Skeletal Muscle Satellite Cells Co-Opt the Tenogenic Gene *Scleraxis* to Instruct Regeneration

Yun Bai[1], Tyler Harvey[1], Minjie Hu[2], Colin Bilyou[1], Chen-Ming Fan[1]*

[1] Department of Embryology, Carnegie Institution for Science, Baltimore, MD, USA

[2] College of Life Sciences, Zhejiang University, 310058 Hangzhou, China

*Correspondence: fan@carnegiescience.edu

Keywords: Scleraxis, muscle stem cell, tendon, muscle differentiation, muscle regeneration, sc-RNA sequencing

Author contributions: T. H and C.-M. F. made the initial observation, collected critical preliminary data, and conceptualized the projects. Y.B. and C.-M.F. designed experiments, analyzed data, made conclusions, and wrote the manuscript. C.B. helped in genotyping the mouse strains, data collection and quantification. Y.B. and M.H. performed bioinformatic analyses. All authors contributed to writing and editing.

Competing interests: The authors declare no competing interests.

Data and/or Code Availability: All data that support the findings of this study and the custom code used in this study are available from the corresponding authors upon reasonable request.

Abstract

Skeletal muscles connect bones and tendons for locomotion and posture. Understanding the regenerative processes of muscle, bone and tendon is of importance to basic research and clinical applications. Despite their interconnections, distinct transcription factors have been reported to orchestrate each tissue's developmental and regenerative processes. Here we show that *Scx* expression is not detectable in adult muscle stem cells (also known as satellite cells, SCs) during quiescence. *Scx* expression begins in activated SCs and continues throughout regenerative myogenesis after injury. By SC-specific *Scx* gene inactivation (ScxcKO), we show that *Scx* function is required for SC expansion/renewal and robust new myofiber formation after injury. We combined single-cell RNA-sequencing and CUT&RUN to identify direct Scx target genes during muscle regeneration. These target genes help explain the muscle regeneration defects of ScxcKO, and are not overlapping with *Scx*-target genes identified in tendon development. Together with a recent finding of a subpopulation of *Scx*-expressing connective tissue fibroblasts with myogenic potential during early embryogenesis, we propose that regenerative and developmental myogenesis co-opt the *Scx* gene via different mechanisms.

Introduction

Regeneration of adult skeletal muscles following injury is initiated by the activation and proliferation of satellite cells (SCs). After extensive proliferation, progenitors undergo differentiation and fusion with each other or existing myofibers to recreate functional muscle tissue[1-3]. The intrinsic and extrinsic factors regulating myogenesis have been extensively investigated. The key transcription factors governing this process are largely the same as those deployed during embryogenesis, including paired-homeodomain proteins Pax3 and Pax7, basic helix-loop-helix (bHLH) myogenic regulatory factors (MRFs) such as Myf5 and MyoD1, and the myocyte enhancer factor 2 (MEF2) family; however, their relative contribution or redundancy vary between the two processes[4]. To date, resident Pax7$^+$ SCs are recognized as the major source of muscle stem cells in adult limb muscles. None of these myogenic transcription factors are known to participate in tendon development or regeneration.

Both muscle and tendon progenitors reside in the somite during embryonic development but are located in different compartments. *Pax3* and *Pax7* are expressed in the dermomyotome, which gives rise to the myotome expressing *Myf5* and/or *MyoD1*. The syndetome, on the other hand, is defined by the expression of the earliest tenogenic progenitor marker *Scx*, and gives rise to tendon and ligament[5]. Like Myf5 and MyoD1, Scx is a bHLH transcription factor, and they all bind to a DNA sequence motif called the E-box[6]. *Scx* expression persists in mature tenocytes, ligaments, and connective tissue fibroblasts (CT)[7]. *Scx* mutant mice have poorly developed tendons with drastically reduced expression of tendon matrix genes[7-9]. In adult tendon regeneration, the Tppp3$^+$Pdgfra$^+$ tendon stem cell population turn on *Scx* for tendon regeneration[10]. Lastly, *Scx* function is required in post-natal tendon growth and regeneration[11-14].

Intriguingly, lineage tracing using a constitutive *Scx*^{Cre} in mouse embryos found descendant cells in cartilage, tendon, ligament, muscle, and muscle interstitial CT[8,15,16], suggesting that *Scx* is expressed either in several distinct musculoskeletal subpopulations, or in a common progenitor that gives rise to different fates. Ablation of embryonic Scx$^+$ cells cause a change in muscle bundling[16], presumably due to the loss of instructive cues from the tendon (or CT) to form proper muscle pattern[17]. In adult muscles, Hic1$^+$ quiescent mesenchymal progenitors (MPs) give rise to Scx$^+$ cells in the muscle interstitial compartment, and ablation of Hic1$^+$ cells negatively impacts muscle regeneration[18]. Muscle interstitial Scx$^+$ cells engrafted into the muscle contribute only to extracellular matrix remodeling[19]. A survey of muscle interstitial CT assigned a sub-population of cells expressing tendon markers (including *Scx*) as paramysial cells - cells lining next to the perimysium that wraps around muscle fascicles[20]. Furthermore, Strenzke and colleagues showed that secretome from Scx overexpressed cells could significantly increase myoblast fusion and metabolic activity *in vitro*[21]. Collectively, these data indicate that while some embryonic Scx$^+$ cells can incorporate into myofibers, adult Scx$^+$ cells contribute to skeletal muscle architecture and repair/regenerative process in a paracrine manner.

Serendipitously, in the ScxGFP transgenic mouse Tg-ScxGFP[22], we observed GFP fluorescence in SCs and regenerating myofibers after injury. We conducted a series of experiments to show that endogenous *Scx* is expressed in activated SC after injury. We show that *Scx* is functionally relevant in muscle regeneration by inactivating *Scx* in *Pax7*$^+$ SCs (ScxcKO). We employed single-cell RNA-sequencing (scRNA-seq) and CUT&RUN to define Scx's target genes during muscle differentiation and fusion. Down-regulation of Scx's target genes such as *Mef2a*, *Cflar*, *Capn2*, and *Myh9,* explains the regenerative defects of ScxcKO mice. In contrast to adult Scx$^+$ muscle

CT and embryonic muscle-forming Scx$^+$ cells, our findings reveal a previously unappreciated role of *Scx* in adult Pax7$^+$ SCs.

Results

ScxGFP transgene is expressed in the regenerative myogenic lineage

When we analyzed tibialis anterior (TA) muscles of the Tg-ScxGFP (ScxGFP) mice, scattered GFP$^+$ cells were found in the interstitial space, but not in quiescent Pax7$^+$ SCs nor in myofibers (Fig. S1A). Unexpectedly, we found GFP signal in injured muscles. In cardiotoxin (CTX) injured TA muscles of ScxGFP mice at 5 days post-injury (dpi) (Fig.1A), we found GFP colocalized with Pax7$^+$ SCs (Fig.1B). GFP also overlaps with committed myogenic progenitor marker MyoD1, myocyte marker Myogenin (Myog), and myosin heavy chain (MyHC) in terminally differentiated myofibers (Fig. 1C; Fig. S1B-D for split channels). When we analyzed muscles administered with 5-ethynyl-2′-deoxyuridine (EdU) for 5 days after CTX injury (Fig. 1A), GFP was found to colocalize with proliferated (EdU$^+$) Pax7$^+$ cells (Fig. 1D). Thus, ScxGFP is expressed, albeit at varying levels, in the myogenic lineage during regenerative process.

We next determined ScxGFP expression in cultured SCs. For this, we employed a four-surface marker fluorescent activated cell sorting (FACS) scheme (Sca1$^-$CD31$^-$CD45$^-$Vcam1$^+$)[23] to purify SCs from hindlimb muscles of ScxGFP mice (Fig. 1E; S1E, F); ~ 98% of isolated cells were Pax7$^+$ (Fig. S1G, H). While Pax7 was detected in the SC immediately after FACS isolation, neither MyoD nor GFP was detected (Fig. 1F). After two days in culture, most cells were Pax7, MyoD and GFP triple positive (Fig. 1G). After switched to differentiation media for 2 days, GFP signal persisted in MyHC$^+$ myotubes (Fig. 1H). We therefore conclude that ScxGFP

expression is initiated after SC becomes activated and continues into differentiated myofibers *in vivo* and *in vitro*.

Endogenous *Scx* is expressed in activated SCs

To ensure that the ScxGFP expression observed in adult regenerative myogenesis is not caused by mis-expression due to transgene insertion site, we utilized $Scx^{CreERT2}$ for tamoxifen (TMX) inducible lineage tracing with a tdTomato (tdT) reporter (R^{tdT})[24]. Two experimental groups with different tamoxifen (TMX) and injury regimen were designed (Fig. 2A): 1) TMX-induced marking before injury, and 2) TMX-induced marking after injury; muscles were harvested at 14 dpi for analysis. Mice treated with TMX before injury showed little to no tdT^+Pax7^+ SCs or tdT^+ regenerative muscle fibers (identified by centrally located nuclei). By contrast, mice treated with TMX after injury showed ~ 30% of $Pax7^+$ SCs and all regenerated myofibers as tdT^+ at 14 dpi (Fig. 2C-D). $Pax7^+$ SC densities were not different between these two groups (Fig. S2A). These data extend the ScxGFP results in that 1) marked interstitial Scx^+ cells prior to injury do not possess myogenic potential, 2) endogenous *Scx* is expressed in activated SCs for regenerative myogenic lineage-marking, and 3) lineage-marked Scx^+ SCs are capable of renewal as $Pax7^+$ SCs at 14 dpi. Examination of two published SC bulk RNA-seq data confirmed *Scx* expression in SCs isolated from wild type[25] and *mdx* mice[26] (Fig. S2B). Re-analysis of a published scRNA-seq data set of SCs[27] also revealed a wide-spread *Scx* expression at 2.5 dpi (Fig. 2E; and more below). We conclude that endogenous *Scx*, like ScxGFP, is expressed by activated SCs and regenerative myogenic cells in adult mice.

***Scx* is required for adult skeletal muscle regeneration**

To determine whether *Scx* plays a direct role in the myogenic lineage during regeneration, we combined floxed *Scx* (Scx^F)[7] and $Pax7^{Cre\text{-}ERT2}$ ($Pax7^{CE}$)[28] to generate ScxcKO mice for TMX-inducible gene inactivation (Fig. S3A); loxP sites flank the first exon of *Scx* (Fig. S3B). Either tdT or YFP (R^{YFP}) reporter (specified in figures and legends) was included for cell marking. Highly efficient and selective removal of exon1 was determined using genomic DNA samples of FACS-isolated control and ScxcKO SCs (Fig. S3A–D).

ScxcKO mice have muscle regeneration defects

Next, we injured control and ScxcKO mice with CTX, and compared their regeneration at 5 and 14 dpi (Fig. 3A); samples shown in Fig. 3 carried the tdT reporter. At 5 dpi, ScxcKO's regenerating myofibers were significantly smaller than those in control mice (Fig. 3B, C). Similar results were obtained in mice carrying the YFP reporter (Fig. S3 J, K). At 14 dpi, regenerated myofibers in ScxcKO mice were still considerably smaller than those of the control (Fig. 3D, E). Thus, *Scx* function is needed in the $Pax7^+$ SC-lineage for robust regeneration of muscle fibers.

ScxcKO mice show reduced SC proliferation and renewal

Considering that *Scx* expression is initiated in activated SCs but not in quiescent SCs, and ScxcKO mice have smaller regenerative myofibers, it stands to reason that *Scx* plays a role in their proliferation. At 5 dpi, we noted a ~ 7-fold reduction in $Pax7^+$ SCs in the ScxcKO samples (Fig. 3F, Fig. S3E, F). To show proliferation defect, we administered EdU and assessed cumulative proliferation index over the first 5 days of injury (Fig. 3A, top panel). Compared to the control, the fraction of $Pax7^+$ ScxcKO SCs that incorporated EdU (i.e., EdU^+Pax7^+) was reduced by ~ 4-fold (Fig. 3G and 3H).

We did not observe appreciable levels of programmed cell death (PCD) in control and ScxcKO at 5 dpi using an anti-active Caspase3 antibody. We also quantified Pax7$^+$ SCs number at 14 dpi (Fig. 3A, bottom panel) and found a ~ 5-fold reduction of renewed SCs in the ScxcKO group (Fig. 3I, Fig. S3H, I). Thus, Scx is autonomously required for SC proliferation and renewal following injury.

Basement membrane-myofiber interaction in control and ScxcKO mice

When we examined laminin (i.e., basement membrane) and MyHC in control 5 dpi samples, we found that the laminin boundary juxtaposed the regenerative myofiber surface (Fig. S3F). By contrast, many small ScxcKO MyHC$^+$ fibers did not fill out to the laminin boundary and the laminin-circled areas were smaller than those in the control (Fig. S3G). We suggest that the laminin pattern in the ScxcKO represent left-over basement membranes of dead myofibers (caused by injury), i.e., the ghost fiber[29]. The collapsed shrunken appearance is likely due to the small new fibers not able to fill up the space. Ghost fibers are thought to be replaced by basement membranes produced by regenerated fibers over time. At 14 dpi, regenerated myofibers in both control and ScxcKO were tightly surrounded by laminin despite their difference in size (Fig. S3H, I), suggesting that ScxcKO regenerated myofibers are capable of making their own new basement membranes.

Scx is needed for robust proliferation of SC in culture

To examine Scx function in the SC without interactions with other cell types in the injured/regenerative environment, we turned to *in vitro* assays using purified SCs (Fig. 4A). To simplify FACS isolation of SC, we utilized either the tdT or the YFP fluorescent reporter (Fig. S4A, B). SC purity was assessed by staining for Pax7 immediately after

FACS. We were surprised that YFP-marked SCs (YFP-SCs) exhibited higher purity than tdT-marked SCs (Fig. 4B). Murach and colleagues have reported exosomal transfer of tdT mRNA from lineage-marked Pax7^{+} cells to other cell types[30]. The difference we observed between these 2 reporters suggests that high levels of tdT mRNA produced by a strong promoter/enhancer leads to more tdT^{+} non-SCs by exosomal transfer.

As such, we opted to use YFP-marked control and ScxcKO SCs in subsequent studies for higher SC purity. Of note, ScxcKO with YFP reporter had similar regenerative defects as that with tdT reporter (Fig. S3J, K). By EdU incorporation assay, we found that cultured ScxcKO SCs displayed reduced proliferation indices at days 2, 3 and 4. Curiously, the number of ScxcKO cells per imaged area (i.e., cell density) barely increased during this time course, despite EdU incorporation (Fig. 4C, D, E). We next carried out live imaging (Fig. S4C) to document the behavior of control and ScxcKO SCs. Consistent with EdU incorporation, control SCs showed a faster increase in cell number/density than ScxcKO SCs (Fig. S4D, E). ScxcKO cells also showed a slightly reduced cell mobility (Fig. S4F). As found *in vivo*, we did not detect a difference of PCD (by anti-cleaved Caspase3) between control and ScxcKO SCs (Fig. 4F). However, we did observe cell loss during live-imaging (Fig. S4G, H): More ScxcKO cells rounded up or appeared necrotic before disappearing. Our results support that Scx acts autonomously in the SC to promote proliferation and migration.

Scx expression by Single-Cell RNA-Sequencing (scRNA-Seq)

To determine the mechanism underlying Scx's role in the SC-lineage, we employed scRNA-seq using the 10x Chromium platform (Fig. S5A). For this, multiple sites of $BaCl_2$ injection were made to TA and gastrocnemius muscles of control and ScxcKO

mice to induce wide-spread injury and activate as many SCs as possible[31]. Because the published sc-RNA-seq data[27] indicated a widespread *Scx* expression at 2.5 dpi (Fig. 2F), we chose this time point for investigation.

YFP-marked control and ScxcKO SCs at 2.5 dpi were FACS-isolated and immediately subjected to scRNA-seq (Fig. 5A). Data were analyzed using the R package Seurat and unsupervised graph-based clustering[32]. After filtering, 11,388 control and 12,844 ScxcKO cells, respectively (with ~23,000 detectable genes), were qualified for analysis. We utilized uniform manifold approximation and projection (UMAP) to display all cells in the unified dataset and performed unsupervised shared nearest neighbor (SNN) clustering to partition cells into 18 (0-17) clusters (Fig. S5B). We annotated the cell types by examining the normalized expression level and frequency of canonical cell type-specific genes. The percentages of cells within each cluster in control and ScxcKO were also calculated (Fig. S5C). Clusters 0-2 and 4-11 contained the majority of cells expressing myogenic genes. Clusters 3 and 12-17 represent non-myogenic cell types, including various immune cells, endothelial cells, and Schwann cells (presumably due to exosomal transfer of YFP mRNA). Cluster 13 was assigned as monocytes/macrophages/platelets, but expressed myogenic genes. They are likely the immunomyoblast proposed by Oprescu et al[33]. Below, we focused on myogenic clusters to investigate the defects associated with ScxcKO.

sc-RNA-seq confirms *Scx* expression during regenerative myogenesis

Of the 11 myogenic cell clusters, we classified them into four categories: early activated SC, activated SC, myocyte, and mature skeletal muscle (Fig. S5B). Within the categories of early activated SC and activated SC, multiple cell clusters were included and numbered as different states. Here, numbers were arbitrarily assigned

and not meant to reflect their temporal sequence. Early activated SC 1 - 3 were represented by clusters 4, 5, 9, and expressed varying levels of *Pax7*, *Myod1*, and *Myf5* (Fig. S5B, D, E). We assigned clusters 0, 1, 2 and 7 as activated SC 1 – 4 respectively as they expressed lower levels of *Pax7* (compared to early activated SC). Further evidencing our assignment as activated SCs, more cells in these clusters expressed *Myod1*, *Myf5*, and *Hspa1a*[34,35]. Cluster 6 represented early myocyte based on increased expression of *Myog* and *Mef2a*. Cluster 8 cells expressed high levels of *Mymk*, indicating that they are competent for fusion. Cluster 11 cells expressed *Myh1* and *Acta1*, representing mature muscle cell. Cluster 10 cells were unknown myogenic cells for they expressed very low levels of myogenic genes. Among these clusters, the level and cell percentage of *Scx* expression were very low in early activated SCs, and gradually increased from activated SCs to early myocytes, fusion competent myocytes, and mature muscle cells (Fig. S5D, F).

