

Multiscale 3D Genome Reorganization during Skeletal Muscle Stem Cell Lineage Progression and Muscle Aging

Huating WANG (✉ huating.wang@cuhk.edu.hk)

The Chinese University of Hong Kong

Yu Zhao

The Chinese University of Hong Kong

Yingzhe Ding

The Chinese University of Hong Kong

Liangqiang He

The Chinese University of Hong Kong

Yuying Li

The Chinese University of Hong Kong

Xiaona Chen

The Chinese University of Hong Kong

Hao Sun

The Chinese University of Hong Kong

Article

Keywords: 3D chromatin architecture, aging, satellite cell lineage progression, Pax7

Posted Date: December 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1190496/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29

**Multiscale 3D Genome Reorganization during Skeletal Muscle Stem Cell Lineage Progression
and Muscle Aging**

Yu Zhao^{1,3,5*}, Yingzhe Ding^{2,4,5}, Liangqiang He^{2,5}, Yuying Li¹, Xiaona Chen¹,
Hao Sun^{2*}, and Huating Wang^{1*}

¹Department of Orthopaedics and Traumatology, Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong, China

²Department of Chemical Pathology, Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong, China

³Molecular Cancer Research Center, School of Medicine, Shenzhen Campus of Sun Yat-sen University, Sun Yat-sen University, Shenzhen, China.

⁴Present address: Centre for Regenerative Medicine and Health, Hong Kong Institute of Science & Innovation, Chinese Academy of Sciences, Hong Kong SAR, China.

⁵These authors contributed equally to this work

*To whom correspondence should be addressed

Address correspondence to: Huating Wang, Room 507A, Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong, China.

Phone: (852) 3763-6047; Fax: (852) 2632-0008; E-mail: huating.wang@cuhk.edu.hk

1 **Abstract**

2 3D genome rewiring is known to influence spatiotemporal expression of lineage-specific genes and
3 cell fate transition during stem cell differentiation and aging processes. Yet it is unknown how 3D
4 architecture remodels and orchestrates transcriptional changes during skeletal muscle stem cell (also
5 called satellite cell, SC) activation, proliferation and differentiation course. Here, using *in situ* Hi-C
6 we comprehensively map the 3D genome topology reorganization at multiscale levels during mouse
7 SC lineage progression and integrate with transcriptional and chromatin signatures to elucidate how
8 3D genome rewiring dictates gene expression program. Specifically, rewiring at compartment level is
9 most pronounced when SC becomes activated. Striking loss in TAD border insulation and chromatin
10 looping also occurs during early activation process. Meanwhile, TADs can also form TAD clusters
11 and super-enhancer containing TAD clusters orchestrate stage-specific gene expression during SC
12 early activation. Furthermore, we elucidate 3D chromatin regulation of key transcription factor,
13 PAX7 and identify cis-regulatory elements that are crucial for local chromatin architecture and *Pax7*
14 expression. Lastly, 3D genome remodeling is profiled in SCs isolated from naturally aging mice,
15 unveiling that geriatric SCs display a prominent gain in long-range contacts and loss of TAD border
16 insulation. Genome compartmentalization and chromatin looping are evidently altered in aged SC
17 while geriatric SC display a more prominent loss in strength of TAD borders. Together, our results
18 implicate 3D chromatin extensively reorganizes at multiple architectural levels and underpin the
19 transcriptome remodeling during SC lineage development and SC aging.

20

21 **Keywords**

22 3D chromatin architecture; aging; satellite cell lineage progression; Pax7

23

1 **Introduction**

2 Skeletal muscle tissue has a remarkable regenerative capacity, which largely depends on the
3 activation of resident muscle stem cells, named satellite cells (SCs)¹. These cells normally reside in a
4 niche beneath the basal lamina of the myofibers in a quiescent state and uniquely marked by the
5 expression of paired box 7 (*Pax7*) protein. Upon injury, SCs rapidly activate to enter into cell cycle,
6 undergo proliferative expansion, differentiate and eventually fuse to form multinucleated myotube
7 cells; these myotubes further mature into myofibers to repair the damaged muscle. A subset of SCs
8 undergo self-renewal and return to the quiescent state to replenish the adult stem cell pool².
9 Deregulated SC activity is linked to the development of many muscle associated diseases, which
10 necessitates the comprehensive understanding of the mechanisms governing SC fate transition³⁻⁵.
11 Gene regulation is the key to understand the molecular mechanisms governing SC function. For
12 example, it is well known that *Pax7* is uniquely expressed in SCs and plays multifaceted roles in
13 regulating SC quiescence, proliferation, and self-renewal⁶. PAX7 protein is abundantly enriched in
14 quiescent SCs but gradually diminishes as cells become activated, proliferating and committed to
15 terminal differentiation. However, the regulatory mechanisms underpinning the temporal expression
16 of *Pax7* are largely elusive.

17 Recent advents in mapping genome-wide 3D chromosomal interactions have led to
18 unprecedented understanding of hierarchical organization of the genome within the 3D nucleus and
19 how gene expression is orchestrated at 3D level⁷⁻¹¹. At intermediate scales (200 kb to 1 Mb), TADs
20 are identified from chromosome contact maps, which correspond to sequences that interact with each
21 other more frequently within the same TAD rather than with sequences outside the TAD. In
22 mammals, TAD borders are often demarcated by CTCF and cohesin complex; perturbation of
23 TAD borders is emerging as a driver of many human diseases¹²⁻¹⁴. At a larger scale (multi-Mb),
24 interactions between TADs are thought to form two main types of compartments termed A and B
25 compartments, which are associated with distant chromatin modifications and gene expression. In
26 contrast to TADs which are largely conserved among different cell types and species, compartments
27 can display high variability. Recent studies also reveal that non-contiguous TAD clusters can act in
28 concert to organize genome topology¹⁵⁻¹⁷. Whereas long-range TAD interactions are enriched in B
29 compartments and correlate with gene repression¹⁵, super-enhancer containing and highly transcribed
30 TADs can also form clusters that span tens of megabases¹⁶. When high-resolution Hi-C data is

1 available, chromatin loops can be further identified in Hi-C contact matrix. In general, two types of
2 chromatin loops are revealed: one is constitutive and demarcates TAD borders; the other is more
3 cell-type specific and is often associated with enhancers. Recent findings suggest the 3D genome is
4 commonly reorganized during stem cell lineage progression and differentiation to ensure the
5 spatiotemporal expression of lineage-specific genes^{7, 10, 18, 19}. For example, during mouse ESC
6 differentiation into cortical neurons, interactions between B compartments increase with a
7 concomitant decrease of interactions within A compartments; promoter-enhancer contacts are
8 established concomitantly with changes in gene expression⁷. Despite the rapidly gained
9 understanding of 3D genome rewiring during stem cell differentiation, the knowledge gaps remain to
10 be filled in SC. In particular, our knowledge on the molecular events governing early steps of
11 quiescence maintenance and activation of SC is scarce because of the difficulty in preserving SCs in
12 its native niche environment^{20, 21}. In fact, the lengthy dissociation steps expose cells to stress and
13 disrupt the physiological niche, which alters SC transcriptome and histone modifications inevitably;
14 fixation prior to isolation has been shown to preserve the quiescent status of SCs^{20, 22}, making it
15 possible to investigate whether the 3D genome organization remodels in the early stages of SC
16 activities and whether the 3D rewiring dictates transcriptional changes.

17 Alterations at 3D levels are associated with aging in many tissues and cells²³⁻²⁶. Skeletal muscle
18 suffers from sarcopenia, i.e. functional and mass decline, during human aging; this can partly be
19 attributed to the decrease in the number and function of SCs^{1, 4, 5, 27, 28}. Elucidating the molecular
20 causes underlying SC aging thus has become the key to better understand sarcopenia and uncover
21 novel approaches for managing and treating muscle weakening conditions during physiological
22 aging. Research from recent years demonstrates that the SC dysfunction during aging process can be
23 ascribed to changes in transcriptional networks and chromatin state^{27, 29-31}. Yet, it is still unknown
24 whether 3D deregulation occurs in aging SC and drives the aging process.

25 Here, we comprehensively map the 3D genome reorganization during SC lineage progression.
26 We reveal the 3D chromatin extensively reorganizes at all scales including compartment, TAD, TAD
27 cluster and chromatin looping levels. Our results demonstrate dynamic chromatin rewiring at
28 different genomic scales, which orchestrates the transcriptome remodeling and SC activities.
29 Furthermore, in depth analysis at *Pax7* locus uncovers entangling enhancer-promoter connections
30 that establish a unique *Pax7* residing TAD and coordinate its expression dynamics during SC

1 differentiation. Deletion of these regulatory elements disrupts 3D genome organization at *Pax7* locus
2 and attenuates *Pax7* expression. Lastly, we map 3D remodeling in SCs isolated from physiologically
3 aging mice and uncover that 3D genome rewiring occurs in aging SCs. Altogether, our findings
4 provide a comprehensive view of 3D genome remodeling during SC lineage progression and SC
5 aging.
6

1 Results

2 Global reorganization of genome compartmentalization during SC lineage development.

3 To delineate the remodeling in genome architecture during SC lineage progression, SCs at various
4 stages were prepared (Fig. 1a and Suppl. Fig. 1a). Freshly isolated SCs (FISCs), were collected from
5 Pax7-nGFP mice³² by fluorescence-activated cell sorting (FACS) sorting as SCs were exclusively
6 marked by nuclear GFP signal in these mice; based on prior report²⁰, these cells were early activating
7 due to the disruption of their niche by the isolation process. To obtain quiescent SCs (QSCs), *in situ*
8 fixation by paraformaldehyde was performed before FACS isolation. Immunofluorescence (IF)
9 staining for PAX7 protein confirmed the purity of the FISCs (91% \pm 4.7% positive for PAX7) and
10 QSCs (94% \pm 8.7% positive for PAX7) (Suppl. Fig. 1b). FISCs were then cultured *in vitro* for 1, 2 or
11 3 days to become activated SCs (ASC-D1), proliferating (ASC-D2) or early differentiating (DSC-D3)
12 cells (Fig. 1a). 3D genome organization in the above cells was then probed by *in situ* Hi-C with two
13 or three biological replicates per sample (Fig. 1a and Suppl. Fig. 1c). Following the standard pipeline
14 of data analysis using HiC-Pro³³, a total of over \sim 2 billion uniquely aligned contacts were obtained
15 from all samples (Suppl. Data 1). The first principal component (PC1) values of the biological
16 replicates of each time point were highly correlated ($R \geq 0.95$) (Suppl. Fig. 1c); and a gradual
17 decrease in PC1 correlation was observed (Suppl. Fig. 1d) as cells progressed from QSC to DSC-D3,
18 suggesting changes in 3D genome organization were highly reproducible for each time point and
19 cumulative during the SC lineage development. The data from the above biological replicates were
20 thus merged for calculation of *cis/trans* interaction ratios; as a result, over 70% of *cis* interactions
21 were observed across all time points (Suppl. Fig. 1e, Suppl. Data 1), reinforcing the high quality of
22 libraries³⁴. A resolution of 5 kb for QSC and FISC, and 15 kb for D1, D2 and D3 were obtained after
23 in-depth sequencing, which rendered us to explore the 3D genome architecture at multi-scales,
24 including compartments, TADs and chromatin loops (Fig. 1a).

25 Overall, comparing the later (D2 and D3) with the early (QSC, FISC and D1) stages we
26 observed a dramatic decrease in the frequency of long-range (> 2 Mb) but an increase in the
27 short-range (< 2 Mb) interactions on all chromosomes (Fig. 1b, c and Suppl. Fig. 1f), implying that
28 from a genome-wide perspective, the 3D genome organization underwent distinct remodeling in the
29 later stages when cells started to proliferate. Genome segregation into A and B compartments was
30 then determined (Suppl. Data 2). As expected, A-type compartments were associated with active

1 histone marks, H3K4me3 and H3K27ac (ChIP-seq data were obtained from Léo Machado *et. al.*²⁰),
2 which were not present in B-type compartments (Suppl. Fig. 1g). Although the overall proportions
3 assigned to A (47%-50%) and B (50%-53%) compartments remained stable throughout the entire
4 lineage progression (Suppl. Fig. 1h), compartmentalization strength (as measured by the average
5 contact enrichment within (A-A, B-B) and between (A-B) compartments) was dynamically altered. A
6 more condensed chromatin state of compartment B was observed at earlier stages (QSC, FISC and
7 D1) with a larger proportion of B-B interactions compared to later stages (D2 and D3) (52% vs 48%)
8 (Fig. 1d) and a significantly higher interaction strength (Fig. 1e); the proportion of A-A interactions
9 and interaction strength, on the contrary, markedly increased at later stages (Fig. 1d-e). For instance,
10 a loss of compartmentalization was observed in FISC vs. QSC as revealed by reduced B-B contact
11 frequencies and increased inter-compartment A-B contacts (Fig. 1f), but recovered A-B compartment
12 segregation was observed at D2 vs. QSC (Fig. 1g). Switching of loci between A and B compartments
13 was examined and only 12% of the genome switched compartment at any two time points with B to
14 A and A to B each accounting for 6% (Suppl. Fig. 1i); interestingly, 32% of these regions were
15 involved in multiple switching events (Suppl. Fig. 1i) and the switched genomic regions exhibited
16 the most substantial PC1 changes (Suppl. Fig. 1j), which is consistent with previous findings⁹.

17 Next, harnessing the existing transcriptomic profiles from QSC²⁰ and the in house generated
18 RNA-seq from FISC, D1, D2 and D3 cells (Suppl. Data 3), the interplay between
19 compartmentalization and gene expression changes was examined (Suppl. Data 3). Consistent with
20 the lineage progression timing, we observed high expression of *Pax7* in QSCs but its level sharply
21 declined in early activating FISCs; *MyoD* was on the other hand absent in QSCs but highly induced
22 in FISCs and maintained the expression at later stages; *MyoG* was only induced when SCs initiated
23 differentiation at D3. *HeyL* was decreased while *Egr1* was induced in FISC vs QSCs (Suppl. Fig. 2a).
24 Overall, the trajectory of genome compartmentalization demonstrated a drastic change occurred at
25 the transition of QSC to FISC, which was not observed on the transcriptomic trajectory (Fig. 1h).
26 *k*-means clustering was then performed based on the PC1 values of the 20% of the genome
27 ($n=2,605$ genes) with compartment switching (Suppl. Data 2). A total of 20 clusters with a broad
28 range of switching dynamics were identified (Fig. 1i, Suppl. Fig. 2b); 6 of them (1II, 1III, 1VII, 2VI,
29 2VIII, and 3IV) demonstrated concomitant changes in compartmentalization and gene expression ($R >$
30 0.8; Fig. 1j), but this correlation was lost at least at one time point in other clusters (Suppl. Fig. 2b).

1 Specifically, 7 out of the 20 clusters displayed compartment changes preceding transcriptional
2 changes, and these clusters involved half of the genes with switched compartment (Suppl. Fig. 2b,
3 cluster 11); only two clusters displayed compartment changes lagging behind gene expression
4 changes whereas 5 clusters displayed both preceding and lagging relationships. For example, as
5 shown in Fig. 1k (top), *Ttn* locus which encodes a contractile and sarcomere structural protein^{35,36},
6 was embedded in B compartment in QSC and switched from B to A at D1 stage which was further
7 strengthened at D3; transcription activation of the *Ttn*, on the other hand, was initiated at D2 and
8 strongly enhanced at D3. In another example, *Chodl* (Fig. 1k, bottom), recently identified as a
9 QSC-signature gene²⁰, showed a sharp decrease of its expression in FISC vs. QSC and completely
10 undetectable at later stages (D1-D3). Accordingly, we noticed a gradual transition from A to B
11 compartment at *Chodl* locus from QSC to FISC; and B-compartment association was enhanced at
12 D1-D3 stages. Furthermore, Gene ontology (GO) analyses demonstrated that genes in all the above
13 20 clusters undergoing compartment switch were associated with metabolic (negative regulation of
14 DNA demethylation, regulation of the apoptotic process, protein auto phosphorylation etc.),
15 secretory functions (extracellular exosome etc.) and developmental processes (cell division & cell
16 cycle, positive regulation of fibroblast proliferation, cell development, cell differentiation etc.)
17 (Suppl. Fig. 2c and Suppl. Data 3). Altogether, the above findings demonstrate that genome
18 compartmentalization is dynamically rewired during SC lineage progression and most pronounced
19 when cells become activated; and compartmentation remodeling is tightly coupled with the
20 transcriptomic reprogramming with compartmentalization changes preceding expression changes in
21 a substantial number of genes.

