1
Hoxa9
Elk1

Figure 3. The transcriptome and chromatin accessibility landscapes of SCs during regeneration (See also Figure S4 and Table S1)

(A) Illustration of the experimental design for isolation of regenerating SCs.

(B and C) Principal component analysis (PCA) plots of ATAC-seq (B) and RNA-seq (C) datasets of uninjured and injured SCs.

(D) The Circos plot of ATAC-seq signal coverage across mouse genome for regenerating SCs.

(E) ATAC-seq signal coverage across transcription start sites (TSS) in regenerating SCs.

(F) Representative genome tracks of ATAC-seq signal coverage across the promoter regions of quiescence genes (*Pax7*) and differentiation genes (*Myog* and *Tnnt1*).

(G) Enrichment of transcription factor (TF) motifs during SC quiescence, activation, and regeneration.

(H) Heatmap showing the expression pattern of TFs during SC quiescence, activation, and regeneration.

The changes in chromatin accessibility during satellite cell self-renewal

To explore the chromatin accessibility changes during SC self-renewal, we performed time-course profiling of RNA-seq and ATAC-seq on regenerating SCs (Figures 3A and S4A; Table S1). The replicates of different injury time-points showed a highly concordant profile (Figures S4B and S4C). The PCA plots of ATAC-seq and RNA-seq datasets both show similar trajectories of the *in vivo* regeneration process (Figures 3B and 3C). A global examination of ATAC-seq signal coverage shows that the chromatin environment becomes highly accessible on full activation (60hpi), gradually decreases the accessibility during regeneration, and re-establishes the compact chromatin environment during self-renewal (28 days post injury (dpi)) (Figures 3D and 3E). Representative genome tracks indicate that promoter regions of typical differentiation-associated genes (e.g., *Myog* and *Tnnt2*) become highly accessible during regeneration (Figure 3F). When examining the relative genomic distribution of ATAC-seq peaks, we found that the proportion of promoter regions is restored during regeneration (Figure S4D). The number of ATAC-seq peaks located within both promoter and distal regions also decreased during regeneration, whereas the peak numbers of distal regions show a more significant change (Figure S4E). Further examination of ATAC-seq data has identified multiple peaks that are differentially accessible during regeneration (Figure S4F).

Analysis of the DNA motif enriched in time-course ATAC-seq samples revealed enrichment of different transcription factor (TF) motifs during different phases of SC activation and regeneration (Figure 3G). We observed a significant enrichment of motifs for TFs from the Jun/Fos family when SCs are fully activated (Figure 3G), corresponding with the high expression level of Jun/Fos family genes in activated SCs (Figure 3H). We have also identified multiple TF motifs from the E-protein family enriched in early activated SCs (Figure 3G). A previous study has suggested that FoxO maintains a genuine quiescent state (García-Prat et al., 2020). Our data also shows that motifs for TFs from the Fox family (e.g., Foxk1, Foxo1, and Foxp2) are specifically enriched in uninjured and regenerated SCs (Figure 3G), and the expression of the Fox genes is high in uninjured SCs (Figure 3H), suggesting their roles in maintaining SC quiescence. Taken together, our data guide in predicting functional TFs during SC quiescence and activation.

Identification of distal regulatory elements in maintaining satellite cell quiescence

Analysis of the genomic distribution of ATAC-seq peaks indicates a dramatic alteration in accessible peaks located in distal regions during SC activation and regeneration (Figures S2E and S4D). A previous study has shown that distal elements are more lineage-specific than promoter regions (Corces et al., 2016). Our hierarchical clustering analysis also indicates that ATAC-seq peaks from distal regions have a higher cluster purity than TSS regions (Figures 4A and 4B). We therefore sought to identify quiescence-specific distal elements. Pax7 is an indispensable regulator for quiescent SCs (Lepper et al., 2011; Von Maltzahn et al., 2013; Oustanina et al., 2004; Sambasivan et al., 2011; Seale et al., 2000). Of interest, we found that close to the *Pax7* locus, two ATAC-seq peaks are significantly enriched in QSCs and early activated SCs, gradually lose their chromatin accessibility after full activation, and become highly accessible again during regeneration (Figure 4C). To test whether these peaks contain potential enhancer elements, we crossed our ATAC-seq data with H3K27ac chromatin immunoprecipitation followed by sequencing (ChIP-seq) data from *in situ* fixed T0 (uninjured) and T3 (3 h after isolation) SCs (Machado et al., 2017), as well as from FISCs and ASCs. Both the ATAC-seq peaks have a high H3K27ac signal in T0 SCs and lose the active enhancer mark in T3 SCs, FISCs, and ASCs (Figure 4C), suggesting that these ATAC-seq peaks contain potential enhancers. Because the chromatin accessibility profiles of these two ATAC-seq peaks correspond with Pax7 expression (Figures 4C and S5A), we name them Pax7 E-intronic and Pax7 E-distal respectively, according to their genomic positions.

To functionally validate the enhancer activity of these distal elements, we generated a luciferase reporter plasmid driven by the minimal promoter, which contains the Pax7 E-intronic and Pax7 E-distal respectively (Figure 4D). The luciferase intensity will increase if the elements we inserted (i.e., E-distal and E-intronic) can

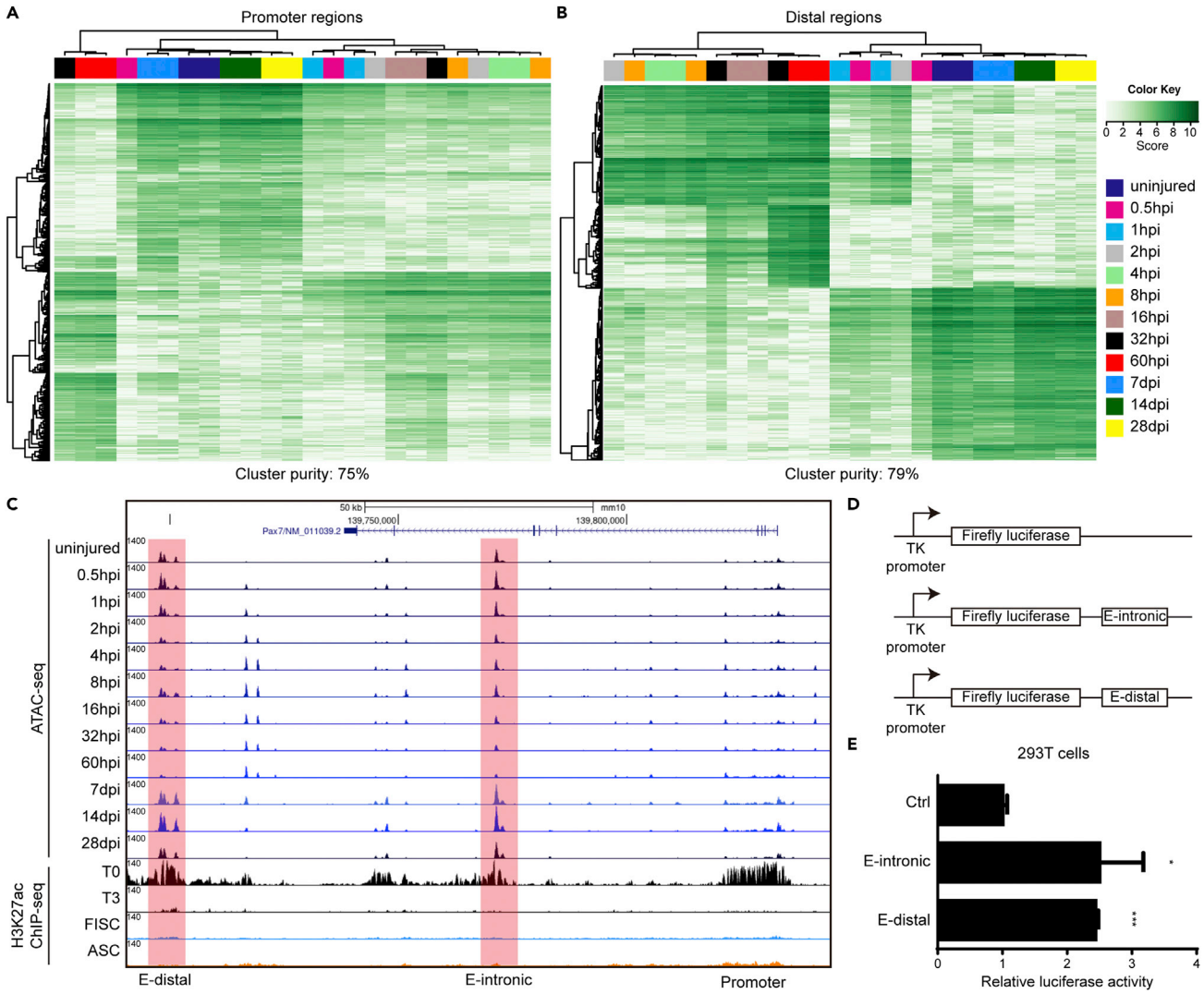


Figure 4. Identification of Pax7 enhancers (See also Figure S5)

(A and B) Hierarchical clustering analysis of differentially enriched ATAC-seq peaks from promoter regions (A) and distal regions (B) across different samples.