We carried out Monocle 2 trajectory analysis to depict the progression of myogenic cell clusters (Fig. 5B). Given that *Scx* expression is very low in the early activated SC category and we observed ScxGFP only in activated SC experimentally, we excluded early activated SC 1-3 from analysis. The trajectory revealed a time line consistent with our assignment, from activated SC 1 to mature muscle cells. Of the 4 activated SC clusters, activated SC 1 cells were distributed throughout the activation time line up to early myocyte stage, activated SC 2 and SC 3 cells were preferentially located in earlier time line, whereas activated SC 4 cells were found in a later time, revealing their different states. Early myocytes, fusion competent myocytes, and mature muscle cells were ordered as expected.

scRNA-seq data help identify myogenic differentiation and fusion defects

To understand the timing of *Scx* action, we compared the relative densities of various cell types/states between the control and ScxcKO cells along the pseudotime (Fig. 5C). Relative to control, a higher density of ScxcKO cells, i.e., peak 2 in Fig. 5C, was noted just before their reduction, i.e., peaks 3 and 4. Peaks 3 and 4 correspond to fusion competent myocytes and mature muscle cells, respectively. This information re-directed us to investigate *Scx* function in fusion and differentiation. For this, control and ScxcKO SCs were isolated, cultured, plated at the same density, and then switched to differentiation medium (DM) (Fig. 5D). They were assessed for expression of MyoG (for differentiation index) and MyHC (for fusion index) daily over 3 days. More control cells expressed MyoG and MyHC when compared to ScxcKO cells at each time point (Fig. 5E, F, G). At day 3, ScxcKO cells caught up in differentiation index (still lower than that of control cells), but were still considerably lower in fusion index. This experimental result, aided by pseudotime analysis, supports a role of *Scx* for regenerative myogenic differentiation.

Molecular pathways governed by *Scx* in regenerative myogenesis

To gain molecular insight, we examined differentially expressed genes (DEGs) between control and ScxcKO cells along the pseudotime-line. We were particularly intrigued by the DEGs in peak 2, as it represents an early time point of difference to capture candidate direct targets of Scx. There were 3956 DEGs in peak 2 – *Scx* exhibited the largest $\log_2$ fold change in ScxcKO (Fig. 5H and Table S1). In particular, cyclin-dependent kinases *Cdk1* and *Cdk2* were down-regulated, and CDK-inhibitors *Cdkn1a* and *Cdkn1c* were up-regulated in ScxcKO cells at, and prior to, peak 2 (Fig. S5G). This helps explain the proliferation defects of ScxcKO cells. However, higher cell density with less proliferation potential is somewhat counter-intuitive. We suggest

that ScxcKO cells not only proliferate slower but also progress slower towards differentiation and fusion, resulting in their stalling and accumulation at the peak 2 transitional juncture (Fig. 5C).

Consistent with the phenotype of ScxcKO, Gene Ontology (GO) term analysis of peak 2 DEGs revealed that control cells showed enrichment of up-regulated genes in the categories of muscle differentiation, growth, and development, among other pathways overlapping with cardiac muscles (Fig. 5I). 17 genes involved in the muscle cell apoptotic process were found, even though we did not detect changes by using anti-cleaved Caspase 3. Unexpectedly, *Mymk* and *Mymx*, indispensable for myocyte fusion, were expressed higher in ScxcKO than control cells (Tables S1, S2), possibly reflecting their compensatory up-regulation due to compromised differentiation/fusion of *Scx*cKO cells. Analyses of peak 3 and 4 DEGs provide additional information about selective differentiation processes being disrupted in ScxcKO cells (see discussion).

Identification of direct targets by CUT&RUN assay

To uncover direct gene targets of Scx that regulate muscle differentiation and/or maturation, we utilized the CUT&RUN[36] assay to determine Scx bindings sites in the genome. To aid this endeavor, a triple-Ty1 tag (3XTy1) was fused to the C-terminus of Scx to create a Scx^{Ty1} allele (Fig. S6A). $Scx^{Ty1/Ty1}$ mice are viable and fertile without apparent tendon abnormality. Ty1 was detected in linearly arrayed patellar tenocytes (Fig. S6B) and in cultured myoblasts (Fig. S6C, D) derived from $Scx^{Ty1/Ty1}$ mice. During differentiation time course over 3 days in culture, the largest fraction of cells with detectable Ty1 presented at day 1 (Fig. S6E).

We performed the CUT&RUN using anti-Ty1 on $Scx^{Ty1/Ty1}$ myoblasts (Scx-CUT&RUN) at 12 h after switching them to DM (Fig. S6F); this time point was chosen

to uncover early targets. We included two controls: ScxGFP myoblasts with anti-Ty1, and *Scx*$^{Ty1/Ty1}$ myoblasts with non-specific IgG. A total of 1003 binding peaks were identified in 862 gene loci with 33.4%, 38.88% and 22.44% located in intergenic regions, introns, and promoters respectively, alongside other genomic regions, respectively (Fig. 6B; Table S3). These peaks were enriched for the bHLH transcription factor binding motif, the E-box: CANNTG (Fig. 6C), indicating high data quality. By integrating the Scx-CUT&RUN data with DEGs in the scRNA-seq data of ScxcKO cells (specifically those in peak 2 of Fig. 5C), we found 207 intersecting genes (Fig. 6D; Table S4). Scx-binding peaks at these gene loci were also enriched for the E-box motif (Fig. 6E), implicating these genes as direct targets. As expected, GO-term of these genes showed enrichment for processes in muscle differentiation, fusion, and myofibril assembly (Fig. 6F; Table S5). We also noted the enrichment for processes of mRNA destabilization, catabolism, poly(A) shortening, etc., suggesting that mRNA metabolism is altered in the ScxcKO (Fig. 6F; Table S5).

Four of the 207 candidate direct target genes provide explanation for the ScxcKO phenotype: *Mef2a*, *Capn2*, *Myh9*, and *Clfar.* The Scx-CUT&RUN peaks at these loci were in either the promoter region or intron (Fig. 6G; Fig. S6F). Knocking-out and knocking-down *Mef2a* led to compromised myoblast differentiation *in vivo* and *in vitro*, respectively [2,37-39]. Reduced *Mef2a* levels explain the compromised myoblast differentiation of ScxcKO cells. Consistently, several *Mef2a* target genes, such as *Hspb7*, *Atp1a2*, *Tmem182*, *Myoz2*[38] were also down regulated (Table S1). Capn2 is a calpain isoform expressed in the skeletal muscle, and the locus harbors 5 E-boxes and 1 MEF-2 binding sites[40]. Knocking-down *Capn2* in C2C12 cells led to compromised cell migration and fusion[41], as observed for ScxcKO cells. *Myh9* was shown to regulate bi-polar cell morphology and alignment during myocyte fusion *in*

vitro[42]. Its down regulation is consistent with the defective fusion of ScxcKO cells. Lastly, *Cflar* were shown to proliferation and prevent apoptosis in vascular smooth muscle cells and T lymphocytes[43-46]. It may act similarly in the SC to explain reduced proliferation and increased cell loss of ScxcKO cells. These 4 genes displayed reduced expression levels at the early part of the pseudotime trajectory (Fig. 6H-K), consistent with them being direct targets. The other 203 genes likely also contribute to aspects of the ScxcKO phenotype in ways yet to be determined. Taken together, Scx directly regulates a set of E-box containing genes, and several of these genes have direct implications to the phenotype observed.

Discussion

Here we show that *Scx* is expressed in activated mouse SCs, and it regulates many aspects of muscle regenerative process, from proliferation, cell survival, migration, to differentiation and fusion. The Scx target genes we identified underscore its function in muscle regeneration.

The multiplicity of Scx-lineage

Since the initial description of the *Scx* gene[6], most efforts have been focused on its role in tendon. Its early expression in the syndetome and the limb mesenchyme eventually becomes realized in tendons, ligaments, and CT[5,22,47,48]. Lineage tracing by Scx^{Cre} confirmed aforementioned descendant cell types alongside other cell types[15,16]. Of relevance, a lineage contribution to myofibers was found. The temporal emergence of Scx^{+} cells with myogenic potential was not provided by constitutive Cre-mediated lineage tracing. On the other hand, TMX-inducible lineage tracing mediated by the CT marker gene *Ors1* (i.e., using an $Ors1^{CreERT2}$) revealed myogenic

incorporation competence that declines towards late embryogenesis[15]. A Prx1$^+$ CT population has also been shown to incorporate into the myofiber near the myotendonous junction (MTJ) at neonatal stages[49]. Consistently, scRNA-seq of embryonic chick limb mesenchyme identified a cell cluster co-expressing CT and myogenic signatures at multiple stages[15]. Whether these bi-potential CT/myogenic cells arise from dermomotyotome, syndetome or a yet-to-be identified origin, remains to be rigorously examined.

We show here that adult SCs express ScxGFP upon injury and culture, and that ScxGFP is co-localized with Pax7, MyoD and MyHC. scRNA-seq data confirm endogenous *Scx* expression in multiple regenerative myogenic clusters/states, in which the other CT markers *Twist2*, *Ors1*, and *Pdgfra* are barely detectable. Moreover, only the lineage-marked Scx$^+$ cells induced after, but not prior to, injury contribute to regenerative muscles and SCs. Together, these results support that muscle interstitial Scx$^+$ CT (lineage-marked prior to injury) have no myogenic potential, whereas activated Pax7$^+$ SCs expressing *Scx* (lineage-marked after injury) can contribute to new muscles and SCs. This is consistent with transplanted Scx$^+$ CT[19] lacking a contribution to muscle. Adult muscle interstitial CT are highly heterogeneous within a muscle group as well as between muscle groups based on scRNA-seq data, and not all CT express *Scx*[20]. Anatomically, adult muscle interstitial *Scx*$^+$ cells are paramysial cells that line the perimysium[20]. Lineage tracing data showed that Scx$^+$ CT and MTJ cells were descendants of Hic1$^+$ MPs, but no myofiber incorporation from the Hic1$^+$ lineage was noted[18]. Whether CT/myogenic bipotential progenitors exist in adult muscle is of considerable interest. Regardless, our results strongly support that SCs express *Scx* after activation and require *Scx* function for efficient regeneration.

Scx function in tendon versus muscle

Scx has been consider a master regulator of tendon (and ligament) development as *Scx* mutant mice develop severely compromised tendons in the limbs and tail[7-9]. *Scx* is required for the expression of multiple tendon matrix protein encoding genes such as *Col1a1*, *Col3a1*, and *Tnmd* [9], but not for tendon progenitor specification. Ablation of embryonic Scx$^+$ cells led to mis-patterned muscle bundles[16], supporting an interdependence between muscle and tendon for connectivity[17]. Retrospectively, the observed muscle mis-pattern by ablating Scx$^+$ cells likely included ablation of CT/myogenic cells and tendon cells. In adult, *Scx* continues to be required for tendon growth and repair after injury[11-14]. By contrast, we focused on *Scx* function in proliferation, migration, differentiation, and fusion within the Pax7$^+$ SC lineage for muscle regeneration. GO-term analyses and literature reviews of DEGs from our scRNA-seq and CUT&RUN data sets have also identified genes in many myogenic processes, instead of genes in tenogenic or CT processes, that help understand the role of *Scx* for muscle regeneration.

Scx downstream target genes in tendon versus muscle

Scx binds to E-box as MyoD1 and Myf5. The initial characterization of Scx showed that it could activate reporter genes driven by an E-box derived from the enhancer of the muscle creatine kinase (*MCK*) gene[6]. Eventually, the tendon-specific *Tnmd* gene, with multiple functional E-boxes in its promoter, was shown to be a direct target of Scx[9,50]. Recently, bulk-RNA-seq and ChIP-seq were combined to define *Scx* target genes in embryonic tenocytes[51]. Although their and our data sets are not age-matched and obtained by different methods, we compared them nonetheless. Overall DEGs (including those without Scx-binding sites) between our and their data yielded minimal

overlap (0.9%, using the criteria of log2FC > 0.5). Two genes, *Htra3* and *Olfml2b*, are overlapping DEGs (48 genes for tenocytes and 207 genes for myoblasts) with Scx-binding sites, and neither gene has been studied in tendon or skeletal muscle. Lastly, the compiled E-box sequences bound by Scx in tenocytes and myoblasts are not different. Thus, the deployment of *Scx* by adult SCs is not a re-use of its function in the tendon. The distinctiveness of Scx target genes between these two tissues is most likely attributed to chromatin accessibility imposed by different epigenomes.

Indirect target genes of Scx further explain defects of the ScxcKO.

Although we emphasized Scx's direct target genes in peak 2 of Fig. 5C, there were many more DEGs that were indirect targets (i.e. without significant Scx-CUT&RUN peaks). Dysregulation of those genes also provide insights to *Scx*-regulated muscle regeneration. For the proliferation defect, we mentioned 4 dysregulated cell cycle regulators in the results section. In addition, *Erk1/2/3* (*Mapk1/3/6*), known for their role in cell growth, also exhibited lower expression levels in ScxCKO cells during the early pseudotime phase (Tables S1, S6). For GO-enrichment in cell migration (20 genes; Table S2), *Itga2* and *Crk* are worth noting as they have been shown to play this role in non-muscle cell contexts[52,53,54,55]. They may mediate myogenic cell migration under the umbrella program of *Scx*. For differentiation and fusion at the later pseudotime phase, there are 1942 and 755 DEGs (Tables S6, S7) in peaks 3 and 4 (Fig. 5C), respectively. GO-term analyses identified 74 (in peak 3) and 47 (in peak 4) genes related to muscle differentiation (Tables S8, S9). Several of them have documented roles in myogenic differentiation, e.g., *Hacd1*[56,57], *Klhl41*[58,59], *Ehd2*[60,61], and *Lmna*[62,63]. As these gene products act in different cellular compartments and mediate distinct processes, *Scx* does not appear to govern a singular process for muscle differentiation.

How these indirect genes come to be dysregulated in the absence of *Scx* remains to be deciphered.

Together with the embryonic CT/myogenic bipotential cells and the Prx1^{+} CT capable of myogenic fusion near the MTJ, our results add an additional layer of complexity and further blur the molecular and cellular boundaries that divide muscle versus tendon/CT identity. The wealth of information of heterogeneous cell types and states obtained by scRNA-seq will continue to break many long-accepted concepts of tissue-restricted functions of transcription factors.

Methods

Mouse strains

Pax7$^{CE/+}$ (*Pax7*$^{Cre\text{-}ERT2}$)[28], *R*^{YFP} (Gt(ROSA)26$^{\text{Sortm19EYFP)Cos/J}}$)[64], *R*^{tdT} (Gt(ROSA)26$^{\text{Sortm14(CAG-tdtomato)Hze/J}}$)[24], *Scx*^{F} (*Scx*^{tm1Stzr})[7], *Scx*^{CreERT2} (*Scx*$^{tm2(cre/ERT2)Stzr}$)[11] and Tg-ScxGFP[22] alleles were obtained from either original investigators or the Jackson Laboratory (JAX). *Scx*^{Ty1} allele was made and characterized by our group, with 3 Ty1 tags (EVHTNQDPLD) inserted upstream of the TGA codon of the *Scx* gene. All animals had mixed background. Genotypes of animals are stipulated in text, figures and legends. For qPCR to determine ScxcKO efficiency, primers are in Table S10 (referenced in Fig. S3 legend). Both sexes were used in all experiments and grouped together, except that only males were used for scRNA–seq. All mice were used between 2-4 month of age. All animal treatment and experiments were approved by the Institutional Animal Care and Use Committee (IACUC) of the Carnegie Institution of Washington (Permit number A3861-01).

TMX and EdU administration

Tamoxifen (TMX; Sigma) was prepared as 20 mg ml^{-1} stock in corn oil (Sigma) and administered by intraperitoneal injection at the mice at 4mg per 40g body weight following regimens in text, figures and legends. For daily *in vivo* proliferation tracing, 5-ethynyl-2′-deoxyuridine (EdU, 0.5 mg ml^{-1} in PBS; Thermo Fisher Scientific) was administered by intraperitoneal injection at 0.1 mg per 20 g body weight per injection. Muscle samples were collected as specified in figures and legends.