22

23 **Dynamic TAD organization during SC lineage development**

24 Next, we examined TAD organization during SC lineage development. First, an average of ~3,000–
25 3,400 borders per time point was identified by using Topdom³⁷ (Fig. 2a and Suppl. Data 4).
26 Partitioning of the genome into TADs was largely stable during the lineage progression as >60%
27 TAD borders (n = 2,155) were detected at all stages (Fig. 2a, invariant borders); a comparable
28 number of TAD borders were stably acquired (704/4,892, 14.4%) or lost (779/4,892, 15.9%), thus
29 resulting in a steady number of borders and conserved average TAD size between 720 to 760 kb
30 (Suppl. Fig. 3a). Nevertheless, border reorganization was observed at the transition of each two

1 stages. For example, 535, 95, 116, 66 and 179 TADs rearranged, merged, split, built, or disappeared
2 at the transition from FISC to D1 (Suppl. Fig. 3b). An insulation score (IS) was then calculated to
3 quantitatively measure the local chromatin insulation strength by TAD borders (Suppl. Data 4). In
4 line with previous reports⁹, the number of CTCF binding motif was negatively correlated with the IS
5 for all borders (Suppl. Fig. 3c), indicating CTCF occupancy increased the insulation strength of TAD
6 borders. Quantitative measurement of the borders that were gained (Suppl. Fig. 3d) or lost (Suppl.
7 Fig. 3e) in each stage relative to QSC revealed that borders gained in later stages (D1-D3) displayed
8 an IS decrease in the early stages (FISC and D1) (Suppl. Fig. 3d). Similarly, borders lost in the later
9 stages already exhibited considerable increase of IS in the beginning stages (Suppl. Fig. 3e).
10 Aggregate interaction matrix centered on borders of each stage revealed a genuine loss of insulation
11 across the TAD borders early in FISC vs. QSC (Fig. 2b). IS also increased drastically in FISC vs.
12 QSC but decreased when SCs activated, proliferated and differentiated (D1, D2, D3 vs. FISC) (Fig.
13 2c). Furthermore, compared with qualitative border changes, quantitative changes in IS of the
14 invariant TADs were more evident: hierarchical clustering identified that 52.8% (1,133/2,155) of all
15 invariant borders showed > 50% difference in between each two time points (Fig. 2d). Altogether, the
16 above findings demonstrate that border re-organization gradually occurs during the entire lineage
17 development. Although distinct TAD re-organization is observed in late vs. early activation stages
18 (D1 vs FISC), the most substantial changes in TAD border strength occur in early activation stages
19 (FISC).

20 To further investigate if TAD reprogramming is correlated with the local gene expression, we
21 found that PCA of IS kinetics (Fig. 2e) showed a progressing trajectory grossly resembling the
22 transcriptome and PC1 trajectories determined earlier (Fig. 1h) despite the qualitative gain or loss of
23 TAD borders did not correlate with overall increased (Suppl. Fig. 3f) or decreased (Suppl. Fig. 3g)
24 changes of local gene expression. When examining transcriptional changes at the most dynamic 545
25 borders that increased (Fig. 2f) or decreased (Suppl. Fig. 3h) in quantitative IS measurement ($|$
26 $(IS_{D3} - IS_{QSC}) / IS_{QSC}| \geq 2$), the majority (81.8%, 446/545) were associated with local gene
27 down-regulation (n = 548) or up-regulation (n = 41) ($\log_2FC > 1.5$ between endpoints) (Fig. 2f, Suppl.
28 Fig. 3h). Unexpectedly, IS increase did not appear to proceed the transcriptional changes, instead
29 concomitant changes were observed on many loci. For example, in QSC, the TAD surrounding *Ier2*
30 locus was insulated by a strong border with a low IS; the IS considerably increased in FISC resulting

1 in a merge with an upstream TAD (Fig. 2g) and the transcriptional level increased concomitantly (Fig.
2 2h). Consistently, Hi-C-derived virtual 4C also revealed the increased interaction between the TAD
3 of *Ier2* with the upstream one in FISC vs. QSC (23.4% vs. 16.3%) (Fig. 2i). As further illustrated on
4 *HeyL* locus (Suppl. Fig. 3i), the associated TAD border weakened with increased IS (Suppl. Fig. 3i),
5 which might result in the concomitant decrease of transcriptional level from QSC to FISC (Suppl.
6 Fig 3j).

7 Next, to determine the self-contacting propensity within a given TAD, we calculated a domain
8 score (D-score) to measure the portion of intra-TAD interactions in all *cis* interactions (Suppl. Data
9 4). As previously noted⁹ D-scores were positively associated with gene expression (Suppl. Fig. 3k).
10 Hierarchical clustering revealed that 72% (n = 594) of the invariant TADs displayed D-score change
11 (> 20% difference between each two time points) (Fig. 2k). Additionally, PCA of D-score kinetics
12 revealed a trajectory (Fig. 2l) resembling those of compartmentalization, transcription, and IS.
13 Overall, The D-score kinetics correlated closely with compartmentalization (PC1) changes (R > 0.84;
14 Fig. 2m). Similar to IS, it was also highly correlated with changes in gene expression (R >0.83, Fig.
15 2n). For example, the D score of the TAD associated with *Egr2* locus increased in FISC vs. QSC (Fig.
16 2o), which led to concomitant transcriptional activation of *Egr2* (Fig. 2p). On *Myc* locus, on the other
17 hand, the D score decrease was accompanied by a transcriptional repression in FISC vs. QSC (Fig.
18 2q-r). Altogether the above results suggest that the 3D genome is dynamically rewired at TAD level
19 and impinges on gene expression during SC lineage development.

20

21 **TADs clusters are dynamically reorganized in FISC vs. QSC.**

22 Increasing evidence underscores the importance of large-scale inter-TAD associations that position
23 chromatin domains in the nucleus^{19, 38-40}. For example, Paulsen, J. *et. al*¹⁵ reported a level of
24 developmental genome organization that involves long-range TAD-TAD interactions into assemblies
25 of linearly non-contiguous TADs defined as TAD cliques. Therefore, to further explore the possible
26 organization of inter-TAD interactions during SC lineage progress, we examined our Hi-C contact
27 matrices and indeed uncovered that TADs engaged in non-contiguous contacts (Fig. 3a, exemplified
28 on chromosome 9 at D1 stage). Utilizing non-central hypergeometric model (NCHG), 3639, 3667,
29 2927, 1843 and 2176 significant intra-chromosomal TAD-TAD interactions were identified from
30 QSC to D3 (Suppl. Data 5), which is exemplified on chromosome 9 at D1 stage (Fig. 3b); the

1 interaction distance between TADs remained constant from QSC to D1 (QSC: 28.6 Mb; FISC: 28.8
2 Mb; D1: 29.2 Mb) but a decrease was observed starting D2 (19.8 Mb) (Fig. 3c). Next, we defined
3 TAD clusters, which encompassed at least 3 TADs that were fully connected pairwise in Hi-C data
4 regardless of interaction distance (Fig. 3d). As a result, a large number of TADs (QSC to D3: 719,
5 563, 386, 260, and 284) were engaged in clusters containing 3 to 11 TADs, representing 31-87% of
6 all TADs (Fig. 3e). A reduced number of TADs in the clusters was observed resulting in smaller
7 clusters as cells progressed from QSC to D3 (Fig. 3e). An alluvial diagram showed that whereas 65%
8 of TAD clusters were maintained during the lineage progression, hundreds of TADs assembled into,
9 or disassembled from clusters; waves of TAD cluster formation and dissociation were observed in
10 FISC, D2 and D3, which resulted in variation of cluster size (Suppl. Fig. 4a). Previously, Paulsen, J.
11 et.al.¹⁵ defined TAD cliques as functionally important TAD assemblies involving long-range
12 TAD-TAD interactions, which are associated with repressed chromatin domain that shapes the
13 genome during cell differentiation and also demonstrated these TAD cliques are enriched in B
14 compartments. Paulsen, J. et. al. ¹⁵, on the other hand, showed that about 30% of the identified TAD
15 cliques in the adipogenesis residue in compartments A. Interestingly, we found the TAD clusters
16 defined in our study were present in both compartments A and B. For example, a TAD cluster (C1)
17 residing in compartment B in QSC and FISC displayed an increase of cluster size in D1 (4 to 6); a
18 second cluster, C3 assembled at D1 in compartment B, while C2 cluster residing in compartment A
19 underwent dynamic assembly/disassembly during the entire lineage progression (Suppl. Fig. 4b).
20 When closely examining frequency of TAD cluster formation, loss, expansion and reduction at each
21 stage compared to QSC, we found the changes occurred in compartments A and B at about the same
22 frequency in FISC vs. QSC; and similar observation was made in D1 vs. QSC. Nevertheless, changes
23 in D2 or D3 vs. QSC occurred more frequently in compartment B vs. A (Suppl. Fig. 4c). Furthermore,
24 we found that interaction distance between the pairwise TAD-TAD interactions significantly
25 decreased in D1 vs FISC and D2 vs D1 (Suppl. Fig. 4d). Interestingly, TADs in short-range
26 interactions (Q1 and Q2, QSC: <18.26Mb; FISC: <17.98Mb; D1: <18.02Mb; D2: <5.1Mb; D3:
27 <10.24Mb) were nearly equally associated with A and B compartments, while those in long-range
28 interactions (Q3 and Q4, QSC: >18.26Mb; FISC: >17.98Mb; D1: >18.02Mb; D2: >5.1Mb;
29 D3: >10.24Mb) displayed a prominent enrichment towards B compartments (Fig. 3f), suggesting
30 these long-range interacting TADs in Q3 and Q4 were reminiscent of the TAD cliques defined in the

1 previous study¹⁵. Next, we examined the gene expression in TAD clusters and found genes in Q1 or
2 Q2 were expressed at significantly higher levels than those in Q3 or Q4 or not in any clusters (Fig.
3 3g), which is consistent with prior findings that TAD clusters involving long-range TAD-TAD
4 interactions are associated with chromatin compaction and gene repression.

5 Lastly, since previous studies revealed that super enhancers (SEs) are enriched in the highly
6 interacting TAD triplets^{16, 17, 41}, we wondered whether SEs-containing TAD clusters are more active
7 in long-range TAD-TAD interactions. Leveraging the H3K27ac ChIP-seq profiles in QSC and
8 FISC²⁰, we defined SEs and typical enhancers (TEs). Indeed, we found 10%-12% TADs in Q3 and
9 Q4 contained SEs and residing in compartment A (Suppl. Fig. 4e), and those SE containing TAD
10 clusters were associated with higher gene expression in both QSC and FISC compared with non-SE
11 clusters (Suppl. Fig. 4f). Overall, in both QSC and FISC, we found SE-containing TAD clusters were
12 engaged in longer interaction distance (35 kb vs 18 kb) (Fig. 3h), a higher number of TAD clusters
13 (Suppl. Fig. 4g) and a larger cluster size (Suppl. Fig. 4h) compared to those containing TEs or no
14 enhancers. In addition, the presence of SE clusters was associated with more evident expression
15 changes, as transcription level in SE not TE containing TAD clusters in both QSC (Fig. 3i) and FISC
16 (Suppl. Fig. 4i) displayed distinct stage-specific expression pattern. As an example, a 3D model of
17 one TAD cluster was developed (Fig. 3j), which contained one of active SEs in QSC. This
18 SE-containing TAD clustered with four other TADs, to form a TAD cluster in QSC; it
19 decommissioned in FISC due to the increased interacting distance between these TADs (Fig. 3k).
20 Consequently, the genes in this TAD cluster (both the SE residing and four interacting TADs)
21 expressed at significantly lower levels in FISC vs. QSC (Fig. 3l). Taken together, our results
22 demonstrate that TAD clusters are present and dynamically reorganized when SC progresses from
23 QSC to FISC.

24

25 **Chromatin loops are dynamically changed during SC early activation**

26 Chromatin loops manifest as foci in high-resolution Hi-C contact maps. Considering the most
27 striking changes observed above at compartment and TAD levels occurred in the early stages (QSC
28 to FISC), we increased the sequence depth for the Hi-C libraries from these two time points to
29 obtain > 500 million uniquely mapped contacts, which enabled us to visualize chromatin looping
30 events at a 5 kb resolution. Using FitHiC2⁴² algorithm, a total of 11,337 and 2412 chromatin loops

1 were identified in QSC and FISC, respectively (Fig. 4a, Suppl. Data 6), indicating a drastic decrease
2 in the number of loops during SC early activation. In addition, a reduction of interaction strength was
3 also observed by aggregated peak analysis (APA) (4.16 vs 3.93) (Fig. 4a). Furthermore, 2,790 and
4 504 loops were further defined as QSC or FISC specific loops, which accounted for a minor
5 proportion of all loops (24.6% for QSC and 20.1% for FISC) (Suppl. Fig. 5a and Suppl. Data 6).
6 Expectedly, QSC-specific loops showed a decreased APA score in FISC (1.23) vs. QSC (4.12), while
7 FISC-specific loops displayed a much higher APA score in FISC (3.57) vs. QSC (1.91) (Suppl. Fig.
8 5b). Interestingly, FISC-specific loops were larger than QSC-specific loops and contained fewer
9 number of genes (Suppl. Fig. 5c). However, in both time points the stage-specific loops localized
10 mainly to compartment A (Suppl. Fig. 5d) and positively correlated with gene expression changes
11 (Suppl. Fig. 5e). Moreover, GO analysis revealed genes involved in the stage-specific loops were
12 enriched for distinct terms (Suppl. Fig. 5f and Suppl. Data 6).

13 Next we investigated subtypes of loops formed between enhancers and promoters (E-P) or two
14 enhancers (E-E). Of all E/P related loops in QSC, 17% were found to be E-P while 22% were E-E
15 interactions; in FISC, 12% were E-P while 21% were E-E interactions (Fig. 4b). Interestingly,
16 enhancers in E-E interactions displayed lower levels of H3K27ac signals in FISC vs QSC, while
17 promoters involved in E-P interactions showed higher levels of H3K4me3 signals, indicating that
18 E-E or E-P loops were epigenetically distinct in QSC and FISC (Suppl. Fig. 5g). The interaction
19 strength of both E-P (APA: 3.73 to 3.55) and E-E loops (APA: 3.71 to 3.52) decreased in FISC vs.
20 QSC (Fig. 4c). Furthermore, we found although enhancer/promoter landscapes were largely
21 unaltered in FISC vs QSC (10,113 shared enhancers/promoters) (Suppl. Fig. 5h), the loop formation
22 was highly dynamic with 95.2% and 64.3% specifically identified in QSC and FISC respectively
23 (Fig. 4d). In addition, stage-specific E-P and E-E loops also displayed more prominent difference in
24 APA score (1.14 vs. 3.62) (Fig. 4e) and correlation with gene expression (Fig. 4f), when compared to
25 all loops (Suppl. Fig. 5e). This phenomenon is illustrated on representative *Hoxc8*, *Myf5* and *Atf3*
26 genes (Fig. 4g). For both *Hoxc8* and *Myf5*, a disappeared E-P loop in FISC vs. QSC was observed
27 between the gene promoter and their distal enhancers which displayed the openness at both stages.
28 As a result, the gene expression was decreased in FISC. On the contrary, on *Atf3* locus, despite the
29 openness of the distal enhancer in QSC, the gene expression was not observed until the E-P loop
30 became established in FISC. Taken together, these results reveal dynamic changes of looping events

1 during early activation of SCs, which may be tightly related to the transcriptional change of pivotal
2 genes associated with this progression.

3

4 **3D regulatory interactions orchestrate Pax7 expression dynamics in SC**

5 PAX7 is the key TF known to regulate the myogenic potential and function of SCs⁴³, the mechanism
6 dictating its high expression in QSC and attenuation after SC becomes activated remains largely
7 unknown. We thus took a close examination at the 3D organization associated with *Pax7* locus. A
8 prominent TAD was identified in QSC encompassing *Pax7* locus only and evident presence of four
9 sub-TADs was also observed (Fig. 5a). Whereas the TAD maintained constant, the four sub-TADs
10 disappeared in FISC due to a loss of interaction strength within the sub-TADs and a gain of
11 interaction strength among them (Fig. 5a-b). In addition, a reduced number of chromatin loops were
12 observed in FISC (3) vs. QSC (9) (Fig. 5b). The above data suggest that *Pax7* locus undergoes
13 drastic reorganization at 3D level when SCs become activate which may account for its expression
14 attenuation.