(C) Genome track of ATAC-seq and H3K27ac ChIP-seq signal coverage across *Pax7* locus.

(D) Experimental design of dual luciferase assay.

(E) Result of the dual-luciferase assay with transfection of indicated plasmid in 293 T cells. *, $p< 0.05$, ***, $p< 0.001$, $n = 3$.

act on the minimal promoter to amplify the promoter activity. The dual-luciferase assay result showed that the addition of both enhancer elements boosts the luciferase activity in 293T cells (Figure 4E). The addition of Pax7 E-intronic can also increase luciferase activity significantly in C2C12 cells (Figure S5B). These results suggest that both Pax7 E-intronic and E-distal regions contain functional enhancer elements.

Pax7 enhancers are essential for satellite cell function

To examine the function of these enhancers in SCs, we generated a tamoxifen (TMX)-inducible, SC-specific dCas9 overexpression mouse strain by crossing the Tg: CAG-LSL-dCas9-SPH mice (Zhou et al., 2018) with $Pax7^{CreERT2/CreERT2}$; $R26R^{YFP/YFP}$ mice (Figure 5A). After 24 h treatment of sgRNAs targeting the enhancer regions, we did not detect a change in Pax7 expression. In contrast, the treatment group with sgRNA targeting Pax7 E-intronic started to show a decrease in cell cycle entry (Figures S6A and S6B). The increase in Pax7 expression is observed after 40h of sgRNA treatment, and both treatment groups have lower EdU incorporation than the control group (Figures 5B and 5C). These observations suggest that the constant

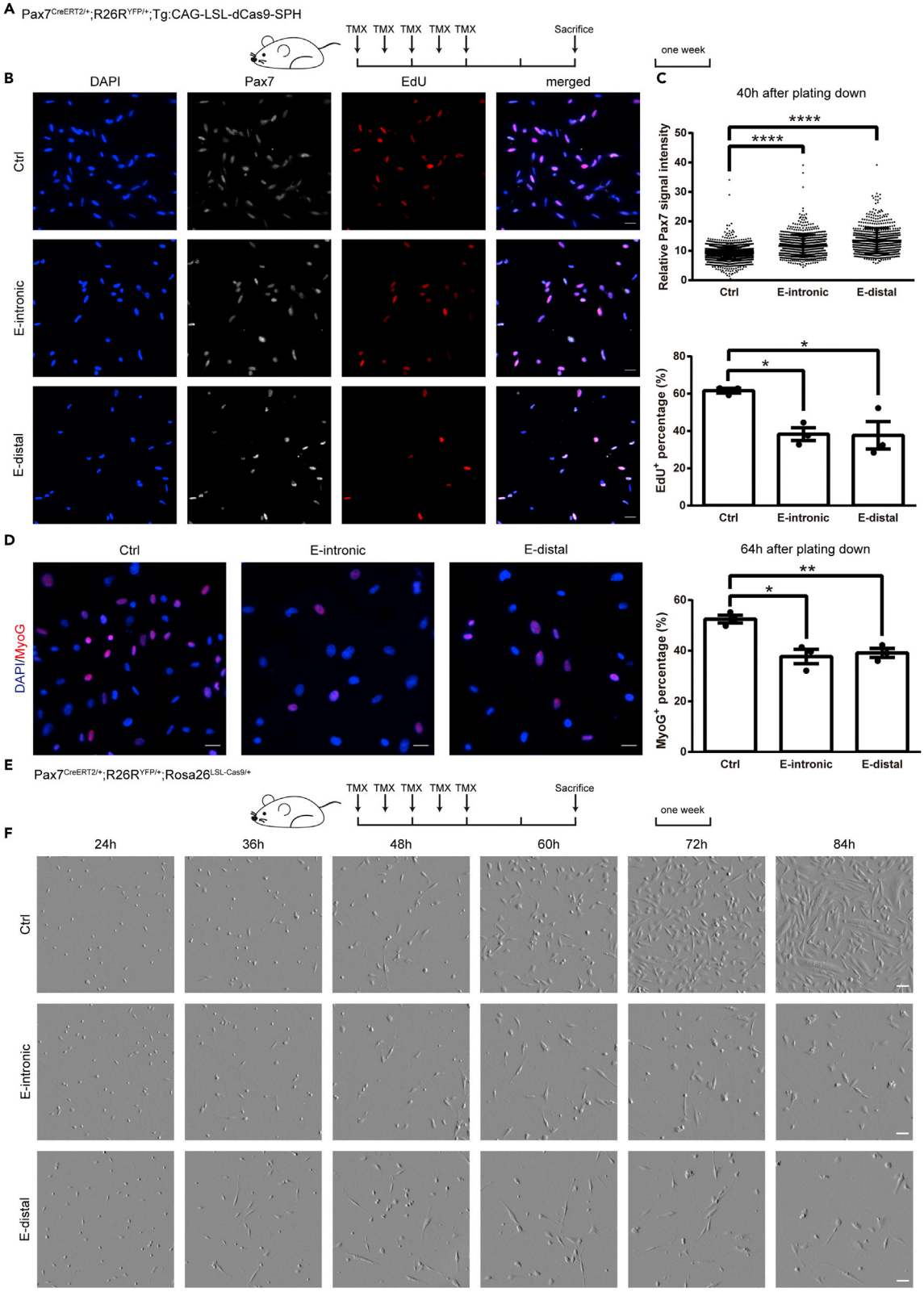

A
Pax7CreERT2/+;R26RYFP/+;Tg:CAG-LSL-dCas9-SPH
TMX TMX TMX TMX TMX
Sacrifice
one week
B
DAPI
Pax7
EdU
merged
Ctrl
E-intronic
E-distal
C
40h after plating down
Relative Pax7 signal intensity
****
****
Ctrl
E-intronic
E-distal
EdU+ percentage (%)
*
*
D
Ctrl
E-intronic
E-distal
DAPI/MyoG
64h after plating down
MyoG+ percentage (%)
**
*
E
Pax7CreERT2/+;R26RYFP/+;Rosa26LSL-Cas9/+
TMX TMX TMX TMX TMX
Sacrifice
one week
F
24h
36h
48h
60h
72h
84h
Ctrl
E-intronic
E-distal

Figure 5. Pax7 enhancers are essential for satellite cell function (See also Figure S6)

(A) Schematic illustration of the experimental design for tamoxifen (TMX)-induced dCas9 expression in SCs.

(B) Staining result of Pax7 and EdU incorporation analysis of dCas9-expressing SCs treated with sgRNAs targeting E-intronic and E-distal elements for 40h. The scale bar represents 20μm.

(C) Quantification results of Pax7 intensity and EdU incorporation percentage in SCs treated with sgRNAs targeting E-intronic and E-distal elements for 40h. *, $p< 0.05$, ****, $p< 0.0001$, $n = 3$.

(D) (Left) Staining result of MyoG of dCas9-expressing SCs treated with sgRNAs targeting E-intronic and E-distal elements for 64h. The scale bar represents 20μm. (Right) Quantification result of $MyoG^{+}$ SCs in different treatment groups. *, $p< 0.05$, **, $p< 0.01$, $n = 3$.

(E) Schematic illustration of the experimental design for TMX-induced Cas9 expression in SCs.

(F) Live-cell imaging result of Cas9-expressing SCs cultured for different time points after treatment with sgRNAs targeting E-intronic and E-distal elements. The scale bar represents 50μm.

activation of Pax7 enhancers prevents SCs from entering the cell cycle. When SCs are cultured for a longer time after sgRNA treatment, we observed a decrease in MyoG expression in the treatment groups (Figure 5D). We also observed fewer cells in the treatment groups (Figure 5D). Our results suggest that constant activation of these enhancers promotes Pax7 expression and prevents SCs from proliferation and differentiation.

To conduct a loss-of-function experiment, we generated a TMX-inducible, SC-specific Cas9 overexpression strain by crossing the $Rosa26^{LSL-Cas9/LSL-Cas9}$ mice with $Pax7^{CreERT2/CreERT2}$; $R26R^{YFP/YFP}$ mice (Figure 5E). After deleting Pax7 enhancers with sgRNA (Figures S6C and S6D), we observed decreased Pax7 expression and EdU incorporation (Figures S6E and S6F). Live-cell imaging showed that loss of these Pax7 enhancers leads to defects in SC activation, essentially preventing SCs from proliferation (Figure 5F and Video S1–S3). These phenotypes reproduce the results reported in the Pax7-deficient SCs (Von Maltzahn et al., 2013), suggesting the significance of these Pax7 enhancers in regulating Pax7 expression.