Muscle injury

For CTX injury, control and experimental mice were anaesthetized by isoflurane/oxygen vapor, tibialis anterior (TA) muscle was injected with 50 µl of 10 µM CTX (Cardiotoxin, Sigma-Aldrich) using an insulin syringe (U-100; BD); For $BaCl_2$ injury, control and experimental mice were anaesthetized with 2,2,2-tribromoethanol (Sigma) which was prepared as a 100% (w/v) stock solution in 2-methyl-2-butanol (Sigma), diluted 1:40 in PBS, This anesthetic was delivered through intraperitoneal injection at 10 µl per 1 g body weight. Muscle injury was administrated by injecting 2-4 ul per site of 1.2% (w/v) barium chloride (Fisher Chemical) into approximately 25 sites in the lower hindlimb muscles. Animals were then harvested at the post-injury time point stated in the text and figure legend.

Muscle sample preparation.

TA muscle samples were collected, fixed for 8 min in ice-cold 4% paraformaldehyde (PFA) (EM Grade, cat, 157-4) in PBS, sequentially incubated in 10, 20 and 30% sucrose/PBS overnight, embedded in OCT compound (Tissue-Tek, #4583), frozen in isopentane (Sigma)/liquid nitrogen and stored at −80 °C until cryosectioning. Cross-sections (10µM) of the mid-belly region of the muscle were stained with haematoxylin and eosin (H&E; Surgipath) or used for immunostaining and EdU reactions.

SC isolation by FACS and myoblast culture.

SCs were isolated according to the protocol described previously[23,65] with slight modifications. Briefly, mouse hindlimb muscles were dissected, minced and digested with collagenase II (1000U/ml, Worthington) in wash medium (10% Horse Serum in Ham's F-10 medium with 1% penicillin/streptomycin) for 1.5h followed by centrifugation and washing. Then, the tissue slurry was further digested by

collagenase II (100U/ml) and dispase (1.1U/ml, Gibco) in wash medium for 0.5h to get single cell suspension for cell sorting. The detailed protocol was described in (Liu et al., 2015). The cell suspension was sorted using a BD ARIA III sorter equipped with 375-nm, 488-nm, 561-nm, and 633-nm lasers. For Fluorescence sorting the YFP^+ (tdT^+) cells are sorted with green fluorescence; FITC channel 488-nm (red fluorescence; PE channel 568-nm). For 4 surface makers labelling, cells were incubated with 4,6-diamidino-2-phenylindole (DAPI) and fluorophore-conjugated antibodies (BioLegend) against CD31, CD45, stem cells antigen-1(Sca-1) and vascular cell adhesion protein 1 (Vcam-1) at 4 °C for 0.5 h. After washing, cells were subjected to FACS ($DAPI^-$, $CD31^-$, $CD45^-$, $Sca\text{-}1^-$ and $Vcam\text{-}1^+$ cells was collected), and data were collected by FACS Diva software v.6.1.3 (BD Biosciences). A small fraction of sorted cells was immunofluorescence staining for the muscle stem cell markers Pax7. For short-time cell culture, freshly sorted mononuclear SCs were plated on Matrigel (catalogue no. 354248; Corning) coated dishes (37 °C for 1 hr) and cultured in SCs culture medium (growth medium: 20% FBS, 5% horse serum, 1% penicillin/streptomycin, 1% GlutaMAX supplement (Gibco) and 2.5 ng/ml FGF (R&D systems) in DMEM (Gibco)) at 37°C in tissue culture incubators with 5% CO_2. Cells were harvested as specified in the text and figure legend; For long-time cell culture to get stable primary myoblast cell line, freshly-isolated satellite cells were cultured in Ham's F10 (F10, Sigma), 10% horse serum (HS, Invitrogen) and Pen-Strep (P/S) for 2 days, then passage the cell into culture medium and expended the cell for 2 more passages. After that the cells were Cryopreserved into liquid nitrogen for later CUT&RUN and myoblast differentiation and fusion assays; For *in vitro* differentiation and fusion assay, freshly sorted cell or frozen cells were thawed and cultured in growth medium for 12 h then changed into differentiation medium (2% horse serum, 1%

penicillin/streptomycin in DMEM(Gibco)) on Matrigel coated plates and harvested as specified in the figure legend; for EdU labelling, 10μM EdU was added to the SC culture medium for 6hr before harvesting for assay.

Live imaging

Freshly isolated SCs were cultured in growth medium on Matrigel-coated 48well dish at 5K cells per well, 3 wells per sample, and 5 locations per well. Images were collected every 10 min for 4 days. A short interval at the end of each day to was used to adjust the focus and add medium to get quality video and keep the cell in a good state. The videos were collected with a Nikon T2i system.

Immunofluorescence staining and detection

Muscle sections were hydrated with PBS, permeabilized with 0.5% Triton X-100 (Sigma-Aldrich)/PBS (0.5% PBT) for 15 min, washed with 0.05% PBT, and blocked with MOM block (Vector Lab) overnight. Sections were washed and incubated in blocking solution (1Xcarbo-free blocking solution (Vector Lab) and 10% goat serum in 0.05% PBT) for 2 h at room temperature, followed by incubation with primary antibodies diluted in blocking solution overnight at 4°C. Sources and dilution for primary antibodies are provided in Table S11, Sections were then washed with 0.05% PBT 3 times and incubated with appropriate Alexa Fluor-conjugated secondary antibodies (1:1,000 for Alexa 488 and Alexa 568 and 1:500 for Alexa 647; Thermo Fisher) in blocking buffer for 1 h at room temperature. Sections were then washed with 0.05% PBT, stained with DAPI (1μg ml−1 in 0.05% PBT) and mounted in anti-fade diamond solution (Invitrogen). tdT fluorescence was preserved, so no antibody staining was used. This protocol was also used for SCs and myoblasts with two

modifications, 1) the cells were fixed for 10 min in 4% PFA and 2) cells were only blocked with blocking buffer (10% goat serum in 0.05% PBT). For EdU detection, the Click-iT Reaction Kit (Thermo Fisher Scientific) was used before blocking according to the manufacturer's recommendations.

CUT&RUN

CUT&RUN experiments were carried out according to the CUTANATM CUT&RUN Protocol vison1.8 with modifications. Briefly, 500k cells were used for each sample and 0.01% digitonin (w/v) was used during the whole process. The antibodies used in the procedure were provided in Table S11. For the library preparation and sequencing, ThruPLEX DNA-seq Kit (Takara) was used to construct the CUT&RUN DNA library for sequencing on an Illumina platform. 5-10 ng purified CUT&RUN-enriched DNA was used for the library preparation. The whole process was performed according the protocol with deviations aiming to preserve short DNA fragments (30-80 bp). After the Library Synthesis step (adaptor ligation), 1.8 x volume of AMPure XP beads was added to the reaction to ensure high recovery efficiency of short fragments. 12 cycles of PCR amplification system was used, then the reaction was cleaned up with 1.2 x volume of AMPure XP beads. The libraries were assayed with a High Sensitivity DNA bioanalyzer (Agilent) for quality control and sequenced in the Illumina NextSeq500. To enable determination of fragment length, paired-end sequencing was performed (2x75 bp, 8 bp index). The data is analyzed by nf-core/CUT&RUN pipeline with –seacr stringent parameters (version 2.0)[66]. The overlapped peaks between replicates were considered as conserved peaks and used for downstream analysis. The motif analysis is carried out with SEA from MEME suit (v. 5.5.0) to identify the enrichment of bHLH family binding motifs in CUT and RUN targets.

Microscopy and image processing.

H&E staining images of TA muscle sections were captured by a Nikon 800 microscope with X20 Plan Apo objectives and with a Canon EOS T3 camera using EOS Utility image acquisition software v.2.10. Fluorescent images of TA muscle sections and cultured myoblasts were either captured by a Nikon Eclipse E800 microscope equipped with X20/0.50 Plan Fluor, X40/0.75 Plan Fluor and Hamamatsu C11440 digital camera using the Meta Morph Microscopy Automation and Image Analysis Software v.7.8.10.0, or captured by a Leica SP5 confocal microscope equipped with a X63/1.4 Plan Apo oil objective using the Leica Application Suite Advanced Fluorescence software version 2.7.3.9723. The same exposure time was used and the images were processed and scored in blinded fashion using ImageJ v.64 (National Institutes of Health (NIH)). If necessary, brightness and contrast were adjusted for an entire experimental image set. Cell number, fiber diameter, fiber number and fiber cross-sectional area were measured with ImageJ v.64.

Single cell RNA sequencing (scRNA-seq)

The lower hindlimb muscles (TA and gastrocnemius muscles) of the ScxcKO and control mice were injured with $BaCl_2$ and recovered for 2.5 day. Pax7 lineage cells were FACS-isolated by YFP fluorescence. Cells were suspended in PBS and counted by hemocytometer into 1000 cells μl^{-1}. Around 17,000 cells per sample were used for single-cell library preparation using the 10x Genomics platform with Chromium Next GEM Single Cell 3′ GEM, Library and Gel Bead Kit v.3.1 (PN-1000121, v.3 chemistry), Single Cell 3′ A Chip Kit (PN-1000009) or Chromium Next GEM Chip G Single Cell Kit (PN-1000127), and i7 Multiplex Kit (PN-120262). We followed the 10x protocol exactly to prepare the scRNA-seq library. In brief, for v.3 chemistry, 16.5µl cell suspension

and 26.7µl nuclease-free water were mixed with 31.8µl reverse transcription master mix. Of this 75µl mix, 70µl was loaded into the Chromium Next GEM Chip G. After barcoding, cDNA was purified and amplified with 11 PCR cycles. The amplified cDNA was further purified and subjected to fragmentation, end repair, A-tailing, adaptor ligation and 14 cycles of sample index PCR. Libraries were sequenced using Illumina NextSeq 500 for paired-end reads.

Analyses of scRNA-seq data

Sequencing reads were processed with the Cell Ranger version 6.0.1 (10X Genomics, Pleasanton, CA) using the mouse reference transcriptome mm10. From the gene expression matrix, the downstream analysis was carried out with R version 4.0.2 (2020-06-22). Quality control, filtering, data clustering and visualization, and the differential expression analysis were carried out using Seurat version 4.0.3 R package[32]. Cells with <1000 UMIs or mitochondrial reads >10% were removed from the analysis. In addition, we removed potential doublets by DoubletFinder (v. 2.0.3)[67]. After log-normalizing the data, the expression of each gene was scaled regressing out the number of UMI and the percentage of mitochondrial genes expressed in each cell. The two datasets were integrated with IntegrateData function from Seurat. We performed PCA on the gene expression matrix and used the first 20 principal components for clustering and visualization. Unsupervised shared nearest neighbor (SNN) clustering was performed with a resolution of 0.6 and visualization was done using uniform manifold approximation and projection (UMAP). The Scx expressed myogenic lineage clusters 0, 1, 2, 6, 7, 8, and 11 were subjected to trajectory analysis by monocle 2 (v. 2.16.0)[68]. To organize cells in pseudotime, we performed new dimension reduction and regressed out mitochondrial effects with reduceDimension

function and unsupervised cluster them into 5 clusters with clusterCells function. The differentially expressed genes were calculated by differentialGeneTest and the top 500 differentially expressed genes are used to order then used by Monocle for clustering and ordering cells using the DDRTree method and reverse graph embedding.

Data availability

Mouse single-cell RNA sequencing data and CUT&RUN data were uploaded to NCBI (PRJNA1050758). Select intermediate RDS objects are available at figshare (https://figshare.com/articles/dataset/Processed_R_objects_for_Scx_scRNA/--).

Quantification and statistical analysis

Statistical analyses were performed in R version 4.0, with tidyverse and ggplot2 packages. The statistical significance of results was determined by unpaired Student's t-test and Two-way ANOVA.

Acknowledgement

We thank Eugenia Dikovsky and mouse facility crew for animal housekeeping, Allison Pinder for technical assistance in scRNA seq, Mahmud Sidiqqi for assistance in microscopy, and L. Yue for assistance with FACS of SCs. This work is supported by the NIH (AR060042 and AR071976) and Carnegie fund to CMF.

Figure legends

Figure 1: Adult regenerative myongenic cells express the transgene ScxGFP

A. Experimental scheme for data in (B-D). Tg-ScxGFP (ScxGFP) mice were injured by CTX to the TA muscle, followed by daily EdU administration for 5 days (5d), and their TA muscles were harvested for analysis at 5 d post-injury (dpi).

B. Muscle samples obtained in (A) were sectioned and stained with Pax7 and GFP (for ScxGFP expression) antibody. Arrows indicate Pax7 and ScxGFP double positive cells; 97.77% $Pax7^+$ SCs were $ScxGFP^+$ (N = 4 mice; n = 1274 cells).

C. Muscle samples obtained in (A) were sectioned and stained in pairs of GFP/Pax7, GFP/MyoD1, GFP/Myog (N = 4 mice). Asterisks indicate cells double-positive for ScxGFP and each respective myogenic marker.

D. Muscle samples obtained in (A) were sectioned and stained for Pax7 and GFP, followed by EdU reaction (N = 4 mice). Arrows indicate Pax7, ScxGFP, and EdU triple-positive SCs.

E. Experimental scheme of SC isolation from Tg-ScxGFP hindlimb muscles using four surface markers ($CD31^-$, $CD45^-$, $Sca1^-$, $Vcam1^+$) by FACS. Isolated SCs were assayed immediately after isolation (D0; data in (F)), after culture in growth media for 2d (D2(GM); data in (G)), or after cultured for 4d in GM followed by 2d in differentiation media (DM) (D6(DM); data in (H)).

F, G. D0 (F) and D2 cultured (G) SCs obtained in (E) were stained with for GFP (i.e. ScxGFP), Pax7, and MyoD. At D0, no $Pax7^+$ cells were GFP^+ or $MyoD^+$. At D2, 95.3% of $Pax7^+$ cells were GFP^+, whereas 99.2% of $MyoD^+$ were GFP^+. (N = 3 mice; n = 1246 cells at D0; n = 1332 cells at D2)

H. D6(DM) cells obtained in (E) were stained for GFP (i.e. ScxGFP) and MyHC. 94.58% MHC^+ were GFP+. (N= 3 mice; n = 1539 nuclei in $MyHC^+$ domain examined)

Nuclei were stained with DAPI (blue); Scale bars = 20 μm.

Figure S1: Addition data for Figure 1

A. TA muscle from uninjured ScxGFP mice were sectioned and stained with Pax7 and GFP (N = 4 mice; n = 203 $Pax7^+$ cells examined and no $Pax7^+GFP^+$ cells found). The arrow indicates the $Pax7^+$ SC; asterisk, $ScxGFP^+$ intramuscular CT; dashed line, outline of a muscle fiber.

B-D. Split channel images of data in Figure 1C; asterisks indicate the same cells.

E, F. FACS strategy (E) and profiles (F) to support Fig. 1E. (F) FACS plot and population hierarchy of SC from 4 surface markers sorting in pseudocolor plots.

G. Freshly isolated SCs from ScxGFP mice were cyto-spun and stained with Pax7 and DAPI.

H. Percentage of Pax7$^+$ cells in freshly isolated ScxGFP SCs by FACS procedures in (E, F). (N = 4 mice; n =1805 cells).

Scale bar = 20 μm. Data are presented with mean ± s.d.

Figure 2: Endogenous *Scx* is expressed by activated but not quiescent SCs.

A. Experimental design for *Scx*^{CreERT2} mediated inducible lineage tracing with the *R*^{tdT} reporter. The two experimental groups are: 1) TMX administrated before injury for 5 days (TMX before injury) and 2) TMX administrated after injury for 5 days (TMX after injury). TA muscles in both groups were harvested at 14d post injury.

B. TA muscles from experiment groups in (A) were stained with Pax7 (green) and Laminin (white) and visualized with tdT (no staining). Open arrowheads indicate Pax7$^+$ SCs; arrows, Pax7$^+$tdT$^+$ SCs.

C. Percentages of Pax7$^+$tdT$^+$ SCs in Pax7$^+$ SCs examined, from data in B. (N = 4 mice per group; n = 252 (Before, TMX before injury) and 297 (After, TMX after injury) Pax7$^+$ cells)

D. Percentages of tdT$^+$ myofibers in regenerated muscle fibers (with centrally located nuclei), from data in B. (N = 4 mice; n = 1956 (Before, TMX before injury) and n = 2024 (After, TMX after injury) regenerated myofibers).

E. UMAP plot of *Scx* expression in SCs isolated from uninjured muscles (left, Freshly sorted) and from muscles at 2.5 dpi in a published dataset[27].

Nuclei were stained with DAPI; Scale bar = 20 μm.

Data are presented with mean ± s.d.; *P* value are indicated. D-E, Unpaired two-tailed Student's *t*-test were applied.

Figure S2: Additional information for Figure 2.

A. Averaged Pax7$^+$ SC number per field (0.06mm^2) from data in B. Mouse and cell numbers are the same as in Figure 2C.

B. Gene expression levels (in FPKM) of *Scx* and *Pax7* in SCs isolated from wild type (WT) and *mdx* mice using bulk-RNA-seq. Data are extracted from published data sets[25,26].