15 To further dissect key regulatory elements governing the above observed 3D re-organization,
16 four putative regulatory elements (REs, R1-R4) were defined based on the looping interactions as
17 well as ATAC-seq, H3K27ac ChIP-seq signals (Fig. 5a). Among them, all four were labeled by high
18 ATAC-seq signals. R4 resided in the *Pax7* promoter region. R2 and R3 were marked with sharp
19 H3K27ac signals in QSC, which disappeared in FISC; R1 resided in the border region. Evident loops
20 between R1/R4, R2/R4, or R3/R4 were observed in QSC but only R1/R4 loop retained in FISC (Fig.
21 5b). To experimentally validate the functionality of these REs in dedicating *Pax7* expression, deletion
22 of each in QSC was achieved by harnessing the muscle specific CRISPR/Cas9/AAV-sgRNA *in vivo*
23 genome editing system recently developed in our group⁴⁴ (Fig. 6c). Briefly, two pairs of sgRNAs
24 targeting the same region (Suppl. Fig. 6a) were applied successively: 4×10^{11} viral genomes
25 (vg)/mouse of AAV9-sgRNA pair #1 was intramuscularly (IM) injected into *Pax7*^{Cas9} mouse at
26 postnatal age of P5 and followed by 8×10^{11} vg /mouse of AAV9-sgRNA pair #2 at P10. The mice
27 were then sacrificed for SC isolation and analysis after four weeks. For the control (Ctrl) group, the
28 same dose of AAV9 virus containing pAAV9-sgRNA backbone without any sgRNA insertion was
29 injected. The variation in deletion efficiency was detected (Suppl. Fig. 6b): R1 (90%) and R4 (76.1%)
30 manifested more efficient deletion than R2 (15.4%) or R3 (17%), which may be due to the inherent

1 difference of protospacer sequences⁴⁵. As a result, the deletion of R1, R2 or R3 led to 20%, 41%, 25%
2 decrease of Pax7 transcripts in FISC (Fig. 5d). Correspondingly, PAX7 protein levels were also
3 decreased (Suppl. Fig. 6c). Interestingly, removal of R4, the *Pax7* promoter region, did not disrupt
4 Pax7 mRNA level (Fig. 5d) despite ~40% decrease of PAX7 protein (Suppl. Fig. 6c).

5 To further determine the potential effect of the above RE deletion on 3D structure at the *Pax7*
6 locus, *in situ* Hi-C was performed in QSC isolated from the above mice to obtain > 220M uniquely
7 aligned contacts for downstream analysis (Suppl. Fig. 6d and Suppl. Data 7). First, comparison of the
8 Hi-C matrices at 40 kb and 10 kb resolution (Suppl. Fig. 6e) demonstrated that the 3D conformation
9 at *Pax7* locus from the *Pax7*^{Cas9} mice used in this experiment was comparable to that from the earlier
10 used Pax7-nGFP mice (R = 0.85 and 0.83 respectively) (Suppl. Fig. 6f). Next, we found that removal
11 of each RE caused significant increase of the inter-TAD interaction strength genome-wide (Fig. 5e);
12 the intra-TAD interaction strength decreased significantly upon loss of R1 but increased upon R4
13 loss and unaltered when R2 or R3 was removed (Fig. 5e). Although no obvious effect was observed
14 on TAD organization surrounding *Pax7* (Suppl. Fig. 6g), close examination of the insulation score at
15 the boundary containing *Pax7* promoter revealed a decrease upon loss of R1, R2 or R4 (Fig. 5g and
16 Suppl. Data 7), indicating strengthened border insulation upon removal of these elements. This was
17 not observed when R3 was removed but its deletion led to fusion of nearby sub-TADs into a new
18 sub-TAD (Suppl. Fig. 6g). Moreover, a TAD cluster was defined and formed through interaction of
19 the *Pax7* residing TAD with four other TADs in Ctrl QSC; the interaction was disrupted upon loss of
20 each RE (Fig. 5h-5i), which was further supported by virtual-4C using *Pax7* promoter as the
21 viewpoint showing decreased interaction frequency among the TADs (Suppl. Fig. 6h). Interestingly,
22 at the loop level, the interaction strength of R1 and R4 anchored intra-TAD loops were both
23 decreased upon R1 or R4 deletion (Fig. 5j, and Suppl. Data 7). Interestingly, we also observed a
24 decreased number of intra-chromosomal loops anchored in *Pax7* TAD upon deletion of each RE
25 (Suppl. Fig. 6i). Altogether, the above results highlight 3D rewiring at *Pax7* locus during SC
26 activation and unveil key REs dictating the 3D remodeling and Pax7 expression.

27

28 **Multiscale analysis of 3D genome during mouse SC aging**

29 It is well accepted that SCs decline in their regenerative capacity with aging²⁸, partly ascribed to the
30 intrinsic changes at both transcriptomic and epigenomic levels. However, it is not known whether 3D

1 genome rewiring occurs in SCs during mouse physiological aging. To fill the gap, we collected
2 FISCs from wild-type mice of 2 (young), 23-24 (aged) and 28-32 (geriatric) months of age (Fig. 6a).
3 Consistent with prior findings²⁷, a decrease in FISC number was detected in the aged (3.3%) vs.
4 young (4.3%) mice and became more prominent at the geriatric stage (1.8%) (Suppl. Fig. 7a-b).
5 Transcriptional changes were profiled by RNA-seq in the above FISCs with at least two replicates
6 from each stage. In line with previous findings⁴⁶, gene expression profiles were remarkably distinct
7 between young and aged SCs ($R = 0.74$), and became more divergent in geriatric cells ($R = 0.65$)
8 (Suppl. Fig. 7c). Notably, *Pax7* expression was steadily upregulated with SC aging, along with stress
9 responsive genes, *Egr1* and *Atf3*; while *Myod* and *Myc*⁴⁴ were highly induced at aged followed by a
10 decline at geriatric phase (Suppl. Fig. 7d). Next, results from pair-wise comparison showed 1,122
11 genes were significantly up- and 1,016 down-regulated in aged vs young SCs; whereas 910 and
12 1,104 were in geriatric vs. aged SCs (Suppl. Fig. 7e and Suppl. Data 8). Nevertheless, across the
13 three stages, gene expression adopted distinct dynamics with only a limited proportion of genes
14 continuously increased ($n = 120$) or decreased ($n = 53$) (Suppl. Fig. 7f). Consistently, GO analysis
15 revealed distinctly enriched terms in aged and geriatric stages (Suppl. Fig. 7g and Suppl. Data 8). For
16 instance, the upregulated genes in aged vs. young SCs were enriched in terms such as “cell adhesion”,
17 “inflammatory response” and “immune response”, while those in geriatric vs. aged SCs were linked
18 with “membrane”, “cell surface” and “extracellular space”.

19 Next, to delineate the changes in 3D genome organization with SCs aging, *in situ* Hi-C was
20 performed at the above stages with three to four replicates for each to obtain a total of >300 million
21 uniquely aligned contacts; high correlation ($R \geq 0.95$) was achieved among the replicates (Suppl.
22 Fig. 8a and Suppl. Data 9). Overall, young SCs exhibited distinct 3D genome organization compared
23 with aged or geriatric SCs; long-range (> 3 Mb) interaction frequencies increased significantly in
24 geriatric SCs compared with young and aged SCs (Fig. 6b-c, Suppl. Fig. 8b). Interestingly, the
25 interaction strength increased between A compartments (A-A) during young to aged transition but
26 remained constant when SCs progressed into geriatric stage; on the contrary, a significant decline in
27 B-B interaction strength was detected during aged to geriatric progression but not young to aged
28 transition (Fig. 6d-e and Suppl. Data 9). Overall, an increase in compartmentalization was observed
29 at aged vs young SCs as revealed by decreased inter-compartment contacts, but then recovered at
30 geriatric stage (Fig. 6d-e). In addition, similar to SC lineage progression (Suppl. Fig. 1i), limited

1 compartment switching was observed: only 6.5% of the genome switched compartment at any two
2 time points with B to A and A to B each accounting for 2.9% and 2.4% respectively (Suppl. Fig. 8c).
3 Overall, the compartmental trajectory showed drastic change at aged vs. young while transcriptomic
4 change occurred late in geriatric stage (Fig. 6k). Whereas the lagged expression change behind the
5 compartment switch was noticed (Suppl. Fig. 8d left), concomitant changes in compartment and gene
6 expression were also identified on some loci (Suppl. Fig. 8d middle and right).

7 At the TAD level, an average of $\sim 3,274 - 3,783$ TAD borders per stage was identified (Suppl.
8 Fig. 8e). Partitioning of the genome into TADs was largely stable during SC aging as $>70\%$ TAD
9 borders ($n = 2,660$) were detected at all stages (Suppl. Fig. 8e, invariant borders). The aggregate
10 interaction matrix revealed a remarkable loss in insulation strength across the TAD borders in
11 geriatric but not aged stage when compared with young SCs (Fig. 6f), which was supported by IS
12 calculation showing a significant increase at geriatric stage (Fig. 6g, Suppl. Fig. 8f and Suppl. Data
13 10). Further, quantitative analysis of IS through hierarchical clustering also showed a greater portion
14 ($n = 995$, 38.3%) of all borders exhibited $> 20\%$ increase in IS, while 18.3% ($n = 475$) showed a loss
15 during SC ageing (Fig. 6h). These data suggested TAD borders largely lost insulation strength when
16 SCs entered into geriatric state. Furthermore, we found that local gene expression underwent
17 significant increase or decrease when TAD borders were gained or lost, respectively (Fig. 6i), which
18 was not observed during SC lineage progression (Suppl. Fig. 3f, 3g). Nevertheless, in agreement with
19 the findings in Fig. 2n, D-scores were positively associated with gene expression (Suppl. Fig. 8g);
20 and hierarchical clustering revealed that $\sim 47\%$ of the TADs ($n = 1,754$) showed D-score increase or
21 decrease ($> 15\%$ difference between endpoints) (Suppl. Fig. 8h and Suppl. Data 10). Additionally,
22 the D-score kinetics was highly correlated with changes in PC1 values (Suppl. Fig. 8i) or gene
23 expression (Suppl. Fig. 8j) during the transition between two stages. For instance, the intra-TAD
24 interactions surrounding *Ier5* (Fig. 6j) and *Hox10* (Suppl. Fig. 8k) loci were largely attenuated at
25 geriatric stage, which might result in the concomitant decrease of transcriptional levels from aged to
26 geriatric FISC (Fig. 6j, Suppl. Fig. 8j). Lastly, we performed PCA analysis for RNA expression,
27 compartmentalization, IS and D score. Interestingly, transcription and IS adopted a very similar
28 trajectory showing distinct changes from young to aged to geriatric stages; D-score was largely
29 similar between young and aged stages but underwent a drastic change from aged to geriatric stage;
30 PC1 on the contrary displayed the most prominent change from young to aged stage (Fig. 6k),

1 suggesting genome compartmentalization was altered at aged stage while TAD connectivity changed
2 later at geriatric stage.

3 Finally, chromatin looping analysis led to identification of a total of 11,180, 10,131 and 11,104
4 loops in young, aged and geriatric FISCs among which 6,703, 6,106, and 7,811 were stage specific
5 (Fig. 6l and Suppl. Data 10). An increase in interaction strength of chromatin loops was observed
6 when SCs aged (APA score = 4.68 in geriatric vs 2.87 in young and 2.82 in aged) (Fig. 6m).
7 Interestingly, expression changes lagged behind their associated loops, as illustrated on the
8 representative *Ier5l* locus (Fig. 6n). The loop between the promoter and a distal enhancer in young
9 cells disappeared in aged cells but decreased gene expression did not occur until the geriatric stage.
10

1 **Discussion**

2 While mechanisms underlying SC lineage progression have been well studied at the transcriptional
3 and epigenetic levels^{1, 20, 29, 47}, how genome architecture reorganization contributes to this process
4 remains enigmatic. Here, we examined the genome conformation in SCs at distinct stages and
5 revealed the characteristics of chromatin 3D landscape at different genomic scales and its association
6 with altered gene expression. First, at the large scale, we identified a shift in the ratio between long-
7 and short-range chromatin interactions and remarkable dynamic compartmentalization during SC
8 lineage progression. Loss of long-range contacts at the late stages (D2 and D3) may reflect a
9 heterogeneous nature of SC populations when a portion of cells enter mitosis during which cells are
10 reported to display a rapid fall-off of chromatin contacts at ~10 Mb^{29, 30}; this was not observed at
11 early stages (QSC, FISC and D1) when SCs mostly reside in G0/G1 phases. In addition, SC lineage
12 progression is accompanied with the dynamic changes in compartmentalization strength (Fig. 1d-e),
13 which could be resulted from the alterations in chromatin states. Previous findings have documented
14 the compartment strength increases concurrently with a gain of heterochromatin marks, H3K27me3¹⁸,
15 ⁴⁸ or H3K9me3⁴⁹. Indeed, epigenomic profiling in SCs uncovered a significant loss of H3K27ac
16 marks during SC isolation procedure²⁰, and a gain of H3K27me3 marks is associated with SC
17 activation^{29, 30}. Hence, it will be of great interest to further dissect the connection between the
18 epigenomic changes with the dynamics of genome compartmentalization during SC lineage
19 progression.

20 Next, at the TAD level, the most striking changes in TAD border strength occur at early
21 activation stage (FISC), which could be a result of the decrease in CTCF expression during the
22 process²⁰ to cause the overall detrition in border strength and looping formation as CTCF is known to
23 play key role in chromatin insulation and TAD boundaries^{50, 51}. Alternatively, CTCF independent
24 mechanism could contribute to the TAD border reorganization in FISC vs. QSC. For example,
25 previous observations have demonstrated TAD boundaries are also enriched for housekeeping genes,
26 RNA polymerase II, and active chromatin marks^{7, 10} thus active transcription could play a role in
27 shaping the chromatin organization⁵². Nevertheless, gain or loss of borders does not correlate with
28 the overall changes in local gene expression in FISC vs. QSC, arguing transcriptional alterations per
29 se may not be the main contributor for insulation dynamics in this process. Interestingly, the
30 weakened insulation of TAD borders gained strength when SCs enter differentiation, which is

1 reminiscent of findings from neuron development⁷, suggesting TAD borders are dynamically
2 re-organized during cellular differentiation.

3 Emerging studies indicated interactions among non-contiguous TADs are prevalent¹⁵⁻¹⁷. In this
4 study, we thus defined TAD clusters as an assembly of ≥ 3 TADs that are fully connected pairwise
5 using a similar approach previously used to annotate TAD cliques¹⁵. Nevertheless, our defining
6 approach is modified by taking into account of the short-range interactions between TADs. In
7 agreement with previous findings¹⁵, the long-range inter-TAD interactions (Q3 and Q4)
8 preferentially reside in B compartments and are associated with repressed gene expression but the
9 short-range ones (Q1 and Q2) do not display such preference. Moreover, we found SE-containing
10 TAD clusters are more active in long-range inter-TAD interactions. Concordantly, 3D modeling
11 revealed SE containing TAD clusters reorganize in nuclear space to orchestrate the stage-specific
12 gene expression. These findings complement recent observations that multi-way interactions occur at
13 gene-dense, active, RNA polymerase II-marked regions or regions containing SEs^{16,17}.

14 Finally, using high-resolution Hi-C datasets to dissect the chromatin looping in FISC and QSC,
15 we revealed the isolation procedure-induced early activation²⁰ is accompanied by a great loss of
16 chromatin looping. Reduction of CTCF expression during this process may provide one reason,
17 considering its dominant effect on chromatin looping⁵⁰. Nevertheless, previous study showed that
18 acute attenuation of CTCF in a short time does not lead to drastic changes in gene expression³⁵,
19 which contradicts to the expression alterations of a vast spectrum of genes observed in FISC vs.
20 QSC²⁰. Hence, CTCF-independent mechanisms may regulate chromatin looping remodeling during
21 the isolation process. Consistent with the overall looping events, the number of enhancer-promoter
22 contacts is largely reduced in FISC vs. QSC, in agreement with the prominent loss in H3K27ac
23 levels during this process²⁰. Intriguingly, among the loci we examined, despite a great portion bear
24 enhancer marks in both QSC and FISC, the interaction is restricted to only one stage to enabling the
25 timely gene activation. Our findings, together with previous reports^{7, 13, 39, 53}, thus highlight that the
26 interactions provide an additional layer of fine-tuned regulation in lineage development, which
27 cannot be simply inferred from the 1D enhancer profiling. In the future, it will be interesting to
28 identify the factors that drive the decommission of enhancer repertoire and rewiring of E-P contacts
29 during SC early activation.

30 In addition to the normal SC lineage progression, we delineated the 3D genome remodeling

1 during SC aging. How SCs deteriorate with aging remains largely unknown^{54, 55}. Here, we have
2 provided a comprehensive view of how genome 3D configuration in SCs undergo massive changes
3 during the natural aging course of mouse. Interestingly, our findings reveal that despite differences
4 between young and old SCs, geriatric cells display a more prominent gain in long-range contacts and
5 loss of TAD insulation. This is in agreement with recent findings suggesting SCs switch from
6 quiescence to an irreversible pre-senescence state at geriatric age²⁷; and a progressive gain of far-cis
7 contacts and loss of insulation has been shown during the onset of various types of cellular
8 senescence²³⁻²⁵. However, our results highlighted a preferential gain of A-A interactions and loss of
9 B-B interactions during SC aging, contrasting with previous findings in either oncogene-induced
10 senescence (OIS) or replicative senescence (RS) scenario²⁵. In OIS the genome organization is
11 dominated by heterochromatin or B-B interactions resulting in a stronger genome
12 compartmentalization, whereas in RS, dampened A-A interactions causes a decrease in
13 compartmentalization²⁵. We speculate these discrepancies may reflect the innate differences between
14 pre-senescence state of geriatric SCs and mature or deep senescent cellular state studied in other
15 works. In addition, SCs need to integrate the information from the deteriorated aging milieu^{27, 30, 56},
16 which may render the state of geriatric SCs more complex and heterogeneous³¹ than the *in vitro*
17 systems used in other studies. Consistent with the notion, PCA analyses revealed distinct trajectories
18 of transcriptome, compartmentalization, TAD insulation and intra-TAD interactions (Fig. 6k).
19 Geriatric SCs overall lose insulation at TAD borders, agreeing with the observations in OIS and RS
20 cells²⁵. Recent reports suggest HMGB1 and HMGB2 proteins act as a rheostat of topological
21 insulation upon OIS senescence entry^{24, 26}. Consistently, we observed decreased expression of
22 HMGB1 and HMGB2 with SC aging (Suppl. Fig. 7d); it will thus be interestingly to investigate the
23 possible role of HMGB proteins in SC aging. In summary, our study represents the first to describe
24 the 3D genome rewiring in adult stem cells of physiologically aged mammalian organism, and we
25 anticipate more studies from other tissues or organs, will enable us to extract the common paradigm
26 controlling stem cell aging.