Old satellite cells exhibit a global increase in chromatin accessibility

SCs preserve a repressive heterochromatin environment for quiescence maintenance and exhibit a profound transition from condensed to open chromatin during activation (Boonsanay et al., 2016). To determine whether SCs exhibit similar chromatin environment transition during activation in aging circumstances, we performed RNA-seq and ATAC-seq on old QSCs, FISCs, and ASCs using $Pax7^{CreERT2/+}$; $R26R^{YFP/+}$ mice (Figures S7A, and S7B; Table S3). Principal component analysis (PCA) shows that old SCs exhibit distinct transcriptomic profiles from young SCs, whereas their chromatin accessibilities resemble more with their young counterparts (Figures 6A, 6B, S7A, and S7B). However, old QSCs have significantly increased chromatin accessibility than young QSCs and resemble most old FISCs (Figures 6C, 6D, and S7C–S7E), suggesting that the repressive chromatin environment is disrupted in old QSCs.

Further examination of ATAC-seq data in young and old SCs has identified multiple peaks that are differentially accessible during aging (Figure S8A). When comparing differentially enriched ATAC-seq peaks in the quiescent state between young and old SCs, we observed a significant enrichment of ATAC-peaks near genes associated with the cell cycle (Figures 6E and 6F), and old QSCs maintain relatively high access to most quiescence genes (Figure 6E). Of interest, KEGG pathway analysis between young and old QSCs shows that, apart from cellular senescence, peaks associated with cancer and metabolic pathways are highly enriched in old QSCs (Figure S8B). We also observed a dramatic reduction in accessibility across the Pax7 promoter region between young QSCs and old QSCs (Figures 6E and 6F), as well as a decrease in chromatin accessibility across Pax7 E-intronic and E-distal regions in old QSCs (Figure S8D). On the contrary, the difference between young and old FISCs or ASCs is less dramatic (Figure S8A). However, there is no strong correlation between chromatin accessibility and gene expression during aging (Figure S8E). Although most of the genes have their chromatin accessibility increased in old SCs, the gene expression does not have an overall increase. Old SCs have been reported to cycle more frequently during homeostasis (Chakkalakal et al., 2012). Our data indicate that this phenomenon can be resulted from increased accessibility of cell cycle genes, leading to precocious activation from a quiescence state.

The relative genomic distribution of ATAC peaks was similar among old SCs (Figure S8F). Notably, old QSCs have an evident increase in ATAC-seq peaks in distal regions (Figures S1D and S8G). Previous studies suggest that the active chromatin states across regulatory elements such as promoters and enhancers contribute to transcriptional bursting, essentially leading to transcriptional noise (Urban and Johnston, 2018). To determine the chromatin environment changes in regulatory elements during aging, we utilized

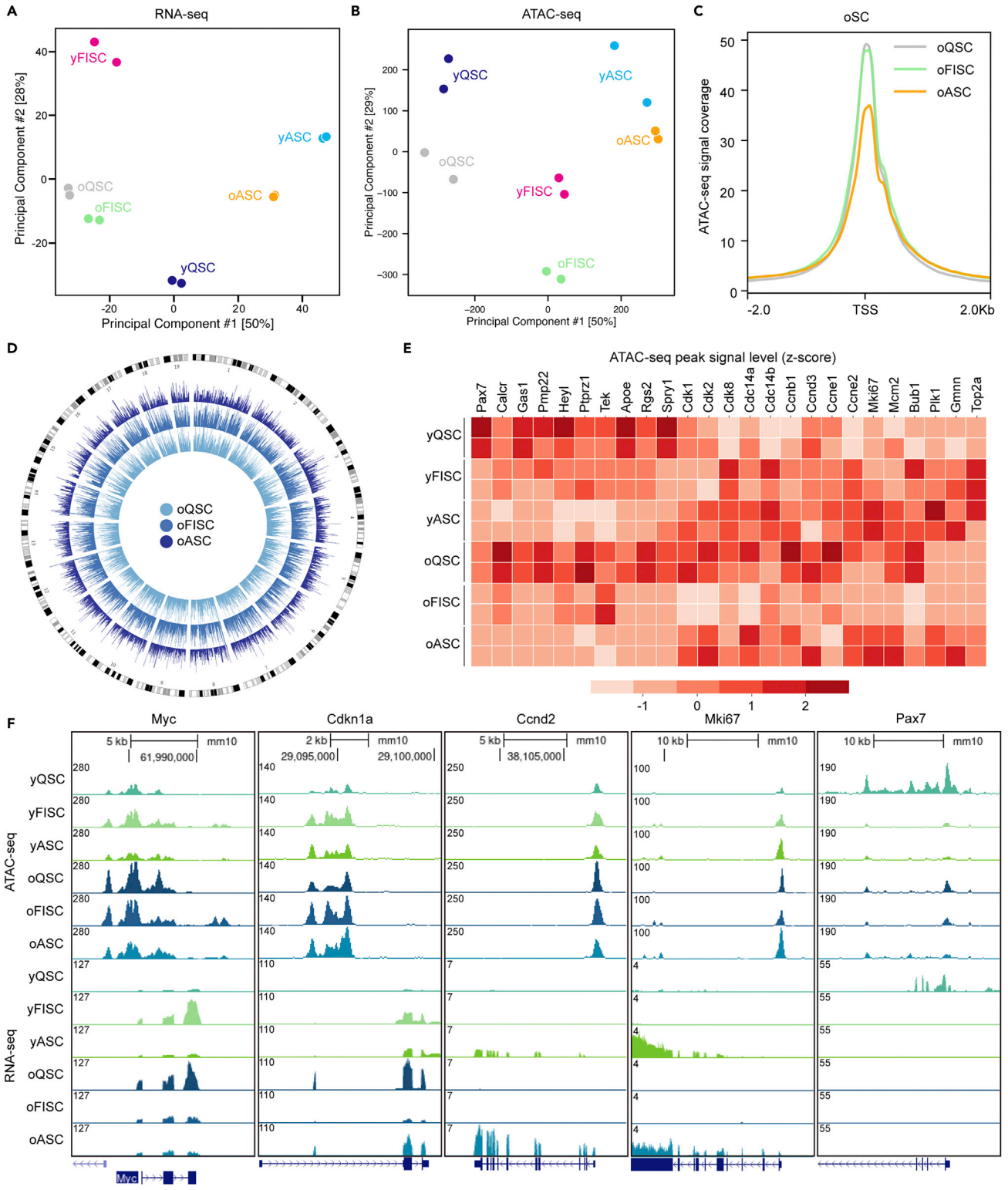


Figure 6. Old satellite cells have a highly open chromatin environment and exhibit a chronic activation chromatin signature (See also Figures S7 and S8, and Table S3)

(A and B) Principal component analysis (PCA) plots of RNA-seq (A) and ATAC-seq (B) datasets of young and old quiescent satellite cells (QSCs), freshly isolated satellite cells (FISCs), and activated satellite cells (ASCs).

Figure 6. *Continued*
(C) ATAC-seq signal coverage across transcription start sites (TSS) in old QSCs, FISCs, and ASCs.
(D) The Circos plot of ATAC-seq signal coverage across mouse genome for old QSCs, FISCs, and ASCs.
(E) Heatmap showing ATAC-seq signal coverage across the promoter regions of quiescence and activation genes in young and old SCs.
(F) Genome tracks of ATAC-seq and RNA-seq signal coverage across different gene loci in young and old SCs.

ChIP-seq data for H3K4me3 and H3K27ac histone marks, which mark active promoters and enhancers respectively (Machado et al., 2017). We extracted enriched regions for these two histone marks in T0 QSCs, FISCs, and ASCs and assessed their chromatin accessibility in young and old SCs (Figures S8H and S8I). The genome accessibility is globally increased, as indicated by upregulated ATAC-seq signals across all H3K4me3 and H3K27ac peaks (Figures S8H and S8I). This result suggests that regulatory elements have pronounced alterations in their chromatin environment during aging. To determine whether there are alterations to the distribution of gene expression, we calculated the distribution of Transcripts Per Kilobase Million (TPM) in young and old QSCs and ASCs and found that old SCs had a significantly higher proportion of genes with low TPM (Figure S8J). This "gene expression leakage" phenomenon is consistent with previous senescence RNA-seq profiles (Zhang et al., 2021), suggesting that dysregulated chromatin accessibility profile in old SCs leads to aberrant gene expression, essentially resulting in transcriptional noise. Moreover, when examining the expression profile of genes highly expressed in yQSCs, we found that the oQSCs resemble more with yFISCs than yQSCs (Figure S8K), indicating that SCs have lost quiescence features during aging. Taken together, our data showed that old SCs have lost the compact chromatin structure in quiescence and exhibit a chronic activation state, which might lead to loss of quiescence signature in transcriptomic profile and essentially result in dysfunction in regeneration.