Data are presented with the mean ± s.d.; the *P* value is indicated. Unpaired two-tailed Student's *t*-test was applied.

Figure 3: Efficient muscle regeneration requires *Scx* function.

A. Experimental designs to compare phenotypes of control (Ctrl) and ScxcKO mice. The R^{tdT} reporter was included (tdT lineage; see Figure S3A for genotypes). TMX was administered before and after the CTX induced injury to maximize gene inactivation. TA muscles were harvested at 5d or 14d after injury.

B-C. (B) Ctrl and ScxcKO TA muscles at 5d post-injury were sectioned and stained with hematoxylin and eosin (H&E) at low (top) and high (bottom) magnifications. (C) Histogram of mean fiber cross-sectional area from data in (B). (N = 5 mice per group).

D-E. (D) Ctrl and ScxcKO TA muscles at 14d post-injury were sectioned and stained with H&E. (E) Histogram of mean fiber cross-sectional area from data in (D). (N = 5 mice per group)

F. Histogram of average $Pax7^+$ SCs number per imaged field ($0.08mm^2$) of TA muscle sections from 5 dpi Ctrl and ScxcKO mice (N = 5 mice per group; n = 1748 Ctrl and n = 442 ScxcKO $Pax7^+$ SCs).

G. After EdU administration, Ctrl and ScxcKO TA muscles at 5d post-injury were sectioned and stained for Pax7, followed by EdU reaction. Arrows indicate $Pax7^+EdU^+$cells, whereas asterisk, $Pax7^+EdU^-$ cells. Nuclei were stained with DAPI; Scale bar = 25um.

H. Percentages of EdU^+ cells within the $Pax7^+$ cell population of Ctrl and ScxcKO, from data in G (N = 5 mice per group; n = 858 Ctrl and n = 325 ScxcKO $Pax7^+$ SCs).

I. Averaged $Pax7^+$ SC number per image field ($0.08mm^2$) in Ctrl and ScxcKO TA muscle sections from 14d post-injury samples (N = 5 mice per group; n = 350 Ctrl and n = 186 ScxcKO $Pax7^+$ SCs).

Data are presented with the mean ± s.d.; *P* values are indicated. (C, E-F, H-I) Unpaired two-tailed Student's *t*-test were applied.

Figure S3: Addition data for Figure 3.

A. Detailed description for the genotypes used as Ctrl and ScxcKO mice in Fig. 3. For tdT lineage, R^{tdT} reporter was included. For YFP lineage, R^{YFP} reporter was included.

B. Depiction of *Scx* gene inactivation by TMX-induced recombination of loxP sites flanking the exon 1 of *Scx* using the $Pax7^{CE}$ allele. PCR primer sets P1, P2, and P3 were used to detect exon 1 (E1), intron 1 (I1) and E2 of the *Scx* gene, respectively. Primer sequences are in Table S10.

C. Experimental scheme to determine the recombination efficiency using FACS-isolated Ctrl and ScxcKO SCs. Freshly sorted SCs were plated down and cultured in growth medium for 3 days and harvested for genomic DNA extraction and qPCR for data in (D).

D. Relative levels of E1, I1, and E2 in control and ScxcKO myoblasts determined by qPCR, followed by $2^{-\Delta\Delta Ct}$ analysis.

E. Experimental scheme as Fig. 3A for 5 dpi data in F-G.

F, G. TA muscle sections (from E) were sectioned and stained with Pax7 and Laminin in (F), and MHC and laminin in (G). Arrows indicate $Pax7^+$ SCs in (F), whereas arrows indicate $Laminin^+MHC^-$ ghost fibers (G) (N = 5 mice in each group; n = 1748 control and n = 442 ScxcKO $Pax7^+$ SCs).

H. Experimental scheme as Fig. 3A for 14 dpi data in (I).

I. TA muscle section from (H) from were sectioned and stained with Pax7 and Laminin (N = 5 mice in each group; n = 1050 control and n = 558 ScxcKO $Pax7^+$ SCs examined).

J. Same experimental scheme as in (E), except that R^{YFP} reporter (YFP lineage), instead of R^{tdT} reporter, was included.

K. Histogram of the fiber CSA from 5 dpi TA muscle sections from (J). (N = 6 mice for control, and N = 5 mice for ScxcKO group)

Nuclei were stained with DAPI. Scale bar = 20 μm.

Data are presented with the mean ± s.d.; adjusted *P* values are shown. (D, K) Unpaired two-tailed Student's *t*-test.

Figure 4: ScxcKO SCs display a proliferation defect.

A. Experimental design to obtain YFP lineage marked $Pax7^+$ SCs for *in vitro* analyses in (B-F).

B. Box plot of percentages of FACS-isolated tdT and YFP marked cells expressing Pax7 by staining immediately after isolation (as D0); each dot represents one image data, 10 image per group, totally n $\geq$ 1000 cells for each group.

C. YFP lineage marked cells were cultured in GM and assayed at days 2, 3, and 4 (D2-D4) intervals. 10 μM EdU was added for 6 h prior to harvesting for EdU detection.

D. Box plot of percentages of EdU^+ cells from data in (C), N=2 mice, each dot represents one image data, 3 well per group, 8 images per well.

E. Box plot of ratios of total cell numbers from data in (C); normalized to the average control cell number at D2 as 1.

F. Box plot of ratios of program cell death assayed by staining for cleaved Caspase3; n $\geq$ 300 cells for each group.

Nuclei were stained with DAPI; Scale bar = 20 μm.

Data are presented with the mean ± s.d.; adjusted *P* values are shown. (B) Two-way ANOVA; (D-F) Unpaired two-tailed Student's *t*-test.

Figure S4: Additional data for Figure 4.

A, B. FACS profiles for isolating tdT (A) and YFP (B) linage marked SCs.

C. Similar experimental scheme as in Fig. 4 A for refence to live imaging data in (D-F).

D. Line plot of relative cell number per frame; normalized to cell number of the Ctrl group at the beginning as 1.

E. Box plot of relative dividing cell ratio. Relative dividing ratio is defined as the divided cell number of each day per group / the beginning cell number of that day. There were no dividing cells detected at D1, each dot represents one image data, 3 well per group.

F. Box plot of cell migration speed based on tracking cell displacement by pixel (0.33 μm / pixel), each dot represents a cell data.

G. Still frames from time lapse show examples of necrotic cells at D3 of the ScxcKO group. Frames 1 - 6 are sequential time lapse images (covering 500 mins). Two cells of focus are indicated by red and black arrows to track their appearance from a healthy state (earlier frames) to a necrotic state (later frames). Scale bar = 20 μm.

H, Box plot of necrotic cell ratio. Necrotic cell was manually identified by morphology in each frame on each day, and normalized to the beginning cell number of that day, each dot represents one image data.

Data are presented with mean ± s.d.; adjusted *P* values are shown. Unpaired two-tailed Student's *t*-test were applied to data collected in each day.

Figure 5: scRNA-seq helps identify the role of *Scx* in myogenic differentiation and fusion.

A. SC scRNA-seq scheme for YFP lineage marked SCs. YFP^+ cells were FACS-isolated from 2.5 dpi $BaCl_2$ injured TA and the gastrocnemius (GA) muscles.

B. Trajectory analysis of the 7 myogenic clusters (complete cell cluster analysis in Fig. S5) indicated to the right. Arrow indicates the direction of pseudotime trajectory.

C. Cell densities of Ctrl and ScxcKO cells along the trajectory in (B). Cell in Peaks 2-4 were used for the DEG analysis; The asterisk indicates Peak 2 as our main focus.

D. *In vitro* differentiation assay scheme. SC-derived myoblasts were cultured in GM for 12 h (D0), switched into DM, and harvested daily for analysis over 3 days (D1 - D3).

E. Myoblasts subjected the scheme in (D) were stained for MyoG (for differentiation index in F) and for MyHC (for fusion index in G). Nuclei were stained with DAPI; Scale bar = 20 μm.

F, G. Box plot of differentiation index (F) and fusion index (G) from data in (E). Each dot represents data from one image. Unpaired two-tailed Student's *t*-tests were applied and adjusted *P* values are shown. (N = 3 mice; 3 wells per group per time point; 10 images per well; in total, 3342 control and 2561 ScxcKO cells examined).

H. Volcano plot of relative gene expression (Log2 fold change) in Ctrl versus ScxKO cells in Peak 2 (in C).

I. GO term enrichment of muscle development related processes from DGEs in (H).

Figure S5: Additional analyses of data in Figure 5.

A. Detailed scRNA-seq procedures diagram as in Fig. 5A for reference to analyses in (B-G).

B. UMAP plot of combined scRNA-seq of Ctrl and ScxcKO cells. A total of 11,388 control cells and 12,844 ScxcKO cells were included for analysis, and 18 cell clusters were delineated (shown by different colors and with assigned cell types indicated to the right). Asterisks denote myogenic clusters subjected to further analyses in Figure 5B, C, H, I.

C. Cell density distribution of each cell cluster in Ctrl and ScxcKO samples. NS, not significant; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

D. Expression of select marker genes in each cell cluster. Dot size represents the percentage of expressed cells within each cluster, whereas color intensity represents relative expression level (keys at bottom).

E. *Pax7* expression levels and cell numbers in each cluster; each dot represents one cell.

F. *Scx* expression levels and cell numbers in each cluster.

G. Relative expression levels (CPM, counts per million UMI) of 4 select cell cycle genes along the pseudotime depicted in Figure 5C.

Figure 6: CUT&RUN and scRNA-seq identify direct targets of Scx.

A. Experimental scheme for CUT&RUN profiling of the Scx binding in the genome of $Scx^{Ty1/Ty1}$ and ScxGFP myoblasts. Primary myoblasts derived from SCs of $Scx^{Ty1/Ty1}$ (experimental group) and ScxGFP (control group) mice were used. They were cultured in GM for 12 h (D0), and switched to DM for 12 h for use. 500,000 (500K) cells per group were subjected to CUT&RUN using an anti-Ty1 antibody or an IgG control antibody, in duplicate.

B. Pie chart for distribution of Scx CUT&RUN peaks in various regions of the genome.

C. Motif enrichment analysis with SEA from MEME suite (v. 5.5.0) identified bHLH protein binding motif (i.e. E-box) in all Scx CUT&RUN binding peaks.

D. Venn diagram of intersecting genes (207 genes) between Scx CUT&RUN target genes (861) and DEGs (3749) in Peak 2 of Fig. 5C.

E. Motif enrichment analysis (as in C) of the 207 genes in (D) also showed enrichment of bHLH protein binding motifs, the E-box.

F. GO term analysis of the 207 genes in (D). GO terms with $P < 0.0001$ were plotted.

G. Genomic snapshots of Scx CUT&RUN peaks on 4 select genes related to muscle differentiation.

H. Expression levels (CPM, counts per million UMI) of the 4 select genes in (G) along the pseudotime trajectory (same trajectory as Fig. 5C).

Figure S6: Additional data to support Figure 6.

A. Diagram of the genomic structure of *Scx^{Ty1}* allele. Three (3X) Ty1 tags were inserted just before the TGA codon of the *Scx* gene.

B. Neonatal patellar tendons of Scx$^{Ty1/Ty1}$ and wild type (WT) mice were fixed, sectioned, and stained for Ty1.

C. *In vitro* myogenic differentiation scheme using Scx$^{Ty1/Ty1}$ SC-derived myoblasts. Myoblasts were cultured in GM for 12 h (D0), and switched to DM for 3 days. Cells in each day were stained for Ty1 (D0-3).

D. Scx$^{Ty1/Ty1}$ and WT myoblasts at D0 were stained for Ty1 as an example for data in (E).

E. Boxplot of percentages of cells with detectable Ty1 signal assayed at different time points listed in (C). Each dot represents one image data, 3 wells per time point, 10 images per well, and totally 2785 cells examined. N = 3 mice. Data are present as the mean ± s.d.; adjusted *P* values are shown. Unpaired two-tailed Student's *t*-test were applied, D0 sample as the refence group.

F, Genomic snapshots of Scx CUT&RUN signals across a 10 Kb region of each of the 4 select genes in Figure 6G. The green lines represent CUT&RUN signal from anti-Ty1 on ScxGFP myoblasts; the magenta and blue lines represent the IgG control and anti-Ty1 antibody CUT&RUN signals detected in Scx$^{Ty1/Ty1}$ primary myoblasts, respectively.

Nuclei were stained with DAPI; Scale bar = 20 μm.

References

1 Yin, H., Price, F. & Rudnicki, M. A. Satellite cells and the muscle stem cell niche. *Physiol Rev* **93**, 23-67, doi:10.1152/physrev.00043.2011 (2013).

2 Liu, N. *et al.* Requirement of MEF2A, C, and D for skeletal muscle regeneration. *Proc Natl Acad Sci U S A* **111**, 4109-4114, doi:10.1073/pnas.1401732111 (2014).

3 Fukada, S. I., Higashimoto, T. & Kaneshige, A. Differences in muscle satellite cell dynamics during muscle hypertrophy and regeneration. *Skelet Muscle* **12**, 17, doi:10.1186/s13395-022-00300-0 (2022).

4 Hernandez-Hernandez, J. M., Garcia-Gonzalez, E. G., Brun, C. E. & Rudnicki, M. A. The myogenic regulatory factors, determinants of muscle development, cell identity and regeneration. *Semin Cell Dev Biol* **72**, 10-18, doi:10.1016/j.semcdb.2017.11.010 (2017).

5 Brent, A. E., Schweitzer, R. & Tabin, C. J. A somitic compartment of tendon progenitors. *Cell* **113**, 235-248, doi:10.1016/s0092-8674(03)00268-x (2003).

6 Cserjesi, P. Scleraxis: a basic helix-loop-helix protein that prefigures skeletal formation during mouse embryogenesis. *Development* **121**, 1099-1110 (1995).

7 Murchison, N. D. *et al.* Regulation of tendon differentiation by scleraxis distinguishes force-transmitting tendons from muscle-anchoring tendons. *Development* **134**, 2697-2708, doi:10.1242/dev.001933 (2007).

8 Yoshimoto, Y. *et al.* Scleraxis is required for maturation of tissue domains for proper integration of the musculoskeletal system. *Sci Rep* **7**, 45010, doi:10.1038/srep45010 (2017).

9 Shukunami, C. *et al.* Scleraxis is a transcriptional activator that regulates the expression of Tenomodulin, a marker of mature tenocytes and ligamentocytes. *Sci Rep* **8**, 3155, doi:10.1038/s41598-018-21194-3 (2018).

10 Harvey, T., Flamenco, S. & Fan, C. M. A Tppp3(+)Pdgfra(+) tendon stem cell population contributes to regeneration and reveals a shared role for PDGF signalling in regeneration and fibrosis. *Nat Cell Biol* **21**, 1490-1503, doi:10.1038/s41556-019-0417-z (2019).

11 Howell, K. *et al.* Novel Model of Tendon Regeneration Reveals Distinct Cell Mechanisms Underlying Regenerative and Fibrotic Tendon Healing. *Sci Rep* **7**, 45238, doi:10.1038/srep45238 (2017).

12 Gumucio, J. P., Schonk, M. M., Kharaz, Y. A., Comerford, E. & Mendias, C. L. Scleraxis is required for the growth of adult tendons in response to mechanical loading. *JCI Insight* **5**, doi:10.1172/jci.insight.138295 (2020).

13 Korcari, A., Muscat, S., McGinn, E., Buckley, M. R. & Loiselle, A. E. Depletion of Scleraxis-lineage cells during tendon healing transiently impairs multi-scale restoration of tendon structure during early healing. *PLoS One* **17**, e0274227, doi:10.1371/journal.pone.0274227 (2022).

14 Sakabe, T. *et al.* Transcription factor scleraxis vitally contributes to progenitor lineage direction in wound healing of adult tendon in mice. *J Biol Chem* **293**, 5766-5780, doi:10.1074/jbc.RA118.001987 (2018).

15 Esteves de Lima, J. *et al.* Unexpected contribution of fibroblasts to muscle lineage as a mechanism for limb muscle patterning. *Nat Commun* **12**, 3851, doi:10.1038/s41467-021-24157-x (2021).

16 Ono, Y. *et al.* Scleraxis-lineage cells are required for correct muscle patterning. *Development* **150**, doi:10.1242/dev.201101 (2023).

17 Kardon, G. Muscle and tendon morphogenesis in the avian hind limb. *Development* **125**, 4019-4032, doi:10.1242/dev.125.20.4019 (1998).