27 The last key contribution of the present study is the elucidation of Pax7 regulation at 3D level.
28 Pax7 represents the uttermost indispensable TF regulating SC maintenance and regenerative
29 capacity^{6, 57, 58}. However, current knowledge regarding the upstream regulation of Pax7 expression
30 remains largely unknown. At the transcriptional level, Pax7 RNA expression is vulnerable to diverse

1 environmental cues and rapidly decreases upon SC activation even in response to isolation
2 procedure²⁰. Recent studies have provided epigenetic clues that the *Pax7* locus is embedded in a SE
3 region in homeostatic quiescence state and the enhancer signals diminish upon activation²⁰, which
4 may account for the acute loss of *Pax7* expression. Complementary to these findings, we here
5 uncovered *Pax7* locus is self-organized into a powerful TAD with profound sub-TAD organization
6 identified in quiescent SCs, and the organization largely collapses upon activation. In addition, we
7 identified significant chromatin loops within the *Pax7* TAD involving four interconnected REs,
8 including promoter, enhancers, and boundary elements. Leveraging our in-house developed *in vivo*
9 genome editing system, we provide *in vivo* evidence to show that deletion of either of the four REs
10 disturbed *Pax7* RNA expression in SCs, and resulted in distinct effects on 3D genome organization at
11 *Pax7* locus. Removal of the downstream boundary region (R1) alone did not cause TAD fusion,
12 possibly due to the enhanced contact newly established between the two convergent CTCF sites^{7, 12}.
13 Intriguingly, it instead led to significant reduction of intra-TAD contacts. Deletion of the downstream
14 enhancer region (R2) largely retained the TAD structure with a marginal loss of intra-TAD contacts,
15 especially the interactions between sub-TADs. In contrast, loss of the intronic enhancer region (R3)
16 resulted in more structured sub-TAD patterns. In contrast to the above deletions resulting in a
17 consistent decreased *Pax7* expression caused by the above deletions, deletion of the promoter region
18 (R4), surprisingly, enhanced *Pax7* RNA expression, which may be caused by the use of an alternative
19 TSS downstream of the enhancer region. Moreover, we showed that the *Pax7* TAD is also involved
20 in inter-TAD interactions to form a TAD cluster. Deletion of each RE within the TAD also abrogated
21 the clustering. Altogether, our findings for the first time unveil that *Pax7* transcription is tightly
22 controlled by the embedded REs and 3D interactions. Nevertheless, we need to point out that the
23 editing efficiency on each RE varied from 15.4% for R2 to 90% for R1, which may be intrinsically
24 dictated by the *in vivo* editing system. It is possible that more evident changes of chromatin
25 architecture and *Pax7* expression can be achieved if the editing efficiency reaches 100% equally for
26 all REs using traditional knockout models.

27 In conclusion, our study provides a comprehensive view of 3D genome organization in SC
28 lineage progression and SC aging. Our findings demonstrate that chromatin rewiring at different
29 genomic scales underpin the transcriptome remodeling and SC activities, underscoring the
30 importance of 3D regulation in stem cell lineage development and stem cell aging. The datasets

1 produced in this study for the first time provide a rich resource of chromatin interactions for the SC
2 community. In addition, our work uncovers the 3D regulation of PAX7, underscoring the importance
3 to connect the 3D genome rewiring to the fine-tuned TF dynamics during lineage development.

4

1 **Methods**

2 **Mice.** All animal experiments were performed following the guidelines for experimentation with
3 laboratory animals set in the Chinese University of Hong Kong and approved by the Animal
4 Experimentation Ethics Committee (AEEC) of the Chinese University of Hong Kong (CUHK). The
5 mice were maintained in animal room with 12 hrs light/12 hrs dark cycles at Animal Facility in
6 CUHK. For SC lineage development study, two-month-old male Pax7-nGFP mice were used to
7 isolate satellite cells. For aging SC study, male wild-type C57BL/6 mice of 2 (young), 23-24 (aged)
8 and 29-31 (geriatric) months of age were used to isolate SCs. For *Pax7* locus dissection, Pax7^{Cas9}
9 mice were generated by crossing homozygous Cre-dependent Rosa26^{Cas9-EGFP} knockin mouse
10 (B6;129-Gt (ROSA)26Sor^{tm1(CAG-cas9*-EGFP) Fezh/J}; stock number 024857, Jackson Laboratory) with a
11 Pax7^{Cre} strain, as previously described⁴⁴.

12 **Satellite cell isolation by FACS.** SCs were sorted by FACS based on established methods⁵⁹. Briefly,
13 hindlimb muscles from Pax7-nGFP mice were collected and digested with collagenase II (1000 U
14 ml⁻¹, Worthington) for 90 min at 37°C, and then the digested muscles were triturated and washed in
15 washing medium (Hams F-10 media (Sigma), 10% HIHS (Gibco), penicillin/streptomycin (1×,
16 Gibco) before SCs were liberated by treating with Collagenase II (1000 U ml⁻¹) and Dispase (11 U
17 ml⁻¹) for 30 min at 37°C. Mononuclear cells were filtered with a 40-µm cell strainer and GFP+ SCs
18 were sorted out by BD FACSAria Fusion cell sorter (BD Biosciences). For SC isolation from
19 C57BL/6 mice, the filtered mononuclear cells were additionally incubated with Vcam1-biotin
20 (105704, BioLegend), CD31-FITC (102506, BioLegend), CD45-FITC (103108, BioLegend), and
21 Sca1-Alxa647 (108118, BioLegend). The Vcam1 signal was amplified with Streptavidin-PE (554061,
22 BD Biosciences). All antibodies were used at a dilution of 1:75. Vcam1+, CD31-, CD45-, Sca1- cells
23 were sorted out by BD FACSAria Fusion cell sorter (BD Biosciences). For QSC isolation, we
24 followed the procedures described previously²⁰. Briefly, the muscles were pre-fixed in ice-cold 0.5%
25 paraformaldehyde (PFA) before cell dissociation. Collagenase II and Dispase were used at 2,000 U
26 mL⁻¹ and 22 U ml⁻¹ for the digestion respectively.

27 **AAV9 virus production, purification and injection.** AAV9 virus was produced as described
28 previously^{44, 60}. In brief, HEK293FT cells were seeded in T75 flask and transiently transfected with
29 AAV9-sgRNA vector (5 µg), AAV9 serotype plasmid (5 µg), and pDF6 (AAV helper plasmid) (10 µg)

1 at a ratio of 1:1:2 using polyethyleneimine (PEI) when the cell reached 80%~90% confluency.
2 Twenty-four hours after transfection, the cells were changed to growth medium (DMEM with 10%
3 FBS, 100 U ml⁻¹ penicillin, 100 µg ml⁻¹ streptomycin and 2 mM L-glutamine) and cultured for
4 another 48 hrs. The cells were harvested and washed with PBS two times. To release the AAV9 virus,
5 the pellet was re-suspended with lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl) followed by
6 three sequential freeze–thaw cycles (liquid nitrogen/37°C). The lysates were treated with Benzonase
7 (Sigma) together with MgCl₂ (final concentration: 1.6 mM) at 37°C for 1 hr followed by
8 centrifugation at 3,000 rpm for 10 min. The supernatant was filtered with 0.45 µm sterile filter and
9 added with equal volume of 1 M NaCl and 20% PEG8000 (w/v) to precipitate the virus at 4 °C
10 overnight. After centrifugation at 12,000 g for 30 min at 4 °C, the pellet was re-suspended with
11 sterile PBS and then subject to centrifugation at 3,000 g for 10 min. Equal volume of chloroform was
12 then added and shaken followed by spin down at 12,000 g for 15 min at 4 °C. The aqueous layer was
13 filtered by 0.22 µm sterile filter and passed through a 100 kDa MWCO (Millipore). The concentrated
14 solution was washed with sterile PBS three times. The titer of the AAV9 virus was determined by
15 qRT-PCR using primers targeting the CMV promoter. Sequences of the used primers are listed in
16 Suppl. Table 1. Heterozygous Pax7^{Cas9} mice were intramuscularly (IM) injected with 4×10¹¹ viral
17 genomes (vg) of AAV9- sgRNA pair #1 at postnatal day 5 (P5) and followed by 8×10¹¹ vg of AAV9-
18 sgRNA pair #2 at P10. SCs were isolated 4 weeks post AAV injection.

19 **Cell culture.** HEK293FT cells were obtained from ATCC (CRL-3216) and maintained in DMEM
20 supplemented with 10% FBS, 100 U ml⁻¹ penicillin and 100 µg of streptomycin in a 5% CO₂
21 humidified incubator at 37°C. Freshly-isolated SCs were cultured in Hams F10 medium (Sigma)
22 with 20% FBS, penicillin/streptomycin (1×) and 2.5 ng ml⁻¹ basic fibroblast growth factor (bFGF,
23 13256, Life Technologies).

24 **Plasmids.** Site-specific sgRNAs were selected using a web tool Crispor⁶¹ (<http://crispor.tefor.net/>).
25 Synthesized oligonucleotides encoding guide sequences were cloned into the AAV9-sgRNA transfer
26 vector (AAV: ITR-U6-sgRNA(backbone)-CMV-DsRed-WPRE- hGHpA-ITR) using Sap I site. To
27 generate dual AAV-sgRNAs expression plasmid, the second sgRNA was constructed into the
28 AAV9-sgRNA vector together with the gRNA cassette and U6 promoter using Xba I and Kpn I sites.
29 Sequences for sgRNAs and primers used to detect deletion efficiency are listed in Suppl. Table 1.

1 **RNA isolation, quantitative RT-PCR and RNA-seq.** For FISC and cultured SCs, total RNAs were
2 extracted using TRIzol reagent (Invitogen). For QSC, total RNAs were isolated from the
3 above-described pre-fixed cells using a miRNeasy FFPE Kit (Qiagen) according to the
4 manufacturer's instruction. For quantitative RT-PCR, cDNAs were prepared using HiScript III 1st
5 Strand cDNA Synthesis Kit (Vazyme, R312-01). Real-time PCR reactions were performed on a
6 LightCycler® 480 Instrument II (Roche Life Science) using Luna Universal qPCR Master Mix
7 (NEB, M3003L). Sequences of all primers used are listed in Suppl. Table 1. For polyA+ mRNA-seq,
8 total RNAs were subject to poly(A) selection (Ambion, 61006) followed by library preparation using
9 NEBNext® Ultra™ II RNA Library Preparation Kit (NEB, E7770S). Libraries were paired-end
10 sequenced with read lengths of 150 bp on an Illumina HiSeq X Ten instrument. For total RNA-seq,
11 total RNAs were isolated from QSC, FISC or cultured SCs; library preparation and paired-end
12 sequencing was performed by Beijing Genomics Institute (BGI, Hong Kong) with read length of 100
13 bp on Nextseq 500.

14 **Immunoblotting and immunofluorescence.** These were performed according to our standard
15 procedures⁶²⁻⁶⁵. Briefly, total proteins were extracted using RIPA lysis buffer. The following dilutions
16 of antibodies were used for each antibody: PAX7 (Developmental Studies Hybridoma Bank; 1:1000),
17 GAPDH (Santa Cruz Biotechnology, sc-137179, 1:2000). For immunofluorescence staining,
18 following antibodies and related dilutions were used: PAX7 (Developmental Studies Hybridoma
19 Bank; 1:100). All fluorescent images were captured with a fluorescence microscope (Leica DM6000
20 B).

21 **Chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq).**
22 Following our standard procedure^{44, 62, 63, 66}, approximately 300,000-600,000 FISCs were
23 cross-linked with 1% formaldehyde for 10 minutes at room temperature and quenched by 125 mM
24 Glycine. Cells were washed and collected by centrifugation at 700 g for 5 min at 4°C, flash-frozen in
25 liquid nitrogen and stored at – 80°C. Cells were lysed in 1 ml of LB1 (50 mM HEPES-KOH pH 7.5,
26 140 mM NaCl, 1 mM EDTA, 0.25% Triton X-100, 0.5% NP-40, 10% glycerol, supplemented with
27 cOmplete protease inhibitors (Roche)) and incubated at 4°C for 10 min. Following centrifugation at
28 1,350 g for 5 min at 4°C, the pellets were washed with 1 ml of LB2 (10 mM Tris-HCl pH 8.0, 200
29 mM NaCl, 1 mM EDTA pH 8.0, 1 mM EGTA pH 8.0, supplemented with complete proteinase

1 inhibitors) by incubating at 4°C for 10 min. Following centrifugation at 1,350 g for 5 min at 4°C,
2 nuclei were rinsed twice with 1 ml of sonication buffer (10 mM Tris-HCl pH 8.0, 0.1% SDS, 1 mM
3 EDTA, supplemented with complete proteinase inhibitors). Then, the nuclei were re-suspended in
4 sonication buffer and sonicated with Covaris S220 (Intensity 140W, Duty Cycle 5%, Cycles per
5 Burst 200, Time 7 mins). The resulting lysate was supplied with NaCl and Triton X 100 to reach a
6 final concentration at 150 mM NaCl and 1% Triton X 100 and then cleared by centrifugation for 20
7 min at 20,000 g and then incubated with washed Dynabeads Protein G (Invitrogen, 10004D) for 2 hrs
8 at 4°C. The cleared lysate was incubated with 0.75 µg H3K27ac (ab4729, Abcam) or IgG control
9 (sc-2027, Santa Cruz) for overnight at 4°C. On the next day, 10 µl Dynabeads Protein G was washed
10 with IP buffer (10 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1% SDS, 1 mM EDTA, 1% Triton X100,
11 supplemented with complete proteinase inhibitors) and then applied to each IP reaction followed by
12 incubation at 4°C for another 2 hrs. Beads were then collected and washed twice with IP buffer, two
13 times with High salt wash buffer (10 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.1% SDS, 1 mM EDTA,
14 1% Triton X100, supplemented with complete proteinase inhibitors), two times with LiCl wash
15 buffer (10 mM Tris-HCl pH 8, 250 mM LiCl, 1 mM EDTA, 0.5% NP-40, 0.5% deoxycholate,
16 supplemented with complete proteinase inhibitors) and one time with cold TE buffer (10 mM
17 Tris-HCl pH8.0, 1 mM EDTA, 50 mM NaCl). Immunocomplexes were eluted in ChIP elution buffer
18 (50 mM Tris-HCl pH8.0, 10 mM EDTA, 1% SDS) with incubation at 65°C for 30 min and eluents
19 were reverse crosslinked by incubating at 65°C for 16 hrs. Immunoprecipitated DNA was treated
20 with RNase A (0.2 mg/ml) at 37°C for 2 hrs, followed by Proteinase K treatment (0.2 mg ml⁻¹) at
21 55°C for 3 hrs. Immunoprecipitated DNA was then subject to phenol:chloroform extraction and
22 ethanol precipitation and finally resuspended in 1x TE buffer. Libraries were prepared using
23 NEBNext® Ultra™ II DNA Library Preparation Kit (NEB, E7645S) and paired-end sequenced with
24 read lengths of 150 bp on an Illumina HiSeq X Ten instrument.