DISCUSSION

Skeletal muscle possesses remarkable regeneration capacity owing to its resident stem cells, muscle stem cells. The transcriptional regulatory network of SC quiescence and activation has been extensively investigated during the past decades (Relaix et al., 2021). Although many efforts have been made in investigating the chromatin organization in muscle stem cells (Dong and Cheung, 2021), the chromatin environment during SC quiescence and activation remains to be explored. This study gained insights into chromatin accessibility and transcriptomic profiles in SC quiescence and aging. We also identified distal regulatory elements for Pax7 expression, which are indispensable for proper SC activation.

Our study revealed the trajectory of chromatin accessibility changes during SC quiescence exit and regeneration. Consistent with the previous finding that the transcriptome changes rapidly on SC activation (Machado et al., 2021), the chromatin accessibility also changes dramatically in as early as 30 min (Figures 2B and 2D). We found that 4 hpi *in vivo* activated SCs have the most open chromatin environment (Figure 2D), suggesting that SCs are fully prepared for activation in as short as 4 h. Notably, FISCs resemble the 4 hpi and 8 hpi *in vivo* activated SCs (Figure S2D), further indicating that FISCs are already activated. However, it takes a much longer time for activated SCs to re-establish the quiescent chromatin environment during self-renewal (Figures 3B, 3D, and 3E). Previous studies have suggested that stimuli-induced activation leaves "epigenetic memory" on resident stem cells (Larsen et al., 2021). It is possible that injury-induced activation will leave this epigenetic memory on SCs for the regeneration process. Examining the epigenetic profiles of SCs after the regeneration is fully completed (i.e., months after activation) will be informative. Another possibility is that SCs are involved during the regeneration process and actively communicate with other cells within the niche environment during the re-constitution of the quiescent niche. Further investigation of the impact of cell-cell communication on the chromatin environment of the resident stem cells will benefit our understanding of the regulation of SC regeneration.

By integrating ATAC-seq and RNA-seq data during SC quiescence exit and regeneration, we have identified multiple TFs that have their motifs accessibility differentially changed in SC quiescence and activation (Figures 3G and 3H). The previous study has shown that FoxO maintains SCs in a genuine quiescent state (García-Prat et al., 2020). Our data also indicates that TFs from the Fox family (e.g., Foxk1, Foxo1, and Foxp2) have high expression levels in QSCs, and their motifs are highly accessible during quiescence (Figures 3G and 3H). This result suggests the role of Fox family members in maintaining SC quiescence, possibly through binding to their target regions. Of interest, not all TFs with motifs accessible in quiescence are highly enriched in QSCs. TFs such as Irf7, Elk4, and Smarca5 only start to express when SCs are activated (Figures 3G and 3H), suggesting that there is a hierarchy of TF networks, and these TFs may benefit from the chromatin remodeling events resulting from primary/pioneer TFs.

SCs preserve a relatively closed heterochromatin structure to maintain quiescence (Boonsanay et al., 2016). Chromatin compaction and the kinetics of heterochromatin opening are rate-limiting factors of transcriptional bursting, which will lead to transcriptional noise (Hendy et al., 2017). Previous studies have examined chromatin accessibility in freshly isolated and injured old SCs (García-Prat et al., 2020; Shcherbina et al., 2020). Our data showed that old QSCs had lost their ability to maintain a compact chromatin structure, as indicated by a dramatic increase in chromatin accessibility (Figure 6). Previous findings indicate that H3K27ac around the promoter may act as a possible regulator for transcriptional bursting (Nicolas et al., 2018). We also observed a significant increase in chromatin accessibility across H3K27ac peaks in old SCs (Figure S8I), indicating that the distal regulatory elements undergo pronounced alterations in their chromatin environment during aging. A previous study has found that old SCs enter the cell cycle more frequently during homeostasis (Chakkalakal et al., 2012). Our data further explored this aspect at the chromatin accessibility level. ATAC-seq peaks near cell cycle-related genes are specifically enriched in oQSCs (Figures 6E and 6F), indicating that the increased chromatin accessibility may be the underlying cause of quiescence interruption. We also observed an enrichment of ATAC-seq peaks associated with cellular senescence (Figure 6E). Of interest, KEGG pathway analysis between young and old QSCs shows that genes associated with cancer and metabolic pathways are also highly accessible in old QSCs (Figure S8B). Further characterization of enriched peaks from cancer pathways shows that they are highly associated with metabolic pathways (Figure S8C). Age-associated declines in stem cell function are characterized by epigenetic and metabolic changes (Ren et al., 2017). Our result further indicates that the high chromatin accessibility of metabolic pathway genes is also a strong indicator of stem cell aging. Taken together, our data suggest that the loss of compact chromatin arrangement and acquisition of chronically activated chromatin states are the hallmarks of stem cell aging, which might be responsible for cellular dysfunction, especially in quiescence maintenance and regeneration.

One of the interesting findings of our study is the identification of two enhancer elements for Pax7 (Figures 4 and 5). Pax7 is indispensable for SC quiescence and activation (Lepper et al., 2011; Von Maltzahn et al., 2013; Oustanina et al., 2004; Sambasivan et al., 2011; Seale et al., 2000). Depletion of Pax7 in SCs leads to cell-cycle arrest and precocious differentiation (Von Maltzahn et al., 2013). Although efforts have been made to identify the regulatory elements for Pax7, little is known about the precise regulation of Pax7 expression. Through combinational analysis of chromatin accessibility and histone modification profiles, we identified two potential enhancer elements for Pax7 (Figure 4C). Using the CRISPR-Cas9 system, we validated the requirement of the enhancer elements for proper SC function (Figure 5). Remarkably, the loss-of-function experiment of these two enhancer elements reproduces the phenotype of Pax7 deficient SCs, suggesting the importance of the enhancer elements in regulating SC activation. Of interest, in the dual-luciferase assay analysis, only the E-intronic region can increase luciferase activity in C2C12 cells (Figure S5B). This result suggests that the activity of the enhancer elements is dependent on cell status and possible help from regulatory proteins. Distal regulatory elements have been reported to be more cell-type-specific than promoter regions (Corces et al., 2016). Further exploration of our ATAC-seq data may lead to the identification of distal regulatory elements for SC quiescence maintenance and activation.

In summary, our study provides a database of the chromatin accessibility and transcriptome profiles of SCs in quiescence and aging. We showed that the chromatin environment of SCs is very compact during quiescence, becomes highly accessible on early activation, and gradually re-establish the compact state after long-term regeneration. Of interest, we found that aged SCs exhibit a chronic activation chromatin structure. We also identified two functional enhancer elements for Pax7 expression, both of which are indispensable for proper SC activation. Taken together, our data demonstrate the dynamics of chromatin accessibility changes in SC quiescence and aging and can be further utilized to identify key distal regulatory elements for SC quiescence maintenance and activation.

Limitations of the study

In the current study, we only described the changes in chromatin accessibility during SC isolation and activation. Further investigation on the mechanism inducing the isolation-sensitive peaks, as well as exploration of the sensitivity of the nucleosome distribution and positioning during the isolation process would benefit our understanding of how chromatin is altered during muscle stem cell isolation. For identifying distal regulatory elements, we only focused on the Pax7 locus as a proof-of-concept example that the ATAC-seq dataset can be utilized to identify potential regulatory elements. Further identification of other regulatory elements can be achieved through a combination of ATAC-seq data and active histone mark

profiles (e.g., H3K27ac mark). Moreover, we only tested the enhancer activity in the *in vitro* system. The *in vivo* manipulation of Pax7 enhancer activity will provide insights into regulations of SC quiescence maintenance.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animals and treatments
  - Cell culture
  - Satellite cell isolation
- METHOD DETAILS
  - RNA-seq
  - ATAC-seq
  - ChIPmentation
  - Plasmid construction
  - Dual-luciferase assay
  - sgRNA transfection
  - EdU incorporation analysis
  - Immunofluorescence
  - ATAC-seq alignment
  - ATAC-seq replicate correlation and peak calling
  - ATAC-seq signal pre-processing
  - ATAC-seq peak annotation
  - ATAC-seq peak fuzzy clustering
  - Union peak set of time course
  - Gene expression analysis in RNA-seq
  - Motif enrichment analysis
  - ChIP-seq analysis
  - Integrating ATAC-seq and RNA-seq data
  - Cluster purity
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104954.