18 Scott, R. W., Arostegui, M., Schweitzer, R., Rossi, F. M. V. & Underhill, T. M. Hic1 Defines Quiescent Mesenchymal Progenitor Subpopulations with Distinct Functions and Fates in Skeletal Muscle Regeneration. *Cell Stem Cell* **25**, 797-813 e799, doi:10.1016/j.stem.2019.11.004 (2019).
19 Giordani, L. *et al.* High-Dimensional Single-Cell Cartography Reveals Novel Skeletal Muscle-Resident Cell Populations. *Mol Cell* **74**, 609-621 e606, doi:10.1016/j.molcel.2019.02.026 (2019).
20 Muhl, L. *et al.* Single-cell analysis uncovers fibroblast heterogeneity and criteria for fibroblast and mural cell identification and discrimination. *Nat Commun* **11**, 3953, doi:10.1038/s41467-020-17740-1 (2020).
21 Strenzke, M. *et al.* Tenogenic Contribution to Skeletal Muscle Regeneration: The Secretome of Scleraxis Overexpressing Mesenchymal Stem Cells Enhances Myogenic Differentiation In Vitro. *Int J Mol Sci* **21**, doi:10.3390/ijms21061965 (2020).
22 Pryce, B. A., Brent, A. E., Murchison, N. D., Tabin, C. J. & Schweitzer, R. Generation of transgenic tendon reporters, ScxGFP and ScxAP, using regulatory elements of the scleraxis gene. *Dev Dyn* **236**, 1677-1682, doi:10.1002/dvdy.21179 (2007).
23 Liu, L., Cheung, T. H., Charville, G. W. & Rando, T. A. Isolation of skeletal muscle stem cells by fluorescence-activated cell sorting. *Nat Protoc* **10**, 1612-1624, doi:10.1038/nprot.2015.110 (2015).
24 Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci* **13**, 133-140, doi:10.1038/nn.2467 (2010).
25 Li, L. *et al.* Muscle stem cell renewal suppressed by Gas1 can be reversed by GDNF in mice. *Nat Metab* **1**, 985-995, doi:10.1038/s42255-019-0110-3 (2019).
26 Madaro, L. *et al.* Macrophages fine tune satellite cell fate in dystrophic skeletal muscle of mdx mice. *PLoS Genet* **15**, e1008408, doi:10.1371/journal.pgen.1008408 (2019).
27 Dell'Orso, S. *et al.* Single cell analysis of adult mouse skeletal muscle stem cells in homeostatic and regenerative conditions. *Development* **146**, doi:10.1242/dev.174177 (2019).
28 Lepper, C., Conway, S. J. & Fan, C. M. Adult satellite cells and embryonic muscle progenitors have distinct genetic requirements. *Nature* **460**, 627-631, doi:10.1038/nature08209 (2009).
29 Webster, M. T., Manor, U., Lippincott-Schwartz, J. & Fan, C. M. Intravital Imaging Reveals Ghost Fibers as Architectural Units Guiding Myogenic Progenitors during Regeneration. *Cell Stem Cell* **18**, 243-252, doi:10.1016/j.stem.2015.11.005 (2016).
30 Murach, K. A. *et al.* Early satellite cell communication creates a permissive environment for long-term muscle growth. *iScience* **24**, 102372, doi:10.1016/j.isci.2021.102372 (2021).
31 Morton, A. B. *et al.* Barium chloride injures myofibers through calcium-induced proteolysis with fragmentation of motor nerves and microvessels. *Skelet Muscle* **9**, 27, doi:10.1186/s13395-019-0213-2 (2019).
32 Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e3529, doi:10.1016/j.cell.2021.04.048 (2021).
33 Oprescu, S. N., Yue, F., Qiu, J., Brito, L. F. & Kuang, S. Temporal Dynamics and Heterogeneity of Cell Populations during Skeletal Muscle Regeneration. *iScience* **23**, 100993, doi:10.1016/j.isci.2020.100993 (2020).
34 Senf, S. M. Skeletal muscle heat shock protein 70: diverse functions and therapeutic potential for wasting disorders. *Front Physiol* **4**, 330, doi:10.3389/fphys.2013.00330 (2013).
35 Francetic T, L. Q. Skeletal myogenesis and Myf5 activation. (2011).

36 Kaya-Okur, H. S., Janssens, D. H., Henikoff, J. G., Ahmad, K. & Henikoff, S. Efficient low-cost chromatin profiling with CUT&Tag. *Nat Protoc* **15**, 3264-3283, doi:10.1038/s41596-020-0373-x (2020).
37 Wang, Y. N. *et al.* Myocyte enhancer factor 2A promotes proliferation and its inhibition attenuates myogenic differentiation via myozenin 2 in bovine skeletal muscle myoblast. *PLoS One* **13**, e0196255, doi:10.1371/journal.pone.0196255 (2018).
38 Estrella, N. L. *et al.* MEF2 transcription factors regulate distinct gene programs in mammalian skeletal muscle differentiation. *J Biol Chem* **290**, 1256-1268, doi:10.1074/jbc.M114.589838 (2015).
39 Seok, H. Y. *et al.* miR-155 inhibits expression of the MEF2A protein to repress skeletal muscle differentiation. *J Biol Chem* **286**, 35339-35346, doi:10.1074/jbc.M111.273276 (2011).
40 Dedieu, S., Mazeres, G., Dourdin, N., Cottin, P. & Brustis, J. J. Transactivation of capn2 by myogenic regulatory factors during myogenesis. *J Mol Biol* **326**, 453-465, doi:10.1016/s0022-2836(02)01310-4 (2003).
41 Honda, M., Masui, F., Kanzawa, N., Tsuchiya, T. & Toyo-oka, T. Specific knockdown of m-calpain blocks myogenesis with cDNA deduced from the corresponding RNAi. *Am J Physiol Cell Physiol* **294**, C957-965, doi:10.1152/ajpcell.00505.2007 (2008).
42 Swailes, N. T., Colegrave, M., Knight, P. J. & Peckham, M. Non-muscle myosins 2A and 2B drive changes in cell morphology that occur as myoblasts align and fuse. *J Cell Sci* **119**, 3561-3570, doi:10.1242/jcs.03096 (2006).
43 Vesely, E. D., Heilig, C. W. & Brosius, F. C., 3rd. GLUT1-induced cFLIP expression promotes proliferation and prevents apoptosis in vascular smooth muscle cells. *Am J Physiol Cell Physiol* **297**, C759-765, doi:10.1152/ajpcell.00213.2009 (2009).
44 Budd, R. C., Yeh, W. C. & Tschopp, J. cFLIP regulation of lymphocyte activation and development. *Nat Rev Immunol* **6**, 196-204, doi:10.1038/nri1787 (2006).
45 Wang, W., Prince, C. Z., Mou, Y. & Pollman, M. J. Notch3 signaling in vascular smooth muscle cells induces c-FLIP expression via ERK/MAPK activation. Resistance to Fas ligand-induced apoptosis. *J Biol Chem* **277**, 21723-21729, doi:10.1074/jbc.M202224200 (2002).
46 Zhang, N. & He, Y. W. An essential role for c-FLIP in the efficient development of mature T lymphocytes. *J Exp Med* **202**, 395-404, doi:10.1084/jem.20050117 (2005).
47 Tozer, S. & Duprez, D. Tendon and ligament: development, repair and disease. *Birth Defects Res C Embryo Today* **75**, 226-236, doi:10.1002/bdrc.20049 (2005).
48 Schweitzer, R. *et al.* Analysis of the tendon cell fate using Scleraxis, a specific marker for tendons and ligaments. *Development* **128**, 3855-3866, doi:10.1242/dev.128.19.3855 (2001).
49 Yaseen, W. *et al.* Fibroblast fusion to the muscle fiber regulates myotendinous junction formation. *Nat Commun* **12**, 3852, doi:10.1038/s41467-021-24159-9 (2021).
50 Shukunami, C., Takimoto, A., Oro, M. & Hiraki, Y. Scleraxis positively regulates the expression of tenomodulin, a differentiation marker of tenocytes. *Dev Biol* **298**, 234-247, doi:10.1016/j.ydbio.2006.06.036 (2006).
51 Li, Y., Wu, T. & Liu, S. Identification and Distinction of Tenocytes and Tendon-Derived Stem Cells. *Front Cell Dev Biol* **9**, 629515, doi:10.3389/fcell.2021.629515 (2021).
52 Ren, D. *et al.* Overexpressed ITGA2 promotes malignant tumor aggression by up-regulating PD-L1 expression through the activation of the STAT3 signaling pathway. *J Exp Clin Cancer Res* **38**, 485, doi:10.1186/s13046-019-1496-1 (2019).

53 Cai, H. *et al.* Overexpressed integrin alpha 2 inhibits the activation of the transforming growth factor beta pathway in pancreatic cancer via the TFCP2-SMAD2 axis. *J Exp Clin Cancer Res* **41**, 73, doi:10.1186/s13046-022-02286-5 (2022).
54 Chuang, Y. C. *et al.* Blockade of ITGA2 Induces Apoptosis and Inhibits Cell Migration in Gastric Cancer. *Biol Proced Online* **20**, 10, doi:10.1186/s12575-018-0073-x (2018).
55 Huang, Y. *et al.* CRK proteins selectively regulate T cell migration into inflamed tissues. *J Clin Invest* **125**, 1019-1032, doi:10.1172/JCI77278 (2015).
56 Lin, X. *et al.* Protein tyrosine phosphatase-like A regulates myoblast proliferation and differentiation through MyoG and the cell cycling signaling pathway. *Mol Cell Biol* **32**, 297-308, doi:10.1128/MCB.05484-11 (2012).
57 Blondelle, J. *et al.* HACD1, a regulator of membrane composition and fluidity, promotes myoblast fusion and skeletal muscle growth. *J Mol Cell Biol* **7**, 429-440, doi:10.1093/jmcb/mjv049 (2015).
58 Paxton, C. W. *et al.* BTB-Kelch protein Krp1 regulates proliferation and differentiation of myoblasts. *Am J Physiol Cell Physiol* **300**, C1345-1355, doi:10.1152/ajpcell.00321.2010 (2011).
59 Ramirez-Martinez, A. *et al.* KLHL41 stabilizes skeletal muscle sarcomeres by nonproteolytic ubiquitination. *Elife* **6**, doi:10.7554/eLife.26439 (2017).
60 Posey, A. D., Jr. *et al.* Endocytic recycling proteins EHD1 and EHD2 interact with fer-1-like-5 (Fer1L5) and mediate myoblast fusion. *J Biol Chem* **286**, 7379-7388, doi:10.1074/jbc.M110.157222 (2011).
61 Doherty, K. R. *et al.* The endocytic recycling protein EHD2 interacts with myoferlin to regulate myoblast fusion. *J Biol Chem* **283**, 20252-20260, doi:10.1074/jbc.M802306200 (2008).
62 Dubinska-Magiera, M., Zaremba-Czogalla, M. & Rzepecki, R. Muscle development, regeneration and laminopathies: how lamins or lamina-associated proteins can contribute to muscle development, regeneration and disease. *Cell Mol Life Sci* **70**, 2713-2741, doi:10.1007/s00018-012-1190-3 (2013).
63 Maggi, L., Carboni, N. & Bernasconi, P. Skeletal Muscle Laminopathies: A Review of Clinical and Molecular Features. *Cells* **5**, doi:10.3390/cells5030033 (2016).
64 S Srinivas 1, T. W., C S Lin, C M William, Y Tanabe, T M Jessell, F Costantini. Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Developmental Biology* (2001).
65 Yue, L., Wan, R., Luan, S., Zeng, W. & Cheung, T. H. Dek Modulates Global Intron Retention during Muscle Stem Cells Quiescence Exit. *Dev Cell* **53**, 661-676 e666, doi:10.1016/j.devcel.2020.05.006 (2020).
66 Ewels, P. A. *et al.* The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol* **38**, 276-278, doi:10.1038/s41587-020-0439-x (2020).
67 McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* **8**, 329-337 e324, doi:10.1016/j.cels.2019.03.003 (2019).
68 Qiu, X. *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* **14**, 979-982, doi:10.1038/nmeth.4402 (2017).

Figure 1

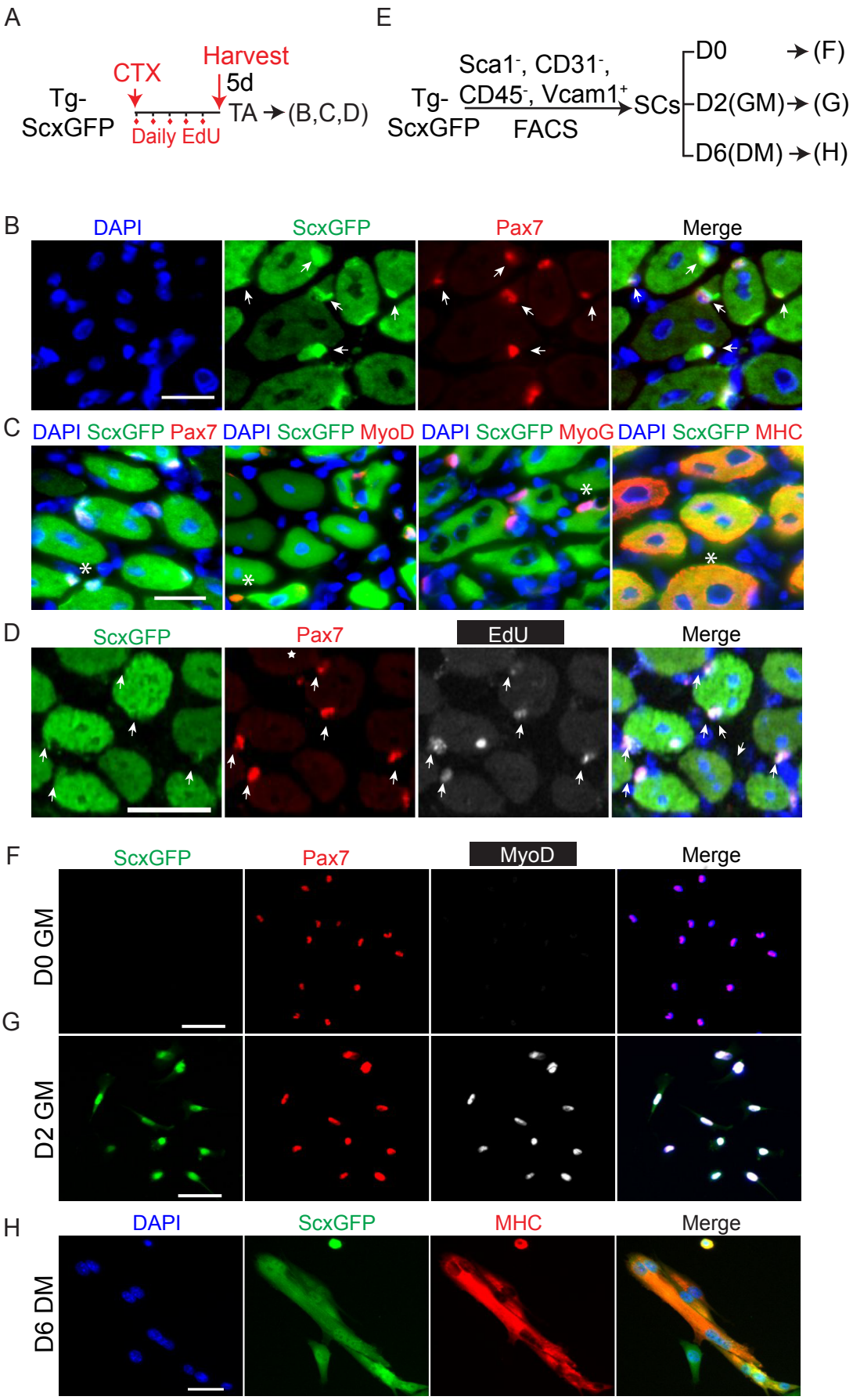

A
Tg-
ScxGFP
CTX
Harvest
5d
Daily EdU
TA → (B,C,D)
E
Tg-
ScxGFP
Sca1-, CD31-,
CD45-, Vcam1+
FACS
SCs
D0 → (F)
D2(GM) → (G)
D6(DM) → (H)
B
DAPI
ScxGFP
Pax7
Merge
C
DAPI ScxGFP Pax7
DAPI ScxGFP MyoD
DAPI ScxGFP MyoG
DAPI ScxGFP MHC
D
ScxGFP
Pax7
EdU
Merge
F
ScxGFP
Pax7
MyoD
Merge
D0 GM
G
D2 GM
H
DAPI
ScxGFP
MHC
Merge
D6 DM