25 ***In situ* Hi-C library preparation and sequencing.** Generation of Hi-C libraries with a low cell
26 number of SCs was performed according to modified protocols⁸. Approximately 100,000 SCs were
27 cross-linked with 1% formaldehyde for 10 minutes at room temperature and quenched by 0.2 M
28 Glycine. The samples were lysed in 500 µL of ice-cold Hi-C lysis buffer (10 mM Tris-HCl pH 8.0,
29 10 mM NaCl, 0.2% Igepal CA630, supplemented with complete proteinase inhibitors) on ice for 15

1 min. Samples were then centrifuged at 2,500 g for 5 min. Pelleted nuclei were washed once with 500
2 μL of $1.25 \times$ NEBuffer 3.1. The supernatant was discarded and the nuclei was re-suspended in 358
3 μL of $1.25 \times$ NEBuffer 3.1 and add 11 μL of 10% SDS followed by incubation at 37°C for 1 hr. After
4 incubating, 75 μL of 10% Triton X-100 was added to quench the SDS and then incubated at 37°C for
5 1 hr. 100 U of DpnII restriction enzyme (NEB, R0543) was added and chromatin were digested at
6 37°C for overnight. Samples were incubated at 62°C for 20 min to inactivate DpnII and then cooled
7 to room temperature. Samples were then centrifuged at 2,500 g for 5 min. Pelleted nuclei were
8 re-suspended in 50 μL of fill-in master mix (3.75 μL of 0.4 mM biotin-14-dATP, 1.5 μL of 1 mM
9 dCTP, 1.5 μL of 1 mM dGTP, 1.5 μL of 1 mM dTTP mix, 2 μL of 5 U μL^{-1} DNA polymerase I, Large
10 (Klenow) Fragment, $1 \times$ NEBuffer 3.1) to fill in the restriction fragment overhangs and mark the
11 DNA ends with biotin. Samples were mixed by pipetting and incubated at 23°C for 4 hrs. Ligation
12 master mix (398 μL of water, 50 μL of $10 \times$ NEB T4 DNA ligase buffer, 1 μL of 50 mg/ml bovine
13 serum albumin, 1 μL of 400 U μL^{-1} T4 DNA ligase) was added and samples were incubated at 16°C
14 for overnight in ThermoMixer C with interval shake. Nuclei were pelleted by centrifugation at 2,500
15 g for 5 min and 380 μL of the supernatant was discarded. Pellets were then re-suspended in the
16 remaining 120 μL ligation mix supplemented with 12 μL of 10% SDS and 5 μL of 20 mg ml^{-1}
17 proteinase K and incubated at 55°C for 2 hrs with shaking at 1,000 rpm. 15 μL of 5M NaCl was
18 added and the reaction was incubated at 65°C for 16 hrs. DNA samples were subjected to phenol:
19 chloroform extraction and ethanol precipitation and finally re-suspended in 130 μL $0.1 \times$ TE buffer.
20 The purified DNA samples were then sheared to a length of ~ 300 bp using Covaris S220 instrument
21 (Intensity 175W, Duty Cycle 10%, Cycles per Burst 200, Time 150 s). The biotinylated DNA was
22 pulled down by 10 μL of 10 mg/ml Dynabeads MyOne Streptavidin C1 beads (Life technologies,
23 65001). The beads were then re-suspended in 23 μL of 10 mM Tris-Cl pH 8.0 and libraries were
24 prepared by on-bead reactions using NEBNext® Ultra™ II DNA Library Preparation Kit (NEB,
25 E7645S). The beads were separated on a magnetic stand and the supernatant was discarded. After
26 washes, the beads were re-suspended in 20 μL of 10 mM Tris buffer and boiled at 98°C for 10 min.
27 The elute was amplified for 10-13 cycles of PCR with Phanta Master Mix (Vazyme, P511-01) and
28 the PCR products were purified using VAHTS DNA Clean Beads (Vazyme, N411-01). The Hi-C
29 libraries were paired-end sequenced with read lengths of 150 bp on an Illumina HiSeq X Ten
30 instrument.

1 **Gene expression analysis of RNA-seq data.** The raw reads of total RNA-seq were processed
2 following the procedures described in our previous publication^{62, 63, 67}. Briefly, the adapter and
3 low-quality sequences were trimmed from 3' to 5' ends for each read and the reads shorter than 50 bp
4 were discarded. The reads that passed the quality control were mapped to mouse genome (mm9) with
5 TopHat2. Cufflinks were then used to estimate transcript abundance in Fragments Per Kilobase per
6 Million (FPKM). Genes were annotated as differentially expressed if the change of expression level
7 is greater than 2 folds between two stages/conditions.

8 **ChIP-seq data analysis.** Raw ChIP-seq reads were processed as previously described⁶⁶. Briefly, the
9 adapter and low-quality sequences were trimmed from 3' to 5' ends by Trimmomatic and the reads
10 shorter than 36 bp were discarded. Subsequently, the preprocessed reads were aligned to the mouse
11 genome (mm9) using Bowtie2. The aligned reads were then converted to bam format using samtools,
12 and the duplicate reads were removed by Picard (<http://broadinstitute.github.io/picard>). Peaks were
13 then identified by MACS2 with q-value equal to 0.01 by using the IgG control sample as background.
14 H3K27ac CHIP-seq is used for typical enhancer (TE) and Super enhancers (SEs) identification. For
15 typical enhancer (TE) identification, blacklisted regions from ENCODE were first excluded, and the
16 peaks overlapping with +/- 2kb regions centered at Transcription Start Site (TSS) from Refseq (June
17 2015) were also discarded. The filtered peaks were defined as TEs and assigned to the expressed
18 transcripts (RPKM >0.5) whose TSS are nearest to the center of the enhancer regions. Super
19 enhancers (SEs) were identified as previously described with minor adjustment⁶⁸. Only the enhancers
20 within 12.5 kb of each another and can be assigned to the same genes were stitched together, then
21 subject to the ROSE algorithm for SE identification.

22 ***In situ* Hi-C data processing.** The *in-situ* Hi-C data was processed with a standard pipeline
23 HiC-Pro³³. First, adaptor sequences and poor-quality reads were removed using Trimmomatic
24 (ILLUMINACLIP: TruSeq3-PE-2.fa:2:30:10; SLIDINGWINDOW: 4:15; MINLEN:50). The filtered
25 reads were then aligned to reference genome (mm9) in two steps: 1) global alignment was first
26 conducted for all pair-end reads, 2) the unaligned reads were split into prospective fragments using
27 restriction enzyme recognition sequence (GATCGATC) and aligned again. All aligned reads were
28 then merged together and assigned to restriction fragment, while low quality (MAPQ<30) or multiple
29 alignment reads were discarded. Invalid fragments including unpaired fragments (singleton),

1 juxtaposed fragments (re-ligation pairs), un-ligated fragments (dangling end), self-circularized
2 fragments (self-cycle), and PCR duplicates were removed from each biological replicate. The
3 remaining validate pairs from all replicates of each stage were then merged, followed by read depth
4 normalization using HOMER and matrix balancing using iterative correction and eigenvector
5 decomposition (ICE) normalization to obtain comparable interaction matrix between different stages.

6 **Identification and analysis of compartments and TADs.** Following previous procedure⁸, to
7 separate the genome into A and B compartments, the ICE normalized intra-chromosomal interaction
8 matrices at 100-kb resolution were transformed to observe/expect contact matrices, and the
9 background (expected) contact matrices were generated to eliminate bias caused by
10 distance-dependent decay of interaction frequency and read depth difference. Pearson correlation
11 was then applied to the transformed matrices and the first principal component (PC1) of these
12 matrices was divided into two clusters. The annotation of genes and the expression profile were used
13 to assign positive PC1 value to gene-rich component as compartment A and negative PC1 value to
14 gene-poor component as compartment B. The compartmentalization strength was measured using
15 cooltools (<https://github.com/mirnylab/cooltools>). Briefly, all 100-kb compartment bins were divided
16 into 50 equal degrees based on the ranking of PC1 values, and the average interaction strength
17 (observed/expected) was calculated between pairs of 100-kb loci arranged by their PC1 and
18 normalized by genomic distance. The dynamics of interaction strength for intra/inter compartment (A
19 vs. A, B vs. B, A vs. B) were measured using top 20% of compartment bins from both compartments
20 A and B based on absolute PC1 value. Normalized contact matrix at 40 kb resolution of each time
21 point was used for TAD identification using TopDom³⁷. In brief, for each 40-kb bin across the
22 genome, a signal of the average interaction frequency of all pairs of genome regions within a distinct
23 window centered on this bin was calculated, thus TAD boundary was identified with local minimal
24 signal within certain window. The false detected TADs without local interaction aggregation were
25 filtered out by statistical testing. Invariant TADs were defined using following criteria: 1) the
26 distance of both TAD boundaries between two conditions is no more than 40 kb; 2) the overlapping
27 between two TADs should be larger than 80%; stage-specific TADs were defined otherwise. The
28 insulation index for each bin was generated based on a previously described method⁶⁹. In brief, the
29 insulation index for each bin was obtained by calculating the number of interaction across a specific

1 bin within certain distance. The insulation score of the identified TAD border was also defined as
2 previously described⁶⁹, which used the local maximum on the outside of TAD to minus the local
3 minimum on the inside of TAD of each boundary bin. The domain score was calculated for each
4 TAD at 40 kb interaction matrix through dividing total intra-TAD contacts by all contacts involving
5 the TAD.

6 **Identification of TAD-TAD interactions and TAD clusters.** The significant TAD-TAD interactions
7 and TAD clusters were identified as previously described⁷⁰. First, the significance of TAD-TAD
8 interactions was tested by applying a non-central hypergeometric (NCHG) model for all pairs of
9 TADs, then corrected by multiple testing with false discovery rate (FDR) < 1% using
10 Benjamini-Hochberg method. Significant TAD-TAD interactions were defined for those TAD pairs
11 with adjusted p-value <0.05 and five-fold enrichment of observed interactions over expected. All
12 significant TAD pairs were then used to identify TAD clusters using the NetworkX Python library
13 (<http://networkx.github.io/>), where each TAD was represented by a node and interactions were
14 represented by edges. Maximal size of TAD clusters was calculated using the Bron-Kerbosch
15 algorithm and allvial package (<https://github.com/mbojan/allvial>) was used to plot the TAD cluster
16 dynamics.

17 **3D modeling of TAD cluster.** The normalized interaction matrix containing TAD cluster as well as
18 their genomic 3D context were used to build 3D modelling following previous method with
19 modification⁷¹. (1) to select key elements contained in the region (that is, TADs in selected TAD
20 cluster); (2) to retrieve the top 1% interactors of each of these elements; (3) to create a network
21 between interactors where edges correspond to top 1% interactions between any of the selected
22 interactions; (4) to group the networks allowing the groups in which the ratio (number of
23 edges)/(number of nodes) smaller than 5 were filtered; and (5) to extract the connected groups that
24 contain the most key elements. Next, the normalized interaction matrices of selected regions were
25 modeled using TADdyn⁷². A similar protocol as previously described was used¹⁴. Briefly, a total of
26 1000 models were generated for each genomic region and dataset. Each ensemble of models was
27 next clustered on the basis of structural similarity. The absence of major structural differences
28 between clusters prompted us to use all of them for further analysis. Next, TADbit⁷³ was used to
29 measure the following features of the models: (1) distance between particles containing genomic

1 regions of interest in the model ensemble; (2) distance distribution between selected pairs or particles;
2 and (3) notable differential distance distributions assessed by the two-sample Kolmogorov–Smirnov
3 statistic. Finally, model images were generated with Chimera⁷⁴.

4 **Identification of chromatin loops.** The loops were identified by FitHIC2⁴²
5 (<https://ay-lab.github.io/fithic/>) at 5 kb resolution of interaction matrix with p-value<1e-5 and
6 required a five-fold enrichment of observed contacts over expected based on genomic distance. To
7 obtain stage-specific loops, fold change (FC) of interaction strength (observed/expected) between
8 two stages was calculated for each identified loop. Then the top 30% dynamics loops with absolute
9 FC > 2 were defined as stage-specific loops.

10 **Motif scanning.** The position-specific probability matrix of motifs was obtained from the
11 TRANSFAC database⁷⁵. Utilizing FIMO, the TAD boundaries as well as enhancers and promoter
12 regions (+/- 2kb) of PAX7 were scanned for motifs of all expressed transcription factors in both QSC
13 and FISC. The motif was kept if the p-value was smaller than 10⁻⁴.

14 **PCA analysis.** Principal component analysis (PCA) was generated using the prcomp package in R
15 (V 3.6.2) with FPKM for RNA-seq, PC1 for A/B compartments and domain score/insulation score
16 for conserved TADs. The eigenvalue for the first two components were used, and the explanatory
17 ratio of each component was calculated using the eigenvalue divided by total eigenvalue.

18 **Data availability.** *In situ* Hi-C and RNA-seq data reported in this paper are deposited in the Gene
19 Expression Omnibus database under accession GSE189842. All used datasets from other
20 publications are summarized in Suppl. Table 2. All other data supporting the findings of this study
21 are available from the corresponding author on reasonable request. The source data for Figs 1h, i,
22 2b-e, k, l, 3b, c, 4a, 5g, j and 6f-h, k-m Suppl. Figs 2a, c, 3a, b 5a, b, f, 7f, g, 8f-h are provided in the
23 Source Data file.

24

1 Reference

- 2 1. Almada, A.E. & Wagers, A.J. Molecular circuitry of stem cell fate in skeletal muscle
3 regeneration, ageing and disease. *Nat Rev Mol Cell Biol* 17, 267-279 (2016).
- 4 2. Yin, H., Price, F. & Rudnicki, M.A. Satellite cells and the muscle stem cell niche. *Physiol*
5 *Rev* 93, 23-67 (2013).
- 6 3. Feige, P., Brun, C.E., Ritso, M. & Rudnicki, M.A. Orienting Muscle Stem Cells for
7 Regeneration in Homeostasis, Aging, and Disease. *Cell Stem Cell* 23, 653-664 (2018).
- 8 4. Hwang, A.B. & Brack, A.S. Muscle Stem Cells and Aging. *Curr Top Dev Biol* 126, 299-322
9 (2018).
- 10 5. Bengal, E., Perdiguero, E., Serrano, A.L. & Munoz-Canoves, P. Rejuvenating stem cells to
11 restore muscle regeneration in aging. *F1000Res* 6, 76 (2017).
- 12 6. Seale, P. et al. Pax7 is required for the specification of myogenic satellite cells. *Cell* 102,
13 777-786 (2000).
- 14 7. Bonev, B. et al. Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell*
15 171, 557-572 e524 (2017).
- 16 8. Rao, S.S. et al. A 3D map of the human genome at kilobase resolution reveals principles of
17 chromatin looping. *Cell* 159, 1665-1680 (2014).
- 18 9. Stadhouders, R. et al. Transcription factors orchestrate dynamic interplay between genome
19 topology and gene regulation during cell reprogramming. *Nat Genet* 50, 238-249 (2018).
- 20 10. Dixon, J.R. et al. Topological domains in mammalian genomes identified by analysis of
21 chromatin interactions. *Nature* 485, 376-380 (2012).
- 22 11. Naumova, N. et al. Organization of the mitotic chromosome. *Science* 342, 948-953 (2013).
- 23 12. Guo, Y. et al. CRISPR Inversion of CTCF Sites Alters Genome Topology and
24 Enhancer/Promoter Function. *Cell* 162, 900-910 (2015).
- 25 13. Kragestein, B.K. et al. Dynamic 3D chromatin architecture contributes to enhancer
26 specificity and limb morphogenesis. *Nat Genet* 50, 1463-1473 (2018).
- 27 14. Miguel-Escalada, I. et al. Human pancreatic islet three-dimensional chromatin architecture
28 provides insights into the genetics of type 2 diabetes. *Nat Genet* 51, 1137-1148 (2019).
- 29 15. Paulsen, J. et al. Long-range interactions between topologically associating domains shape
30 the four-dimensional genome during differentiation. *Nat Genet* 51, 835-843 (2019).
- 31 16. Beagrie, R.A. et al. Complex multi-enhancer contacts captured by genome architecture
32 mapping. *Nature* 543, 519-524 (2017).
- 33 17. Quinodoz, S.A. et al. Higher-Order Inter-chromosomal Hubs Shape 3D Genome
34 Organization in the Nucleus. *Cell* 174, 744-757 e724 (2018).
- 35 18. Du, Z. et al. Polycomb Group Proteins Regulate Chromatin Architecture in Mouse Oocytes
36 and Early Embryos. *Mol Cell* 77, 825-839 e827 (2020).
- 37 19. Hug, C.B., Grimaldi, A.G., Kruse, K. & Vaquerizas, J.M. Chromatin Architecture Emerges
38 during Zygotic Genome Activation Independent of Transcription. *Cell* 169, 216-228 e219
39 (2017).
- 40 20. Machado, L. et al. In Situ Fixation Redefines Quiescence and Early Activation of Skeletal
41 Muscle Stem Cells. *Cell Rep* 21, 1982-1993 (2017).
- 42 21. van Velthoven, C.T.J., de Morree, A., Egner, I.M., Brett, J.O. & Rando, T.A. Transcriptional
43 Profiling of Quiescent Muscle Stem Cells In Vivo. *Cell Rep* 21, 1994-2004 (2017).
- 44 22. Yue, L., Wan, R., Luan, S., Zeng, W. & Cheung, T.H. Dek Modulates Global Intron Retention