ACKNOWLEDGMENTS

We thank S. Tajbakhsh for providing Pax7-nGFP mice. We thank T.W. Fung for the assistance with the Graphical Abstract. This work was supported by research grants from the Hong Kong Research Grant Council (GRF16102319, GRF16102420, C6018-19G, C6027-19G, AoE/M-604/16, T13-605/18W), the Lee Hysan Foundation (LHF17SC01), and the Hong Kong Epigenome Project (Lo Ka Chung Charitable Foundation). This study was supported in part by the Innovation and Technology Commission (ITCPD/17-9). T.H.C. is the S. H. Ho Associate Professor of Life Science at the Hong Kong University of Science and Technology.

AUTHOR CONTRIBUTIONS

Conceptualization, A.D. and T.H.C.; Experiments, A.D., K.L., W.Z., and W.K.S., and S. H.; Computational analysis, A.D., J.L., and S. H.; Funding Acquisition, T.H.C.; Writing – Original Draft, A.D.; Writing – Review and Editing, A.D., K.L., S. H., and T.H.C.; Supervision, T.H.C.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 11, 2021
Revised: June 30, 2022
Accepted: August 12, 2022
Published: September 16, 2022

REFERENCES

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Boonsanay, V., Zhang, T., Georgieva, A., Kostin, S., Qi, H., Yuan, X., Zhou, Y., and Braun, T. (2016). Regulation of skeletal muscle stem cell quiescence by suv4-20h1-dependent facultative heterochromatin formation. Cell Stem Cell *18*, 229–242.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

Buske, F.A., Bodén, M., Bauer, D.C., and Bailey, T.L. (2010). Assigning roles to DNA regulatory motifs using comparative genomics. Bioinformatics *26*, 860–866.

Chakkalakal, J.V., Jones, K.M., Basson, M.A., and Brack, A.S. (2012). The aged niche disrupts muscle stem cell quiescence. Nature *490*, 355–360.

Chen, X., Shen, Y., Draper, W., Buenrostro, J.D., Litzenburger, U., Cho, S.W., Satpathy, A.T., Carter, A.C., Ghosh, R.P., East-Seletsky, A., et al. (2016). ATAC-see reveals the accessible genome by transposase-mediated imaging and sequencing. Nat. Methods *13*, 1013–1020.

Cheung, T.H., and Rando, T.A. (2013). Molecular regulation of stem cell quiescence. Nat. Rev. Mol. Cell Biol. *14*, 329–340.

Chow, L.S., Bosnakovski, D., Mashek, D.G., Kyba, M., Perlingeiro, R.C.R., and Magli, A. (2021). Chromatin accessibility profiling identifies evolutionary conserved loci in activated human satellite cells. Stem Cell Res. *55*, 102496.

Collins, C.A., Olsen, I., Zammit, P.S., Heslop, L., Petrie, A., Partridge, T.A., and Morgan, J.E. (2005). Stem cell function, self-renewal, and behavioral heterogeneity of cells from the adult muscle satellite cell niche. Cell *122*, 289–301.

Corces, M.R., Buenrostro, J.D., Wu, B., Greenside, P.G., Chan, S.M., Koenig, J.L., Snyder, M.P., Pritchard, J.K., Kundaje, A., Greenleaf, W.J., et al. (2016). Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat. Genet. *48*, 1193–1203.

Dong, A., and Cheung, T.H. (2021). Deciphering the chromatin organization and dynamics for muscle stem cell function. Curr.Opin. Cell Biol. *73*, 124–132.

Dong, A., Preusch, C.B., So, W.K., Lin, K., Luan, S., Yi, R., Wong, J.W., Wu, Z., and Cheung, T.H. (2020). A long noncoding RNA, LncMyoD, modulates chromatin accessibility to regulate muscle stem cell myogenic lineage progression. Proc. Natl. Acad. Sci. USA *117*, 32464–32475.

Frankish, A., Diekhans, M., Ferreira, A.M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res. *47*, D766–D773.

Frazee, A.C., Pertea, G., Jaffe, A.E., Langmead, B., Salzberg, S.L., and Leek, J.T. (2015). Ballgown bridges the gap between transcriptome assembly and expression analysis. Nat. Biotechnol. *33*, 243–246.

García-Prat, L., Perdiguero, E., Alonso-Martín, S., Dell'Orso, S., Ravichandran, S., Brooks, S.R., Juan, A.H., Campanario, S., Jiang, K., Hong, X., et al. (2020). FoxO maintains a genuine muscle stem-cell quiescent state until geriatric age. Nat. Cell Biol. *22*, 1307–1318.

Hendy, O., Campbell, J., Weissman, J.D., Larson, D.R., and Singer, D.S. (2017). Differential context-specific impact of individual core promoter elements on transcriptional dynamics. Mol. Biol. Cell *28*, 3360–3370.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat. Methods *12*, 357–360.

Köster, J., and Rahmann, S. (2012). Snakemake—a scalable bioinformatics workflow engine. Bioinformatics *28*, 2520–2522.

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

Larsen, S.B., Cowley, C.J., Sajjath, S.M., Barrows, D., Yang, Y., Carroll, T.S., and Fuchs, E. (2021). Establishment, maintenance, and recall of inflammatory memory. Cell Stem Cell *28*, 1758–1774.e8.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS Comput. Biol. *9*, e1003118.

Lepper, C., Partridge, T.A., and Fan, C.M. (2011). An absolute requirement for Pax7-positive satellite cells in acute injury-induced skeletal muscle regeneration. Development *138*, 3639–3646.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics *25*, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; Subgroup, 1000 Genome Project data processing (2009). the sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079.

Liu, L., Cheung, T.H., Charville, G.W., and Rando, T.A. (2015). Isolation of skeletal muscle stem cells by fluorescence-activated cell sorting. Nat. Protoc. *10*, 1612–1624.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550–621.

Machado, L., Esteves de Lima, J., Fabre, O., Proux, C., Legendre, R., Szegedi, A., Varet, H., Ingerslev, L.R., Barrès, R., Relaix, F., and Mourikis, P. (2017). In situ fixation redefines quiescence and early activation of skeletal muscle stem cells. Cell Rep. *21*, 1982–1993.

Machado, L., Geara, P., Camps, J., Dos Santos, M., Teixeira-Clerc, F., Van Herck, J., Varet, H., Legendre, R., Pawlotsky, J.M., Sampaolesi, M., et al. (2021). Tissue damage induces a conserved stress response that initiates quiescent muscle stem cell activation. Cell Stem Cell *28*, 1125–1135.e7.

Von Maltzahn, J., Jones, A.E., Parks, R.J., and Rudnicki, M.A. (2013). Pax7 is critical for the normal function of satellite cells in adult skeletal muscle. Proc. Natl. Acad. Sci. USA *110*, 16474–16479.

Mauro, A. (1961). Satellite cell of skeletal muscle fibers. J. Biophys. Biochem.Cytol. *9*, 493–495.

Nicolas, D., Zoller, B., Suter, D.M., and Naef, F. (2018). Modulation of transcriptional burst frequency by histone acetylation. Proc. Natl. Acad. Sci. USA *115*, 7153–7158.

Oustanina, S., Hause, G., and Braun, T. (2004). Pax7 directs postnatal renewal and propagation of myogenic satellite cells but not their specification. EMBO J. *23*, 3430–3439.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.-C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. *33*, 290–295.

Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat. Protoc. *11*, 1650–1667.

Probst, A.V., Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. Nat. Rev. Mol. Cell Biol. *10*, 192–206.

Relaix, F., and Zammit, P.S. (2012). Satellite cells are essential for skeletal muscle regeneration: the cell on the edge returns centre stage. Development *139*, 2845–2856.

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

Relaix, F., Bencze, M., Borok, M.J., Der Vartanian, A., Gattazzo, F., Mademtzoglou, D., Perez-Diaz, S., Prola, A., Reyes-Fernandez, P.C., Rotini, A., and Taglietti, , 5th. (2021). Perspectives on skeletal muscle stem cells. Nat. Commun. *12*, 692–711.

Ren, R., Ocampo, A., Liu, G.-H., Carlos, J., and Belmonte, I. (2017). Regulation of stem cell aging by metabolism and epigenetics. Cell Metabol. *26*, 460–474.

Rumman, M., Dhawan, J., and Kassem, M. (2015). Concise Review: quiescence in adult stem cells: biological significance and relevance to tissue regeneration. Stem Cell. *33*, 2903–2912.