Figure S1

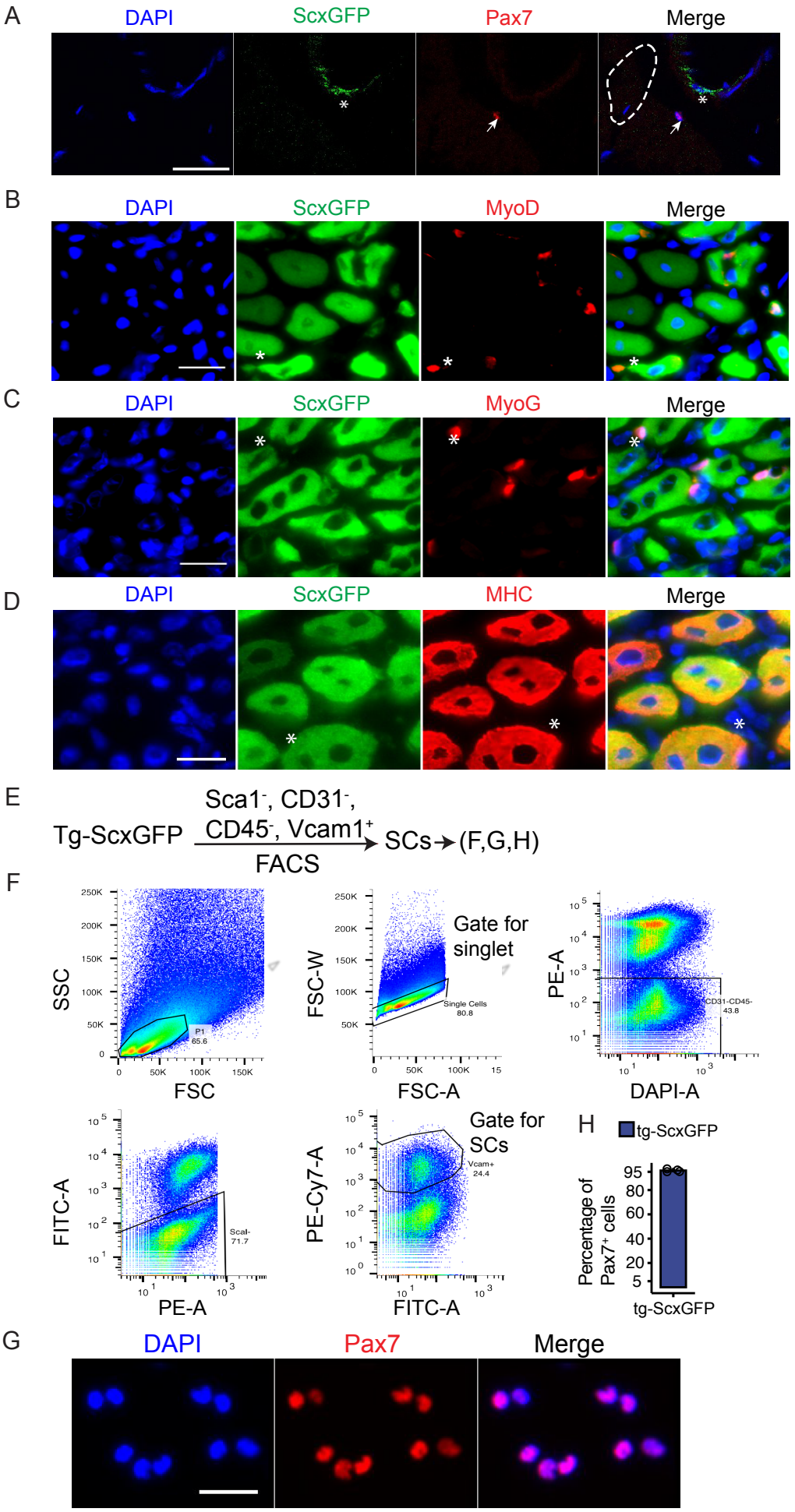

A
DAPI
ScxGFP
Pax7
Merge
B
DAPI
ScxGFP
MyoD
Merge
C
DAPI
ScxGFP
MyoG
Merge
D
DAPI
ScxGFP
MHC
Merge
E
Tg-ScxGFP
Sca1-, CD31-,
CD45-, Vcam1+
FACS
SCs→(F,G,H)
F
SSC
FSC
P1
65.6
FSC-W
FSC-A
Gate for
singlet
Single Cells
80.8
PE-A
DAPI-A
CD31-CD45-
43.8
FITC-A
PE-A
Sca1-
71.7
PE-Cy7-A
FITC-A
Gate for
SCs
Vcam+
24.4
H
tg-ScxGFP
Percentage of
Pax7+ cells
95
80
60
40
20
5
tg-ScxGFP
G
DAPI
Pax7
Merge

Figure 2

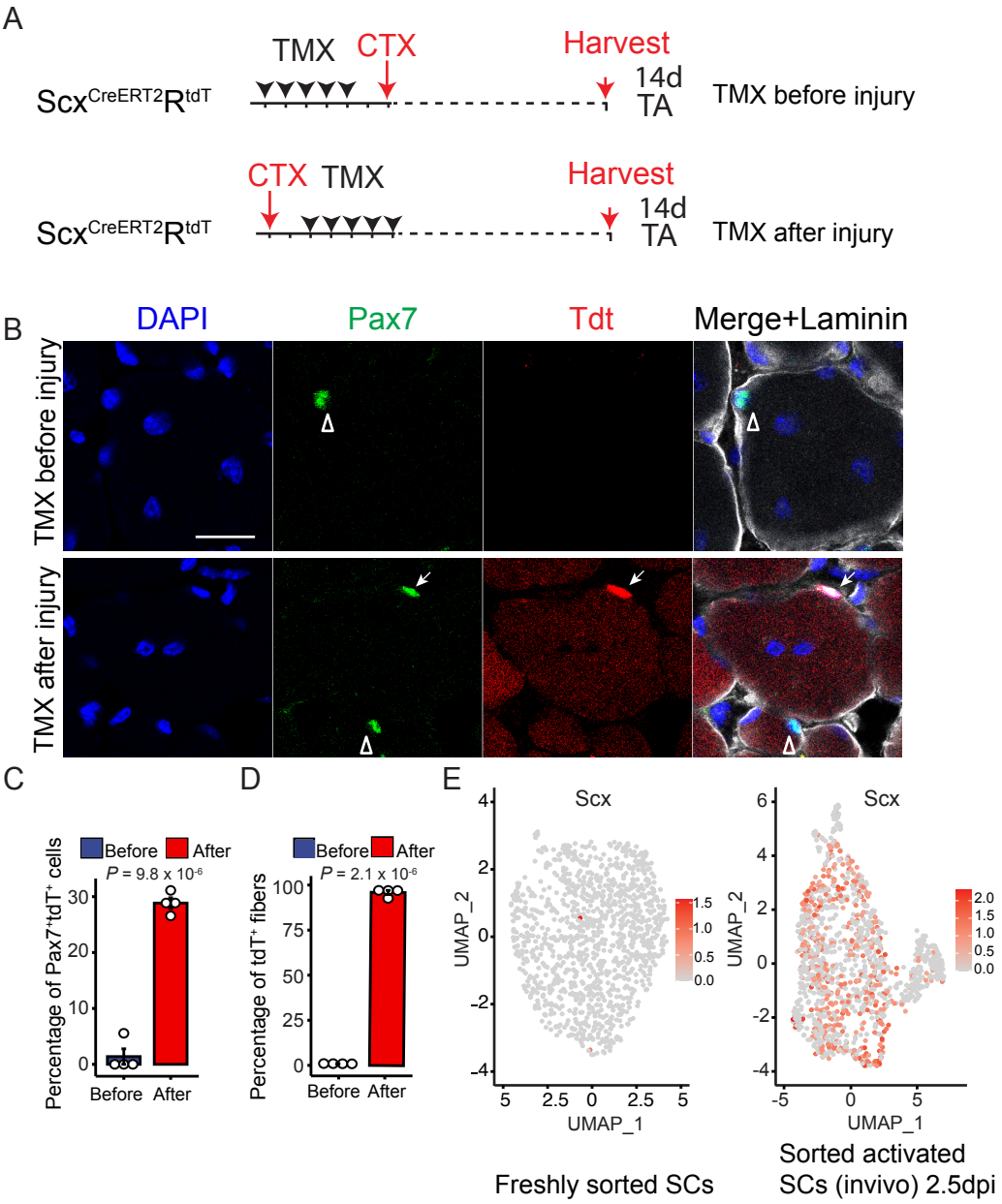
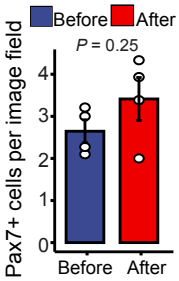

A
TMX
CTX
Harvest
ScxCreERT2RtdT
14d
TA
TMX before injury
CTX
TMX
Harvest
ScxCreERT2RtdT
14d
TA
TMX after injury
B
DAPI
Pax7
Tdt
Merge+Laminin
TMX before injury
TMX after injury
C
Before
After
P = 9.8 x 10-6
Percentage of Pax7+tdT+ cells
Before
After
D
Before
After
P = 2.1 x 10-6
Percentage of tdT+ fibers
Before
After
E
Scx
UMAP_2
UMAP_1
Scx
UMAP_2
UMAP_1
Freshly sorted SCs
Sorted activated
SCs (invivo) 2.5dpi

Figure S2

A



B

FPKM / Gene symble	WT SCs	MDX SCs
Scx	5.15 ± 0.62	9.37 ± 0.89
Pax7	59.63 ± 7.63	10.65 ± 0.14

Figure 3

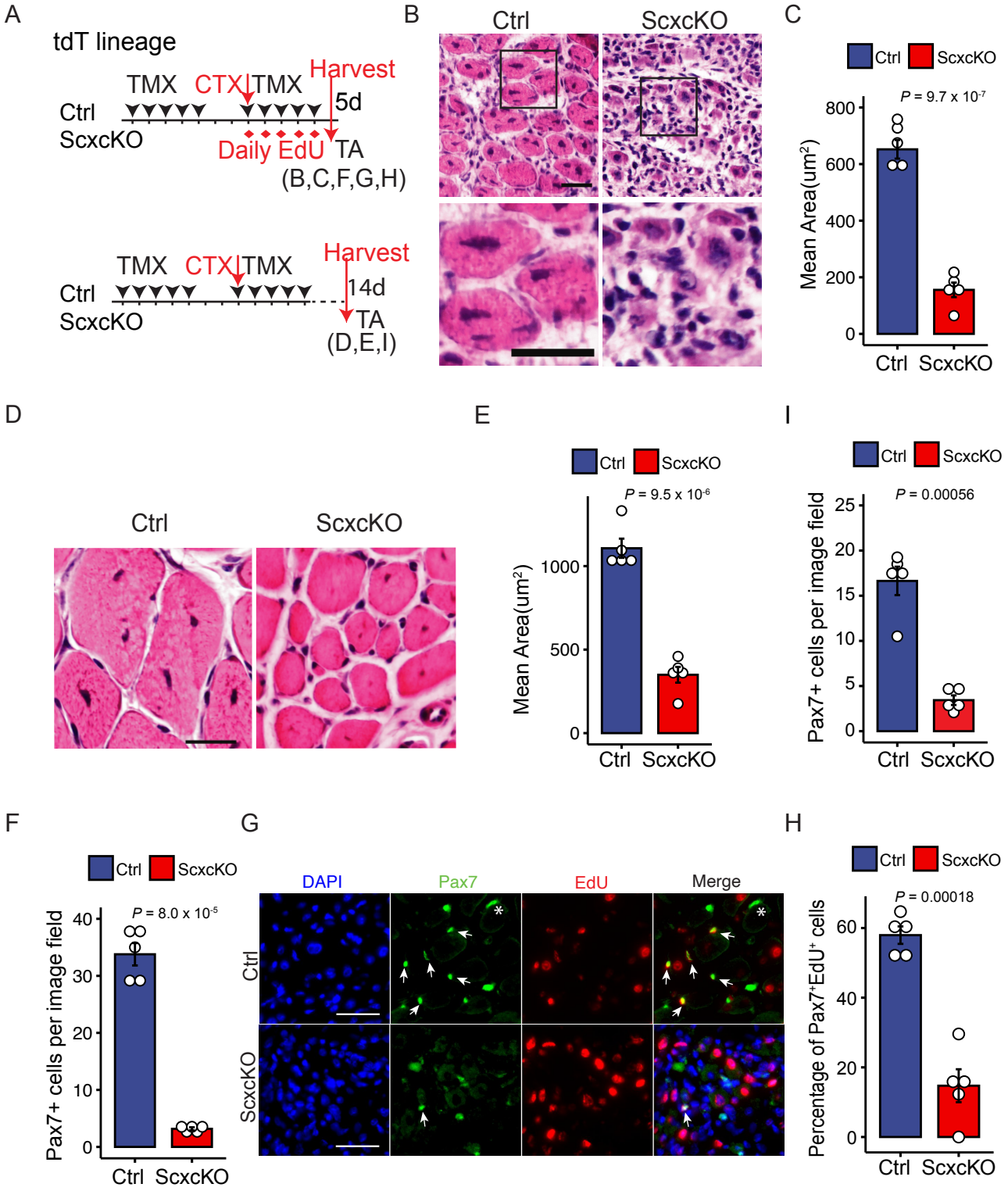

A
tdT lineage
TMX
CTX
TMX
Harvest
5d
Ctrl
ScxcKO
Daily EdU
TA
(B,C,F,G,H)
TMX
CTX
TMX
Harvest
14d
Ctrl
ScxcKO
TA
(D,E,I)
B
Ctrl
ScxcKO
C
Ctrl
ScxcKO
P = 9.7 x 10-7
Mean Area(um2)
800
600
400
200
0
Ctrl
ScxcKO
D
Ctrl
ScxcKO
E
Ctrl
ScxcKO
P = 9.5 x 10-6
Mean Area(um2)
1000
500
0
Ctrl
ScxcKO
I
Ctrl
ScxcKO
P = 0.00056
Pax7+ cells per image field
25
20
15
10
5
0
Ctrl
ScxcKO
F
Ctrl
ScxcKO
P = 8.0 x 10-5
Pax7+ cells per image field
40
30
20
10
0
Ctrl
ScxcKO
G
DAPI
Pax7
EdU
Merge
Ctrl
ScxcKO
H
Ctrl
ScxcKO
P = 0.00018
Percentage of Pax7+EdU+ cells
60
40
20
0
Ctrl
ScxcKO

Figure S3

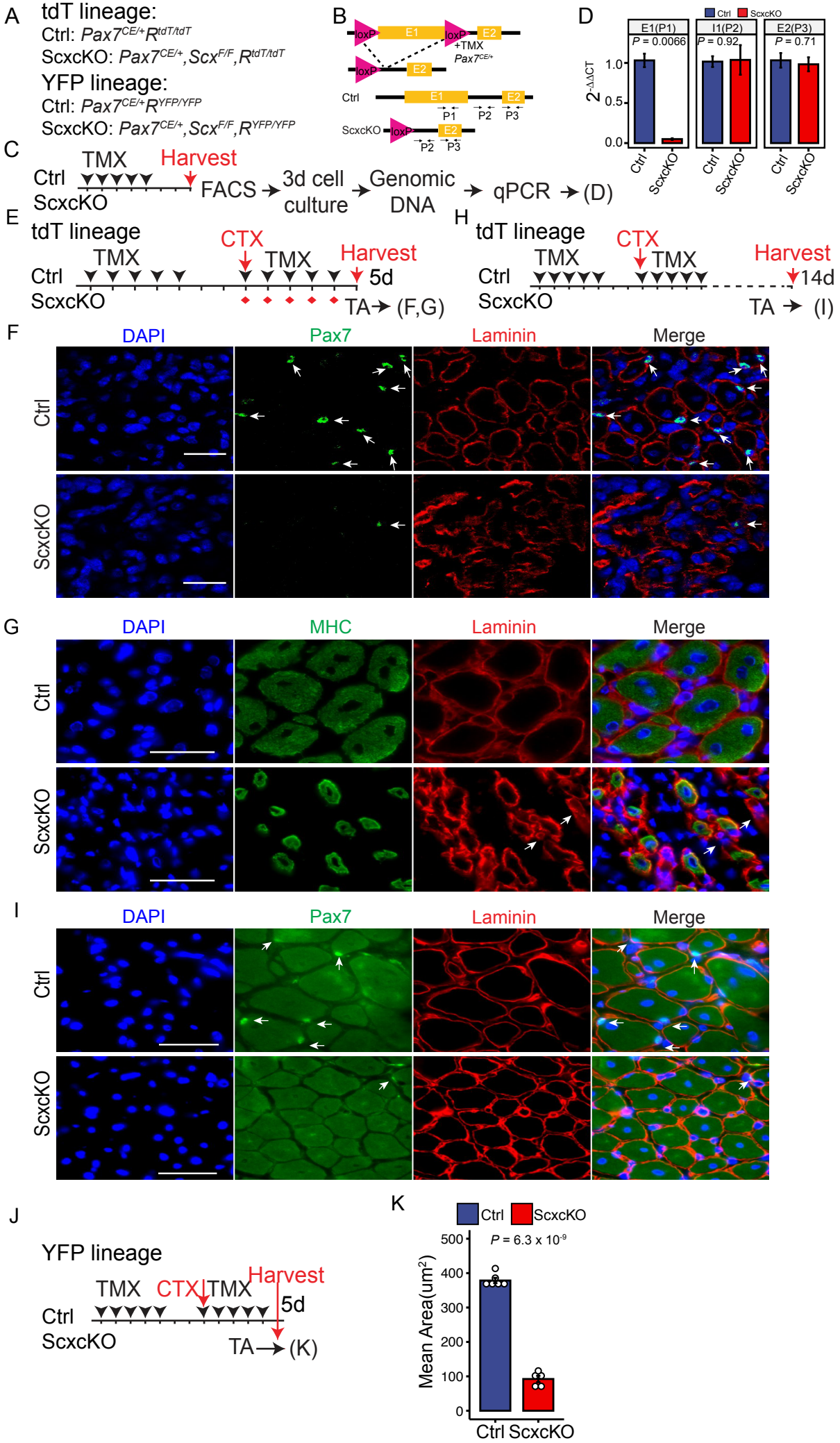

A
tdT lineage:
Ctrl: Pax7CE/+RtdT/tdT
ScxcKO: Pax7CE/+,ScxF/F,RtdT/tdT
YFP lineage:
Ctrl: Pax7CE/+RYFP/YFP
ScxcKO: Pax7CE/+,ScxF/F,RYFP/YFP
B
loxP
E1
E2
+TMX
Pax7CE/+
Ctrl
ScxcKO
P1
P2
P3
D
Ctrl
ScxcKO
E1(P1)
I1(P2)
E2(P3)
P = 0.0066
P = 0.92
P = 0.71
2-ΔΔCT
1.0
0.5
0.0
C
TMX
Harvest
FACS
3d cell culture
Genomic DNA
qPCR
(D)
E
tdT lineage
CTX
5d
TA
(F,G)
H
14d
(I)
F
DAPI
Pax7
Laminin
Merge
G
MHC
I
J
YFP lineage
K
Mean Area(um2)
P = 6.3 x 10-9
500
400
300
200
100
0