- 1 during Muscle Stem Cells Quiescence Exit. *Dev Cell* 53, 661-676 e666 (2020).
- 2 23. Chandra, T. et al. Global reorganization of the nuclear landscape in senescent cells. *Cell Rep*
3 10, 471-483 (2015).
- 4 24. Zirkel, A. et al. HMGB2 Loss upon Senescence Entry Disrupts Genomic Organization and
5 Induces CTCF Clustering across Cell Types. *Mol Cell* 70, 730-744 e736 (2018).
- 6 25. Sati, S. et al. 4D Genome Rewiring during Oncogene-Induced and Replicative Senescence.
7 *Mol Cell* 78, 522-538 e529 (2020).
- 8 26. Sofiadis, K. et al. HMGB1 coordinates SASP-related chromatin folding and RNA
9 homeostasis on the path to senescence. *Mol Syst Biol* 17, e9760 (2021).
- 10 27. Sousa-Victor, P. et al. Geriatric muscle stem cells switch reversible quiescence into
11 senescence. *Nature* 506, 316-321 (2014).
- 12 28. Sousa-Victor, P., Garcia-Prat, L., Serrano, A.L., Perdiguero, E. & Munoz-Canoves, P. Muscle
13 stem cell aging: regulation and rejuvenation. *Trends Endocrinol Metab* 26, 287-296 (2015).
- 14 29. Liu, L. et al. Chromatin modifications as determinants of muscle stem cell quiescence and
15 chronological aging. *Cell Rep* 4, 189-204 (2013).
- 16 30. Schworer, S. et al. Epigenetic stress responses induce muscle stem-cell ageing by Hoxa9
17 developmental signals. *Nature* 540, 428-432 (2016).
- 18 31. Hernando-Herraez, I. et al. Ageing affects DNA methylation drift and transcriptional
19 cell-to-cell variability in mouse muscle stem cells. *Nat Commun* 10, 4361 (2019).
- 20 32. Rocheteau, P., Gayraud-Morel, B., Siegl-Cachedenier, I., Blasco, M.A. & Tajbakhsh, S. A
21 subpopulation of adult skeletal muscle stem cells retains all template DNA strands after cell
22 division. *Cell* 148, 112-125 (2012).
- 23 33. Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing.
24 *Genome Biol* 16, 259 (2015).
- 25 34. Nagano, T. et al. Comparison of Hi-C results using in-solution versus in-nucleus ligation.
26 *Genome Biol* 16, 175 (2015).
- 27 35. Rivas-Pardo, J.A. et al. Work Done by Titin Protein Folding Assists Muscle Contraction. *Cell*
28 *Rep* 14, 1339-1347 (2016).
- 29 36. Linke, W.A. Titin Gene and Protein Functions in Passive and Active Muscle. *Annu Rev*
30 *Physiol* 80, 389-411 (2018).
- 31 37. Shin, H. et al. TopDom: an efficient and deterministic method for identifying topological
32 domains in genomes. *Nucleic Acids Res* 44, e70 (2016).
- 33 38. Wang, Q., Sun, Q., Czajkowsky, D.M. & Shao, Z. Sub-kb Hi-C in *D. melanogaster* reveals
34 conserved characteristics of TADs between insect and mammalian cells. *Nat Commun* 9, 188
35 (2018).
- 36 39. Javierre, B.M. et al. Lineage-Specific Genome Architecture Links Enhancers and Non-coding
37 Disease Variants to Target Gene Promoters. *Cell* 167, 1369-1384 e1319 (2016).
- 38 40. Li, L. et al. Widespread rearrangement of 3D chromatin organization underlies
39 polycomb-mediated stress-induced silencing. *Mol Cell* 58, 216-231 (2015).
- 40 41. Novo, C.L. et al. Long-Range Enhancer Interactions Are Prevalent in Mouse Embryonic
41 Stem Cells and Are Reorganized upon Pluripotent State Transition. *Cell Rep* 22, 2615-2627
42 (2018).
- 43 42. Kaul, A., Bhattacharyya, S. & Ay, F. Identifying statistically significant chromatin contacts
44 from Hi-C data with FitHiC2. *Nat Protoc* 15, 991-1012 (2020).

- 1 43. Ancel, S., Stuelsatz, P. & Feige, J.N. Muscle Stem Cell Quiescence: Controlling Stemness by
2 Staying Asleep. *Trends Cell Biol* 31, 556-568 (2021).
- 3 44. He, L. et al. CRISPR/Cas9/AAV9-mediated in vivo editing identifies MYC regulation of 3D
4 genome in skeletal muscle stem cell. *Stem Cell Reports* (2021).
- 5 45. van Overbeek, M. et al. DNA Repair Profiling Reveals Nonrandom Outcomes at
6 Cas9-Mediated Breaks. *Mol Cell* 63, 633-646 (2016).
- 7 46. Garcia-Prat, L. et al. FoxO maintains a genuine muscle stem-cell quiescent state until
8 geriatric age. *Nat Cell Biol* 22, 1307-1318 (2020).
- 9 47. Evano, B. & Tajbakhsh, S. Skeletal muscle stem cells in comfort and stress. *NPJ Regen Med*
10 3, 24 (2018).
- 11 48. Zheng, H. et al. Resetting Epigenetic Memory by Reprogramming of Histone Modifications
12 in Mammals. *Mol Cell* 63, 1066-1079 (2016).
- 13 49. Bian, Q., Anderson, E.C., Yang, Q. & Meyer, B.J. Histone H3K9 methylation promotes
14 formation of genome compartments in *Caenorhabditis elegans* via chromosome compaction
15 and perinuclear anchoring. *Proc Natl Acad Sci U S A* 117, 11459-11470 (2020).
- 16 50. Nora, E.P. et al. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome
17 Domains from Genomic Compartmentalization. *Cell* 169, 930-944 e922 (2017).
- 18 51. Fudenberg, G. et al. Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep* 15,
19 2038-2049 (2016).
- 20 52. Heinz, S. et al. Transcription Elongation Can Affect Genome 3D Structure. *Cell* 174,
21 1522-1536 e1522 (2018).
- 22 53. Kim, Y.H. et al. Rev-erbalpha dynamically modulates chromatin looping to control circadian
23 gene transcription. *Science* 359, 1274-1277 (2018).
- 24 54. Munoz-Canoves, P., Neves, J. & Sousa-Victor, P. Understanding muscle regenerative decline
25 with aging: new approaches to bring back youthfulness to aged stem cells. *FEBS J* 287,
26 406-416 (2020).
- 27 55. Chen, W., Datzkiw, D. & Rudnicki, M.A. Satellite cells in ageing: use it or lose it. *Open Biol*
28 10, 200048 (2020).
- 29 56. Garcia-Prat, L. et al. Autophagy maintains stemness by preventing senescence. *Nature* 529,
30 37-42 (2016).
- 31 57. Sambasivan, R. et al. Pax7-expressing satellite cells are indispensable for adult skeletal
32 muscle regeneration. *Development* 138, 3647-3656 (2011).
- 33 58. von Maltzahn, J., Jones, A.E., Parks, R.J. & Rudnicki, M.A. Pax7 is critical for the normal
34 function of satellite cells in adult skeletal muscle. *Proc Natl Acad Sci U S A* 110,
35 16474-16479 (2013).
- 36 59. Liu, L., Cheung, T.H., Charville, G.W. & Rando, T.A. Isolation of skeletal muscle stem cells
37 by fluorescence-activated cell sorting. *Nat Protoc* 10, 1612-1624 (2015).
- 38 60. Grieger, J.C., Choi, V.W. & Samulski, R.J. Production and characterization of
39 adeno-associated viral vectors. *Nat Protoc* 1, 1412-1428 (2006).
- 40 61. Haeussler, M. et al. Evaluation of off-target and on-target scoring algorithms and integration
41 into the guide RNA selection tool CRISPOR. *Genome Biol* 17, 148 (2016).
- 42 62. Chen, F. et al. YY1 regulates skeletal muscle regeneration through controlling metabolic
43 reprogramming of satellite cells. *EMBO J* 38 (2019).
- 44 63. Zhao, Y. et al. MyoD induced enhancer RNA interacts with hnRNPL to activate target gene

- 1 transcription during myogenic differentiation. *Nat Commun* 10, 5787 (2019).
- 2 64. Li, Y. et al. Long noncoding RNA SAM promotes myoblast proliferation through stabilizing
3 Sugt1 and facilitating kinetochore assembly. *Nat Commun* 11, 2725 (2020).
- 4 65. Chen, X. et al. Translational control by DHX36 binding to 5'UTR G-quadruplex is essential
5 for muscle stem-cell regenerative functions. *Nat Commun* 12, 5043 (2021).
- 6 66. Peng, X.L. et al. MyoD- and FoxO3-mediated hotspot interaction orchestrates super-enhancer
7 activity during myogenic differentiation. *Nucleic Acids Res* 45, 8785-8805 (2017).
- 8 67. Huang, Y. et al. Large scale RNA-binding proteins/LncRNAs interaction analysis to uncover
9 lncRNA nuclear localization mechanisms. *Brief Bioinform* 22 (2021).
- 10 68. Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. *Cell* 155, 934-947
11 (2013).
- 12 69. Crane, E. et al. Condensin-driven remodelling of X chromosome topology during dosage
13 compensation. *Nature* 523, 240-244 (2015).
- 14 70. Paulsen, J. et al. Chrom3D: three-dimensional genome modeling from Hi-C and nuclear
15 lamin-genome contacts. *Genome Biol* 18, 21 (2017).
- 16 71. Stik, G. et al. CTCF is dispensable for immune cell transdifferentiation but facilitates an
17 acute inflammatory response. *Nat Genet* 52, 655-661 (2020).
- 18 72. Di Stefano, M. et al. Transcriptional activation during cell reprogramming correlates with the
19 formation of 3D open chromatin hubs. *Nat Commun* 11, 2564 (2020).
- 20 73. Serra, F. et al. Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals
21 structural features of the fly chromatin colors. *PLoS Comput Biol* 13, e1005665 (2017).
- 22 74. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and
23 analysis. *J Comput Chem* 25, 1605-1612 (2004).
- 24 75. Wingender, E., Dietze, P., Karas, H. & Knuppel, R. TRANSFAC: a database on transcription
25 factors and their DNA binding sites. *Nucleic Acids Res* 24, 238-241 (1996).

26

27

28

29

30

31

32

33

34

35

36

37

38

1 **Acknowledgments**

2 We thank Prof. Zhenyu Ju (Jinan University, China) for providing the aged mice and Prof. Danny
3 Leung (HKUST, Hong Kong, China) for assistance on the Hi-C sequencing. This work was
4 supported by General Research Funds (GRF) from the Research Grants Council (RGC) of the Hong
5 Kong Special Administrative Region (14116918, 14120420, and 14120619 to H.S.; 14115319,
6 14100018, 14100620, 14106117 and 14106521 to H.W.); the National Natural Science Foundation of
7 China (NSFC) to H.W. (Project code: 31871304); Collaborative Research Fund (CRF) from RGC to
8 H.W. (C6018-19GF); NSFC/RGC Joint Research Scheme to H.S. (Project code: N_CUHK 413/18);
9 Hong Kong Epigenomics Project (EpiHK) Fund to H.W. and H.S.; Area of Excellence Scheme (AoE)
10 from RGC (Project number: AoE/M-402/20).

11

12 **Author Contributions**

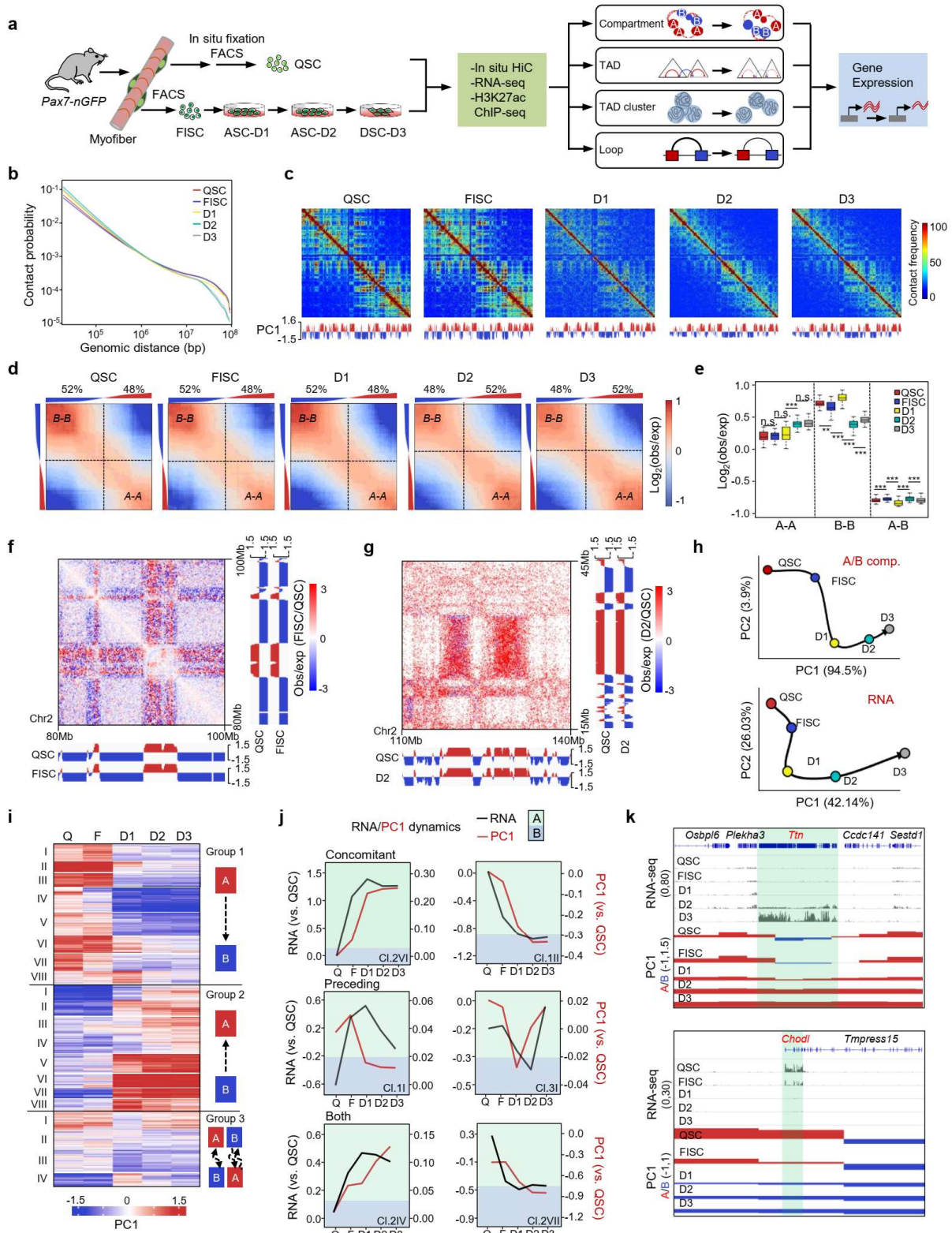
13 H.W., Y.Z., Y.D. and H.S. designed the project; Y.Z., L.H., Y.L., and X.C. performed the research;
14 Y.D. performed bioinformatic analyses of sequencing data; Y.Z., Y.D., L.H. and H.W. wrote the paper.
15 All authors read and approved the final manuscript.