Sambasivan, R., Yao, R., Kissenpfennig, A., van Wittenberghe, L., Paldi, A., Gayraud-Morel, B., Guenou, H., Malissen, B., Tajbakhsh, S., and Galy, A. (2011). Pax7-expressing satellite cells are indispensable for adult skeletal muscle regeneration. Development *138*, 3647–3656.

Schep, A.N., Buenrostro, J.D., Denny, S.K., Schwartz, K., Sherlock, G., and Greenleaf, W.J. (2015). Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. Genome Res. *25*, 1757–1770.

Schmidl, C., Rendeiro, A.F., Sheffield, N.C., and Bock, C. (2015). ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. Nat. Methods *12*, 963–965.

Seale, P., Sabourin, L.A., Girgis-Gabardo, A., Mansouri, A., Gruss, P., and Rudnicki, M.A. (2000). Pax7 is required for the specification of myogenic satellite cells. Cell *102*, 777–786.

Shcherbina, A., Larouche, J., Fraczek, P., Yang, B.A., Brown, L.A., Markworth, J.F., Chung, C.H., Khaliq, M., de Silva, K., Choi, J.J., et al. (2020). Dissecting murine muscle stem cell aging through regeneration using integrative genomic analysis. Cell Rep. *32*, 107964.

Sitbon, D., Podsypanina, K., Yadav, T., and Almouzni, G. (2017). Shaping chromatin in the nucleus: the bricks and the architects. Cold Spring Harb.Symp. Quant. Biol. *82*, 1–14.

Stark, R., and Version, G.B. (2011). DiffBind: differential binding analysis of ChIP-Seq peak data. R Packag. Version *100*, 4.3.

Urban, E.A., and Johnston, R.J. (2018). Buffering and amplifying transcriptional noise during cell fate specification. Front. Genet. *9*, 591.

van Velthoven, C.T.J., de Morree, A., Egner, I.M., Brett, J.O., and Rando, T.A. (2017). Transcriptional profiling of quiescent muscle stem cells in vivo. Cell Rep. *21*, 1994–2004.

Weissman, I.L. (2000). Stem cells: units of development, units of regeneration, and units in evolution. Cell *100*, 157–168.

Yu, G., Wang, L.-G., and He, Q.-Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics *31*, 2382–2383.

Yue, L., and Cheung, T.H. (2020). Protocol for isolation and characterization of in situ fixed quiescent muscle stem cells. STAR Protoc. *1*, 100128.

Yue, L., Wan, R., Luan, S., Zeng, W., and Cheung, T.H. (2020). Dek modulates global Intron retention during muscle stem cells quiescence exit. Dev. Cell. *53*, 661–676.e6.

Zammit, P.S., Golding, J.P., Nagata, Y., Hudon, V., Partridge, T.A., and Beauchamp, J.R. (2004). Muscle satellite cells adopt divergent fates. J. Cell Biol. *166*, 347–357.

Zhang, X., Liu, X., Du, Z., Wei, L., Fang, H., Dong, Q., Niu, J., Li, Y., Gao, J., Zhang, M.Q., et al. (2021). The loss of heterochromatin is associated with multiscale three-dimensional genome reorganization and aberrant transcription during cellular senescence. Genome Res. *31*, 1121–1135.

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

Zhou, H., Liu, J., Zhou, C., Gao, N., Rao, Z., Li, H., Hu, X., Li, C., Yao, X., Shen, X., et al. (2018). In vivo simultaneous transcriptional activation of multiple genes in the brain using CRISPR–dCas9-activator transgenic mice. Nat. Neurosci. *21*, 440–446.

Zhu, L.J., Gazin, C., Lawson, N.D., Pagès, H., Lin, S.M., Lapointe, D.S., and Green, M.R. (2010). ChIPpeakAnno: a bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinf. *111*, 1–10.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-rabbit H3K27ac	Active Motif	Cat# 39133; RRID:AB_2561016
Anti-mouse Pax7	This Study	N/A
Anti-mouse Myogenin	BD Biosciences	Cat# 556358; RRID:AB_396383
Chemicals, peptides, and recombinant proteins		
40,6-diamidino-2-phenylindole (DAPI)	Thermo Fisher Scientific	Cat# D1306; RRID:AB_2629482
Tamoxifen	Sigma	Cat# T5648
DharmaFECT 1 Transfection Reagent	Dharmacon	Cat# T-2001-03
NEBNext High-Fidelity 2x PCR Master Mix	New England Biolabs	Cat# M0541S
Agencourt AMPure XP	Beckman Coulter	Cat# A63881
Critical commercial assays		
Dual-Luciferase Reporter Assay System	Promega	Cat# e1960
miRNeasy FFPE Kit	Qiagen	Cat# 217504
Deposited data		
Raw and processed fixed cells' ATAC-seq and RNA-seqdata	This study	GEO: GSE189074
QSC, FISC, and ASC RNA-seqdata	Yue et al. (2020)	GEO: GSE113631
FISC and ASC ATAC-seqdata	Dong and Cheung (2021)	GEO: GSE129745
Raw and processedold SC ATAC-seq and RNA-seqdata	This study	GEO: GSE189074
Raw and processed FISC and ASC H3K27ac ChIP-seq data	This study	GEO: GSE189074
Experimental models: cell lines		
Mouse: C2C12	ATCC	CRL-1772; RRID: CVCL_0188
Human: HEK-293T	ATCC	CRL-3216; RRID: CVCL_0063
Satellite cell culture	This study	N/A
Experimental models: Organisms/strains		
Mouse: C57BL/6	Laboratory Animal Facility of HKUST	N/A
Mouse: Tg: Pax7-nGFP	A kind gift from Shahragim Tajbakhsh (Institut Pasteur)	N/A
Mouse: Tg: CAG-LSL-dCas9-SPH	Shanghai Model Organisms Center, Inc.	NM-TG-00025
Mouse: Rosa26$^{LSL-Cas9}$	JAX	JAX: 026175
Mouse: Pax7^{CreERT2}	JAX	JAX: 017763
Mouse: R26R^{YFP}	JAX	JAX: 006148

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oligonucleotides		
Sequences of primers for recombinant DNA constructs	See Table S4	N/A
SgRNA	Dharmacon	N/A
Sequences of sgRNAs	See Table S4	N/A
Recombinant DNA		
pGL3.0-Basic	Promega	Cat# E1751
pGL3.0-TK	This study	N/A
pGL3.0-TK-E-intronic	This study	N/A
pGL3.0-TK-E-distal	This study	N/A
Software and algorithms		
GraphPad Prism 6	GraphPad Software	N/A
ZEN software (blue edition)	Carl Zeiss	N/A
ImageJ	N/A	https://imagej.nih.gov/ij/
Picard v2.9.2	N/A	https://broadinstitute.github.io/picard/
Trimmomatic v0.36	Bolger et al. (2014)	https://github.com/timflutre/trimmomatic
UCSC bedGraphToBigWig v332	N/A	https://github.com/ENCODEDCC/kentUtils
Bedtools v2.26.0	Quinlan and Hall, 2010	http://bedtools.readthedocs.io/en/latest/
Bowtie2 v2.3.2	Langmead and Salzberg (2012)	http://bowtiebio.sourceforge.net/bowtie2/index.shtml
HISAT2 v2.1.0	Kim et al. (2015)	https://ccb.jhu.edu/software/hisat2/index.shtml
bwa v0.7.17	Li and Durbin (2009)	http://bio-bwa.sourceforge.net/
MACS2 v2.1.1	Zhang et al. (2008)	https://github.com/taoliu/MACS/
Samtools v1.3	Li and Durbin (2009)	http://samtools.sourceforge.net/
Snakemake v4.6.0	Köster and Rahmann, 2012	https://bitbucket.org/snakemake/snakemake/wiki/Home
StringTie v1.3.3	Pertea et al. (2015)	https://ccb.jhu.edu/software/stringtie/
Ballgown v2.2.0	Frazee et al., 2015	https://github.com/alyssafrazee/ballgown
DiffBind v3.4.0	Stark and Version, 2011	https://bioconductor.org/packages/release/bioc/html/DiffBind.html
ChIPpeakAnno v3.26.0	Zhu et al.(2010)	http://bioconductor.org/packages/release/bioc/html/ChIPpeakAnno.html
GenomicAlignments v1.28.0	Lawrence et al. (2013)	https://bioconductor.org/packages/release/bioc/html/GenomicAlignments.html
GenomicRanges v1.44.0	Lawrence et al. (2013)	https://bioconductor.org/packages/release/bioc/html/GenomicRanges.html
ChIPSeeker v1.30.0	Yu et al. (2015)	https://bioconductor.org/packages/release/bioc/html/ChIPseeker.html
Fuzzy-c-means v1.6.3	N/A	https://git.io/fuzzy-c-means
DESeq2 v1.32.0	Love et al. (2014)	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
MEME v5.0.2	Buske et al. (2010)	https://meme-suite.org/meme/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Tom H. Cheung (tcheung@ust.hk).