Figure 4

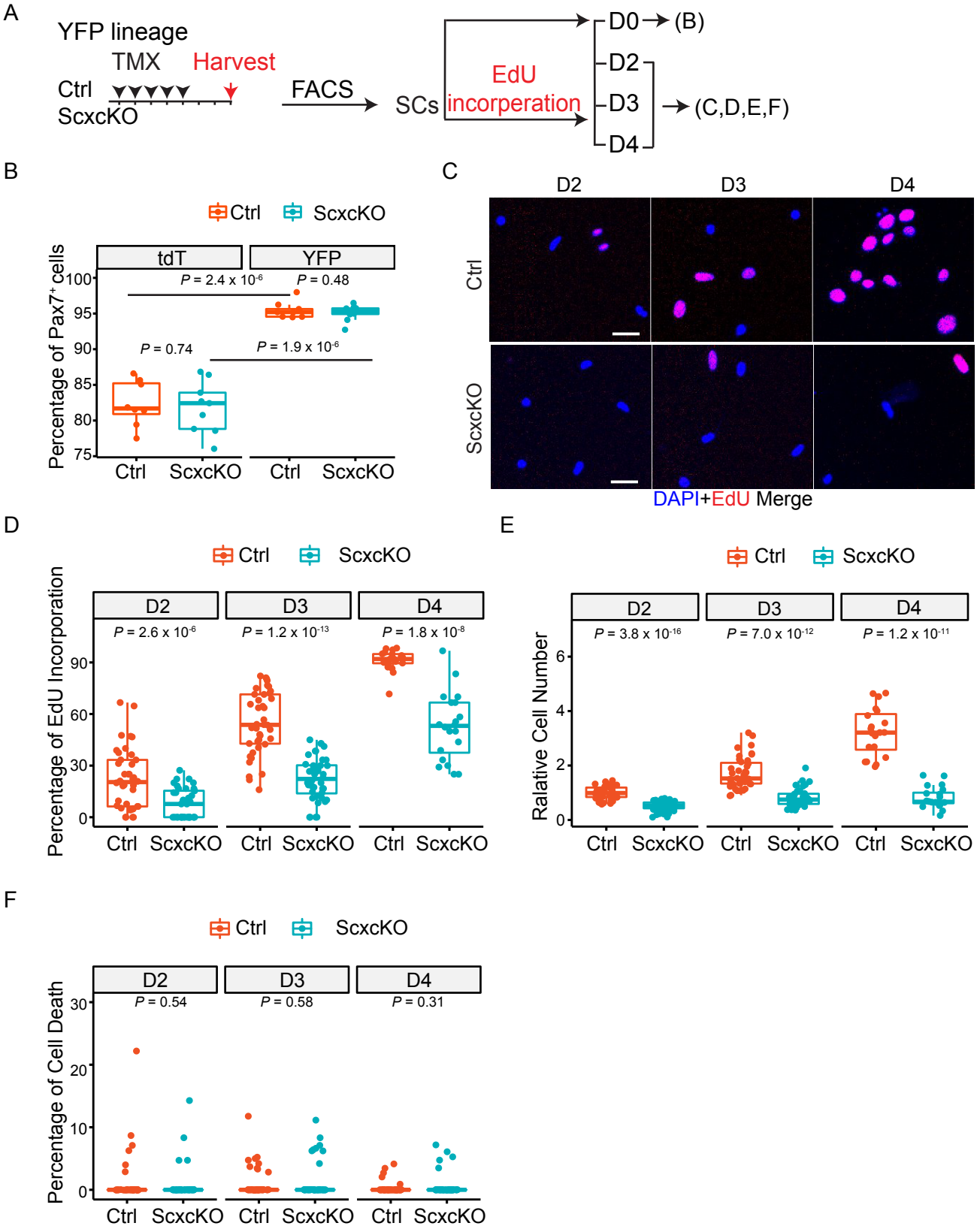

A
YFP lineage
TMX
Harvest
Ctrl
ScxcKO
FACS
SCs
EdU incorperation
D0
(B)
D2
D3
D4
(C,D,E,F)
B
Ctrl
ScxcKO
tdT
YFP
Percentage of Pax7+ cells
P = 2.4 x 10-6
P = 0.48
P = 0.74
P = 1.9 x 10-6
C
D2
D3
D4
Ctrl
ScxcKO
DAPI+EdU Merge
D
Percentage of EdU Incorporation
P = 2.6 x 10-6
P = 1.2 x 10-13
P = 1.8 x 10-8
E
Ralative Cell Number
P = 3.8 x 10-16
P = 7.0 x 10-12
P = 1.2 x 10-11
F
Percentage of Cell Death
P = 0.54
P = 0.58
P = 0.31

Figure S4

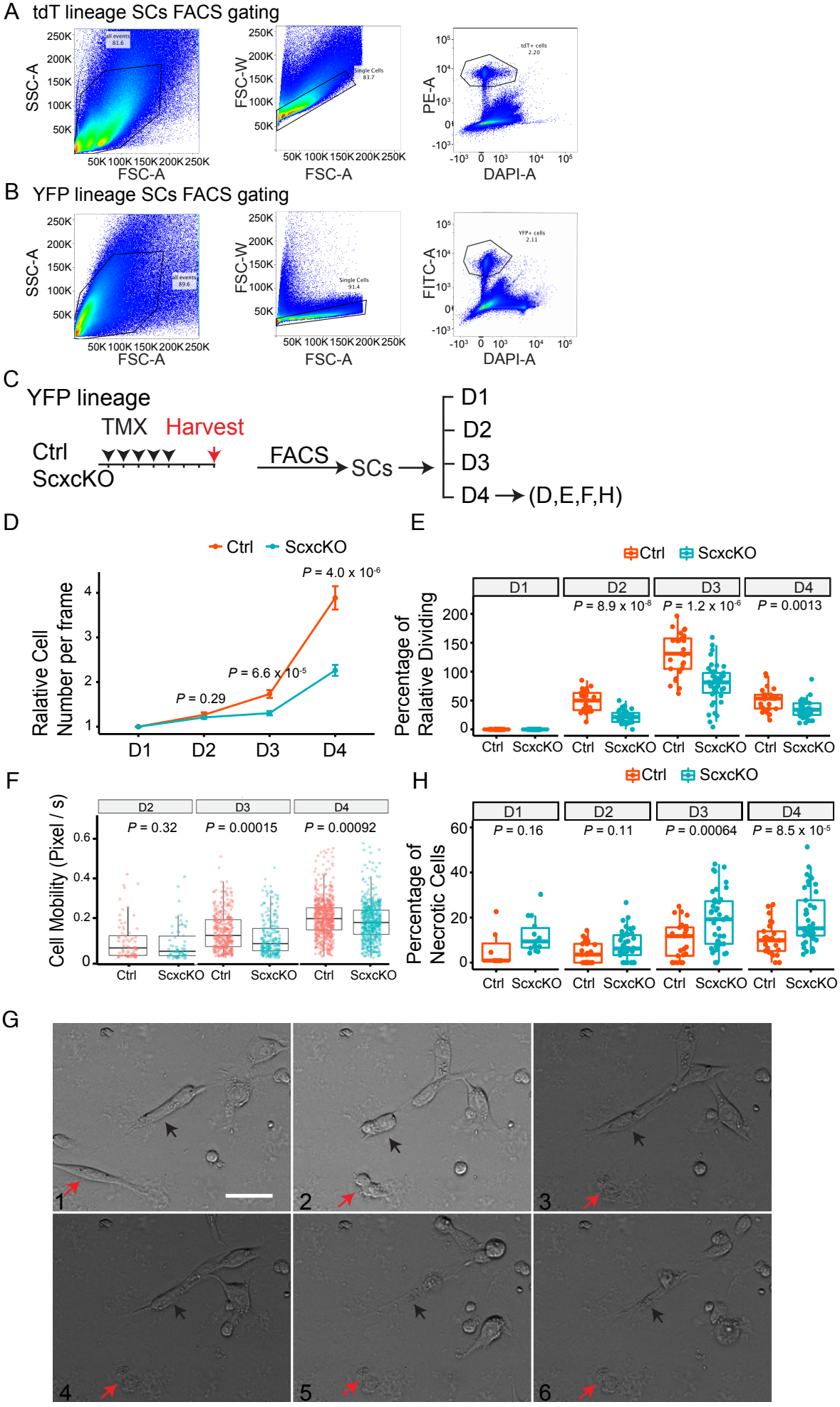

A tdT lineage SCs FACS gating
SSC-A
FSC-A
FSC-W
PE-A
DAPI-A
B YFP lineage SCs FACS gating
FITC-A
C
YFP lineage
TMX
Harvest
Ctrl
ScxcKO
FACS
SCs
D1
D2
D3
D4
(D,E,F,H)
D
Ctrl
ScxcKO
Ralative Cell Number per frame
P = 4.0 x 10-6
P = 6.6 x 10-5
P = 0.29
E
Percentage of Ralative Dividing
P = 8.9 x 10-8
P = 1.2 x 10-6
P = 0.0013
F
Cell Mobility (Pixel / s)
P = 0.32
P = 0.00015
P = 0.00092
H
Percentage of Necrotic Cells
P = 0.16
P = 0.11
P = 0.00064
P = 8.5 x 10-5
G
1
2
3
4
5
6

Figure 5

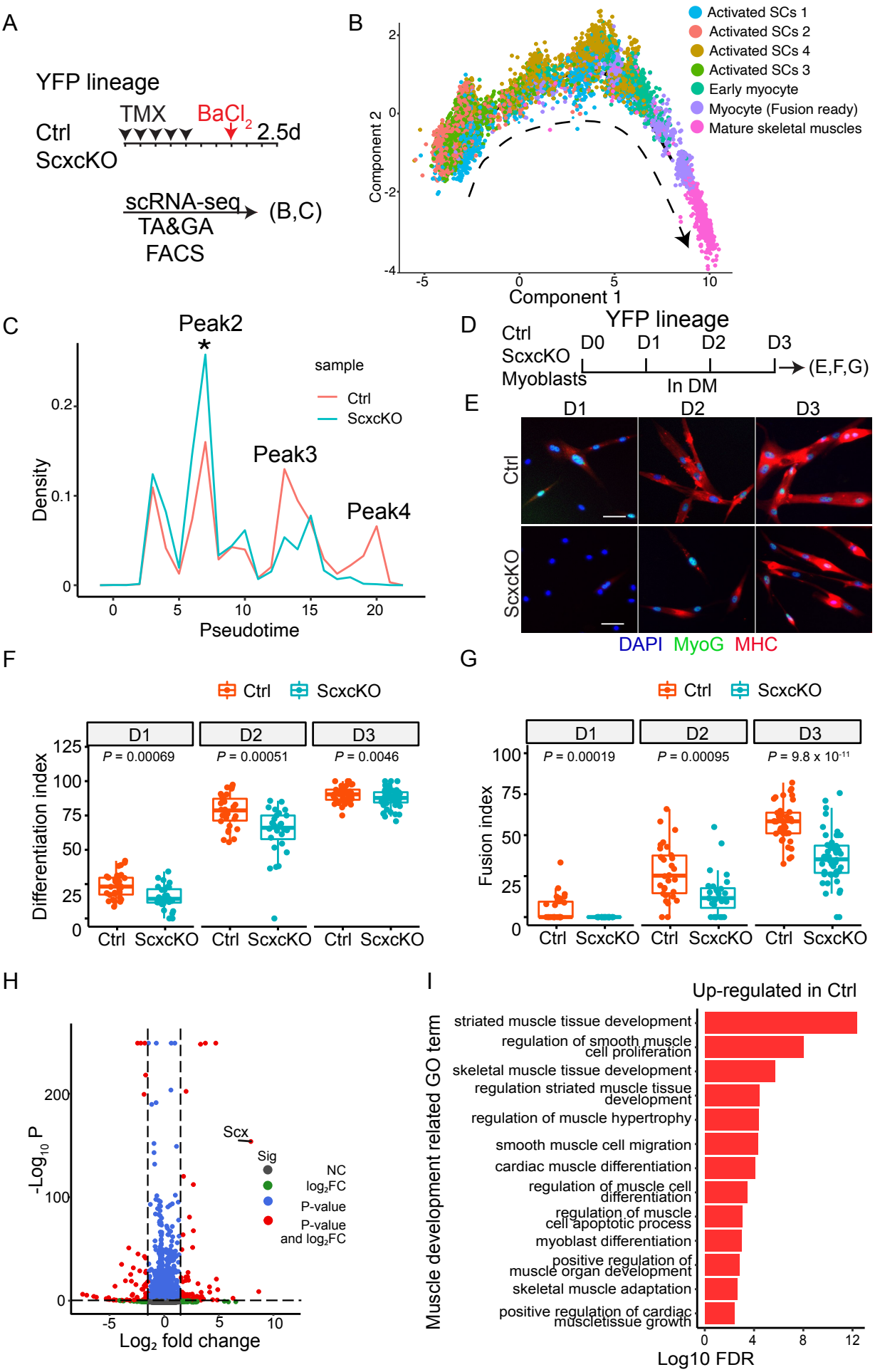
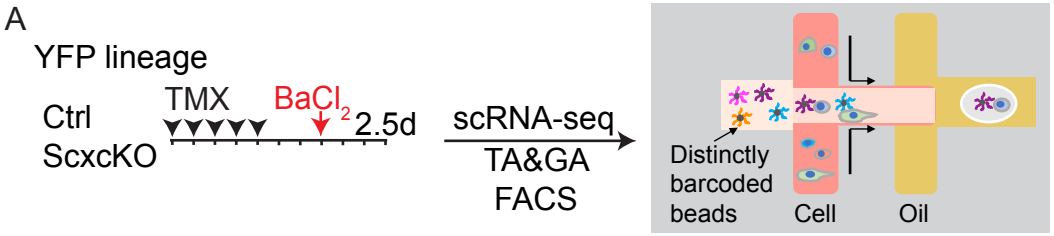
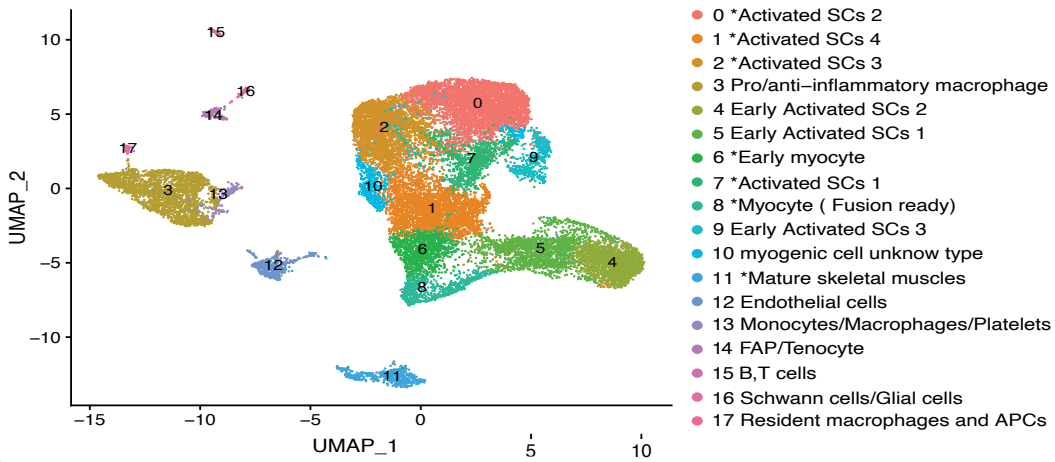

A
YFP lineage
Ctrl
ScxcKO
TMX
BaCl2
2.5d
scRNA-seq
TA&GA
FACS
(B,C)
B
Activated SCs 1
Activated SCs 2
Activated SCs 4
Activated SCs 3
Early myocyte
Myocyte (Fusion ready)
Mature skeletal muscles
Component 2
Component 1
C
Peak2
Peak3
Peak4
sample
Ctrl
ScxcKO
Density
Pseudotime
D
YFP lineage
Ctrl
ScxcKO
Myoblasts
D0
D1
D2
D3
In DM
(E,F,G)
E
D1
D2
D3
Ctrl
ScxcKO
DAPI MyoG MHC
F
Ctrl
ScxcKO
D1
D2
D3
P = 0.00069
P = 0.00051
P = 0.0046
Differentiation index
G
Ctrl
ScxcKO
D1
D2
D3
P = 0.00019
P = 0.00095
P = 9.8 x 10-11
Fusion index
H
Scx
Sig
NC
log2FC
P-value
P-value and log2FC
-Log10 P
Log2 fold change
I
Up-regulated in Ctrl
striated muscle tissue development
regulation of smooth muscle cell proliferation
skeletal muscle tissue development
regulation striated muscle tissue development
regulation of muscle hypertrophy
smooth muscle cell migration
cardiac muscle differentiation
regulation of muscle cell differentiation
regulation of muscle cell apoptotic process
myoblast differentiation
positive regulation of muscle organ development
skeletal muscle adaptation
positive regulation of cardiac muscletissue growth
Muscle development related GO term
Log10 FDR