16

17 **Competing interests**

18 The authors declare no competing interests.

19



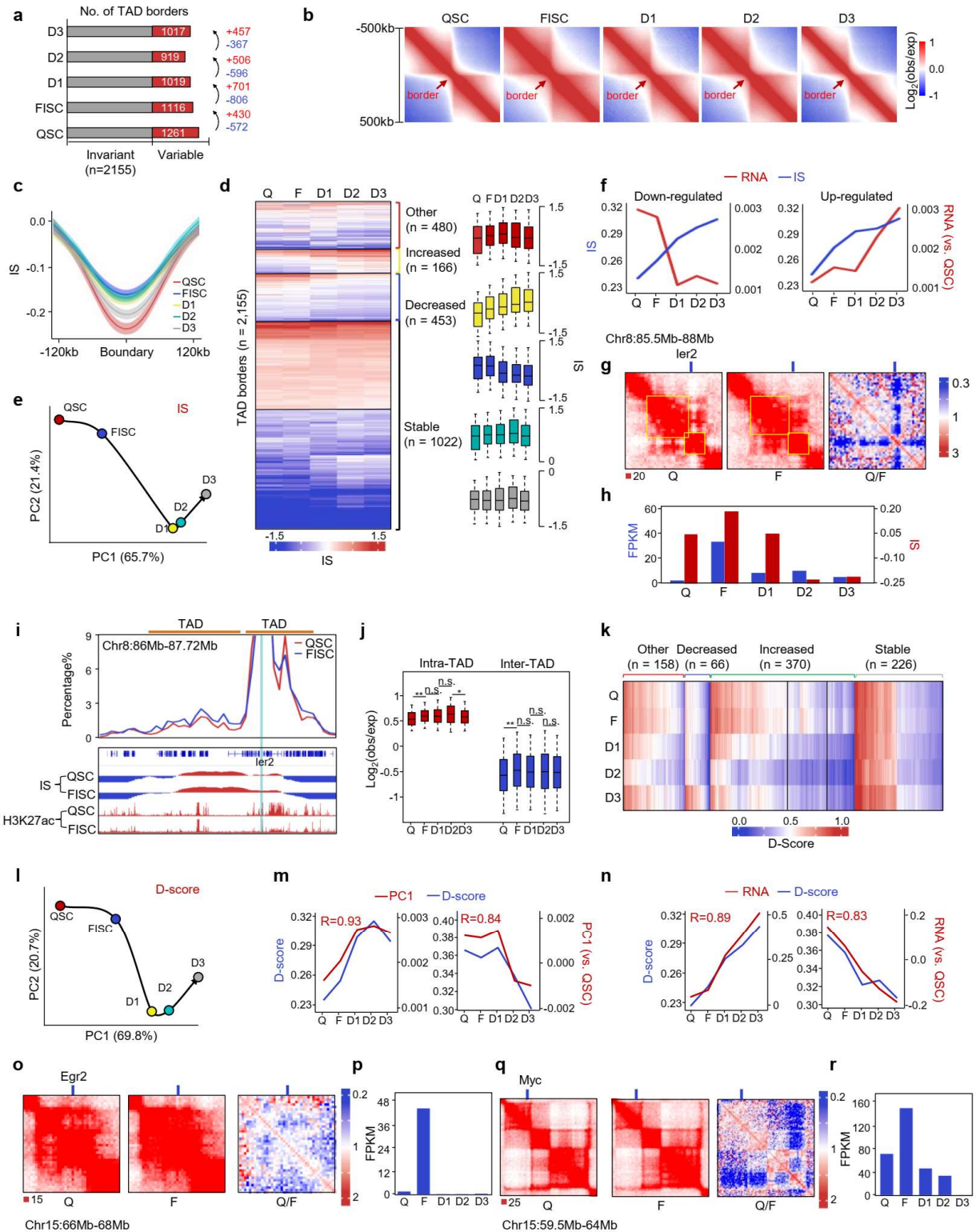
1

2 **Fig. 1 Global reorganization of genome compartmentalization during SC lineage development.**

3 a Schematic overview of the study design. SCs were isolated from 2 month old Pax7-nGFP mouse

1 by FACS sorting with or without the *in situ* fixation to obtain quiescent (QSCs) or freshly isolated
2 SCs (FISCs) respectively. FISCs were cultured in vitro for various time (D1, D2 and D3) to collect
3 activated (ASC) or differentiating (DSC) SCs. The collected cells were subject to *in situ* Hi-C and
4 RNA-seq. RNA-seq data from QSC and H3K27Ac ChIP-seq data from QSC and FISC were
5 obtained from Léo Machado et. al.²⁰ The above datasets were integrated to elucidate 3D genome
6 remodeling at compartment, TAD, TAD cluster, and chromatin loop levels and how 3D rewriting
7 instructs gene expression dynamics during the SC lineage progression. **b** Contact probability was
8 calculated as a function of genomic distance for interactions within individual chromosome across
9 the whole genome at each stage. **c** Observed contact matrices for chr1 at 500-kb resolution and the
10 first eigenvector at 100-kb resolution during the course of SC lineage progression. **d** Saddle plots
11 showing the genome-wide compartmentalization over time. Percentage of compartment A/B and PC1
12 value of each 100-kb loci arranged by eigenvector are shown on top. **e** Contact enrichment of 100-kb
13 loci between compartments of the same (A vs A or B vs B) and different (A vs B) types. For all box
14 plots, center lines denote medians; box limits denote upper and lower quartiles; whiskers denote 1.5×
15 interquartile range. P-values were calculated by a Wilcoxon signed-rank test. **f** Differential heatmap
16 for 20Mb regions on chromosome 2 showing increased A-B interactions in FISC vs QSC. PC1 signal
17 at each stage is displayed on the bottom of the heatmap. **g** Differential heatmap for a 30 Mb region
18 on chromosome 2 showing an increased A-A interaction in FISC vs. QSC. **h** PCA of PC1 values and
19 gene expression dynamics during SC lineage progression. Black arrows denote the hypothetical
20 trajectory. **i** *k*-means clustering (*k*=20) of PC1 values for 100-kb genomic bins that switch
21 compartment at any time point during SC lineage progression. **j** Examples of individual switching
22 clusters with concomitant PC1 and gene expression changes (VI/group2), and clusters with PC1
23 changes preceding expression changes (II/group1), clusters with expression changes preceding PC1
24 changes (I/group, I/group3) or clusters with both IV/group2 and VII/group2 occurring. **k**
25 Genome-browser view of *Osbpl6-Ttn-Sestd1* (top) and *Chodl-Tmpress15* (bottom) loci. PC1 and
26 RNA-seq values at each stage are shown. Green shadings denote B-A-switch (top) and A-B-switch
27 (bottom) regions.

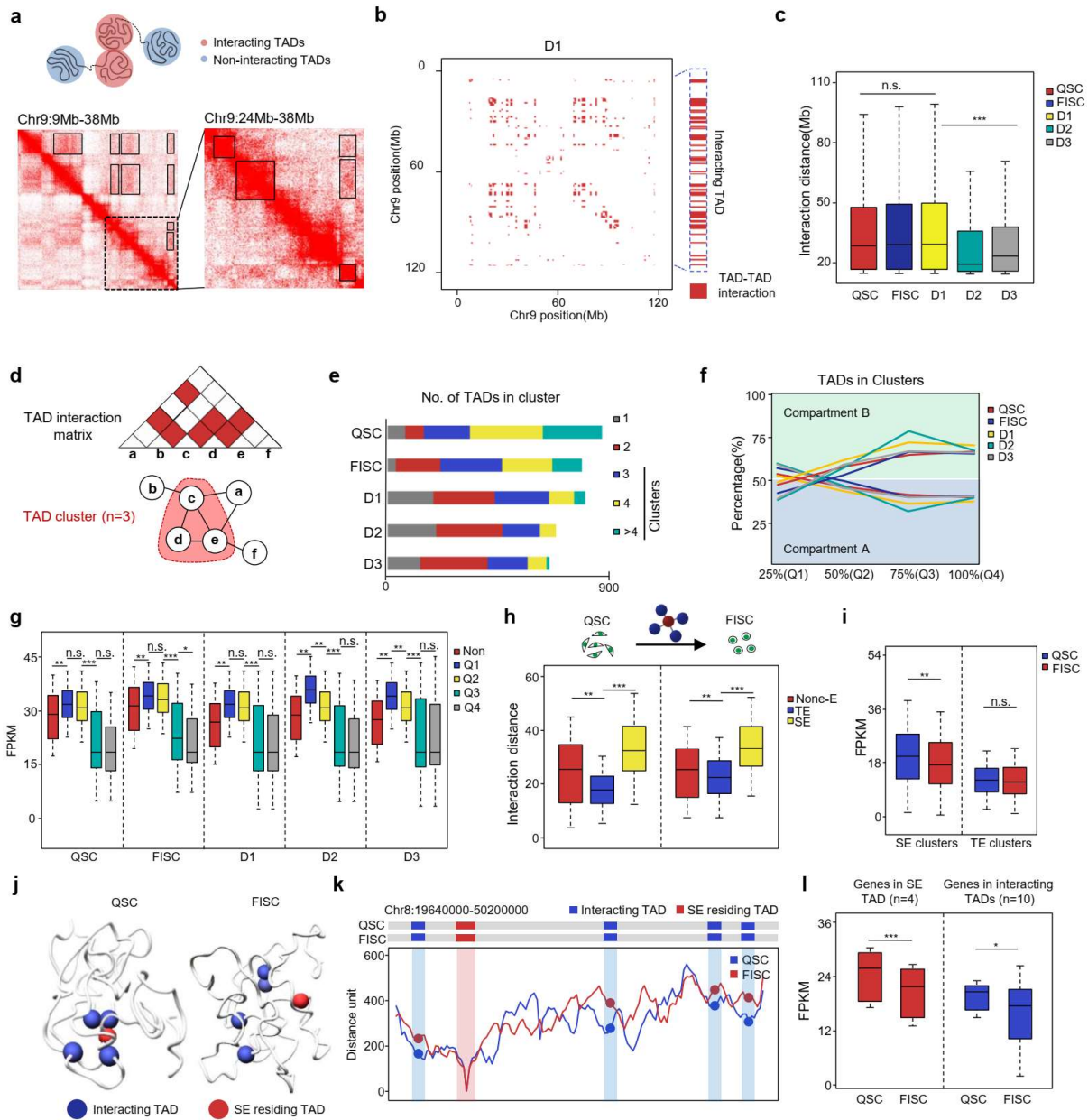
28



1

2 **Fig. 2 Dynamic TAD organization during SC lineage development.** a The number of TAD
 3 borders identified per time point. Invariant borders are present at all stages while variable borders are

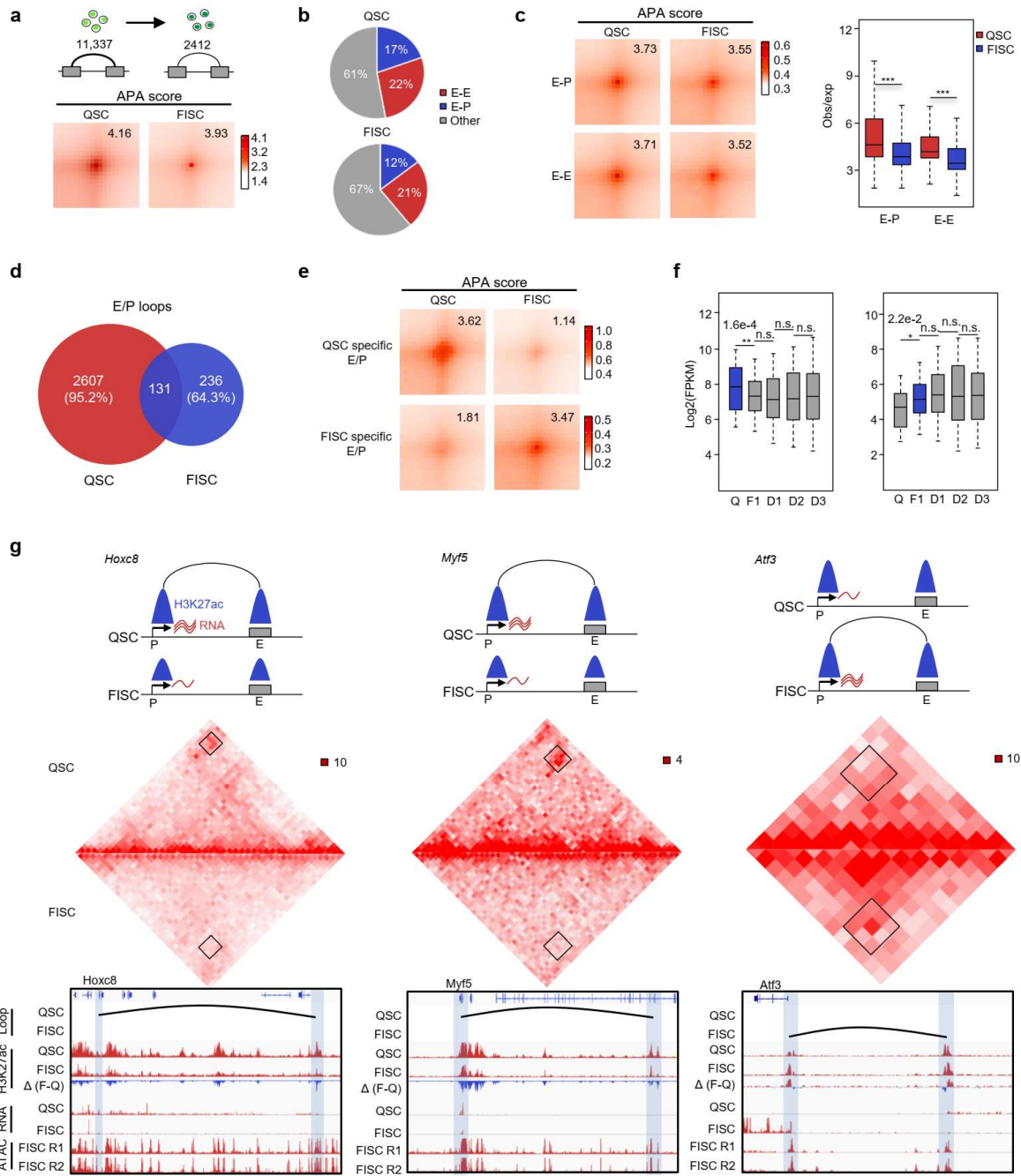
1 lost or acquired in between two adjacent stages. **b** Aggregate Hi-C maps in a 1Mb region centered on
2 the invariant TADs boundaries. Data are presented as the log ratio of observed and expected contacts
3 in 40kb bins. **c** Average insulation score (IS) in a 240 kb region centered on invariant TADs
4 boundaries. Lines denote the mean values, while the shaded ribbons represent the 95% confidence
5 interval (CI). **d** *k*-means clustering ($k = 20$) of IS. Bar graphs show kinetics for groups that increased
6 ($n = 1,258$), decreased ($n = 527$), transiently increased ($n = 254$) or did not change ($n = 2,632$) IS
7 during the lineage progression. **e** PCA of IS during SC lineage progression. Hypothetical trajectory is
8 shown as black arrow. **f** Kinetics of mean expression level (FPKM) changes at dynamic borders
9 harboring up- or down-regulated genes with IS increased. **g** Illustration of *in situ* Hi-C contact maps
10 (40-kb resolution) on the *Ier2* locus. **h** IS kinetics of *Ier2* residing TAD border and its expression
11 dynamics. **i** (top) Representative virtual 4C analysis on *Ier2* locus in QSC and FISC. *Ier2* residing
12 and upstream TADs are indicated. (bottom) Genome-browser view of IS and H3K27ac ChIP-seq are
13 shown. **j** Contact enrichment of intra- and inter-TAD during SC lineage progression. Data is
14 represented as a boxplot based on the intra- and inter-TAD values per TAD. For all box plots, center
15 lines denote medians; box limits denote upper and lower quartiles; whiskers denote $1.5\times$ interquartile
16 range. P-values were calculated by a Wilcoxon signed-rank test. **k** *k*-means clustering ($k = 20$) of D
17 scores. Bar graphs show D-score kinetics for groups that increased ($n = 2,172$), decreased ($n = 467$),
18 or stable ($n = 1,456$). **l** PCA of D-score during SC lineage progression. Hypothetical trajectory is
19 shown as black arrow. **m, n** Average D-score and PC1 kinetics (**m**) or gene expression (**n**) analysis
20 for clusters of TADs that were gained ($n = 2,172$, left) or lost ($n = 467$, right) in D3 vs QSC. Pearson
21 correlation coefficients (R) are indicated. **o** Illustration of *in situ* Hi-C contact maps (40-kb resolution)
22 on the *Egr2* locus on Chr15. **p** The bar graphs show the FPKM kinetics of *Egr2* expression. **q**
23 Illustration of *in situ* Hi-C contact maps (40-kb resolution) on the *Myc* locus on Chr15. **r** The bar
24 graphs show the FPKM kinetics of *Myc* expression.



1

2 **Fig. 3 TADs clusters are dynamically reorganized in FISC vs. QSC.** a (top) Illustration of
 3 TAD-TAD interaction (top). (bottom left) a region of TAD-TAD interactions on chromosome 9 (D1)
 4 are visualized in the Hi-C data. Boxes, TAD-TAD interactions. the dashed region is enlarged on the
 5 right. b Matrix of significant TAD-TAD interactions (red pixels) in chromosome 9 (D1); an example
 6 of interacting TADs is highlighted (red pixels on the right side); c Boxplots showing the average
 7 distance between pair-wise TAD interactions at each time point. Center lines denote medians; box
 8 limits denote upper and lower quartiles; whiskers denote $1.5\times$ interquartile range. P-values were
 9 calculated by a two-sided Wilcoxon test. d (top) illustration of significant interactions between TADs

1 in a matrix (red pixels). (bottom) A TAD cluster (pink) is identified when ≥ 3 TADs are fully
2 connected with each other. Edges indicating interaction between two TADs while nodes indicating
3 TADs. **e** Bar charts depicting the number of TADs in clusters as a function of cluster size at each
4 time point. **f** TADs in clusters were stratified by quartile of distances between pair-wise TAD
5 interactions. Line charts depicting the fraction of TADs in clusters of each quartile in A and B
6 compartments over time of SC development. **g** Gene expression levels outside clusters (Non) and in
7 clusters of each quartile (Q1, Q2, Q3 or Q4) over time of SC development. Box-and-whisker plots
8 show the median FPKM; numbers refer to the number of outliers, including non-protein-coding
9 transcripts. **h** Boxplots showing the average distance between pair-wise TAD interactions that TE or
10 SE engaged in at QSC or FISC stage. **i** Transcription levels in the above SE- or TE-engaged TAD
11 clusters. **j** Top, 3D chromatin conformation model of a SE-engaged TAD cluster identified in QSC
12 and its reorganization in FISC. The SE containing core TAD is marked in red and other interacting
13 TADs are in blue. **k** line plot at 5 kb resolution displaying the median distance distribution between
14 SE-containing TAD and other interacting TADs in the above cluster at QSC and FISC stages. **l**
15 Expression levels of genes ($n=5$) in the above SE-containing TAD or in interacting TADs ($n = 28$).
16 Data in **g**, **h**, **i** and **l** are presented in boxplots. Center line, median; box limits, upper and lower
17 quartiles; whiskers, $1.5\times$ interquartile range. Statistical analyses were done by Wilcoxon signed-rank
18 test; * $P < 0.01$, ** $P < 0.001$, *** $P < 0.0001$.