Materials availability

This study did not generate new unique reagents.

Data and code availability

ATAC-seq, RNA-seq, and ChIP-seq data have been deposited at GEO and are publicly available. Accession numbers are listed in the key resources table.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals and treatments

Mice used in this study were C57BL/6 strains around 2 to 6-month-old (young) or around 24-month-old (old), housed and maintained in the Laboratory Animal Facility (LAF) of the Hong Kong University of Science and Technology (HKUST). Mice used in each experiment were age- and gender-matched. C57BL/6 mice were obtained from the LAF of HKUST. Labeling of the old SCs was achieved using the $Pax7^{CreERT2/+}$;$R26R^{YFP/+}$ mice. The mice were injected with tamoxifen when 2 months old for 2 weeks and sacrificed around 24 months old. Tamoxifen injection for Cre recombinase activation was performed via intraperitoneal injection to the mouse in a dosage of 0.2 mg of tamoxifen per gram of mouse, with a frequency of one injection every 3 days, for a total of 5 injections. For perfusion, mice were first anesthetized using 2,2,2-Tribromoethanol (250 mg/kg, Sigma Aldrich) through intraperitoneal injection, then perfusion was performed with a 25-gauge needle (Terumo) attached to a 10 mL syringe (Terumo) through the left cut ventricle. The anesthetized animal was first transcardially washed with 30 mL pre-chilled PBS and then fixed with 30 mL pre-chilled 0.5% PFA (diluted in PBS) for 5 min. Afterward, 30 mL pre-chilled 2 M glycine (Sigma, diluted in PBS) was perfused to quench fixation. For muscle injury, mice were anesthetized using 2,2,2-Tribromoethanol (250 mg/kg, Sigma Aldrich) through intraperitoneal injection. Muscle injury was induced by injecting 50 μL of 1.2% barium chloride (Sigma) into the lower hindlimb muscles for each hindlimb. All animal experiments were approved by the HKUST Animal Ethics Committee.

Cell culture

C2C12 myoblasts, HEK-293T and HeLa cells were cultured in Dulbecco's modified Eagle medium (DMEM; Fisher Scientific) supplemented with 10% fetal bovine serum (FBS, Invitrogen) and penicillin-streptomycin (P/S, Invitrogen). Freshly isolated satellite cells (SCs) were cultured in Ham's F10 (F10, Sigma) supplemented with 10% horse serum (HS, Invitrogen) and P/S.

Satellite cell isolation

Satellite cells (SCs) were isolated as previously described (Liu et al., 2015). In brief, mouse hindlimb muscles were dissected and minced carefully. Muscles were digested with collagenase II (800 U/mL) in Ham's F-10 with 10% horse serum (washing medium) at 37°C for 90 min. Digested muscles were then washed and incubated with dispase (1.1 U/mL) and collagenase II (100 U/mL) at 37°C for 30 min for the second digestion. Muscle suspension was syringed 10 times through a 20-G needle and subsequently filtered through a 40-μm strainer. SCs were purified from the cell suspension through nGFP/YFP reporter using an Influx cell sorter.

METHOD DETAILS

RNA-seq

To get high-quality RNA from fixed cells, the total RNA from fixed cells was isolated using a miRNeasy FFPE Kit (Qiagen) according to the manufacturer's instructions with one optimization that the digestion of RNA by proteinase K was performed at 56°C for 1 h, following the protocol in Thomsen et al. (2016). RNA concentrations were determined using a Qubit RNA HS Assay Kit (Life Technologies) on a Qubit 2.0 Fluorometer (Life Technologies). RNA (1 ng) was used for first-strand cDNA preparation according to the SMART-seq v4 UltraLow Input RNA Kit for Sequencing (Clontech) protocol and processed for library generation using Tn5-mediated library generation. Sequencing was carried out in 2 × 100 cycle paired-end mode on a DNBSEQ-G400 sequencer (BGI).

ATAC-seq

The ATAC-seq protocol for fixed SCs was adapted from Chen et al. (2016). Approximately 1×10^5 fixed SCs were collected by centrifugation at 500 × g for 5 min at 4°C and then washed in 50 μL cold PBS. Cells were collected by centrifugation at 500 × g for 5 min at 4°C and resuspended in 50 μL cold lysis buffer (10mM

NaCl, 3mM MgCl2, 0.1% NP-40, and 10mM Tris-Cl [pH 7.4]). The cell lysate was collected by centrifugation at 500 × g for 10 min at 4°C. Next, 35 μL UltraPure Distilled Water, 5 μL 5x Buffer, and 5 μL TDE 1 (Vazyme) were added to the cell lysate for tagmentation reaction. Samples were then incubated at 37°C for 30 min, followed by overnight reverse-crosslinking in 50mM Tris-Cl, 1mM EDTA, 1% SDS, 0.2M NaCl, 5ng/ml proteinase K at 65°C. DNA was then purified using an EconoSpin micro-volume DNA/RNA spin column (Epoch Life Science) in a final elution volume of 10 μL. The subsequent PCR was set up with 10 μL transposed DNA, 25 μL NEBNext High-Fidelity 2× PCR Master Mix (NEB), and 1.25 μM ATAC-seq oligos. The PCR reaction was performed as described by Buenrostro et al. (2013) for five cycles and processed for cycle number determination. The amplified libraries were purified using an EconoSpin micro-volume DNA/RNA spin column (Epoch Life Science), followed by incubation with Agencourt AMPure XP beads (Beckman Coulter) for size selection ranging from 100-1000 bp. Libraries were sequenced in 2 × 100 cycle paired-end mode on a DNBSEQ-G400 sequencer (BGI).

ChIPmentation

ChIP-seq was performed according to the ChIPmentation protocol described by Schmidl et al. (2015) with minor modifications. In brief, 8 × 10^4 FACS-isolated SCs were fixed in 0.5% formaldehyde with 10 mg/mL BSA for 10 min at room temperature with rotation to cross-link chromatin proteins and DNA. Cross-linking was stopped by incubation with 0.125 M glycine for 5 min at room temperature. The cells were washed twice with ice-cold phosphate-buffered saline (PBS) containing 10 mg/mL bovine serum albumin (BSA) and lysed in sonication buffer (50 mM Tris-HCl [pH 8.0], 5 mM EDTA, 0.5% SDS, and protease inhibitor cocktail). The lysates were sonicated by a Covaris S220 Focused-ultrasonicator (Covaris) to shear genomic DNA into 200-600 bp fragments. The lysates were centrifuged to remove debris and were then diluted 1:4 in SDS top-up buffer (1.25% Triton X-100, 0.13% sodium deoxycholate, 175 mM NaCl, and protease inhibitor cocktail). Sonicated chromatin was pre-cleared with Dynabead Protein A (Invitrogen) at 4°C for 2 h. Antibody-protein A complexes were prepared by incubating Dynabead Protein A with antibodies against H3K27ac (Active Motif) at 4°C for 3 h. The complexes were incubated with the pre-cleared chromatin overnight at 4°C with rotation to immunoprecipitate the acetylated histones. The immuno-complexes were washed for 5 min at 4°C with the following buffers: RIPA-low-salt wash buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA, 140 mM NaCl, 0.1% SDS, 0.1% sodium deoxycholate, and 1% Triton X-100), RIPA-high-salt wash buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA, 500 mM NaCl, 0.1% SDS, 0.1% sodium deoxycholate, and 1% Triton X-100), RIPA-LiCl wash buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA, 250 mM LiCl, 0.5% Nonidet P-40, and 0.5% sodium deoxycholate), and TE buffer (10 mM Tris-HCl [pH 8.0] and 1 mM EDTA). To perform tagmentation directly on the bead-bound immunoprecipitated DNA, immuno-complexes were tagmented in a 10 μL tagmentation reaction containing 1 μL Tn5 (Vazyme) at 55°C for 10 min. Tn5 transposase was used to cleave double-stranded DNA and ligate adaptors at both ends. Tagmentation was stopped by adding 1 μL 1% SDS. Cross-linking was reversed and the tagmented DNA was eluted in elution buffer (10 mM Tris-HCl [pH 8.0], 5 mM EDTA, 300 mM NaCl, 0.4% SDS, and proteinase K) at 55°C for 1 h, followed by 6-hour elution at 65°C. The eluted DNA was purified using Agencourt AMPure beads (Beckman Coulter). To prepare input DNA for each ChIP sample, the corresponding pre-cleared DNA also underwent cross-link reversal, and 2.5 ng DNA was tagmented at 55°C for 5 min in a 5-μL tagmentation reaction containing 1 μL 10× diluted Tn5. To prepare ChIP and input libraries, tagmented immunoprecipitated DNA and input DNA were amplified by PCR, each with a unique index incorporated. Libraries were selected by size using Agencourt AMPure beads. Fragments from 300-700 bp were pooled and sequenced on a NextSeq 500 sequencer (Illumina).