Figure S5

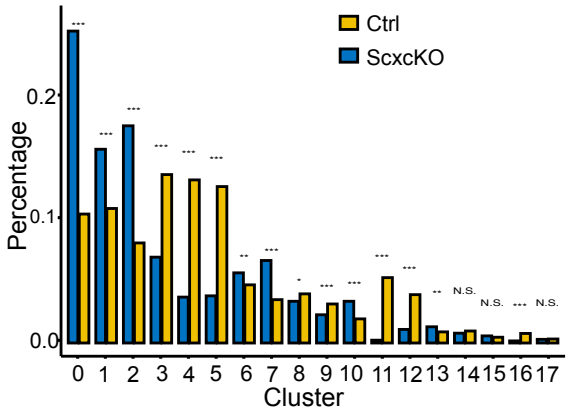
A



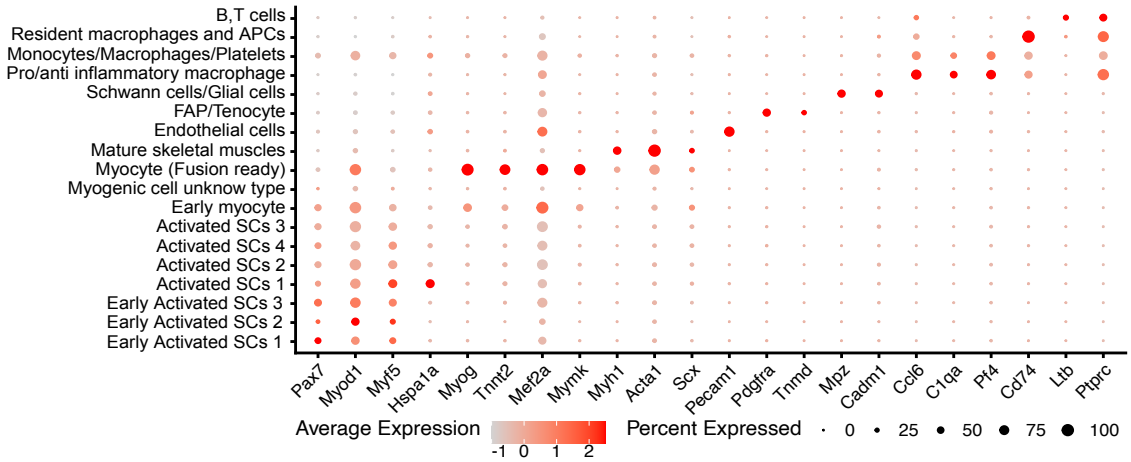
B



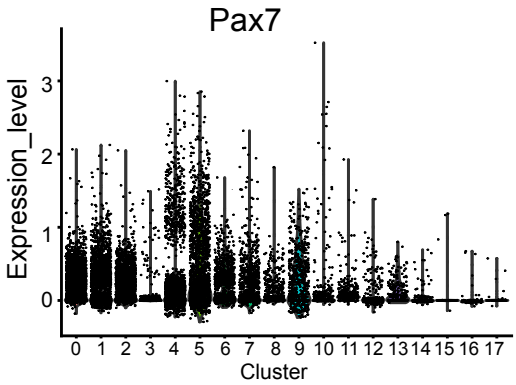
C



D



E



F

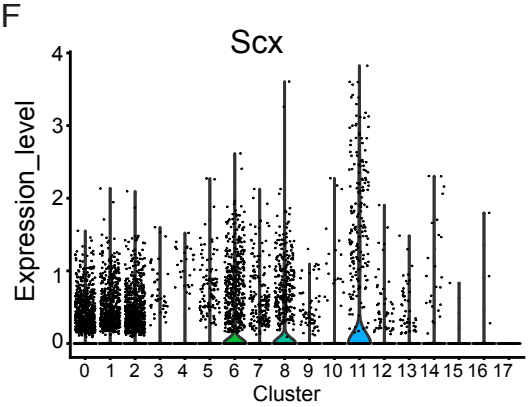


Figure S5

G

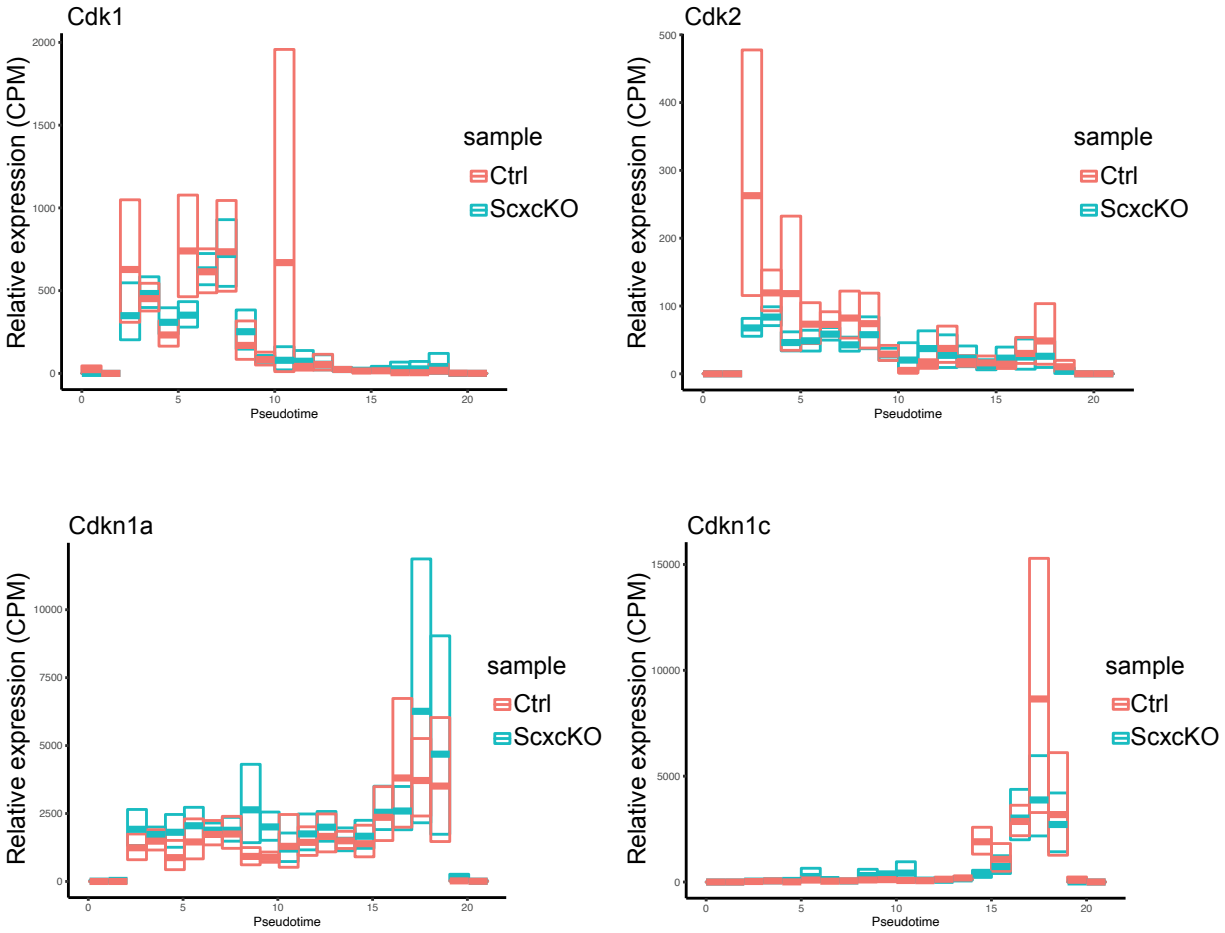
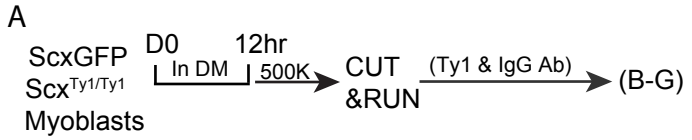
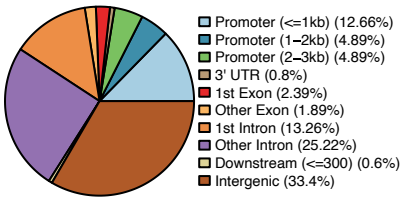

Cdk1
Relative expression (CPM)
Pseudotime
sample
Ctrl
ScxcKO
Cdk2
Relative expression (CPM)
Pseudotime
sample
Ctrl
ScxcKO
Cdkn1a
Relative expression (CPM)
Pseudotime
sample
Ctrl
ScxcKO
Cdkn1c
Relative expression (CPM)
Pseudotime
sample
Ctrl
ScxcKO

Figure 6

A



B

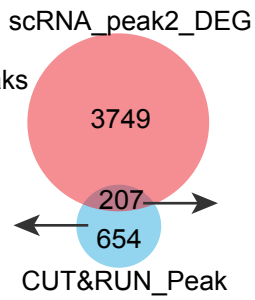


C

bHLH Motif enrichment of CUT&RUN Peaks

Name	Motif	*P*-value
Scx		3.76 x 10^{-19}
Scx		2.80 x 10^{-19}

D

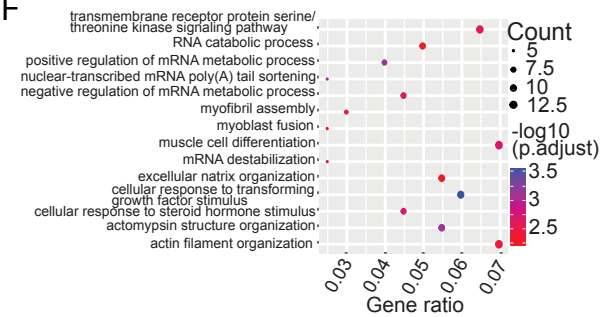


E

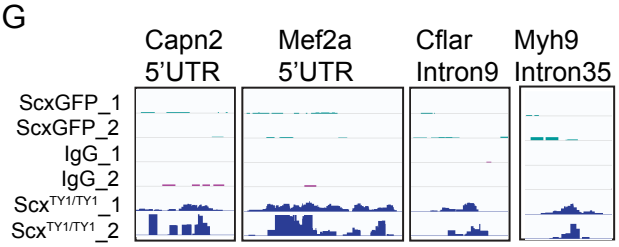
bHLH Motif enrichment of commen Peaks

Name	Motif	*P*-value
Scx		8.04 x 10^{-6}
Scx		3.76 x 10^{-6}

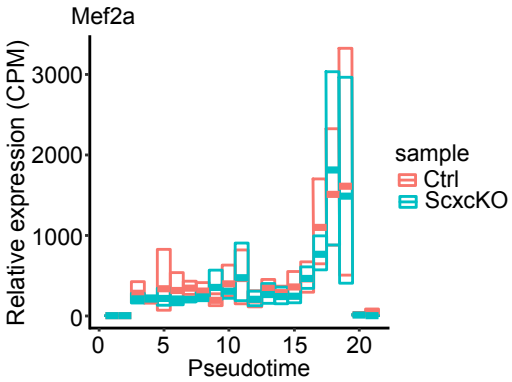
F



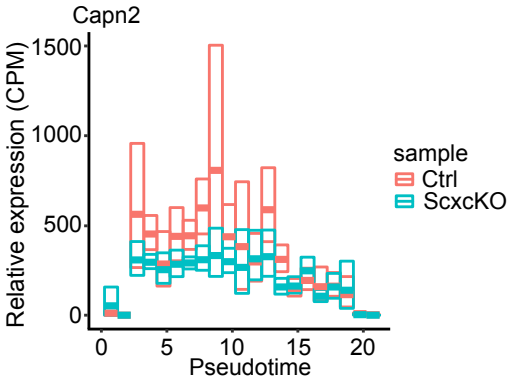
G



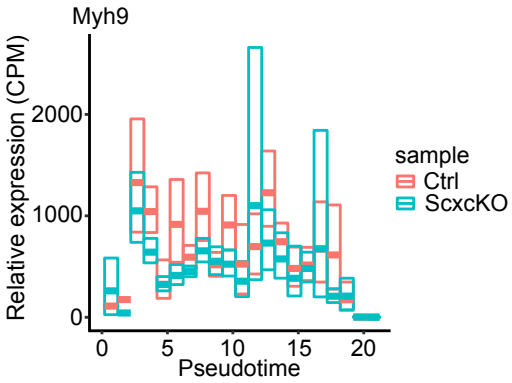
H



I



J



K

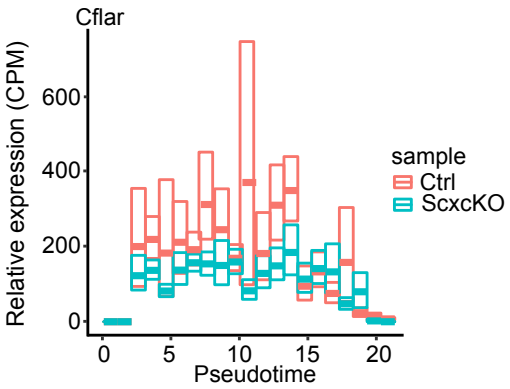
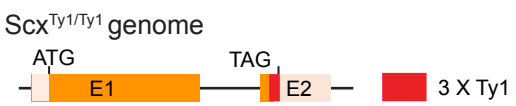
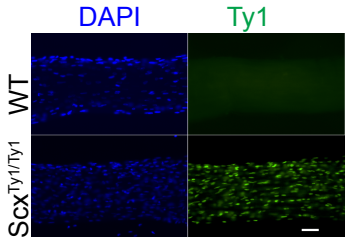


Figure S6

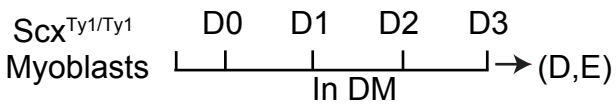
A



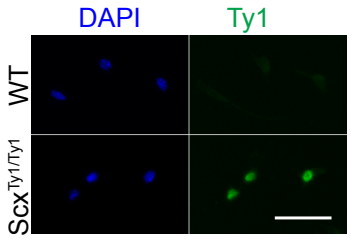
B



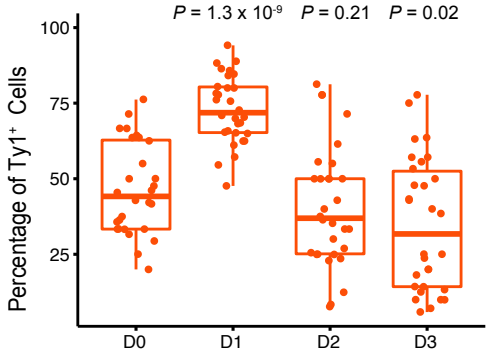
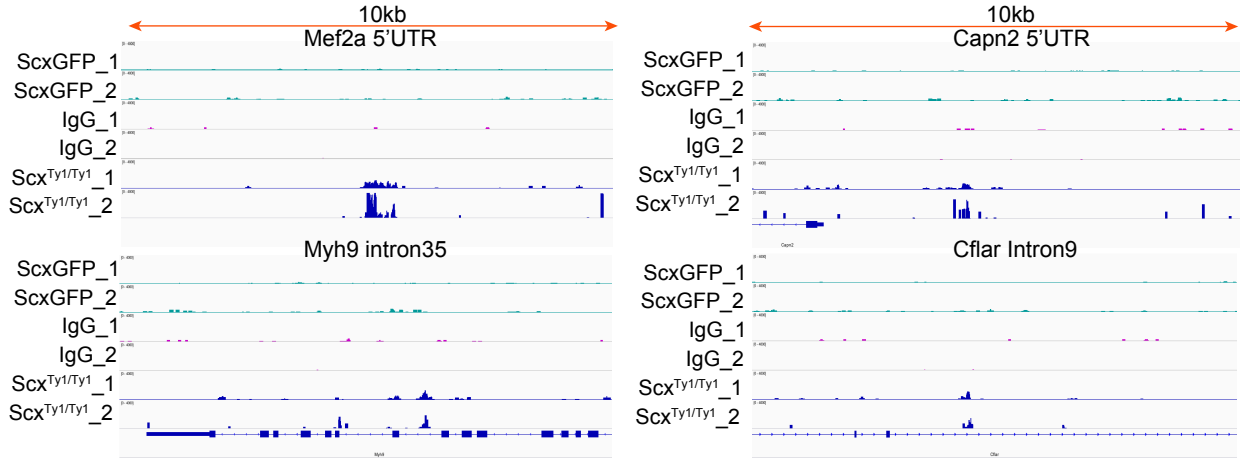
C



D



E



F