1

2

3

4

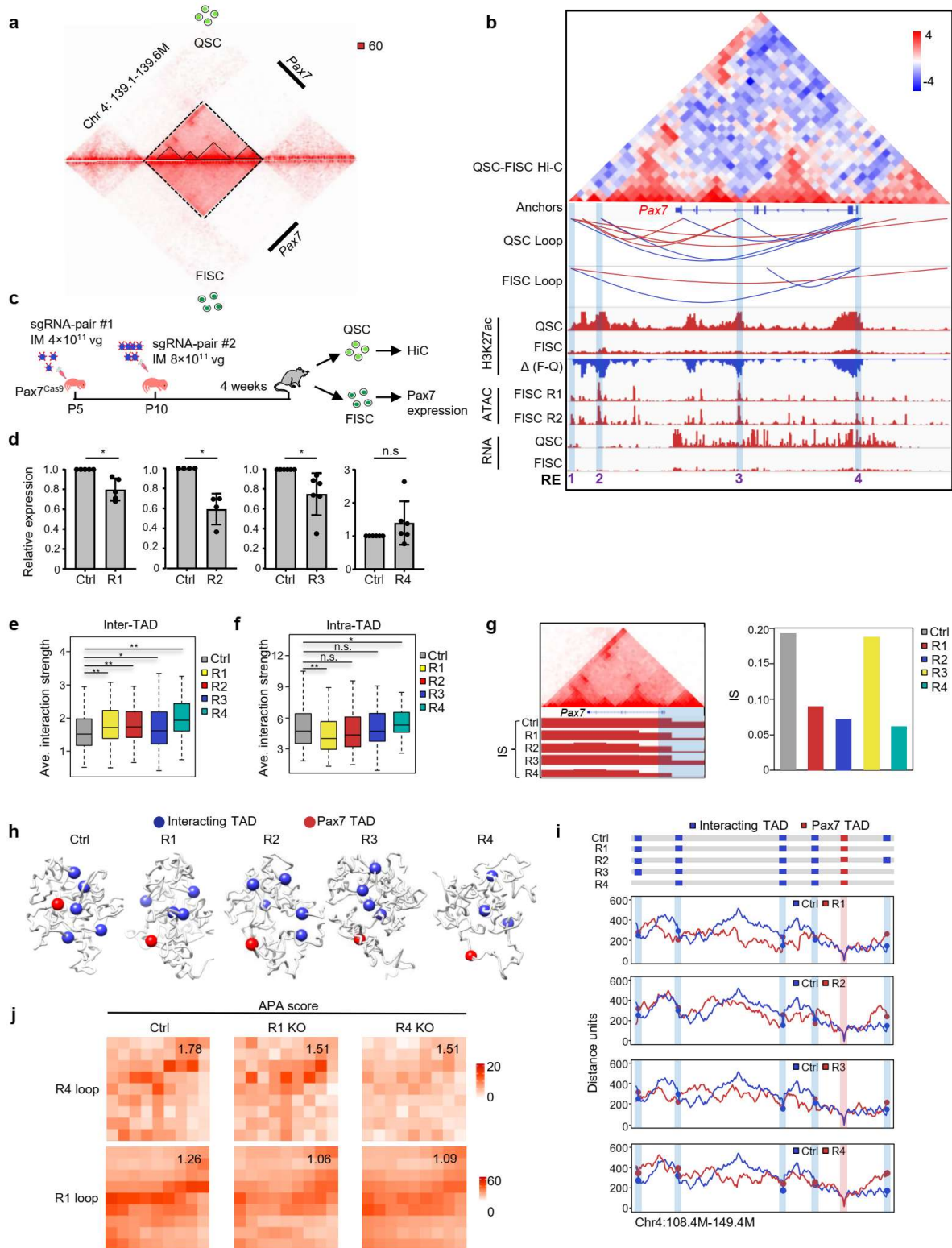
5

6

7

Fig. 4 Chromatin loops are dynamically changed during SC early activation. **a** (top) Illustration of identified chromatin loops in QSC and FISC. (Bottom) Aggregated peak analysis (APA) analysis of chromatin loops detected at QSC (n = 2,790) or FISC (n = 504). Bin size, 5 kb. **b** Distribution of enhancer-enhancer (E-E), Enhancer-Promoter(E-P) and other types of loops in QSC and FISC. **c** (left) APA analysis of E-E and E-P chromatin loops detected in QSC and FISC. (right) Box plot depicting interaction intensity (Obs/exp) of the above E-E and E-P loops. **d** Overlapping of E/P loops between

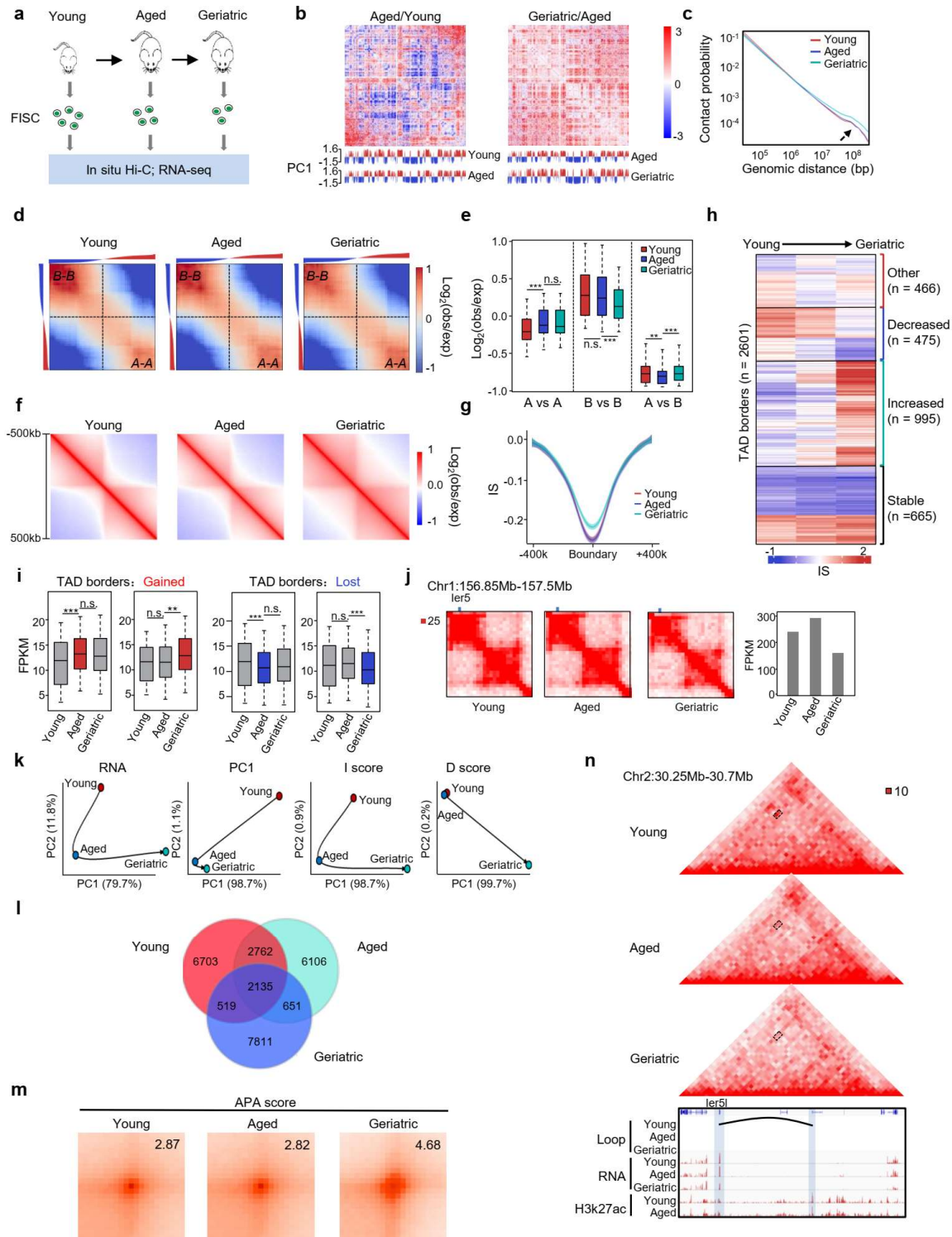
1 QSC and FISC. The number and relative percentage of overlapped and specific E/P loops are shown.
2 e APA analysis of interaction intensity of the above QSC or FISC specific E/P loops. f Box plot
3 depicting expression dynamics of genes within QSC-specific (left, n = 1,874 genes) or FISC-specific
4 (right, n = 469 genes) loops. g (top) Illustration of the dynamics of formation of E-P loop, H3K27ac
5 signal and gene expression on the *Hoxc8* (left), *Myf5* (center) and *Atf3* (right) loci in QSC and FISC.
6 (middle) Hi-C contact maps (5-kb resolution) on the *Hoxc8*, *Myf5*, and *Atf3* loci. (bottom) E-P loops
7 are marked in black curve. Gene promoters and interacting regions are highlighted in blue boxes.
8 Genome browser tracks showing H3K27ac CHIP-seq and RNA-seq profiles in QSC and FISC,
9 ATAC-seq signals of two replicates in FISC are also shown. Data in c and f are presented in boxplots.
10 Center line, median; box limits, upper and lower quartiles; whiskers, 1.5× interquartile range.
11 Statistical analyses were done by Wilcoxon signed-rank test; *P < 0.01, **P < 0.001, ***P < 0.0001.



1

2 **Fig. 5 3D regulatory interactions orchestrate Pax7 expression dynamics in SC. a** Comparison of
 3 contact frequencies at the TAD encompassing *Pax7* locus between QSC (Upper) and FISC (Lower).

1 The dashed and solid lines indicate TAD and sub-TADs respectively. **b** Identification of the four
2 regulatory elements (RE, 1, 2, 3 and 4) within the *Pax7* residing TAD. Subtraction of FISC from
3 QSC contact matrices shows changes in contact frequencies at the TAD. Chromatin loops identified
4 by FitHiC2, dynamics of RNA expression, and H3K27ac signals surrounding the *Pax7* locus in QSC
5 and FISC, ATAC-seq signal in FISC are shown below. REs are highlighted in green box. **c** Schematic
6 illustration of *in vivo* deletion of each RE in SCs by the CRISPR/Cas9/AAV9-sgRNA system.
7 4×10^{11} gc/mouse of AAV9-sgRNA pair #1 were intramuscularly (IM) injected into *Pax7*^{Cas9} mice at
8 postnatal day 5 (P5) and followed by 8×10^{11} gc/mouse of AAV9-sgRNA pair #2 at P10. The mice
9 were sacrificed for FISC collection and analysis after four weeks. For the control (Ctrl) group, the
10 same dose of AAV9 virus containing pAAV9-sgRNA backbone without any sgRNA insertion was
11 injected. **d** *Pax7* expression was detected by qRT-PCR in the above isolated SCs. The *Pax7*
12 expression level was normalized to *Hprt1* mRNA level and presented as mean \pm s.d. n = 4-6 mice per
13 group. **e-f** *In situ* Hi-C was performed in QSC from Ctrl or each mutant (R1-R4) mouse. Boxplots
14 showing inter- (e) or intra-TAD (f) interactions. **g** (left) Hi-C contact maps (40-kb resolution) of
15 *Pax7* TAD in Ctrl(top) and IS (bottom) of *Pax7* TADs in Ctrl and sgR1-sgR4. *Pax7* promoter is
16 highlighted in blue box. The TAD boundary covering the *Pax7* promoter is highlighted. (right) Bar
17 plot showing the IS value of the above boundary. **h** 3D chromatin conformation models for the TAD
18 cluster identified with *Pax7* residing TAD and other interacting TADs in Ctrl or each mutant QSC. **i**
19 Line plots display distance distribution between *Pax7* residing TAD and interacting TADs.
20 Interacting TADs in Ctrl and sgR1-sgR4 are indicated in blue rectangle (top) and highlighted in blue
21 box (bottom). **j** Aggregated peak analysis (APA) of R4 (top) and R1 (bottom) anchored intra-TAD
22 loops that identified in Ctrl. Numbers on each top right indicate average loop strength. *P<0.05,
23 ***P<0.001. ns, no significance. Data in e and f are presented in boxplots. Center line, median; box
24 limits, upper and lower quartiles; whiskers, 1.5 \times interquartile range. Statistical analyses were done
25 by T-test and Wilcoxon signed-rank test; *P < 0.01, **P < 0.001, ***P < 0.0001.



1

2 **Fig. 6 Multiscale analysis of 3D genome during mouse SC aging.** a Schematic illustration of the
3 aging SC study. FISCs were collected from Young (2 months), Aged (23-24 months) and Geriatric
4 (28-32 months) male Pax7-nGFP mice by FACS sorting. *In situ* Hi-C and RNA-seq were performed
5 in the above cells. b The Hi-C data was analyzed and differential heatmaps in Aged/Young (left) or

1 Geriatric/Aged (right) are shown for chromosome 1. The color maps are displayed on the same scale
2 for each comparison. **c** Contact probability relative to genomic distance at each stage. **d** Saddle plots
3 showing the genome-wide compartmentalization over time. Percentage of compartment A/B and PC1
4 value of each 100-kb loci arranged by eigenvector are shown on top. Corners represent
5 intra-compartment(A-A/B-B) interactions. **e** Contact enrichment of 100-kb loci between
6 compartments of the same (A vs A or B vs B) and different (A vs B) types. P values were calculated
7 by a Wilcoxon signed-rank test. **f** TAD boundaries were identified and aggregate Hi-C maps in a
8 1Mb region centered on the invariant boundaries during aging. Data are presented as the log ratio of
9 observed and expected contacts in 40kb bins. **g** Average insulation score in an 800 kb region centered
10 on invariant TAD boundaries. Lines denote the mean values, while the shaded ribbons represent the
11 95% CI. **h** k-means clustering ($k = 20$) of IS. Bar graphs show IS kinetics for groups that increased
12 ($n = 1,146$), decreased ($n = 528$), transiently increased ($n = 312$) or did not change ($n = 674$). **i**
13 Boxplots depicting gene expression (RPKM) dynamics for border regions that were gained (left) or
14 lost (right) during SC aging. Stages that gain or lost TAD borders are indicated in blue and right
15 respectively. **j** Illustration of in situ Hi-C contact maps (40-kb resolution) on the *Ier5* locus during
16 SC aging. The bar graphs show the RNA expression of the *Ier5*. **k** PCA of RNA expression, PC1
17 values, I-score and D-score during SC aging. Black arrows denote the hypothetical trajectory. **l**
18 Chromatin loops were identified and Venn diagrams showing the overlapping of loops at each stage.
19 **m** APA analysis of chromatin loops detected at young, aged and geriatric FISCs. Bin size, 10 kb.
20 Numbers indicate average loop strengthen. **n** (top) Representative Hi-C contact maps of Chr2:
21 30.25Mb-30.7Mb depicting the changes in chromatin looping frequencies between *Ier5l* promoter
22 and a putative enhancer during SC aging. Dashed box, chromatin loops between *Ier5l* promoter and
23 putative enhancer. Genome browser tracks showing corresponding RNA-seq, and H3K27ac signals.
24 Light blue box showing the *Ier5l* promoter and a potential enhancer region. Black curve denotes the
25 chromatin loop. Data in e and i are presented in boxplots. Center line, median; box limits, upper and
26 lower quartiles; whiskers, $1.5\times$ interquartile range. Statistical analyses were done by Wilcoxon
27 signed-rank test; * $P < 0.01$, ** $P < 0.001$, *** $P < 0.0001$

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Suppl.Data1BasicstatisticsofInsituHiCdata.xlsx](#)
- [Suppl.Data2CompartmentanalysisduringSClineageprogression.xlsx](#)
- [Suppl.Data3RNAseqanalysisduringSClineageprogression.xlsx](#)
- [Suppl.Data4TADanalysisduringSClineageprogression.xlsx](#)
- [Suppl.Data5SignificantTADTADinteractionduringSClineageprogression.xlsx](#)
- [Suppl.Data6LoopanalysisfromQSCtoFISC.xlsx](#)
- [Suppl.Data7Pax7KOanalysisinQSC.xlsx](#)
- [Suppl.Data8RNAseqanalysisinagingSCs.xlsx](#)
- [Suppl.Data9CompartmentanalysisduringSCaging.xlsx](#)
- [Suppl.Data10TADXloopanalysisduringSCaging.xlsx](#)
- [Suppl.Table1.xlsx](#)
- [Suppl.Table2.xlsx](#)
- [Suppl3DgenomeinSC.pdf](#)