Plasmid construction

Pax7 E-intronic and E-distal regions were cloned from the genomic DNA of C2C12 cells. The minimal TK promoter was first inserted into pGL3.0 Basic plasmid (Promega). The pGL3.0-TK plasmid was digested by the BamHI restriction enzyme. Pax7 E-intronic and E-distal regions were inserted in pGL3.0-TK using Gibson Assembly with the primers listed in Table S4.

Dual-luciferase assay

The dual-luciferase assay was performed according to the manufacturer's instructions (Promega). In brief, C2C12 myoblasts and HEK-293T cells were collected 48 h after transfection and lysed with 200 μL passive lysis buffer. Next, 20 μL lysate was transferred onto a luminometer plate, and luciferase activity was measured using a luminometer (Berthold Technologies). After adding 50 μL LAR II, Firefly luciferase activity will be measured. After dispensing 50 μL Stop & Glo Reagent, Renilla luciferase activity can be measured

and used as the internal control. By analyzing the relative luciferase activity, the induction of Firefly luciferase could be calculated.

sgRNA transfection

All sgRNAs were designed using the Dharmacon CRISPR DNA Region Designer tool and obtained from Dharmacon. FACS-sorted SCs were transfected with sgRNA listed in Table S4 targeting E-intronic and E-distal one hour after plating down for culture using Dharmafect reagent (Dharmacon) according to the manufacturer's instructions. Cells were cultured for different time points, followed by fixation and staining for further analysis.

EdU incorporation analysis

Freshly isolated SCs were constantly supplied with 10 μM EdU during culture until harvest. Cells were then fixed and stained using the Click-iT EdU Imaging Kit (Invitrogen) according to the manufacturer's instructions.

Immunofluorescence

For immunofluorescence microscopy, cells were grown on chamber slides (SPL Life Science). Cells were rinsed in PBS and fixed in 2% paraformaldehyde for 5 min. Samples were then rinsed in PBS, permeabilized in 0.1% PBST, and blocked with 5% goat serum in 0.1% PBST for 30 min. Samples were then incubated with primary antibodies in 0.1% PBST overnight at 4°C. On the following day, samples were washed in 0.1% PBST and incubated with DAPI and secondary antibodies in 0.1% PBST at room temperature for 30 min. Samples were then washed in 0.1% PBST and mounted for microscopic examination.

ATAC-seq alignment

The Fastq files generated by the sequencer were trimmed using Trimmomatic (Bolger et al., 2014) with parameters 'PE CROP: 50'. After adapter trimming, the reads were mapped to the mm10 genome using Bowtie2 (Langmead and Salzberg, 2012) alignment software with the following settings:–minins 30–maxins 2000–dovetail-k 4–very-sensitive-local. The MarkDuplicates function in Picard was used to mark reads identified as PCR and optical duplicates. All unmapped and duplicate reads as well as reads with a mapping quality <30 were removed from each sample using Samtools (Li et al., 2009) with the following parameters:-F 1804-f 2-q 30.

ATAC-seq replicate correlation and peak calling

Peak calling was performed on each replicate from the previous section using Genrich (v0.6, available at https://github.com/jsh58/Genrich) with the following parameters:-j-y-r-e chrM-v-f. The correlation and PCA plots of replicates were performed using DiffBind (Stark and Version, 2011). After confirming the correlation of the replicates, the reads from different replicates were downsampled to the read depth of the replicate with the minimum reads followed by another round of peak calling using Genrich with the same parameters listed above. The differential MA plots were performed using downsampled replicates with DiffBind. The aligned reads from a sample's two replicates were then combined into one file. Peak calling was done again on the combined reads using Genrich with the same parameters listed above. The combined reads were used for genome track generation and NucleoATAC analysis (Schep et al., 2015).

ATAC-seq signal pre-processing

ATAC-seq read counts for each biological replicate were normalized by calculating RPKM (Reads Per Kilobase per Million), and then the peaks with RPKM greater than 1 in all replicates of at least one time point were kept for downstream analysis. The RPKM of each time point was obtained by averaging the RPKM of all replicates. The coefficient of variation among all time points was calculated on the RPKM of time points.

ATAC-seq peak annotation

The relative genomic region for each peak (Promoters, Intron, Exon, etc.) was determined by overlapping each peak with features defined in the UCSC mm10 known genes using ChIPseeker (Yu et al., 2015). Promoters were defined as 1 kb upstream and downstream of the TSS.

ATAC-seq peak fuzzy clustering

To avoid the impact of the unchanged peaks during the time course, the peaks with the top 30% coefficient of variation were treated as the changing peaks for clustering, and the remaining 70% low variation peaks were filtered out in the clustering and labeled as stable regions. The z-score transformation was applied to the RPKM value of the changing peaks across the time course before the fuzzy c-means clustering, which was performed by the fuzzy-c-means module with the fuzziness parameter set to 1.25.

Union peak set of time course

The merged peak set was generated from the merged ATAC-seq signals of all replicates at each time point. Limiting the merging gap to 100nt, a union peak set was created by merging all peaks of each time point in activation or regeneration time course using the reduce function from GenomicRanges (Lawrence et al., 2013) package, and then summarizeOverlaps function from GenomicAlignments (Lawrence et al., 2013) package was applied to count the reads within each peak. By using TSS as feature type and setting the output to nearestLocation in annotatePeakInBatch of ChIPpeakAnno (Zhu et al., 2010), the peaks in the union peak set were annotated as the genes whose TSS is nearest to the middle of the peak.

Gene expression analysis in RNA-seq

The Fastq files were mapped to the mm10 genome using HISAT2 (Kim et al., 2015) with the default parameters. After read mapping, StringTie (Pertea et al., 2015) was used with the default parameters (Pertea et al., 2016). The GTF files produced by StringTie for every sample were then merged using StringTie in merge mode, again with the default parameters. This merged GTF file was then used with StringTie with the-eB setting to quantify expression levels for all transcripts in each sample with the Gencode vM21 annotation (Frankish et al., 2019). Expression level estimates from StringTie were then used to determine differentially expressed genes between different samples using DESeq2 (Love et al., 2014). The raw read counts of the genes were first filtered when there were no reads in all samples and then normalized by the variance stabilizing transformation in DESeq2. The PCA plot of the normalized count was also drawn by DESeq2.

Motif enrichment analysis

The motif enrichment of each peak cluster was computed by MEME AME (Buske et al., 2010), with shuffle letters as the control sequence when the random seed was 42 and other parameters were set to default. The enriched motifs of a cluster were selected with adjusted *p*-value less than 10^{-6} and presented in more than 20% peak regions of the cluster (true positive rate greater than 20%) while not falsely classified in less than 20% peak regions (false positive rate less than 20%), and the mRNA expression level (TPM) of the transcription factor gene was required to be larger than 3 for all replicates in at least one time point.

ChIP-seq analysis

ChIP and input reads were separately aligned to the mm10 genome using bwa_mem (Li and Durbin, 2009) and then filtered by Picard to remove duplicated reads. Narrow peaks were called with –SPMR and –nomodel and bigwig files were generated using MACS2 (Zhang et al., 2008) and then visualized on the UCSC genome browser. All packages were run with default settings if not specified.

Integrating ATAC-seq and RNA-seq data

The peaks from DiffBind (Stark and Version, 2011) were labeled as significantly changed or not, by FDR<0.01, and then the median value of significantly changed peaks and insignificantly changed ones around a gene were obtained separately, as the differential chromatin accessibility related to a gene. The differential expressed genes were calculated by DESeq2 (Love et al., 2014) and filtered by FDR<0.05. The change in the chromatin accessibility and the mRNA expression level of a gene was drawn in a scatterplot to show the relationship between the chromatin accessibility and the mRNA expression level.

Cluster purity

The cluster purity calculation is adapted from Corces et al. (2016). Briefly, 12 clusters were defined as branches of the dendrogram. Each cluster was assigned to a specific cell state that was most frequent in the cluster. For each cluster, the accuracy of this assignment was measured by counting the number of correctly assigned cell states. The sum of the number of correctly assigned cell states was divided by the total number of cell states to get the cluster purity.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were performed using GraphPad Prism 6 (GraphPad Software). All the replicates in this manuscript are biological replicates. Unless otherwise noted, all error bars indicate standard deviation (SD). No statistical methods were used to predetermine the sample size. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment.