

## **FoxO maintains a genuine muscle stem-cell quiescent state until geriatric age**

Laura García-Prat<sup>1,2^#</sup>, Eusebio Perdiguero<sup>1^</sup>, Sonia Alonso-Martín<sup>2^&</sup>, Stefania Dell’Orso<sup>3+</sup>, Srikanth Ravichandran<sup>4+</sup>, Stephen R. Brooks<sup>5+</sup>, Aster H. Juan<sup>3</sup>, Silvia Campanario<sup>1,2</sup>, Kan Jiang<sup>5</sup>, Xiaotong Hong<sup>2</sup>, Laura Ortet<sup>1</sup>, Vanessa Ruiz-Bonilla<sup>1</sup>, Marta Flández<sup>2§</sup>, Victoria Moiseeva<sup>1</sup>, Elena Rebollo<sup>6</sup>, Mercè Jardí<sup>1</sup>, Hong-Wei Sun<sup>5</sup>, Antonio Musarò<sup>7</sup>, Marco Sandris, Antonio del Sol<sup>4,9,10</sup>, Vittorio Sartorelli<sup>3\*</sup> and Pura Muñoz-Cánoves<sup>1,2,11\*</sup>

<sup>1</sup>Department of Experimental and Health Sciences, Pompeu Fabra University (UPF), CIBER on Neurodegenerative diseases (CIBERNED), E-08003 Barcelona, Spain; <sup>2</sup>Spanish National Center on Cardiovascular Research (CNIC), E-28029 Madrid, Spain; <sup>3</sup>Laboratory of Muscle Stem Cells and Gene Regulation, National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), NIH Bethesda, MD 208292, USA; <sup>4</sup>Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 4362, Luxembourg; <sup>5</sup>Biodata Mining and Discovery Section, NIAMS, NIH Bethesda, MD 20892 USA; <sup>6</sup>Advanced Fluorescence Microscopy Unit, Molecular Biology Institute of Barcelona (IBMB-CSIC), E-08028 Barcelona, Spain; <sup>7</sup>DAHFMO-Unit of Histology and Medical Embryology, Sapienza University of Rome, Laboratory Affiliated to Istituto Pasteur Italia-Fondazione Cenci Bolognetti, 00161 Rome, Italy; <sup>8</sup>Department of Biomedical Science, University of Padova, 35100 Padova, Italy; <sup>9</sup>ICREA, E-08010 Barcelona, Spain; <sup>10</sup>CIC bioGUNE, E-48160 Derio, Spain; <sup>11</sup>IKERBASQUE, Basque Foundation for Science, 48009 Bilbao, Spain

<sup>^</sup>Equal contribution

<sup>+</sup>Equal contribution

\*Corresponding authors:

Pura Muñoz-Cánoves (pura.munoz@upf.edu); Vittorio Sartorelli (sartorev@mail.nih.gov)

#Present address:

Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 2M9, Canada

&Present address:

Neurosciences Area, Biodonostia Health Research Institute, E-20014, Donostia-San Sebastián, Spain

§Present address:

Grupo de Investigación en Oncología Clínico Traslacional, Instituto de Investigación Hospital 12 de Octubre, E-28041 Madrid, Spain

## Summary

Tissue regeneration declines with ageing, but little is known about whether this arises from changes in stem-cell heterogeneity. Here, in homeostatic skeletal muscle, we identify two quiescent stem-cell states distinguished by relative CD34 expression: CD34<sup>High</sup>, with stemness properties (genuine state), and CD34<sup>Low</sup>, committed to myogenic differentiation (primed state). Unexpectedly, the genuine-quiescent state is preserved into later life, succumbing only in extreme old age due to the acquisition of primed-state traits. Niche-derived IGF1-dependent Akt activation debilitates the genuine stem-cell state by imposing primed-state features via FoxO inhibition. Interventions to neutralize Akt and promote FoxO activity drive primed-to-genuine state conversion, whereas FoxO inactivation deteriorates the genuine state at a young age, causing muscle regenerative failure, as occurs in geriatric mice. These findings reveal transcriptional determinants of stem-cell heterogeneity that resist ageing more than previously anticipated and are only lost in extreme old age, with implications for the repair of geriatric muscle.

**One Sentence Summary:** Skeletal muscle contains a subpopulation of quiescent stem cells that is preserved into late life but succumbs in extreme geriatric age

Stem cell ageing is a nodal cause of the regenerative capacity decline of most tissues over time, and a vivid example is provided by skeletal muscle<sup>1,2</sup>. Its stem cells (also known as satellite cells) are quiescent for most of their life. In response to injury, these quiescent cells activate, expand and acquire distinct fates, with some differentiating and forming new myofibres and others self-renewing to replenish the homeostatic quiescent stem-cell pool<sup>3,4</sup>. Mouse transgenesis, *in vivo* labelling, and recent single-cell transcriptomic analyses suggest heterogeneity among quiescent satellite cells (QSCs)<sup>5-12</sup>; however, whether this heterogeneity is affected by age and whether it influences muscle stem-cell progeny, fate and regenerative function with ageing is not known.

The present study addresses the dynamics of muscle stem-cell quiescence diversity throughout the entire postnatal life of an animal. We show that quiescent stem cells expressing high CD34 levels emerge during juvenile/young age and are set to self-renew upon injury, whereas those expressing low CD34 levels are primed and restricted to the myogenic fate. Based on this classification, we term these states “genuine” and “primed”. We further show that the genuine stem-cell state is preserved into old age through active repression of gene programs characteristic of the primed state. However, in extreme old (geriatric) age, inappropriate expression of primed-state genes causes regenerative failure.

## Results

### Differential CD34 expression distinguishes genuine and primed quiescent stem-cell states

Our first goal was to achieve a molecular and functional characterization of quiescent satellite cells (QSCs) and to identify cell-surface receptors apt for their isolation by flow cytometry. For this, we performed a meta-analysis of the transcriptomic data generated in five laboratories from quiescent and activated satellite cells using ImaGEO<sup>13</sup>, and obtained a ‘quiescence core signature’ of genes upregulated in quiescence (Fig. 1a; Extended Data Fig. 1a; Table S1). This set of upregulated genes included *Cd34*, a classical surface marker on many stem-cell types, including satellite cells<sup>14-16</sup> (Fig. 1a; Table S1). By coupling CD34 to the  $\alpha 7$ -integrin ( $\alpha 7$ )-based fluorescence-activated cell sorting (FACS) protocols<sup>5,17,18</sup>, we confirmed that CD34 expression was associated with QSCs (Fig. 1b)<sup>14-16,19,20</sup>. While there was a continuum of expression, CD34 expression in QSCs was non-homogeneous (Fig. 1c). We isolated QSCs with the 15% highest (H) and 15% lowest (L) CD34 expression levels (CD34<sub>H</sub> and CD34<sub>L</sub>) and subjected them to RNA-sequencing (RNA-seq)-coupled bioinformatics analysis. QSC fractions are molecularly distinct (Fig. 1d; Table S1), with 361 genes upregulated in CD34<sub>H</sub> cells (including *Dcn*, *Igfbp6*, *Notch3*, *Ryr3* and *Cd34*) and 99 genes upregulated in CD34<sub>L</sub> cells (including *Myog* [Myogenin], *Tmem8c* [*Mymk*, Myomaker], and *Actc1*). Single-cell RNA-seq of QSCs confirmed an association of *Myod1* and *Myog* transcripts with cells expressing low *Cd34* levels (Extended Data Fig. 1b). Through RNAscope, we validated the existence of two quiescent satellite-cell fractions based on *Cd34* expression in intact muscle of *Pax7<sup>CRE-ER</sup>;ROSA2<sup>tdTomato</sup>* mice (where Pax7-expressing satellite cells are permanently labelled): cells expressing high *Cd34* levels also expressed high *Notch3* levels, while those expressing low *Cd34* levels expressed high *Myog* levels (Fig. 1e). Compared to activated cells, both CD34<sub>H</sub> and CD34<sub>L</sub> satellite-cell fractions showed reduced size and no proliferative traits (Extended Data Fig. 1c), consistent with full quiescence. Thus, quiescent satellite cells exist in distinct transcriptional states *in vivo* and can be distinguished by CD34 relative expression.

Compared with freshly-isolated CD34<sub>L</sub> QSCs, CD34<sub>H</sub> QSCs had delayed activation kinetics and completion of the first-division cycle, increased dye-retention (PKH26), and displayed greater clonogenic and self-renewal capacities (Fig. 1f-i). Competitive transplantation experiments with CD34<sub>H</sub> and CD34<sub>L</sub> QSCs from GFP-expressing or tdTomato-expressing mice, revealed that CD34<sub>H</sub> cells expanded more efficiently and had a repopulation advantage over CD34<sub>L</sub> cells (Fig. 1j). CD34<sub>H</sub> QSCs also gave rise to more regenerated myofibres and self-renewing satellite cells than CD34<sub>L</sub> cells (Fig. 1k). To validate these findings, we used *Pax7<sup>CRE-ER</sup>;ROSA2<sup>6YFP</sup>* mice, in which satellite cells are permanently YFP-labelled (Fig. 2a). Immunostaining of snap-frozen resting muscle confirmed differential CD34-protein fluorescence intensity in YFP<sup>+</sup> QSCs (Extended Data Fig. 1d), while proteins corresponding to genes upregulated in CD34<sub>H</sub> cells, such as *Ryr3*, were more expressed in satellite cells with higher CD34 expression (Extended Data Fig. 1e), thus excluding influences of the isolation/sorting protocol on the CD34-based stem-cell classification. *Pax7<sup>CRE-ER</sup>;ROSA2<sup>6YFP</sup>* mice also contained a small (2%) YFP<sup>+</sup>/CD34-negative (YFP<sup>+</sup>/CD34<sub>N</sub>) satellite-cell fraction (Fig. 2a). RNA-seq/bioinformatics analyses of these three YFP-labelled subsets (YFP<sup>+</sup>/CD34<sub>H/L/N</sub>) revealed similar transcriptomic profiles in YFP<sup>+</sup>/CD34<sub>H</sub> and

$\alpha 7^{+}/CD34^{H}$  and also in  $YFP^{+}/CD34^{L}$  and  $\alpha 7^{+}/CD34^{L}$  cells (Extended Data Fig. 1f-i). *Ex vivo* and *in vivo* experiments confirmed that  $\alpha 7^{+}/CD34^{H}$  and  $YFP^{+}/CD34^{H}$  QSC fractions were functionally similar, both displaying higher stemness potential than  $\alpha 7^{+}/CD34^{L}$  and  $YFP^{+}/CD34^{L/N}$  cells (Fig. 2b,c). Finally, in *Pax7-nGFP* mice (where *Pax7* satellite cells expressed nuclear GFP (nGFP)), nGFP<sup>+</sup>/CD34<sup>H</sup> satellite cells showed higher stemness potential than nGFP<sup>+</sup>/CD34<sup>L/N</sup> cells, which was retained after muscle re-injury (Extended Data Fig. 2a-d).

We next compiled the transcriptomic data from all young mice (2-4 months old) (see below; n=15) and interrogated them for 19,074 gene-sets<sup>21</sup>. Compared with CD34<sup>L</sup> QSCs, the CD34<sup>H</sup> QSC transcriptome showed lower cellular activities, including at the myogenic, metabolic, transcriptional and translational levels (Extended Data Fig. 2e), reminiscent of genuine (or more naïve) long-term stem cells<sup>22</sup>, with high engrafting and self-renewing capacities. In contrast, CD34<sup>L</sup> QSCs are in a primed state, more prone to enter cell cycle and committed to myogenesis. Based on this functional separation, we termed the CD34<sup>H</sup> quiescent satellite-cells “genuine” (QSC<sub>genuine</sub>) and the CD34<sup>L</sup> cells “primed” (QSC<sub>primed</sub>). Transcriptomics of all CD34<sup>H</sup> and CD34<sup>L</sup> QSCs datasets determined that the QSC<sub>genuine</sub> state core signature was composed of 219 genes, and included *Thy1* (CD90), *Tek* (Tie-2), *Ryr3* and *Notch3*<sup>23,24</sup>, while the QSC<sub>primed</sub> core signature was composed of 168 genes, including *Myog*, *Tmem8c*, *Tnnt2*, and *Cdk2* (Fig. 2d,e; Table S2).

### **The genuine stem-cell state persists into advanced old age, only collapsing at geriatric age**

We next investigated the origin and development of the quiescent stem-cell populations from birth to advanced age. In neonatal mice (8 days post-birth), satellite cells were mostly CD34 negative (Fig. 3a), reflecting their highly proliferative status<sup>25-27</sup>. CD34 expression was first detected in juvenile mice (21 days post-birth), coinciding with entry of satellite cells into quiescence<sup>27</sup>, and witnessing the emergence of the QSC<sub>genuine</sub> state (Fig. 3a). In young mice (2-4 months-old), CD34 expression increased further, concurring with satellite cells in the QSC<sub>genuine</sub> state reaching full quiescence<sup>25,26</sup>. Surprisingly, the QSC<sub>genuine</sub> state persisted in old mice (18-22 months-old), despite the decline in satellite-cell numbers<sup>5,17,18,28,29</sup> (Fig. 3a; Extended Data Fig. 3a). Moreover, based on CD34 expression, the QSC<sub>genuine</sub> state was preserved even in extreme old (geriatric) age (28-30 months-old) (Fig. 3a). Neonatal satellite cells were transcriptionally close to juvenile cells and both were distant from young cells (Fig. 3b). Importantly, neonatal and juvenile satellite cells were transcriptionally closer to the QSC<sub>primed</sub> state, whereas young cells were close to the QSC<sub>genuine</sub> state (Fig. 3b). To our surprise, old satellite cells were transcriptionally closer to young QSC<sub>genuine</sub> than to young QSC<sub>primed</sub> (Fig. 3b).

We next performed RNA-seq of CD34<sup>H</sup> and CD34<sup>L</sup> cells from young, old and geriatric mice. Hierarchical clustering indicated that the QSC<sub>genuine</sub> state in old and geriatric mice is transcriptionally distant from the QSC<sub>primed</sub> state in this age range, and closer to the young QSC<sub>genuine</sub> (Fig. 3c,d). PCA analysis confirmed that the transcriptome distance between QSC<sub>genuine</sub> and QSC<sub>primed</sub> states was maintained in old mice (Fig. 3c). However, in geriatric mice, QSC<sub>genuine</sub> and QSC<sub>primed</sub> transcriptomes drifted from the corresponding states in young cells, and the transcriptional distance between them was reduced (Fig. 3d). In transplantation assays, QSC<sub>genuine</sub> from young or old mice showed a similar capacity to regenerate myofibres, which was impaired in geriatric QSC<sub>genuine</sub> (Fig. 3e). The age-dependent functional difference

between young/old and geriatric QSC<sub>genuine</sub> was reproduced at the level of the whole stem-cell population (Extended Data Fig. 3b). Thus, the QSCs that remain in old muscles are age-resistant, preserving a high regenerative potential, but this potential is no longer maintained in extreme-old, geriatric age. In contrast, the endogenous muscle regenerative capacity (in absence of transplanted cells) was already attenuated in old mice (Extended Data Fig. 3c), but was accentuated in geriatric mice, despite similar reduction in total satellite-cell number in old and geriatric mice (Extended Data Fig. 3a), as reported previously<sup>18</sup>. Notably, this endogenous regenerative decline in old mice was largely due to a numerical decline in the QSC<sub>primed</sub> state, whereas geriatric mice showed a similar numerical decline in both QSC states (Extended Data Fig. 3d).

Pathway enrichment analysis of the RNA-seq data revealed that two molecular networks identified as hallmarks of the young QSC<sub>primed</sub> state (protein translation and TCA/OxPhos, both indicative of higher cellular activities) (Extended Data Fig. 2e) were enriched in the geriatric (but not the old) QSC<sub>genuine</sub> state (Extended Data Fig. 3e). Thus, the clear demarcation between the genuine and primed stem-cell states is lost in extreme old age, due to a blurring of the transcriptome features that earlier define the young and old QSC<sub>genuine</sub> state clusters.

### **The genuine stem-cell state is maintained by FoxO signalling**

To investigate the molecular specification of the QSC<sub>genuine</sub> state, we searched for the most enriched transcription factor (TF) binding motifs in the genes belonging to QSC<sub>genuine</sub> and QSC<sub>primed</sub> core signatures. This search revealed enrichment for the occupancy of Forkhead box O (FoxO) family DNA recognition motif (Fig. 4a; Extended Data Fig. 4a). We next performed transposase-accessible chromatin high-throughput sequencing (ATAC-seq) in freshly-sorted CD34<sub>H</sub> and CD34<sub>L</sub> cells (and in the minor CD34<sub>N</sub> cell fraction; Fig. 2a) and analysed accessibility of genes defining QSC<sub>genuine</sub> and QSC<sub>primed</sub> core signatures, respectively. Despite a comparable chromatin accessibility in all quiescent populations, QSC<sub>genuine</sub> core genes were more accessible in CD34<sub>H</sub> cells, while QSC<sub>primed</sub> core genes were more accessible in CD34<sub>L/N</sub> cells (Fig. 4b). Open-chromatin peaks in the three cell subsets contained FoxO consensus motifs (Extended Data Fig. 4b). The number of FoxO motifs per peak was elevated in the promoters of QSC<sub>genuine</sub> core genes in all three QSC fractions but was the highest in CD34<sub>H</sub> cells (Extended Data Fig. 4b), thus highlighting the potential control of the genuine satellite-cell state by FoxO.

Intracellular FACS analysis revealed higher expression of FoxO TFs in the QSC<sub>genuine</sub> than QSC<sub>primed</sub> state (Fig. 4c). FoxO3a (FoxO3) showed strong nuclear localization, FoxO1 appeared in both nucleus and cytoplasm and FoxO4 was cytoplasmic (Fig. 4d; Extended Data Fig. 4c). We used the *Pax7<sup>CRE-ER</sup>* driver to specifically ablate FoxO3a in satellite cells after tamoxifen administration to *FoxO3a<sup>ΔPax7ER</sup>* mice (Extended Data Fig. 4d,e). *FoxO3a<sup>ΔPax7ER</sup>* mice showed no differences in satellite-cell number compared to *FoxO3a<sup>WT</sup>* mice, in agreement with a previous study<sup>30</sup>, or in the relative cell abundance in QSC<sub>genuine</sub> and QSC<sub>primed</sub> states (based on CD34 expression) (Extended Data Fig. 4f), consistent with lack of muscle regenerative phenotype after one injury round<sup>30</sup>. We also found no regenerative phenotype in *FoxO1<sup>ΔPax7ER</sup>* mice (Extended Data Fig. 4g). To circumvent potential compensatory effects among FoxO TFs, we generated a triple *FoxO1,3a,4*-mutant line (*FoxO1,3a,4<sup>ΔPax7ER</sup>* mice) (Extended Data Fig. 4h). Tamoxifen administration to young *FoxO1,3a,4<sup>ΔPax7ER</sup>* mice caused full disappearance of the QSC<sub>genuine</sub> state, as indicated by loss of CD34<sub>H</sub> quiescent cells after 30 days (Fig. 4e), and spontaneous satellite-cell activation and differentiation (Fig. 4f;

Extended Data Fig. 4i). These results were validated by lineage tracing in *FoxO1,3a,4<sup>ΔPax7ER</sup>;ROSA26<sup>YFP</sup>* mice (Fig. 5a; Extended Data Fig. 4j,k), as abrogation of all three FoxO genes (FoxOs) triggered loss of CD34<sup>H</sup> cells and impacted the total YFP<sup>+</sup> QSC population (30% reduction) (Fig. 5a). BrdU labelling demonstrated that these FoxO1,3a,4-deficient cells were able to incorporate BrdU and spontaneously fuse to uninjured myofibres (Fig. 5b). Furthermore, *FoxO1,3a,4<sup>ΔPax7ER</sup>* satellite cells showed reduced clonogenic capacity *ex vivo* (Fig. 5c) and severely impaired capacity to form new myofibres and self-renew upon transplantation (Fig. 5d), consistent with the loss of QSC<sub>S<sup>genuine</sup></sub>. The numerical and functional satellite-cell deficit in *FoxO1,3a,4<sup>ΔPax7ER</sup>* mice was evident in impaired muscle regeneration even 15 days post-injury (Fig. 5e).

## Genetic loss of FoxO imposes the primed-state gene programme

To investigate how FoxO signalling impacts both stem-cell states, we carried out RNA-seq on satellite cells from *FoxO1,3a,4<sup>ΔPax7ER</sup>* and *FoxO1,3a,4<sup>WT</sup>* mice 30 days post-tamoxifen treatment. Whereas 743 genes were downregulated in FoxO1,3a,4<sup>ΔPax7ER</sup> satellite cells (Fig. 6a), more genes (1000) were surprisingly upregulated (Fig. 6a), suggesting that FoxOs mediate both transcriptional activation and repression in quiescent satellite cells. PCA showed that FoxO1,3a,4<sup>ΔPax7ER</sup> satellite cells are transcriptionally closer to QSC<sub>S<sup>primed</sup></sub> and distant from QSC<sub>S<sup>genuine</sup></sub> (Fig. 6b). Many upregulated genes in FoxO1,3a,4<sup>ΔPax7ER</sup> satellite cells were also upregulated in QSC<sub>S<sup>primed</sup></sub>. Moreover, this shared gene-list was enriched in skeletal-muscle-differentiation genes, including *Myog* and *Tnnt2* among 43 QSC<sub>S<sup>primed</sup></sub> core-signature genes and mitochondrial/OxPhos genes such as *Cox8b* or *Mtx2* (Extended Data Fig. 5a). In contrast, 49 QSC<sub>S<sup>genuine</sup></sub> core-signature genes were downregulated in absence of FoxOs, including *Cd34* and *Igfbp4* (Extended Data Fig. 5b). FoxO signalling thus maintains expression of important QSC<sub>S<sup>genuine</sup></sub>-state molecules, and FoxO silencing correspondingly causes disappearance of major QSC<sub>S<sup>genuine</sup></sub> core genes and inappropriate expression of genes characteristic of the QSC<sub>S<sup>primed</sup></sub> state.

To identify early, and thus likely direct, FoxO-target genes, we performed RNA-seq of satellite cells isolated only 4 days after the first tamoxifen administration to *FoxO1,3a,4<sup>WT</sup>;ROSA26<sup>YFP</sup>* and *FoxO1,3a,4<sup>ΔPax7ER</sup>;ROSA26<sup>YFP</sup>* mice. *FoxO1,3a,4* genes were deleted at this early post-tamoxifen stage in YFP<sup>+</sup> cells (Extended Data Fig. 5c). Compared with the analysis at 30 days post-tamoxifen treatment, significantly fewer genes showed dysregulated expression at 4 days after FoxOs deletion, forming a set of candidate direct FoxO-target genes (Fig. 4a; Extended Data Fig. 5d,e). Notably, downregulation of QSC<sub>S<sup>genuine</sup></sub> and upregulation of QSC<sub>S<sup>primed</sup></sub> core-gene signatures was already evident early after FoxOs deletion, including altered expression of *Myog* and *Cd34* (Fig. 6c; Extended Data Fig. 5d,e). Furthermore, GSEA analysis revealed alteration to muscle differentiation pathways, but not mitochondria or OxPhos pathways (Extended Data Fig. 5f).

We next mapped accessible chromatin regions in FoxO1,3a,4<sup>WT</sup> and FoxO1,3a,4<sup>ΔPax7ER</sup> satellite cells by ATAC-seq. PCA revealed distinct patterns of accessibility (Fig. 6d), and accessible regions were analysed in the key genes *Cd34* and *Myog*. Peaks in the WT *Cd34* proximal promoter and in intragenic and intergenic regions were diminished or absent in FoxO1,3a,4<sup>ΔPax7ER</sup> cells, correlating with reduced *Cd34* expression (Extended Data Fig. 5g). In contrast, new peaks appeared in intragenic and intergenic regions of the *Myog* gene, correlating with increased *Myog* expression in FoxO1,3a,4<sup>ΔPax7ER</sup> cells (Extended Data Fig. 5g). Similar accessibility results were obtained for the promoters of *Igfbp4* and *Tnnt2*

(Extended Data Fig. 5g), matching the reduced and increased expression, respectively, in FoxO1,3a,4-null satellite cells (Fig. 6a; Extended Data Fig. 5g).

Analysis of the *Myog* promoter identified a putative FoxO DNA-binding site (TGTTTAC) at -62 to -49 bp from the transcription start, and ChIP confirmed promoter association of FoxO3a (Extended Data Fig. 6a). ChIP also confirmed that *Cd34* is a direct FoxO3a target (Extended Data Fig. 6a), in agreement with loss of CD34 expression in the absence of FoxOs. These results are consistent with increased and decreased acetylation of *Myog* and *Cd34* gene promoters, respectively, in FoxO1,3a,4 $\Delta$ Pax7<sup>ER</sup> cells, shown by histone H4 acetylation by ChIP (Fig. 6e). FoxOs ablation thus provokes de-repression of *Myog* and premature activation of the myogenic-differentiation gene program, with concomitant loss of *Cd34*. ChIP also demonstrated association of FoxO3a with *Igfbp4* and *Tnnt2* promoters (Extended Data Fig. 6a), matching their reduced and increased expression, respectively, in cells lacking FoxOs (Fig. 6a). Finally, delivery of short hairpin RNA (shRNA) targeting *Myog* mRNA (*Myog* shRNA) into YFP<sup>+</sup> satellite cells from *FoxO1,3a,4 $\Delta$ Pax7<sup>ER</sup>;ROSA26<sup>YFP</sup>* mice (to silence *Myog*) significantly restored their capacity to engraft and form new regenerating myofibres after *in vivo* transplantation (Fig. 6f; Extended Data Fig. 6b). Transcription from the *Myog* promoter was increased in the absence of FoxO, as shown by the induction of *Myog*-promoter-driven luciferase activity in FoxOs-silenced satellite cells, whereas FoxO3a overexpression repressed *Myog* transcription in differentiating cells (Fig. 6g; Extended Data Fig. 6c,d). Thus, FoxO signalling maintains the QSC<sub>genuine</sub> state, while its loss imposes the QSC<sub>primed</sub> gene-expression program and depletes the genuine state, with deleterious consequences on regeneration.

### **Establishment and life persistence of the genuine stem-cell state requires FoxO activity**

To investigate the relevance of FoxO signalling to the initial establishment and evolution of the QSC<sub>genuine</sub> state, we generated a new mouse line, *FoxO1,3a,4 $\Delta$ Pax7*, with constitutive deletion of the three FoxOs in the Pax7 satellite-cell lineage (Extended Data Fig. 6e). The gradual acquisition of CD34 expression and emergence of QSC<sub>genuine</sub> in WT juvenile mice (Fig. 3a) did not occur in *FoxO1,3a,4 $\Delta$ Pax7* mice (Fig. 7a); indeed, FoxO1,3a,4 $\Delta$ Pax7 satellite cells were CD34-negative (Fig. 7a) and continued to express cell-cycle and myogenic differentiation-related genes (Extended Data Fig. 6f), thus resembling QSC<sub>Sprimed</sub>. *In vivo*-BrdU labelling from postnatal day 21 to 30 detected BrdU-positive satellite cell nuclei in myofibres of *FoxO1,3a,4*-deficient but not WT mice (Extended Data Fig. 6g), demonstrating that, in absence of FoxOs, satellite cells fail to enter quiescence and establish the genuine state, and this failure persisted throughout life (Fig. 7a) and was exacerbated in old age (Fig. 7b). Thus, in absence of FoxO signalling the quiescent satellite-cell population is never established. Transcriptomic analysis showed that juvenile and old FoxO1,3a,4-deficient satellite cells failed to induce the QSC<sub>genuine</sub> core signature while showing persistent activation of the QSC<sub>primed</sub> core signature (Fig. 7c), and enrichment in QSC<sub>primed</sub>-related pathways, such as cell cycle, OxPhos and myogenic differentiation (Extended Data Fig. 7a).

We next tested for changes in FoxOs expression in the genuine and primed states in the two most distant ages. In young QSCs, FoxO3a and FoxO4 levels were higher in the genuine than primed state (Fig. 7d; Extended Data Fig. 7b); a similar trend was observed for FoxO1 (Extended Data Fig. 7b). In geriatric QSC<sub>Sgenuine</sub>, however, FoxO3a levels were lower than in their young counterparts (Fig. 7d), suggesting that reduced FoxO signalling in geriatric QSC<sub>Sgenuine</sub> may account for the stem-cell regenerative decline in extreme-old age (Fig. 3e;

Extended Data Fig. 3c,d). Furthermore, acute genetic impairment of FoxOs in *FoxO1,3a,4<sup>APax7ER</sup>;ROSA26<sup>YFP</sup>* geriatric mice did not exacerbate the already-pronounced functional deficits of geriatric satellite cells (Fig. 7e; Extended Data Fig. 7c), despite affecting CD34 levels (Fig. 7f; Extended Data Fig. 7d). In agreement with these results, no significant differential enrichment of the genuine and primed signatures was detected between WT and FoxO1,3a,4-null geriatric satellite cells in contrast to juvenile, young and old ages (Extended Data Fig. 7e; Fig. 7c; Fig. 6c). Thus, young QSCs, equipped with higher FoxO levels, are more sensitive than geriatric cells to FoxO1,3a,4 genetic depletion.

### **Computational prediction of Igf/Akt-pathway inhibition for reinforcement of the genuine stem-cell state**

We next aimed to identify signalling molecules emanating from the niche that might maintain the two quiescent stem-cell-state phenotypes throughout life, and which might impact FoxOs. For this, we implemented a computational method, NicheHotSpotter, based on a probabilistic model of sustained signal transduction, previously developed to predict hotspots (key molecules) in signalling pathways that are constantly activated/inhibited by the niche to robustly maintain distinct cellular states<sup>31,32</sup>. For this predictive modelling, we used RNA-seq data for QSC<sub>Sgenuine</sub> and QSC<sub>Sprimed</sub> from young and geriatric mice. The signalling molecules predicted to be active or inactive in young genuine or primed states are shown in Figure 8a. In young QSC<sub>Sgenuine</sub>, Akt signalling (Akt1 and Akt2) was predicted to be inactive (or inhibited) (Fig. 8a; Extended Data Fig. 8a). Akt is activated through phosphorylation<sup>33</sup> and consistent with Akt inactivation, the Akt phosphatase PPP2CA, known to dephosphorylate Akt in myogenic cells<sup>34</sup>, was predicted to be active in young QSC<sub>Sgenuine</sub> (Fig. 8a). In contrast, Akt signalling might be active in young QSC<sub>Sprimed</sub> (Fig. 8a). Higher levels of phosphorylated-Akt (p-Akt) were found in young QSC<sub>Sprimed</sub> than QSC<sub>Sgenuine</sub> (Extended Data Fig. 8b).

Consistent with Akt being inhibitory for FoxO transcriptional activity<sup>35</sup>, NicheHotSpotter-derived networks revealed that FoxOs target several differentially-expressed TFs, which are directly inhibited by Akt1 in young QSC<sub>Sgenuine</sub> (Fig. 8a; Extended Data Fig. 8a). This prediction is consistent with higher FoxO3a protein levels in young QSC<sub>Sgenuine</sub> than QSC<sub>Sprimed</sub> (Fig. 7d). Furthermore, NicheHotSpotter predicted that the signalling network centred around Akt is controlled upstream by several growth factors/receptors, including insulin (Ins)/insulin-growth-factor (Igf) and receptor (Igf1r) (Fig. 8a; Extended Data Fig. 8a), suggesting that niche-derived Ins/Igf-driven Akt signalling may inhibit FoxO in young QSC<sub>Sprimed</sub>. To test this prediction, we administered the Akt inhibitor wortmannin to young mice for 2 weeks, which resulted in reduced p-Akt and increased nuclear FoxO3a in QSC<sub>Sprimed</sub> (Extended Data Fig. 8b,c). QSC<sub>Sgenuine</sub> and QSC<sub>Sprimed</sub> from wortmannin-treated young mice both showed increased engraftment capacity *in vivo* and clonogenic capacity *in vitro* (Fig. 8b; Extended Data Fig. 9a). Furthermore, we observed enrichment of the “QSC<sub>Sgenuine</sub> signature” in CD34<sup>L</sup> QSCs from wortmannin-treated mice (Fig. 8c). Thus, Akt inhibition, by activating FoxO, reinforces the genuine stem-cell state, and this has a benefit on the poorly-functional QSC<sub>Sprimed</sub> population.

NicheHotSpotter predicted the Igf/Igfr/Akt pathway to be more active in geriatric than young QSC<sub>Sgenuine</sub> (Fig. 8d; Extended Fig. 9b), consistent with the stronger FoxO inhibition in geriatric satellite cells (Fig. 7d). The method also predicted that activation of the Igf/Ins receptors by extrinsic ligands would have negative consequences on young satellite cells related to Akt/FoxO (Fig. 8a; Extended Data Fig. 8a). To test this prediction, we used transgenic mice that overexpress a locally-secreted form of IGF1 (Tg-IGF1) in skeletal

muscle<sup>36</sup>, and found decreased nuclear FoxO3a (Extended Data Fig. 9c), correlating with increased p-Akt, which was reflected in altered satellite-cell clonogenic potential (Extended Data Fig. 10a,b), thus partially phenocopying FoxO1,3a,4-deficient satellite cells. Consistently, CD34<sup>H</sup> QSCs from Tg-IGF1 mice showed partial enrichment of the “QSC<sub>primed</sub> signature” (Extended Data Fig. 10c). Thus, Akt-signalling inhibition and FoxO-nuclear relocation enhance stemness traits of young QSC<sub>primed</sub> and may constitute a promising strategy to rejuvenate geriatric QSC<sub>genuine</sub> that have acquired primed-state traits.

## Discussion

Despite studies showing quiescent muscle stem-cell heterogeneity<sup>5,7,8,37</sup>, the demonstration of this heterogeneity in unperturbed skeletal muscle, the dynamics of heterogeneity throughout life and its requirement during tissue repair remain unclear. Through orthogonal approaches, including RNAScope, we demonstrate the existence of two stem-cell states in undisturbed muscle tissue. We further provide a high-resolution reconstruction of the molecular events that maintain functional muscle stem-cell heterogeneity in quiescence and that will determine the generation of diverse progeny upon stress and how these properties change over time.

Stem cell quiescence is progressively lost with ageing due to alterations affecting the intrinsic and niche factors<sup>5,17,18</sup>. Through the relative expression of CD34, we show that quiescent satellite cells (QSCs) encompass a genuine state (QSC<sub>genuine</sub>), with inferred stemness features and low metabolism, and a primed state (QSC<sub>primed</sub>), more prone to myogenic differentiation and metabolic activity. Opposing the concept of a general malfunctioning of aged stem cells, our study surprisingly reveals that the genuine stem-cell state is age-resistant, since old QSC<sub>genuine</sub> perform as well as young counterparts *in vitro* and during muscle regeneration. Of note, acquisition of terminal cell fate in yeast also appears to be determined quite early in life<sup>38</sup>. Nonetheless, in extreme-old age, the genuine stem-cell state undergoes a steep functional decline. Molecular features of the QSC<sub>primed</sub> gene-signature emerge in the geriatric, but not the old, QSC<sub>genuine</sub> state, indicating that the identity of the young QSC<sub>genuine</sub> state prevails in old age, but becomes blurred in extreme-old age. Indeed, although geriatric QSC<sub>genuine</sub> maintain CD34 expression, they concomitantly upregulate genes involved in OxPhos metabolism. The acquisition of these inappropriate traits specifically in geriatric satellite cells likely explains the previously described sharp muscle-regeneration decline in extreme old age<sup>17,18</sup>.

Our results show that FoxO TFs act as a sentinel of the genuine-stem-cell state (Extended Data Fig. 10d). Thus, the separation between QSC<sub>genuine</sub> and QSC<sub>primed</sub> states is disrupted by FoxOs deletion. Consistent with this, FoxOs-deficient satellite cells largely phenocopied QSCs in the primed state, exemplified by the shared high expression of the master-regulator *Myog*, and myogenesis-related genes, and downregulation of classical stemness- and quiescence-related genes such as *Notch3* and *Calcr*<sup>14,39-41</sup>. We therefore propose that long-lived-quiescent stem cells may rely on FoxOs to preserve diversity and stemness, thus extending FoxO functions beyond cellular differentiation<sup>42,43</sup>.

It is unclear why QSC<sub>genuine</sub> are more responsive to FoxO signalling. The newly implemented computational method, NicheHotSpotter, predicted that niche-produced Igf and Akt signalling regulate the primed-state by inhibiting FoxO activity. This was confirmed by the conversion of QSC<sub>primed</sub> into QSC<sub>genuine</sub> upon Akt blockade, while IGF1 overexpression in the muscle niche causes Akt activation and FoxO inhibition and deteriorates the genuine-stem-cell state. The enhanced activity of the Akt inhibitory pathway could therefore underly the reduced FoxO activity of geriatric QSC<sub>genuine</sub>, and provoke acquisition of primed features by QSC<sub>genuine</sub> in this extreme-old age. From conceptual and technical standpoints, despite reports suggesting the non-equivalence of muscle stem-cell populations<sup>5-9,44</sup>, it is worth mentioning that no previous study has achieved the separation of distinct quiescent states in the absence of transgenesis or labelling, likely due to lack of cell-surface markers. Our study shows that CD34 reports stemness in quiescence.

In conclusion, our findings set forth the notion that quiescent stem cells in low-turnover tissues are already pre-programmed to adopt divergent future fates in response to regenerative demands. This molecular partitioning becomes less demarcated in geriatric age, with detrimental functional consequences. We anticipate that understanding this molecular diversity and how it changes throughout life will be critical to harnessing the potential of stem cells for regenerative medicine in sarcopenia.

## References

- 1 Hwang, A. B. & Brack, A. S. Muscle Stem Cells and Aging. *Curr Top Dev Biol* **126**, 299-322, doi:10.1016/bs.ctdb.2017.08.008 (2018).
- 2 Sousa-Victor, P., Garcia-Prat, L., Serrano, A. L., Perdiguero, E. & Munoz-Canoves, P. Muscle stem cell aging: regulation and rejuvenation. *Trends Endocrinol Metab* **26**, 287-296, doi:10.1016/j.tem.2015.03.006 (2015).
- 3 Brack, A. S. & Rando, T. A. Tissue-specific stem cells: lessons from the skeletal muscle satellite cell. *Cell Stem Cell* **10**, 504-514, doi:10.1016/j.stem.2012.04.001 (2012).
- 4 Feige, P., Brun, C. E., Ritso, M. & Rudnicki, M. A. Orienting Muscle Stem Cells for Regeneration in Homeostasis, Aging, and Disease. *Cell Stem Cell* **23**, 653-664, doi:10.1016/j.stem.2018.10.006 (2018).
- 5 Chakkalakal, J. V., Jones, K. M., Basson, M. A. & Brack, A. S. The aged niche disrupts muscle stem cell quiescence. *Nature* **490**, 355-360, doi:10.1038/nature11438 (2012).
- 6 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).
- 7 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).
- 8 Kuang, S., Kuroda, K., Le Grand, F. & Rudnicki, M. A. Asymmetric self-renewal and commitment of satellite stem cells in muscle. *Cell* **129**, 999-1010, doi:10.1016/j.cell.2007.03.044 (2007).
- 9 Rocheteau, P., Gayraud-Morel, B., Siegl-Cachedenier, I., Blasco, M. A. & Tajbakhsh, S. A subpopulation of adult skeletal muscle stem cells retains all template DNA strands after cell division. *Cell* **148**, 112-125, doi:10.1016/j.cell.2011.11.049 (2012).
- 10 Scaramozza, A. *et al.* Lineage Tracing Reveals a Subset of Reserve Muscle Stem Cells Capable of Clonal Expansion under Stress. *Cell Stem Cell*, doi:10.1016/j.stem.2019.03.020 (2019).
- 11 Der Vartanian, A. *et al.* PAX3 Confers Functional Heterogeneity in Skeletal Muscle Stem Cell Responses to Environmental Stress. *Cell Stem Cell*, doi:10.1016/j.stem.2019.03.019 (2019).
- 12 De Micheli, A. J. *et al.* Single-Cell Analysis of the Muscle Stem Cell Hierarchy Identifies Heterotypic Communication Signals Involved in Skeletal Muscle Regeneration. *Cell Rep* **30**, 3583-3595 e3585, doi:10.1016/j.celrep.2020.02.067 (2020).
- 13 Toro-Dominguez, D. *et al.* ImaGEO: integrative gene expression meta-analysis from GEO database. *Bioinformatics* **35**, 880-882, doi:10.1093/bioinformatics/bty721 (2019).

- 14 Sidney, L. E., Branch, M. J., Dunphy, S. E., Dua, H. S. & Hopkinson, A. Concise review: evidence for CD34 as a common marker for diverse progenitors. *Stem Cells* **32**, 1380-1389, doi:10.1002/stem.1661 (2014).
- 15 Beauchamp, J. R. *et al.* Expression of CD34 and Myf5 defines the majority of quiescent adult skeletal muscle satellite cells. *J Cell Biol* **151**, 1221-1234, doi:10.1083/jcb.151.6.1221 (2000).
- 16 van Velthoven, C. T. J., de Morree, A., Egner, I. M., Brett, J. O. & Rando, T. A. Transcriptional Profiling of Quiescent Muscle Stem Cells In Vivo. *Cell Rep* **21**, 1994-2004, doi:10.1016/j.celrep.2017.10.037 (2017).
- 17 Garcia-Prat, L. *et al.* Autophagy maintains stemness by preventing senescence. *Nature* **529**, 37-42, doi:10.1038/nature16187 (2016).
- 18 Sousa-Victor, P. *et al.* Geriatric muscle stem cells switch reversible quiescence into senescence. *Nature* **506**, 316-321, doi:10.1038/nature13013 (2014).
- 19 Alfaro, L. A. *et al.* CD34 promotes satellite cell motility and entry into proliferation to facilitate efficient skeletal muscle regeneration. *Stem Cells* **29**, 2030-2041, doi:10.1002/stem.759 (2011).
- 20 Lee, J. Y. *et al.* Clonal isolation of muscle-derived cells capable of enhancing muscle regeneration and bone healing. *J Cell Biol* **150**, 1085-1100, doi:10.1083/jcb.150.5.1085 (2000).
- 21 Reimand, J. *et al.* Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc* **14**, 482-517, doi:10.1038/s41596-018-0103-9 (2019).
- 22 Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487-492, doi:10.1016/j.stem.2009.05.015 (2009).
- 23 Abou-Khalil, R. *et al.* Autocrine and paracrine angiopoietin 1/Tie-2 signaling promotes muscle satellite cell self-renewal. *Cell Stem Cell* **5**, 298-309, doi:10.1016/j.stem.2009.06.001 (2009).
- 24 Kitamoto, T. & Hanaoka, K. Notch3 null mutation in mice causes muscle hyperplasia by repetitive muscle regeneration. *Stem Cells* **28**, 2205-2216, doi:10.1002/stem.547 (2010).
- 25 Evano, B. & Tajbakhsh, S. Skeletal muscle stem cells in comfort and stress. *NPJ Regen Med* **3**, 24, doi:10.1038/s41536-018-0062-3 (2018).
- 26 Mourikis, P. & Tajbakhsh, S. Distinct contextual roles for Notch signalling in skeletal muscle stem cells. *BMC Dev Biol* **14**, 2, doi:10.1186/1471-213X-14-2 (2014).
- 27 White, R. B., Bierinx, A. S., Gnocchi, V. F. & Zammit, P. S. Dynamics of muscle fibre growth during postnatal mouse development. *BMC Dev Biol* **10**, 21, doi:10.1186/1471-213X-10-21 (2010).
- 28 Day, K., Shefer, G., Shearer, A. & Yablonka-Reuveni, Z. The depletion of skeletal muscle satellite cells with age is concomitant with reduced capacity of single progenitors to produce reserve progeny. *Dev Biol* **340**, 330-343, doi:10.1016/j.ydbio.2010.01.006 (2010).
- 29 Liu, W. *et al.* Loss of adult skeletal muscle stem cells drives age-related neuromuscular junction degeneration. *Elife* **6**, doi:10.7554/eLife.26464 (2017).
- 30 Gopinath, S. D., Webb, A. E., Brunet, A. & Rando, T. A. FOXO3 promotes quiescence in adult muscle stem cells during the process of self-renewal. *Stem Cell Reports* **2**, 414-426, doi:10.1016/j.stemcr.2014.02.002 (2014).
- 31 Kalamakis, G. *et al.* Quiescence Modulates Stem Cell Maintenance and Regenerative Capacity in the Aging Brain. *Cell* **176**, 1407-1419 e1414, doi:10.1016/j.cell.2019.01.040 (2019).

- 32 Ravichandran, S., Hartmann, A. & Del Sol, A. SigHotSpotter: scRNA-seq-based computational tool to control cell subpopulation phenotypes for cellular rejuvenation strategies. *Bioinformatics*, doi:10.1093/bioinformatics/btz827 (2019).
- 33 Alessi, D. R. *et al.* Mechanism of activation of protein kinase B by insulin and IGF-1. *EMBO J* **15**, 6541-6551 (1996).
- 34 Yun, B. G. & Matts, R. L. Hsp90 functions to balance the phosphorylation state of Akt during C2C12 myoblast differentiation. *Cell Signal* **17**, 1477-1485, doi:10.1016/j.cellsig.2005.03.006 (2005).
- 35 Stitt, T. N. *et al.* The IGF-1/PI3K/Akt pathway prevents expression of muscle atrophy-induced ubiquitin ligases by inhibiting FOXO transcription factors. *Mol Cell* **14**, 395-403, doi:10.1016/s1097-2765(04)00211-4 (2004).
- 36 Ascenzi, F. *et al.* Effects of IGF-1 isoforms on muscle growth and sarcopenia. *Aging Cell* **18**, e12954, doi:10.1111/accel.12954 (2019).
- 37 Dell'Orso, S. *et al.* Correction: Single cell analysis of adult mouse skeletal muscle stem cells in homeostatic and regenerative conditions (doi: 10.1242/dev.174177). *Development* **146**, doi:10.1242/dev.181743 (2019).
- 38 Li, Y. *et al.* A programmable fate decision landscape underlies single-cell aging in yeast. *Science* **369**, 325-329, doi:10.1126/science.aax9552 (2020).
- 39 Baghdadi, M. B. *et al.* Reciprocal signalling by Notch-Collagen V-CALCR retains muscle stem cells in their niche. *Nature* **557**, 714-718, doi:10.1038/s41586-018-0144-9 (2018).
- 40 Bjornson, C. R. *et al.* Notch signaling is necessary to maintain quiescence in adult muscle stem cells. *Stem Cells* **30**, 232-242, doi:10.1002/stem.773 (2012).
- 41 Mourikis, P. *et al.* A critical requirement for notch signaling in maintenance of the quiescent skeletal muscle stem cell state. *Stem Cells* **30**, 243-252, doi:10.1002/stem.775 (2012).
- 42 Webb, A. E. *et al.* FOXO3 shares common targets with ASCL1 genome-wide and inhibits ASCL1-dependent neurogenesis. *Cell Rep* **4**, 477-491, doi:10.1016/j.celrep.2013.06.035 (2013).
- 43 Wilhelm, K. *et al.* FOXO1 couples metabolic activity and growth state in the vascular endothelium. *Nature* **529**, 216-220, doi:10.1038/nature16498 (2016).
- 44 Tierney, M. T., Stec, M. J., Rulands, S., Simons, B. D. & Sacco, A. Muscle Stem Cells Exhibit Distinct Clonal Dynamics in Response to Tissue Repair and Homeostatic Aging. *Cell Stem Cell* **22**, 119-127 e113, doi:10.1016/j.stem.2017.11.009 (2018).

## Acknowledgements

We are indebted to Drs. C. Keller and M. Capecchi for Pax-Cre mouse lines, A. Webb and A. Brunet for FoxO3-lentivirus; E. Andrés, J. Segalés, V. Lukesova, A. Martín, A. Pollán, A. Navarro, J.M. Ballesteros and M. Raya for their technical contributions; Drs. A. Serrano, J. Isern and PMC lab members for helpful discussions; J. Martín-Caballero and S. Rodríguez-Colilla (PRBB/CNIC Animal Facilities), O. Fornas and M. Montoya (UPF/CRG/CNIC FACS Facilities), A. Dopazo (CNIC-Genomics Facility), J. Rodríguez (qGenomics), Leica Microsystems for IIBM-imaging-platform support, S. Bartlett for editing, and W. Link (IIB-CSIC) for FoxO-activation advise. Authors acknowledge funding from MINECO-Spain (RTI2018-096068), ERC-2016-AdG-741966, LaCaixa-HEALTH-HR17-00040, MDA,

UPGRADE-H2020-825825, AFM and DPP-Spain to PMC; María-de-Maeztu-Program for Units of Excellence to UPF (MDM-2014-0370), Severo-Ochoa-Program for Centers of Excellence to CNIC (SEV-2015-0505). Work was also supported by NIAMS IRP through NIH grants AR041126 and AR041164 to V.S. and utilized computational resources of NIH HPC Biowulf cluster (<http://hpc.nih.gov>); by ASI, Ricerca Finalizzata, Ateneo Sapienza to A.M.; by AIRC (23257), ASI (MARS-PRE, DC-VUM-2017-006); H2020-MSCA-RISE-2014 (645648) to M.S; and FNR CORE grant (C15/BM/10397420) to AdS. L.G.P. was partially supported by FPI fellowship and EMBO Fellowship (ALTF 420-2017), and S.C, X.H and V.M. by FI, Severo-Ochoa and PFI Fellowships (Spain), respectively.

## Main Figures

### Fig. 1. CD34 identifies quiescent satellite-cell states with distinct engraftment and self-renewal capabilities.

**a**, Venn-diagram showing cell-surface-proteins/receptors in the “Quiescence Core” gene-list (Fig. S1A; Table S1). *Calcr* and *Cd34* are included. *Cd34* is highlighted. **b**, CD34 Mean-Fluorescence Intensity (MFI) by flow-cytometry of freshly-sorted satellite cells from WT mice in basal conditions and 3-days post-cardiotoxin injection (3d CTX) (2-4 mo, n=4 mice). **c**, FACS to isolate CD34<sup>High</sup> and CD34<sup>Low</sup> (CD34<sup>H</sup> and CD34<sup>L</sup>) subpopulations within  $\alpha$ 7-integrin<sup>+</sup>/CD34<sup>+</sup> quiescent satellite cells (QSCs) from young mice. **d**, Volcano-plot with differentially-expressed genes by RNA-seq between CD34<sup>H</sup> and CD34<sup>L</sup> from mice (2-3 mo, n=4 samples, 3-mice pooled/sample). **e**, Dual-RNAscope for *Myog* (red)/*Cd34* (white) or *Notch3* (red)/*Cd34* (white) probes in resting muscles from *Pax7<sup>CRE-ER</sup>;ROSA26<sup>tdTomato</sup>* mice. Satellite cells stained with anti-RFP antibody (green) and nuclei with DAPI. Quantification of *Cd34/Myog* puncta or *Cd34/Notch3* puncta in each cell from mice (2-4 mo, n=5 mice, 222 cells) and representative pictures. Scale bar, 5 $\mu$ m. **f**, Videomicroscopy of time (h) to first-division (activation time) of CD34<sup>H</sup> and CD34<sup>L</sup> QSCs (CD34<sup>H</sup> n=77 and CD34<sup>L</sup> n=60 cells) from mice (2-4 mo, n=2 mice). Mann-Whitney test. **g**, As in (f), PKH26-dye retention; PKH26-fluorescence-intensity after 24 h in culture (2-4 mo, n=2 mice, Bulk n=429, CD34<sup>H</sup> n=210 and CD34<sup>L</sup> n=115 cells). **h**, CD34<sup>H</sup> and CD34<sup>L</sup> QSCs clonogenic-assay, number of colonies and cells (2-4 mo, n=16 mice, 4-12 replicates/mouse). **i**, CD34<sup>H</sup> and CD34<sup>L</sup> QSCs self-renewal *in vitro*. Cells were expanded 3-days in GM and cultured for 72-96 h in DM to obtain ‘reserve QSC’ (self-renewal; Pax7<sup>+</sup>Ki67<sup>-</sup>-cells from total DAPI<sup>+</sup> cells) (2-4 mo, n=4 mice, 2 replicates/mouse). Paired t-test. **j**, Representative image and quantification of CD34<sup>H</sup> and CD34<sup>L</sup> QSCs from *ROSA26-tdTomato* and  *$\beta$ -actin-GFP* mice (2-4 mo, n=3 mice, 9 transplanted muscles) transplanted into pre-injured tibialis anterior (TA) of recipient-immunodeficient mice for 3-days. Equal cell numbers were used. Scale bar, 75 $\mu$ m. **k**, CD34<sup>H</sup> and CD34<sup>L</sup> QSCs from mice (2-4 mo, n=10 mice), GFP-lentivirus transduced and transplanted for 21-days into pre-injured-TA of immunodeficient mice. Equal numbers of cells were used. Representative images and quantification of GFP<sup>+</sup>-fibres (16-20 transplanted muscles) and self-renewal (GFP<sup>+</sup>Pax7<sup>+</sup>-cells) (7-9 transplanted muscles). Scale bar, 100 $\mu$ m. Means  $\pm$  s.e.m.; two-tailed unpaired t-test with Welch’s correction unless otherwise indicated.

### Fig. 2. CD34 identifies distinct quiescent stem-cell states in genetically-marked satellite cells.

**a**, Sorting strategy used to identify and isolate CD34<sup>H</sup>, CD34<sup>L</sup> and CD34<sup>N</sup> subpopulations within YFP<sup>+</sup> QSCs from young *Pax7<sup>CRE-ER</sup>;ROSA26<sup>YFP</sup>* mice. **b**, Clonogenic assay of freshly-sorted CD34<sup>H</sup>, CD34<sup>L</sup> and CD34<sup>N</sup> satellite cells from *Pax7<sup>CRE-ER</sup>;ROSA26<sup>YFP</sup>* mice, number of colonies and cells were quantified (2-4 mo, n=13 mice, 4-8 replicates/mouse). Paired t-test. **c**, Equal number of freshly-isolated quiescent CD34<sup>H</sup>, CD34<sup>L</sup> and CD34<sup>N</sup> satellite cells from young *Pax7<sup>CRE-ER</sup>;ROSA26<sup>YFP</sup>* mice (2-4 mo, n=20 mice) were transplanted for 21 days into pre-injured TA muscle of recipient immunodeficient mice. Representative images and quantification of YFP<sup>+</sup>-fibres (13-24 transplanted muscles) and self-renewal (YFP<sup>+</sup>Pax7<sup>+</sup>-cells) (3-6 transplanted muscles). Scale bar, 100  $\mu$ m. **d**, Volcano plot showing the most differentially expressed genes identified by RNA-seq between CD34<sup>H</sup> and CD34<sup>L</sup> from all young WT mice (2-4 mo, n=15 mice). **e**, Heatmap illustrating differential gene expression profiles between CD34<sup>H</sup> and CD34<sup>L</sup> QSCs from all young WT mice (2-4 mo, n=15 mice)

constituting the genuine and primed signatures (Table S2). Means  $\pm$  s.e.m.; two-tailed unpaired t-test with Welch's correction unless otherwise indicated.

**Fig. 3. Analysis of the quiescent stem-cell states from birth to old age.**

**a**, Representative flow cytometry plots and quantification of CD34 MFI (bottom) of  $\alpha$ 7-integrin<sup>+</sup> cells in resting muscle of WT mice at different ages: neonatal (N) (8d, n=5 mice), juvenile (J) (21d, n=5 mice), young (Y) (2-3 mo, n=39 mice), old (O) (22-24 mo, n=33 mice) and geriatric (G) (>28 mo, n=17 mice). **b**, PCA analysis of the full transcriptome of QSCs ( $\alpha$ 7-integrin<sup>+</sup>CD34<sup>+</sup>) freshly isolated from neonatal (8d, n=3 mice), juvenile (21d, n=3 mice), young (2-4 mo, n=4 mice) and old (22-24 mo, n=5 mice) WT mice and QSC<sub>genuine</sub> and QSC<sub>primed</sub> freshly isolated from young WT mice (2-3 mo, n=3 mice) (Bottom). Heatmap illustrating the expression of differentially-expressed genes between QSC<sub>genuine</sub> and QSC<sub>primed</sub> freshly isolated from young WT mice at different ages (Top). **c**, PCA analysis of the full transcriptome of QSC<sub>genuine</sub> and QSC<sub>primed</sub> freshly isolated from young (2-3 mo, n=4 mice) and old (22 mo, genuine n=3 mice and primed n=4 mice) WT mice (Bottom). Heatmap shows the expression of differentially expressed genes between QSC<sub>genuine</sub> and QSC<sub>primed</sub> freshly isolated from WT mice at different ages (Top). **d**, As in (c), from young (2-3 mo, n=4 mice) and geriatric (>28 mo, n=3 mice) WT mice. **e**, QSC<sub>genuine</sub> and QSC<sub>primed</sub> isolated from young (2-4 mo, n=6 mice), old (22-24 mo, n=8 mice) and geriatric (>28 mo, n=8 mice) WT mice were transduced with a GFP-lentivirus and transplanted for 21 days into pre-injured TA muscle of recipient immunodeficient mice. Equal number of freshly isolated cells were used. Representative images and quantification of GFP<sup>+</sup> fibres are shown (4-12 transplanted muscles). Scale bar, 100  $\mu$ m. Means  $\pm$  s.e.m.; two-tailed unpaired t-test with Welch's correction unless otherwise indicated.

**Fig. 4. The QSC<sub>genuine</sub> state is maintained by the FoxO family of transcription factors.**

**a**, Transcription factor binding sites enrichment analysis in QSC<sub>genuine</sub> and QSC<sub>primed</sub> gene signatures, using position weight matrices from the whole UCSC genome browser (through Enrichr). FoxO binding site enrichment is highlighted in red. **b**, ATAC-seq libraries were generated from YFP<sup>+</sup>CD34<sup>H</sup>, CD34<sup>L</sup> and CD34<sup>N</sup> QSCs isolated from young *Pax7<sup>CRE-ER</sup>; ROSA26<sup>YFP</sup>* mice (2-3 mo, n=6 mice). Tag density plots around the TSS for genes involved in the QSC<sub>genuine</sub> and QSC<sub>primed</sub> signatures are shown. One sample t-test *p*-value for QSC<sub>genuine</sub> genes: CD34<sup>H</sup> vs CD34<sup>L</sup>, 1.87-51; CD34<sup>H</sup> vs CD34<sup>N</sup>, 3.87-77; CD34<sup>L</sup> vs CD34<sup>N</sup>, 4.25-42. One sample t-test *p*-value for QSC<sub>primed</sub> genes: CD34<sup>H</sup> vs CD34<sup>L</sup>, 1.38-09; CD34<sup>H</sup> vs CD34<sup>N</sup>, 2.02-26; CD34<sup>L</sup> vs CD34<sup>N</sup>, 1.37-17. **c**, Quantification of the FoxO1, 3a and 4 MFI analyzed by flow cytometry of freshly-isolated QSC<sub>genuine</sub> and QSC<sub>primed</sub> (2-4 mo, n=5 mice). **d**, Representative immunofluorescence images of FoxO1, FoxO3a and FoxO4 in QSCs freshly isolated from WT mice (2-4 mo). Scale bar, 10  $\mu$ m. This experiment was repeated 3 times. **e**, Representative flow cytometry analysis and quantification of the percentage of  $\alpha$ 7-integrin<sup>+</sup>CD34<sup>+</sup> satellite cells within the alive cell population from *FoxO1,3a,4<sup>WT</sup>* and *FoxO1,3a,4<sup>APax7<sup>ER</sup></sup>* mice (2-4 mo, n=3 mice) after 30 days from the last tamoxifen injection. **f**, Representative immunofluorescence images and quantification of MyoD<sup>+</sup> and Myogenin<sup>+</sup> satellite cells on muscle sections from *FoxO1,3a,4<sup>WT</sup>* and *FoxO1,3a,4<sup>APax7<sup>ER</sup></sup>* mice (2-4 mo, n=4 mice) after 7 days from the last tamoxifen injection. Scale bar, 50  $\mu$ m and 10  $\mu$ m (magnification). Means  $\pm$  s.e.m.; two-tailed unpaired t-test with Welch's correction unless otherwise indicated.

**Fig. 5. Loss of FoxO signalling causes spontaneous activation and reduced regenerative potential of satellite cells.**

**a**, (Left) Quantification of CD34 MFI analysed by flow cytometry (2-4 mo, n=4 mice). (Right) Representative immunofluorescence images and quantification of YFP<sup>+</sup> satellite cells from *FoxO1,3a,4<sup>WT</sup>;ROSA26<sup>YFP</sup>* and *FoxO1,3a,4<sup>ΔPax7ER</sup>;ROSA26<sup>YFP</sup>* mice after 30 days of tamoxifen administration (2-4 mo, n=3-6 mice) by flow cytometry. Scale bar, 10 μm. **b**, Representative immunofluorescence images and quantification of BrdU<sup>+</sup> nuclei in single myofibres isolated from young *FoxO1,3a,4<sup>WT</sup>* and *FoxO1,3a,4<sup>ΔPax7ER</sup>* mice (2-4 mo, n=3 mice) treated with BrdU in drinking water for 1 month after tamoxifen administration. Scale bar, 20 μm. **c**, Clonogenic assay of freshly-sorted satellite cells from *FoxO1,3a,4<sup>WT</sup>* and *FoxO1,3a,4<sup>ΔPax7ER</sup>* mice. Number of colonies and cells were quantified (2-4 mo, n=10 mice, 8 replicates/mouse). **d**, Equal numbers of freshly-isolated QSC from *FoxO1,3a,4<sup>WT</sup>;ROSA26<sup>YFP</sup>* and *FoxO1,3a,4<sup>ΔPax7ER</sup>;ROSA26<sup>YFP</sup>* mice (2-4 mo, n=3-5 mice) were transplanted for 21 days into pre-injured TA muscle of recipient immunodeficient mice. Representative immunofluorescence images and quantification of GFP<sup>+</sup> fibres and Pax7<sup>+</sup>/GFP<sup>+</sup> self-renewing satellite cells are shown (4 transplanted muscles). Scale bar, 100 μm. **e**, (Left) Representative images of eMHC immunofluorescence in cryosections of regenerating TA muscle from *FoxO1,3a,4<sup>WT</sup>* and *FoxO1,3a,4<sup>ΔPax7ER</sup>* mice, 15 days after CTX-induced injury. Scale bar, 75 μm. (Right) Self-renewal analysis by Pax7 immunostaining on muscle sections was quantified (2-3 mo, n=3 mice). Means ± s.e.m.; two-tailed unpaired t-test with Welch's correction unless otherwise indicated.

**Fig. 6. FoxO prevents inappropriate expression of myogenic differentiation genes in the QSC<sub>genuine</sub> state.**

**a**, Volcano plot showing the most differentially expressed genes identified by RNA-seq between *FoxO1,3a,4<sup>WT</sup>* and *FoxO1,3a,4<sup>ΔPax7ER</sup>* mice (2-4 mo, n=3 mice) after 30 days from the last tamoxifen injection. **b**, RNA-seq PCA analysis comparing data from (a) to CD34<sup>H</sup> and CD34<sup>L</sup> satellite cells from WT mice. Average distance analysis between sets is shown (right). **c**, GSEA analysis of the QSC<sub>genuine</sub> (top) and QSC<sub>primed</sub> (bottom) signatures in the RNAseq data of satellite cells from WT and *FoxO1,3a,4<sup>ΔPax7ER</sup>* mice at 4 (2-4 mo, n=2 mice) and 30 days (2-3 mo, n=3 mice) after the last tamoxifen injection. **d**, ATAC-seq PCA analysis of normalized peak tags in promoters of satellite cells from *FoxO1,3a,4<sup>WT</sup>* and *FoxO1,3a,4<sup>ΔPax7ER</sup>* mice (2-4 mo, n=4 mice) after 30 days from the last tamoxifen injection. **e**, Quantification of H4-Acetylation (H4Ac) enrichment at the *Cd34* and *Myog* promoters by chromatin immunoprecipitation (ChIP) assays of satellite cells freshly isolated from *FoxO1,3a,4<sup>WT</sup>* and *FoxO1,3a,4<sup>ΔPax7ER</sup>* mice (2-4 mo, n=3 mice) after 30 days from the last tamoxifen injection. **f**, Equal number of QSCs freshly isolated from *FoxO1,3a,4<sup>WT</sup>;ROSA26<sup>YFP</sup>* and *FoxO1,3a,4<sup>ΔPax7ER</sup>;ROSA26<sup>YFP</sup>* mice (2-4 mo, n=3-5 mice) were transfected for 3 h with siRNA-Myog or siRNA-scramble (as control) and transplanted for 21 days into injured TA muscle of recipient immunodeficient mice. Representative images and quantification of GFP<sup>+</sup> fibres are shown (3-7 transplanted muscles). Scale bar, 50 μm. **g**, Satellite cells were transfected with a Myogenin-Luciferase plasmid (p-Myog-Luc) (or pLuc as control), in the absence or presence of a FoxO3a expression plasmid or a GFP empty plasmid, in GM or DM. Luciferase activities are expressed relative to the activity found for basal pMyog-Luc in GM (2-4 mo, n=3-8 mice). Means ± s.e.m.; two-tailed unpaired t-test with Welch's correction unless otherwise indicated.

**Fig. 7. Early establishment and life persistence of the genuine stem-cell state requires FoxO activity.**

**a**, Representative flow cytometry plots and quantification of CD34 MFI of  $\alpha7$ -integrin<sup>+</sup> cells from WT and *FoxO1,3a,4 $\Delta$ Pax7* mice at different ages: neonatal (N) (8d, n=4-5 mice), juvenile (J) (21d, n=3-6 mice), young (Y) (2-3 mo, n=4 mice) and old (O) (22-24 mo, n=3 mice). **b**, Percentage of satellite cells ( $\alpha7$ -integrin<sup>+</sup> for neonatal and  $\alpha7$ -integrin<sup>+</sup>/CD34<sup>+</sup> for other ages) within Lin<sup>-</sup> cells quantified by flow cytometry from WT and *FoxO1,3a,4 $\Delta$ Pax7* mice at different ages: neonatal (N) (8d, n=3 mice), juvenile (J) (21d, n=3-6 mice), young (Y) (2-3 mo, n=4 mice) and old (O) (22-24 mo, n=3-5 mice). **c**, GSEA analysis of the QSC<sub>genuine</sub> (left) and QSC<sub>primed</sub> (right) signatures in the RNA-seq data of satellite cells from WT and *FoxO1,3a,4 $\Delta$ Pax7* mice at different ages. **d**, Quantification of the FoxO3a MFI analysed by flow cytometry of freshly isolated QSC<sub>genuine</sub> and QSC<sub>primed</sub> from young (2-4 mo, n=4-5 mice) and geriatric (>28mo, n=5 mice) mice. **e**, Clonogenic assay of freshly sorted satellite cells from WT and *FoxO1,3a,4 $\Delta$ Pax<sup>ER</sup>7* mice at young and geriatric ages. Number of colonies and cells were quantified (5-6 replicates/mouse, n=3 mice). **f**, Quantification of the percentage of  $\alpha7$ -integrin<sup>+</sup>CD34<sup>+</sup> satellite cells within the YFP<sup>+</sup> population from geriatric (>28mo, n=3 mice) *FoxO1,3a,4<sup>WT</sup>;ROSA26<sup>YFP</sup>* and *FoxO1,3a,4 $\Delta$ Pax<sup>7ER</sup>;ROSA26<sup>YFP</sup>* mice after tamoxifen administration, and quantification of CD34 MFI analysed by flow cytometry. Means  $\pm$  s.e.m.; two-tailed unpaired t-test with Welch's correction unless otherwise indicated.

**Fig. 8. A Markov-chain-based computational method predicts reinforcement and rejuvenation of the genuine state by Igf/Akt pathway inhibition.**

**a**, Summary of steps involved in NicheHotSpotter for the prediction of niche-induced signalling hotspots. A simplified representation of NicheHotSpotter derived Akt centric network that controls FoxO and other TFs (see extended Fig. 9a for full network). Predictions of active (compatibility score > 0.5, in green) and inactive (compatibility score < 0.5, in red) signalling hotspots for young QSC<sub>genuine</sub> in comparison to QSC<sub>primed</sub>. **b**, Equal numbers of QSC<sub>genuine</sub> and QSC<sub>primed</sub> isolated from young mice (4 mo, n=4 mice) previously treated with wortmannin or vehicle for 2 weeks (see Methods for details) were stained with DIL and transplanted for 5 days into pre-injured TA muscle of recipient immunodeficient mice. Quantification of DIL<sup>+</sup> cells/mm<sup>2</sup> and representative immunostaining images are shown (12 transplanted muscles). Scale bar, 50  $\mu$ m (bottom), 10  $\mu$ m (top). **c**, GSEA analysis showing significant (FDR q-value<0.05) enrichment of the "QSC<sub>primed</sub> signature" in CD34<sup>L</sup> (QSC<sub>primed</sub>) satellite cells isolated from WT mice treated with vehicle (CTRL) (4 mo, n=5 mice) compared to CD34<sup>L</sup> (QSC<sub>primed</sub>) satellite cells isolated from WT mice treated with wortmannin (+Wort) (4 mo, n=4 mice). **d**, As in (a), predictions of signalling hotspots for young QSC<sub>genuine</sub> in comparison to geriatric QSC<sub>genuine</sub>. Means  $\pm$  s.e.m.; two-tailed unpaired t-test with Welch's correction.

## **Author contributions**

L.G.-P. and P.M.-C. conceived the study, conceptualized ideas, and wrote the manuscript. L.G.-P., E.P. and S.A.-M. designed and performed most experiments, helped in interpreting results and revising/editing the manuscript. S.D'O., A.H.J., and L.G.P. performed RNA-seq, ChIP and ATAC-seq analyses. S.R.B., K.J., and H-W.S. designed bioinformatics pipelines and performed bioinformatics analyses for RNA-seq and ATAC-seq. L.G.-P. and E.P. partially contributed to RNA-seq bioinformatics. S.R. and A.d.S. designed and generated the NicheHotSpotter model and performed computation analysis of RNA-seq data. S.C. performed RNAscope analysis. V.M performed single-cell RNA-seq. L.O., V.R.-B., M.F., X.H., and M.J. performed some cell/mouse experiments and/or provided technical support. E.R. provided technical support in microscopy. A.M. and M.S. provided essential mouse lines. L.G.-P. and E.P. prepared the final figures. V.S. and P.M.-C. discussed and interpreted results, and supervised, managed and obtained funding for the study.

## SUPPLEMENTARY MATERIALS

### METHODS

#### Animals

Male mice C57BL/6J (wild-type, WT), *ROSA26-tdTomato*<sup>1</sup>, *β-actin-GFP*, *Pax7-nGFP*<sup>2</sup>, *Pax7<sup>CRE-ER</sup>;ROSA26<sup>YFP</sup>*, *Pax7<sup>CRE-ER</sup>;ROSA26<sup>Tomato</sup>* the offspring of intercrossing *FoxO3a<sup>fl/fl</sup>* or *FoxO1,3a4<sup>fl/fl</sup>* with *Pax7<sup>Cre</sup>*, *Pax7<sup>CreER</sup>*, *ROSA26<sup>YF</sup>* or *Tg-IGF1 (MLC-IGF1)*<sup>2</sup> lines were used at indicated ages. Mice with *FoxO3a* or *FoxO1,3a,4* genes deleted in satellite cells (SCs), inducible or constitutive were generated by breeding *FoxO3a<sup>fl/fl</sup>* or *FoxO1,3a,4<sup>fl/fl</sup>* (previously described in<sup>3</sup> with *Pax7<sup>Cre</sup>* or *Pax7<sup>CreER</sup>* lines (provided by C.Keller and M.Capecchi). *FoxO3a<sup>ΔPax7ER</sup>* and *FoxO1,3a,4<sup>ΔPax7ER</sup>* mice were intercrossed with *ROSA26<sup>YFP</sup>* mice to generate *FoxO3a<sup>ΔPax7ER</sup>;ROSA26<sup>YFP</sup>*. Fox Chase SCID *CB17/Icr-Prkdc<sup>scid</sup>/IcrIcoCrl* (Strain Code: 236) mice from Charles River. Mouse experiments were not randomized and not blinded. No statistical methods were used to predetermine sample size. Cre activity was induced by intraperitoneal injection (1-injection/day, 4 days) with 2mg/30g bodyweight tamoxifen (Sigma; 20mg/mL in corn-oil). Animal experiments were approved by the Catalan/Madrid Government using sex-, age- and weight-matched littermate animals.

#### In vivo treatments

BrdU treatment for 1 month in drinking-water (Sigma; 1mg/mL, 2.5% sucrose) every 3-days. Wortmannin (*Penicillium fumiculosum*), administrated intraperitoneally (3-injections/week, 2-weeks) with 0.03mg/30g body weight (Sigma; 1 mg/mL in DMSO), and DMSO (10% in PBS) in control-animals.

#### Muscle regeneration

Mice were anaesthetized with ketamine/xylazine (80/10 mg kg<sup>-1</sup>, intraperitoneally) or isoflurane. Muscle regeneration was induced by injection of cardiotoxin (CTX, Latoxan; 10μM) in the *Tibialis anterior* (TA) as described<sup>4</sup>. At indicated times post-injury, mice were sacrificed and muscles were dissected, frozen in liquid nitrogen-cooled isopentane, and stored at -80 °C. For GFP immunostaining, muscle was prefixed 2 h in PFA 2% at 4°C, embedded in sucrose 15% O/N at 4°C, and frozen in liquid-nitrogen cooled-isopentane.

#### Satellite cell (SC) isolation by FACS

Muscles were mechanically disaggregated and dissociated in Ham's F10 media containing Liberase (0.1mg/g muscle-weight) (Roche; 5mg/mL) and Dispase 0.3% at 37 °C for 2h and then filtered. Cells were incubated in lysis buffer (BD Pharm Lyse) for 10min on ice, re-suspended in PBS with 2% goat serum (FACS buffer) and counted. PE-Cy7-conjugated anti-CD45 (Biolegend 103114), anti-Sca-1 (Biolegend 108114) and PE-Cy7-conjugated anti-CD31 (Biolegend 102418) antibodies were used for lineage-negative selection, and Alexa Fluor 647-conjugated anti-CD34 (BD Pharmigen 560230) and PE-conjugated anti-α7-integrin (AbLab AB10STMW215) were used for double-positive staining of quiescent satellite cells (QSCs). Cells were sorted using a FACS Aria II (BD). QSCs were used either for RNA extraction or cultured in Ham's F10 with 20% FBS and bFGF (0.0025 μg mL<sup>-1</sup>) (GM) for proliferation assays, or plated in glass (Thermo Scientific) or plastic (μ-Slide Angiogenesis ibiTreat: Ibi) slides for immunostaining.

#### Single myofibre isolation

*Extensor digitorum longus* (EDL) were dissected and digested in Collagenase type I (Sigma-Aldrich) solution for 1.5h as described<sup>5</sup>, manually rocked before individual fibres were

harvested with a fire-polished Pasteur pipet, pre-flushed with 0.5% horse serum, and fixed for immunofluorescence.

### **Satellite cell transplantation**

FACS-isolated SCs were collected, re-suspended in GM and plated in 6-well plates coated with Collagen (3mg/mL) (40-50,000 cells/well) overnight as in<sup>6,7</sup>, and transferred to a 6-well plate coated with Retronectin (Takara); lentiviral particles were added, the plate centrifuged at 50g for 5 min and left overnight. Cells were recovered the day after, centrifuged at 900 g for 15 min and resuspended (100,000-cells/100 $\mu$ L), and injected with a Hamilton syringe into recipient mice using 10,000 cells (10  $\mu$ L/muscle).

In siRNA experiments, SCs were transfected with siRNA for 3h on ice and transplanted as indicated for GFP-labelling. For transfections, Lipofectamine RNAi Max kit (Invitrogen) was used, with 20nM non-targeting siRNA control-pool or SMART mouse siMyog-pool oligomers (ThermoScientific Dharmacon<sup>TM</sup>).

For short-time engraftments, avoiding culture effect on freshly-isolated satellite cells, cells were stained with 1/200 Vybrant<sup>TM</sup> Dil cell-labelling in FACS-buffer (Invitrogen). When less than 3,000 satellite cells, 20,000 C2C12 myoblasts (Supplemental Table S3) were added. TA from recipient *CB17/Icr-Prkdc<sup>scid</sup>/IcrIcoCrl* mice were injured by freeze-crush one-day before transplantation. For second-injury experiments, we used *Pax7-nGFP* mice. Different CD34-sorted subpopulations were grafted into pre-injured TAs; 21-days later we induced a second injury, and muscle was recovered after 21-days. 3 days (satellite-cell expansion) or 21 days (muscle regeneration) after satellite-cell transplantations, muscles were processed for histology. Results were expressed as number of labelled fibres per muscle-section or SC number per mm<sup>2</sup> damaged area.

### **Clonogenic assays**

SCs were plated on collagen-coated dishes at low density for clonal analysis (50-75 cells/well in 96-well plates), and cultured in proliferation medium: DMEM Glutamax containing 20% FBS, 10% HS, 1% penicillin–streptomycin, 1% glutamine, 1% HEPES, 1% sodium pyruvate, and bFGF (0.005  $\mu$ g mL<sup>-1</sup>) for three-days. siRNA was transfected for 3 h for transplantation experiments, and plated. Cells were finally fixed, 10 min in PFA 4%, and washed with PBS. Cells in well (4-8 wells/sample) were quantified. A colony is considered when  $\geq 3$  cells are together.

### **Reserve cells**

SCs were plated on collagen-coated dishes at a density of 5-8,000 cells/cm<sup>2</sup>, cultured in GM for three-days, then: i) in experiments with CD34<sup>H</sup>, CD34<sup>L</sup> and CD34<sup>N</sup> QSCs, medium was switched into differentiation medium (DM, 5% HS) for four-five extra days; ii) WT and FoxO1,3a,4 $\Delta$ <sup>Pax7ER</sup> QSCs, cells were plated at 12-13,000 cells/cm<sup>2</sup> for 24 h and then switched into DM for four-five extra days. siRNA transfection occurred for 3 h in sorted SCs. PAX7/KI67 immunostaining was performed, and PAX7<sup>+</sup>KI67<sup>-</sup> cells counted in five-random fields/well.

### **RT-qPCR: RNA extraction, cDNA synthesis and PCR**

Total RNA from FACS-isolated SCs was obtained using RNeasy Micro kit (Qiagen) and analysed by RT-qPCR. DNase digestion was performed with Qiagen. Complementary DNA (cDNA) was synthesized from 100 ng total RNA using First-Strand cDNA Synthesis kit (Roche). Real-time PCR was performed on LightCycler 480 System using LightCycler 480 SYBR Green I Master reaction mix (Roche Diagnostic) and specific primers (Supplemental Table S3). Thermocycling conditions were: initial step of 10 min at 95 °C, 50 cycles of 15 s

denaturation at 94 °C, 10 s annealing at 60 °C and 15 s extension at 72 °C. Reactions were run in triplicate, and automatically-detected threshold cycle (Ct) values compared between samples. Transcripts of *L7*, *Tbp* and *Hprt* were used as housekeeping genes, with each unknown sample normalized to their content.

### **Chromatin Immunoprecipitation (ChIP)**

ChIPs for FOXO3a were performed from 5x10<sup>6</sup> cells/experiment. SCs were expanded in GM, and incubated in 20 μM LY294002 (Calbiochem) for 2h. Cells were crosslinked with 1% formaldehyde for 10 min at room-temperature, and crosslinking stopped with 0.125 M glycine for 5 min. Cross-linked cells were washed twice with cold PBS, resuspended in 600 μL of SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM TrisHCl, pH 8.1, plus protease inhibitors) and incubated on ice for 10 min. Cell lysate was sonicated to shear chromatin to an average length of 200-500 bp using a Diagenode Bioruptor. Sonicated-samples were centrifuged at 13,000 rpm for 10 min and chromatin concentration was quantified. 70 μg of chromatin were diluted 10-fold in ChIP-dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 167 mM NaCl, 16.7 mM Tris-HCl, pH 8.1, plus protease inhibitors) and immunoprecipitated overnight at 4 °C with 10 μg of FOXO3A antibody (Supplemental Table X) or 10 μg of rabbit IgG (Supplemental Table S3). 20 μL of Magna ChIP Protein A + G Magnetic Beads (Millipore) were added and incubated for 1 h at 4 °C. Beads were washed with 1 mL of different buffers: low-salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8.1, 150 mM NaCl), high-salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8.1, 500 mM NaCl), LiCl immune complex wash buffer (0.25 M LiCl, 1% NP40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris-HCl, pH 8.1), and TE (1 mM EDTA, 10 mM Tris-HCl, pH 8.0). To elute immunocomplexes, beads were incubated twice in elution buffer (1% SDS, 100 mM NaHCO<sub>3</sub>) for 15 min at room-temperature. The crosslinking was reverted by incubating DNA at 65 °C overnight. DNA was phenol/chloroform-extracted, ethanol-precipitated and resuspended in 80 μL H<sub>2</sub>O for analysis. Bound fraction and input were analysed by qPCR using specific primers.

ChIPs for PanH4Ac were performed from 105 freshly-sorted SCs from *FoxO1,3a,4<sup>wT</sup>* (n=5) and *FoxO1,3a,4<sup>ΔPax7ER</sup>* (n=4) mice. Cells were cross-linked in 1% formaldehyde for 10 min, quenched by adding glycine (0.125 M) afterwards, lysed in RIPA buffer (1x PBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) and centrifuged at 2,000 rpm for 5 min. The chromatin fraction was sheared by sonication and collected in 1.5 mL siliconized Eppendorf tubes. The sheared chromatin samples were cleared for 1-h and immunoprecipitated overnight with 5 μg Histone H4ac (Supplemental Table S3) or 5 μg rabbit IgG. Immunoprecipitated-fragments were collected by magnetic proteinG dynabeads (Invitrogen), washed with the following buffers: buffer I (20 mM TrisHCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), buffer II (20 mM TrisHCl pH 8.0, 500 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), buffer III (10 mM TrisHCl pH 8.0, 250 mM LiCl, 1% NP-40; 1% sodium deoxycholate, 1 mM EDTA) and Tris-EDTA (pH 8.0). All washes were performed at 4 °C for 5 min. Cross-linking was reversed in elution buffer (100 mM sodium bicarbonate (NaHCO<sub>3</sub>), 1% SDS) at 65 °C overnight. Used primers are listed in Supplemental Table S3.

### **Luciferase assay**

pMyog-Luc plasmid containing Myogenin promoter followed by luciferase gene was kindly provided by Dr. S.Tapscott. SCs from *FoxO1,3a,4<sup>ΔPax7ER</sup>* mice were transduced with Adv-GFP or Adv-Cre (Vector Biolabs) 1x10<sup>10</sup> pfu/mL (1:1500) for 48 h. Cells were transfected with 0.25μg reporter DNA using Lipofectamine 3000 (Invitrogen). As control, cells were co-transfected with 0.2 μg β-Galactosidase plasmid.

SCs were transduced with LV-GFP or LV-FoxO3a (1:500 from a titre 3-10<sup>10</sup> VP/mL). Doxycycline was added the next day (1:1000 to final 2 µg/mL) for 48 h. Then, cells were transfected as described above. Following transfection, cells were cultured in GM for 24 h, or 24 h in GM and 48 h in DM; cells were harvested and luciferase activity determined with luciferase assay kit (Promega), and normalized to β-galactosidase activity. For each experimental group, a minimum of four independent transfections (in triplicate) were performed.

### **Western blotting**

Satellite-cell lysates and Western blotting was performed as in<sup>9</sup>. Used antibodies are listed in Supplemental Table S3.

### **Histology and immunohistochemistry in muscle cryosections**

TA and EDL were frozen in liquid-nitrogen-cooled isopentane and stored at -80 °C. 8-10 µm sections were stained with haematoxylin/eosin (HE) or immunostained with mouse monoclonal primary antibodies using tperoxidase or fluorescein M.O.M kit staining (Vector Laboratories). Double-immunostaining was performed by sequential addition of each primary and secondary antibody using positive and negative controls. Sections were air-dried, fixed on PFA 2-4%, PBS-washed and incubated with primary antibodies (Supplemental Table S3) after blocking for 1 h at room temperature with a high protein-containing PBS solution (Vector Laboratories). Subsequently, the slides were PBS-washed and incubated with appropriate secondary antibodies and labelling dyes. For immunofluorescence, secondary antibodies were coupled (Supplemental Table S3) and nuclei stained with DAPI (Invitrogen). Tissue sections were mounted with Fluoromount G (SouthernBiotech).

### **Immunofluorescence**

Immunofluorescence in isolated SCs was performed in glass slides (Thermo Scientific), while in single myofibres it was performed in suspension, as in <sup>10</sup> (Supplemental Table S3).

### **RNAscope multiplex fluorescence *in situ* hybridization**

*In situ* hybridization was performed using the RNAscope Multiplex Fluorescent Reagent kit v2 (Advanced Cell Diagnostics) following manufacturer's protocol with some modifications. TA from *Pax7<sup>CRE-ER</sup>; ROSA26<sup>tdTomato</sup>* mice was dissected and fixed in 2% PFA for 2 h at 4 °C, frozen and processed into 10 µm longitudinal-sections. Slices were digested with Protease Plus. Probe hybridization was performed by incubation of mRNA target probes for 2 h at 40 °C. Slices were hybridized with two probe-sets of dual probes using *Cd34-C3* as common probe in each set and *Myog-C2* or *Notch3-C2* as companion probes. *Cd34-C3* probe was conjugated to Opal 690 fluorophore while *Myog-C2* and *Notch3-C2* were conjugated to Opal 570 fluorophore (Akoya Biosystems). For SC immunofluorescence, sections were blocked in 3% BSA for 1 h at room-temperature prior to incubation overnight with anti-RFP primary antibody (Rockland, 600-401-379, 1:200) at 4 °C. Anti-rabbit Alexa Fluor 488 secondary antibody (A11008, Life Technologies, 1:500) and DAPI (Invitrogen) were added for 1 h at room-temperature. After washing, sections were mounted with Prolong Gold Antifade Mountant (Thermo Fisher). Confocal images were taken with Zeiss LSM-700 confocal-system with a Plan-Apochromat 40 × /1.4 NA oil objective.

### **Videomicroscopy**

Between 5,000-10,000 freshly-sorted SCs were plated on 8-well glass slides (Lab-Tek) coated with collagen and recorded for 96 h. A cell was considered to have divided when cytokinesis

was completed. With same settings, cells were labelled with the red-fluorescent membrane dye PKH26 (Sigma).

### **Digital image acquisition**

Digital images were acquired using: (1) an upright microscope DMR6000B (Leica) with DFC300FX camera for immunohistochemical colour pictures and a Hamamatsu ORCA-ER camera for immunofluorescence pictures; (2) a Thunder imager 3D live-cell microscope (Leica-Microsystems) with AFC (hardware autofocus control) and a Leica DFC9000 GTC sCMOS camera, using HC PL FLUOTAR 10x/0.32 PH1  $\infty$ /0.17/ON257C and HC PL FLUOTAR 20x/0.4 CORR PH1  $\infty$ /0-2/ON25/C objectives; (3) confocal images of muscle sections or SCs were taken using a Leica SP5 confocal laser-scanning microscope with HCX PL Fluotar 40x/0.75 and 63x/0.75 objectives or Nikon ECLIPSE Ti-TimeLapse with Apo  $\lambda$  20x/0.75 and Fluor ELWD 40x/0.6 Ph2 DM objectives. The different fluorophores (3 to 4) were excited using the 405, 488, 568 and 633 excitation-lines; (4) videomicroscopy was taken using Zeiss Cell Observer HS with 20x Air objective and Zeiss AxioCam MrX camera. Acquisition was performed using Leica Application or LAS X software (Leica) or Zeiss LSM software Zen.

### **Digital image processing and automated counting**

Image processing and quantification was carried out in Fiji<sup>11</sup> using macro language to create these pipelines: i) For myofibre size, individual fibres were automatically segmented using a machine-learning approach (trainable weka segmentation) complemented by a manual-correction macro; cross-sectional area (CSA) was measured on the resulting mask, and ii) Transplanted-cells labelled with GFP or Dil were automatically segmented, and fluorescence intensity of selected proteins or PKH26 dye quantified for each cell; the average of relative fluorescence was expressed by mean or sum of fluorescence intensities. The macros and their detailed explanation are available at: <https://github.com/MolecularImagingPlatformIBMB>.

### **RNA-sequencing and ATAC-sequencing sample preparation**

Total RNA was isolated from either FACS-isolated SCs using RNeasy Micro kit (Qiagen) for regular RNAseq or low input RNAseq. For low-cell number (100-400), samples were collected directly into either 96-conic well-plates or 200  $\mu$ l-PCR conic-Eppendorf, containing 10uL lysis buffer (SMART-Seq<sup>TM</sup> v4 Ultra<sup>TM</sup> Low Input RNA Kit for Sequencing), spun-down, and frozen at -80 °C before processing.

RNA-seq of CD34<sup>H</sup>, CD34<sup>L</sup> SCs from young WT mice, or CD34<sup>H</sup>, CD34<sup>L</sup>, CD34<sup>N</sup> satellite cells from young *Pax7<sup>CreER</sup>;ROSA26<sup>YFP</sup>* mice, or satellite cells from *FoxO1,3a,4<sup>WT</sup>* and *FoxO1,3a,4 <sup>$\Delta$ Pax7<sup>ER</sup></sup>* mice were completed on a HiSeq2000 instrument (Illumina) using cDNA libraries from poly A+ purified mRNA and sequenced using 50 bp single-end reads.

For RNA-seq of CD34<sup>H</sup>, CD34<sup>L</sup> or CD34<sup>N</sup> satellite cells, 0.2-0.4 ng total RNA (or 400 cells) were used to amplify the cDNA with SMART-Seq v4 Ultra Low Input RNA Kit (Clontech-Takara). 1-ng amplified cDNA was used to generate barcoded-libraries with Nextera XT DNA-library preparation kit (Illumina). cDNA was fragmented, adapters added in a single-reaction, amplified and cleaned-up. Libraries' size was checked using Agilent 2100 Bioanalyzer High-Sensitivity DNA chip and the concentration determined with Qubit<sup>®</sup> fluorometer (ThermoFisher Scientific). Libraries were sequenced on a HiSeq2500 (Illumina) to generate 61-bases single-end reads. Single-cell RNAseq from mononucleated cells from young muscle was performed as described<sup>12</sup>. For ATAC-seq, freshly-sorted SCs were processed as described<sup>13</sup>.

### **RNA-sequencing statistical analysis**

1. For all samples (except those in point 2), from the demultiplexed, single-end read fastq files, the following steps were sequentially performed to generate matrices of read counts: 1) Adapter sequences were trimmed with TrimGalore; 2) Trimmed reads were mapped to the mm10 *Mus musculus* reference genome with TopHat (TopHat calls the mapper Bowtie). The following options were selected: --library-type fr-firststrand for stranded libraries (dUTP protocol) and --library-type fr-unstranded for unstranded libraries (low input RNA-seq); 3) Filter out reads with MQ<3; 4) Data-quality check: all samples were initially checked for raw number of sequenced reads, duplication levels and number of final mapped reads. Duplicate reads were not removed for downstream analysis; 5) Count reads was performed with HTSeq; 6) A matrix of counts was generated for each experiment, with one row for every feature (genes) and one column for every sample. This data structure was used as input for the subsequent data analysis in R. R pipeline for statistical analysis. Each experiment was analysed with the R package Deseq2<sup>14</sup>. This package assumes a negative binomial distribution to generate lists of differentially expressed (DE) genes. For each experiment, a matrix containing all the experiment's relevant information (variables, potential batches, etc.) was constructed. A model was defined for each experiment based on the experimental design to accept or reject the alternative hypothesis. Before running the statistical tests, all experimental data was subjected to an exploratory analysis. To this end, variance across all median values within samples was stabilized with the rlog transformation (regularized log transformation, returns log<sub>2</sub> values, normalized for library size and in a way that minimizes variation in genes with small/large number of counts). Rlog transformed data was the subject to PCA (Principal Component Analysis) and/or clustering to visualize sample-to-sample distances. Samples were also visualized before/after the removal of known batch effects through the application of the Limma's removeBatchEffect function and PCA/clustering visualization. In all cases, to avoid the unwanted effects of potentially unknown batch effects, all experiments were subject to surrogate variable analysis (SVA). In the event surrogate variables (SV) were identified, those were added to the deseq2 model and data were once more visualized through PCA/clustering analysis before/after surrogate variable removal. If the addition of SV into the model resulted in better sample clustering/separation, SVs were kept in the model for statistical testing. Lists of DE genes were generated through the results function in Deseq2. DE genes were corrected for multiple testing (an adjusted *p*-value is generated through the Benjamini & Hochberg correction). Gene counts for DE genes were visualized with the pheatmap function in R.

2. For satellite cells isolated from young WT and *Tg-IGF* mice; and young WT treated with DMSO (control) or Wortmannin for two weeks: mRNA expression-profiling was performed in NIAMS Genome Core Facility at NIH. Illumina Hiseq 3000 runs were demultiplexed and converted to FastQ format using bcl2fastq2 (Illumina). FastQ reads were mapped to mm10 using Tophat 2.1.1. Expression values (RPKM) and fold-changes were calculated and analyzed with the Partek Genomics Suite (<http://www.partek.com>). Analysis of variance was performed on Log<sub>2</sub> RPKM (with a 0.1 offset).

### **Pathway enrichment analysis**

Pathway +-enrichment analysis was performed using the following protocol<sup>15</sup>. A score to rank genes from top up-regulated to down-regulated was calculated using the formula  $-\text{sign}(\log\text{FC}) * -\log_{10}(\text{pvalue})$ . The rank file from each comparison was used in Gene Set Enrichment analysis (GSEA) analysis<sup>16</sup> with 2,000 permutations and default parameters against a pathway database containing Msigdb c2 and c3, NCI, IOB, NetPath, HumanCyc, GO BP, Reactome, WikiPathways, and Panther (<http://baderlab.org/GeneSets>, version June 2018).

EnrichmentMap version 2.1.0 in Cytoscape 3.4.0 was used to visualize enriched gene-sets at FDR  $\leq$  0.05 with a Jaccard coefficient set to 0.25.

### **Transcription-factor binding analysis**

Comprehensive gene-set enrichment analysis for transcriptional machinery was carried out using Enrichr<sup>17</sup> library of published position weight matrices (PWMs) from UCSC genome browser. Z-score was used to represent the TF lists, defined as the likelihood that the number of TFBS nucleotides detected for the included target genes/sequences is significant as compared with the number of TFBS nucleotides detected for the background set. Z-score was expressed in units of magnitude of the standard deviation. Alternatively, transcription factor targets were interrogated in GSEA<sup>16</sup> using the TFT library included in MSigDB database v6.2.

### **ATAC-sequencing analysis**

Libraries were sequenced on a HiSeq2500 (Illumina) to generate 50-bases paired-end reads. A published ATAC-Seq data-analysis method<sup>18</sup> was followed to generate ATAC-Seq peaks from fastq files. Briefly, redundant paired-end (PE) reads were removed using fastuniq<sup>19</sup>. PE reads of 50 bases were aligned to the mouse genome build mm10 with Bowtie 1.1.1, following the guidelines presented by<sup>58</sup>. Customized python scripts were used to calculate fragment length of each pair of uniquely mapped PE reads and to identify reads from fragments of less than 175 bases, which were then used for peak-calling with MACS1.4.2<sup>20</sup>. HOMER<sup>21</sup> and BEDtools<sup>22</sup> were used to generate raw and peak bigWig files. The read density profile around TSS was made using ngs.plot<sup>23</sup>.

### **Computational analysis by NicheHotSpotter**

NicheHotSpotter is based on a method previously developed to identify hotspots in signalling pathways from single-cell RNAseq data<sup>31,32</sup>. Here, we have extended the approach to be applicable to bulk RNA-sequencing by considering that steady-state protein levels are more likely to have high correlation with the respective steady-state mRNA levels. Importantly, we focused on the effect of signalling molecules via TFs that are not differentially expressed but controlled by post-translational modifications to exert their regulatory effect on downstream differentially expressed TFs. Detailed description of the method is available elsewhere<sup>32</sup>. Briefly, the method involves two-main steps (1) Markov chain model-based identification of signalling molecules that exhibit the high probability of signal flow through them and (2) topological characterization based assessment of compatibility with differential expression status of the downstream transcription factors (TFs), to infer the activity status of such molecules. The differentially-expressed TFs were based on the analyses that defined signatures of CD34 cells. Those molecules that exhibit high-signalling probability and also high (or low) compatibility are identified as active (or inactive) signalling hotspots. To build the signalling networks around the predicted hotspots, we traced all shortest paths from the Niche node to the predicted hotspot (Akt1) and from the hotspot to the downstream differentially-expressed TFs. This integrated network (Extended Data Fig. 8-9) of signalling and TF molecules serves as model to understand the potential mechanisms of how niche cues could regulate downstream TFs to stably maintain the cellular phenotype. NicheHotSpotter is available at:

<https://gitlab.com/srikanth.ravichandran/signalingfactorscd34positive>.

### **Statistical analysis**

For mouse experiments, no specific blinding method was used but mice in each sample-group were selected randomly. The sample size (n) of each experimental group is described in each corresponding Figure legend, and all experiments were repeated at least with three biological

replicates unless otherwise indicated. GraphPad Prism software was used for all statistical analyses except for sequencing-data analysis. Quantitative data displayed as histograms are expressed as means  $\pm$  standard error of the mean (represented as error bars). Results from each group were averaged and used to calculate descriptive statistics. Two-tailed unpaired t test with Welch's correction was used for pairwise comparisons among groups unless otherwise indicated. Statistical significance was set at a *p*-value  $<0.05$ .

### Data availability

RNA-seq and ATAC-seq data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE155642. All other data supporting the findings of this study are available from the corresponding author on request.

### Code availability

Codes used for immunostaining analysis are available at: <https://github.com/MolecularImagingPlatformIBMB>. Codes for RNA-seq and ATAC-seq analysis (GSE155642) are described in previous sections and available from the corresponding author on request.

### References

- 1 Alvarez, S. *et al.* Replication stress caused by low MCM expression limits fetal erythropoiesis and hematopoietic stem cell functionality. *Nat Commun* **6**, 8548, doi:10.1038/ncomms9548 (2015).
- 2 Sambasivan, R. *et al.* Distinct regulatory cascades govern extraocular and pharyngeal arch muscle progenitor cell fates. *Dev Cell* **16**, 810-821, doi:10.1016/j.devcel.2009.05.008 (2009).
- 3 Paik, J. H. *et al.* FoxOs are lineage-restricted redundant tumor suppressors and regulate endothelial cell homeostasis. *Cell* **128**, 309-323, doi:10.1016/j.cell.2006.12.029 (2007).
- 4 Suelves, M. *et al.* uPA deficiency exacerbates muscular dystrophy in MDX mice. *J Cell Biol* **178**, 1039-1051, doi:10.1083/jcb.200705127 (2007).
- 5 Moyle, L. A. & Zammit, P. S. Isolation, culture and immunostaining of skeletal muscle fibres to study myogenic progression in satellite cells. *Methods Mol Biol* **1210**, 63-78, doi:10.1007/978-1-4939-1435-7\_6 (2014).
- 6 Sousa-Victor, P. *et al.* Geriatric muscle stem cells switch reversible quiescence into senescence. *Nature* **506**, 316-321, doi:10.1038/nature13013 (2014).
- 7 Sacco, A. *et al.* Short telomeres and stem cell exhaustion model Duchenne muscular dystrophy in mdx/mTR mice. *Cell* **143**, 1059-1071, doi:10.1016/j.cell.2010.11.039 (2010).
- 8 Weintraub, H. *et al.* The myoD gene family: nodal point during specification of the muscle cell lineage. *Science* **251**, 761-766, doi:10.1126/science.1846704 (1991).
- 9 Perdiguero, E., Ruiz-Bonilla, V., Serrano, A. L. & Munoz-Canoves, P. Genetic deficiency of p38alpha reveals its critical role in myoblast cell cycle exit: the p38alpha-JNK connection. *Cell Cycle* **6**, 1298-1303 (2007).

- 10 Garcia-Prat, L., Munoz-Canoves, P. & Martinez-Vicente, M. Monitoring Autophagy in Muscle Stem Cells. *Methods Mol Biol* **1556**, 255-280, doi:10.1007/978-1-4939-6771-1\_14 (2017).
- 11 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).
- 12 De Micheli, A. J. *et al.* Single-Cell Analysis of the Muscle Stem Cell Hierarchy Identifies Heterotypic Communication Signals Involved in Skeletal Muscle Regeneration. *Cell Rep* **30**, 3583-3595 e3585, doi:10.1016/j.celrep.2020.02.067 (2020).
- 13 Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**, 1213-1218, doi:10.1038/nmeth.2688 (2013).
- 14 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
- 15 Reimand, J. *et al.* Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc* **14**, 482-517, doi:10.1038/s41596-018-0103-9 (2019).
- 16 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).
- 17 Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90-97, doi:10.1093/nar/gkw377 (2016).
- 18 Shih, H. Y. *et al.* Developmental Acquisition of Regulomes Underlies Innate Lymphoid Cell Functionality. *Cell* **165**, 1120-1133, doi:10.1016/j.cell.2016.04.029 (2016).
- 19 Xu, H. *et al.* FastUniq: a fast de novo duplicates removal tool for paired short reads. *PLoS One* **7**, e52249, doi:10.1371/journal.pone.0052249 (2012).
- 20 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).
- 21 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589, doi:10.1016/j.molcel.2010.05.004 (2010).
- 22 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).
- 23 Shen, L., Shao, N., Liu, X. & Nestler, E. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* **15**, 284, doi:10.1186/1471-2164-15-284 (2014).

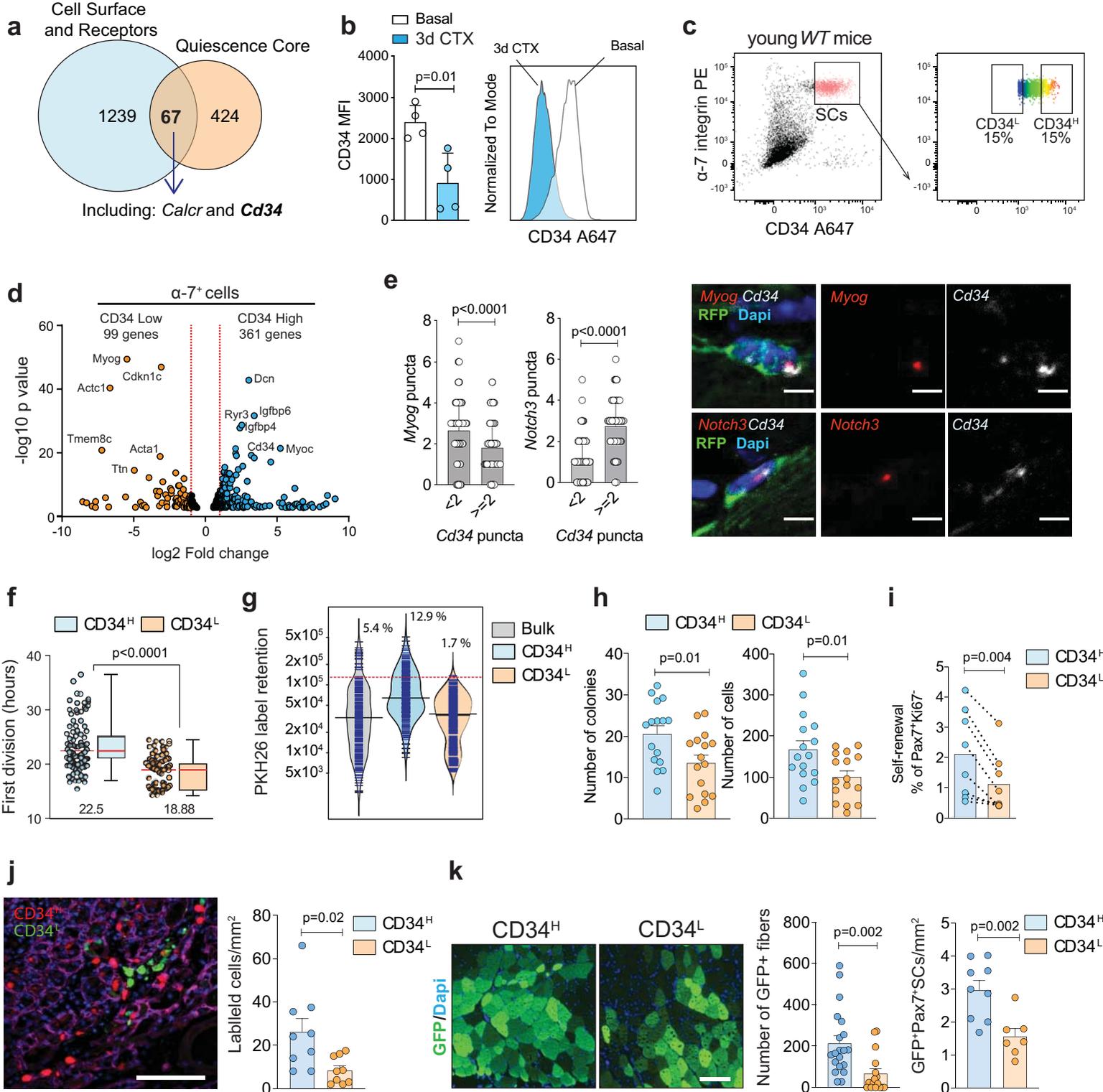


Figure 1

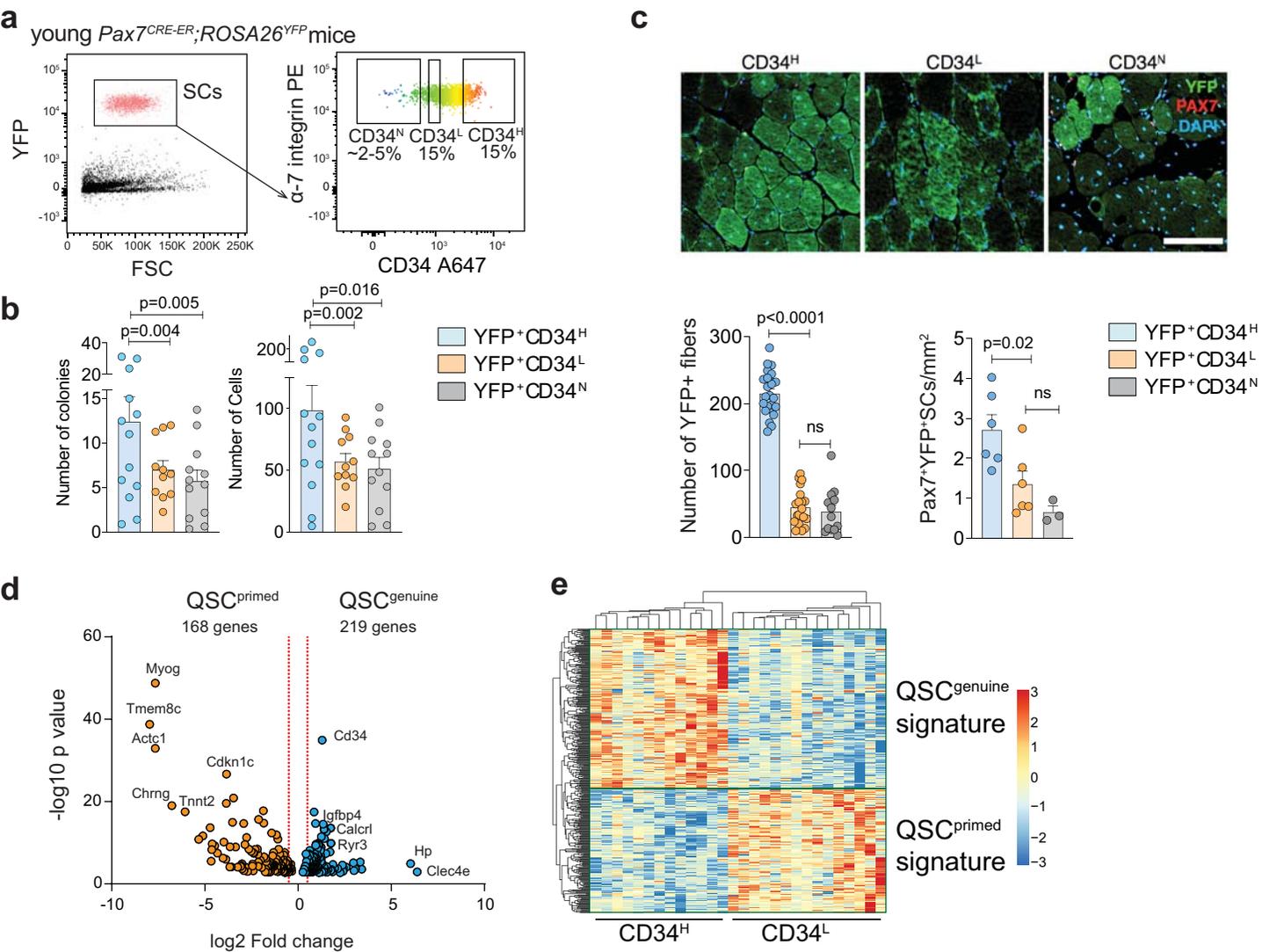


Figure 2

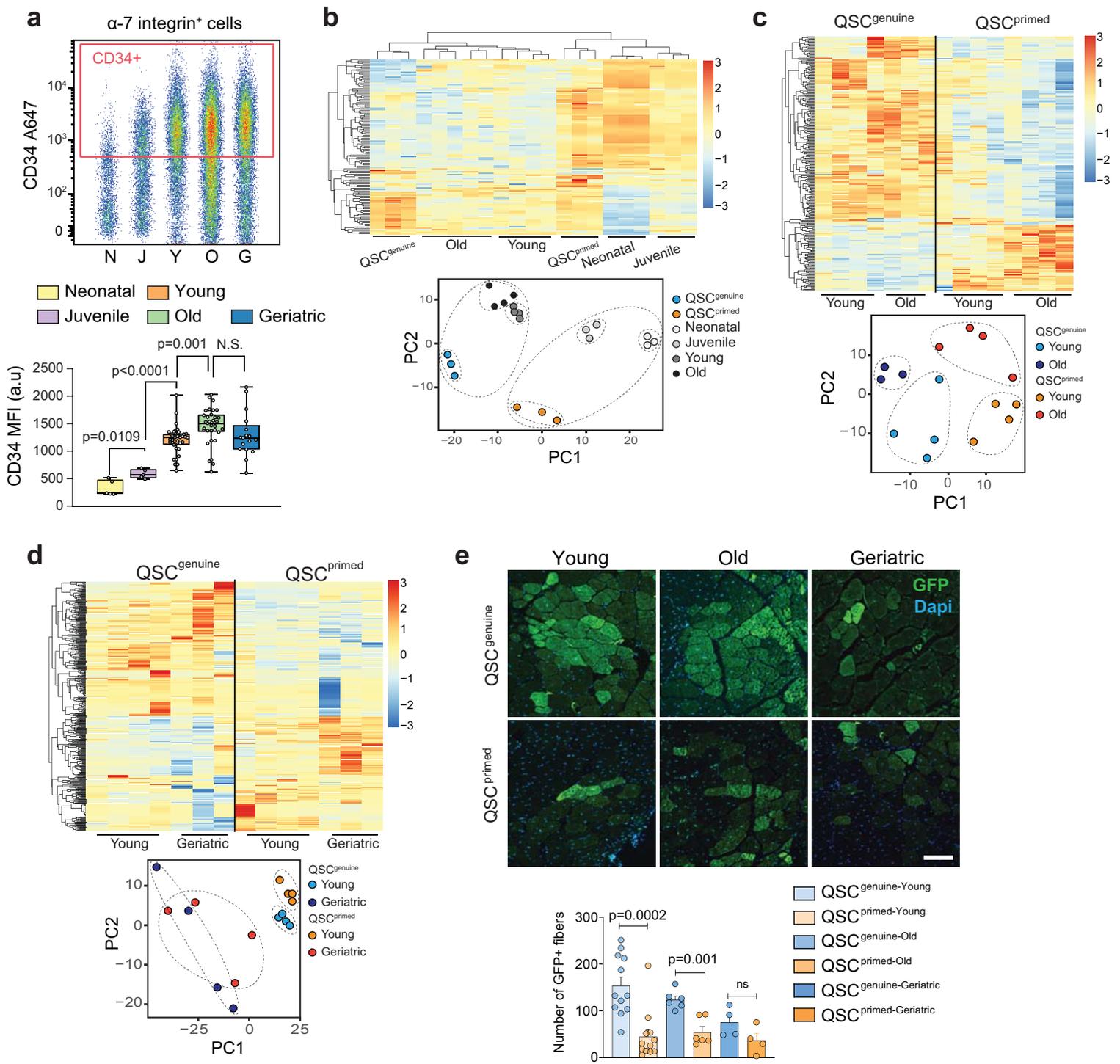


Figure 3

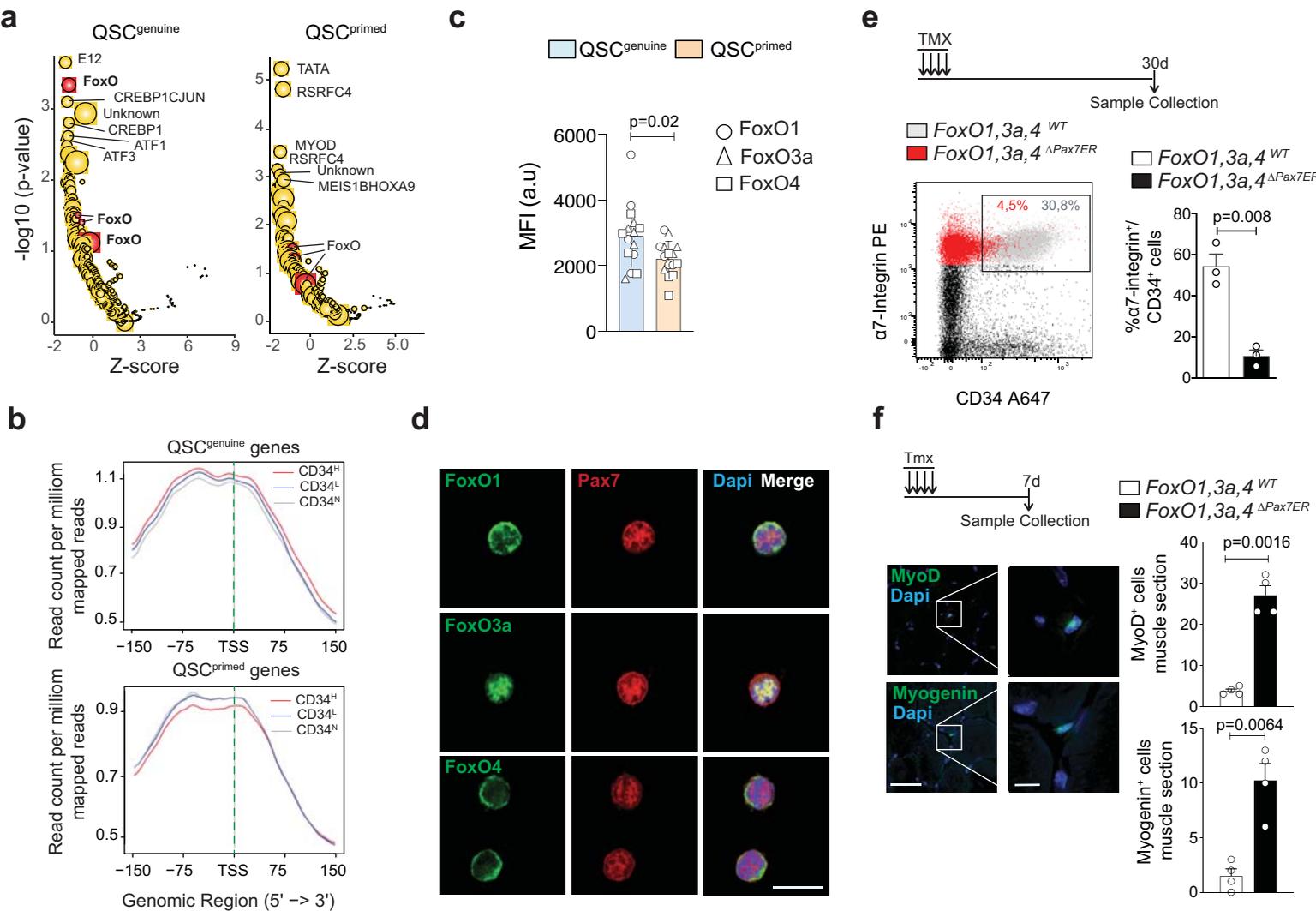


Figure 4

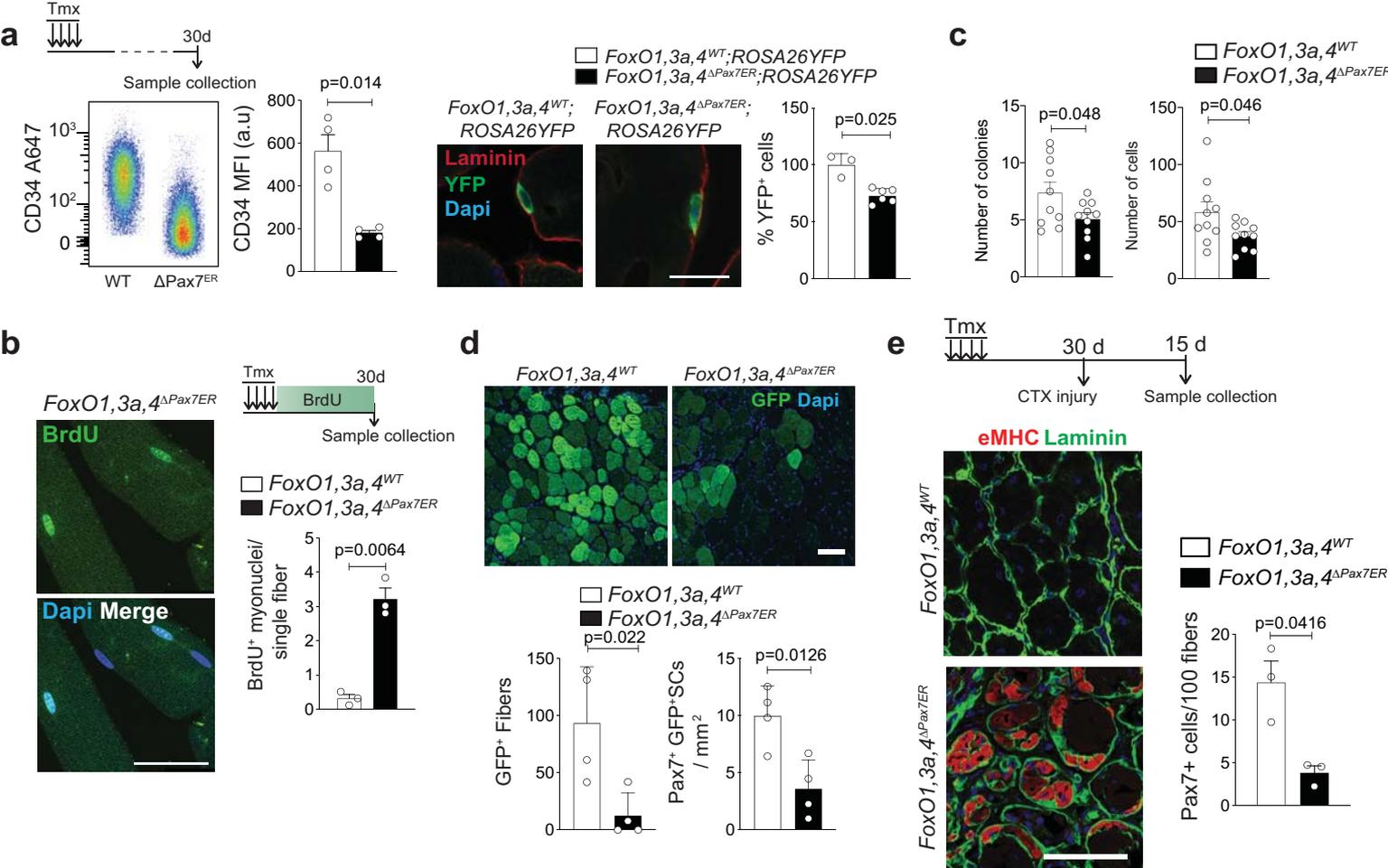


Figure 5

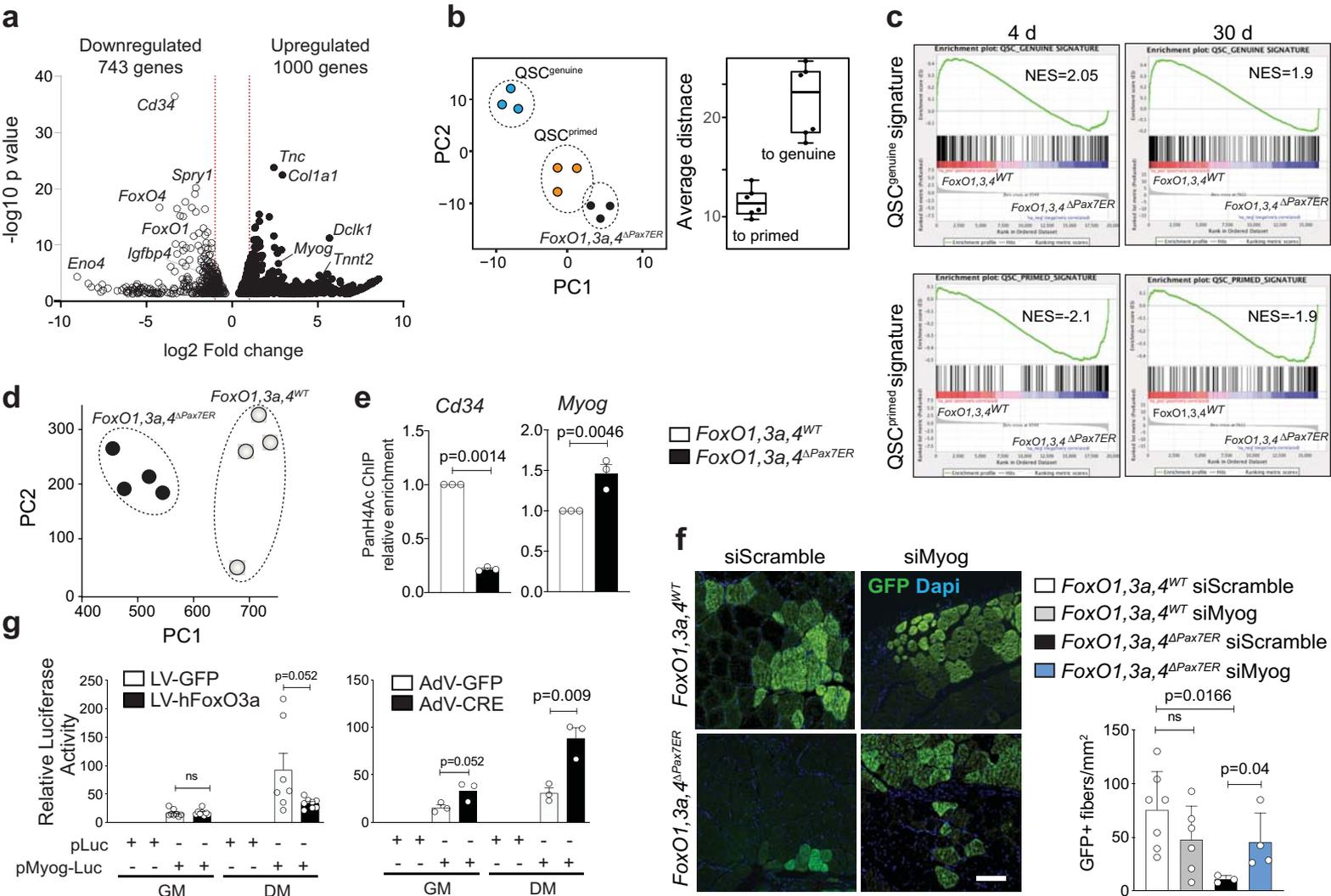


Figure 6

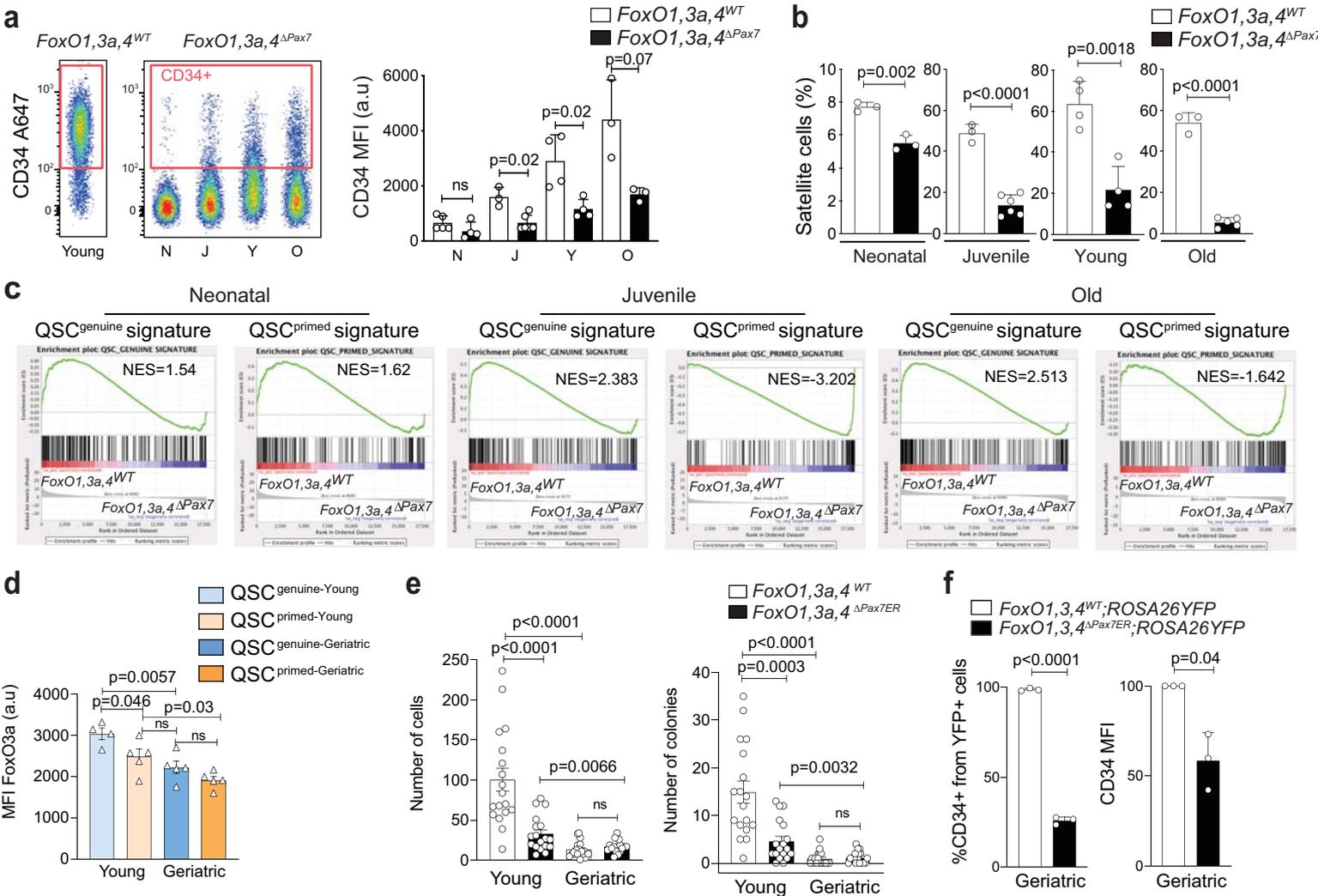


Figure 7

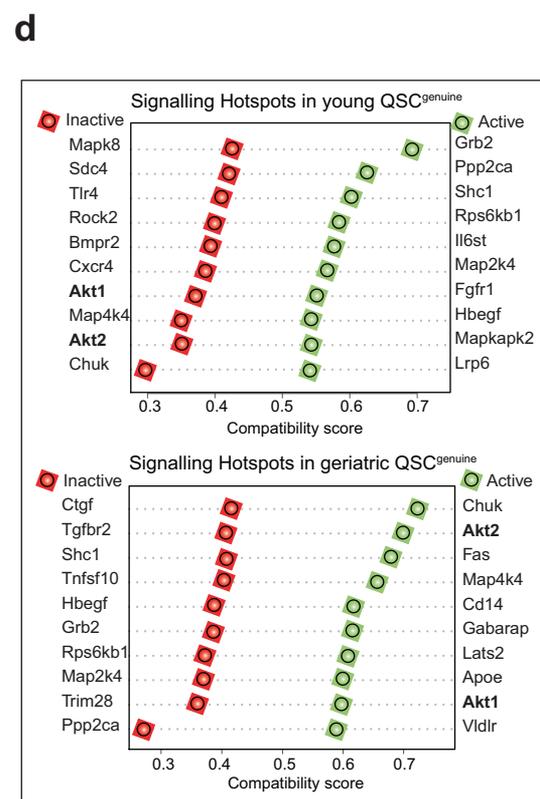
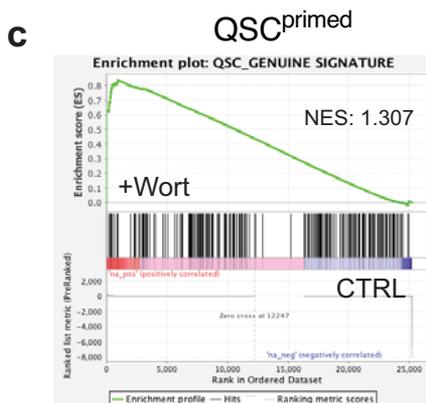
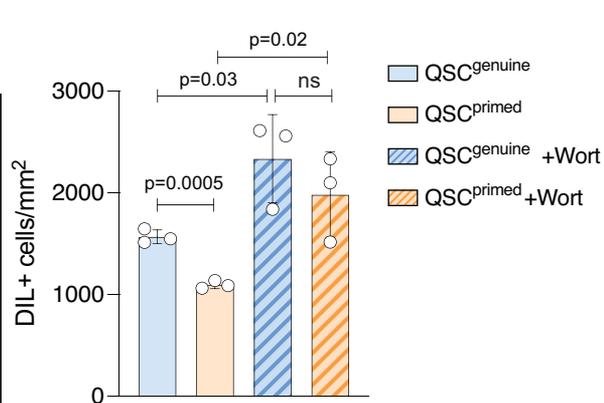
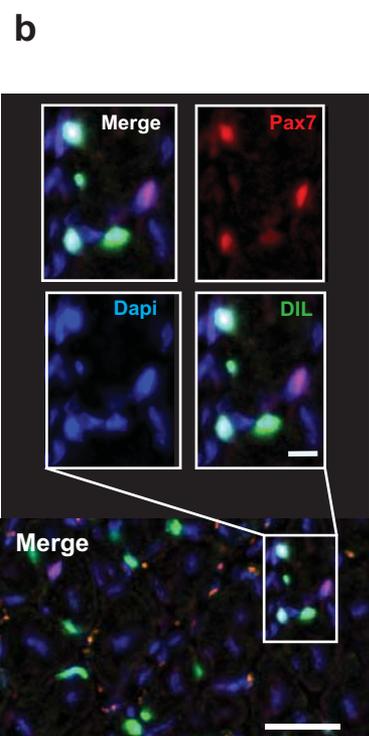
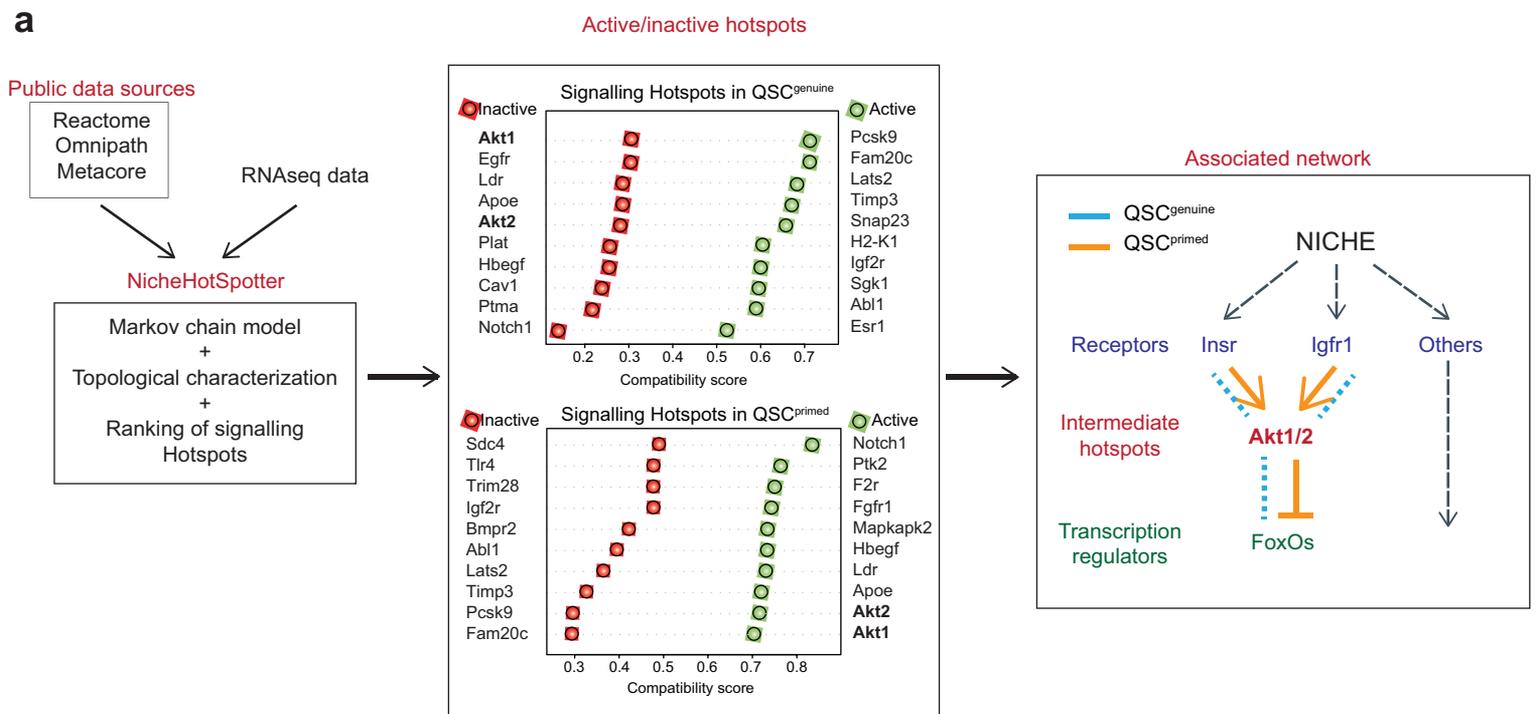
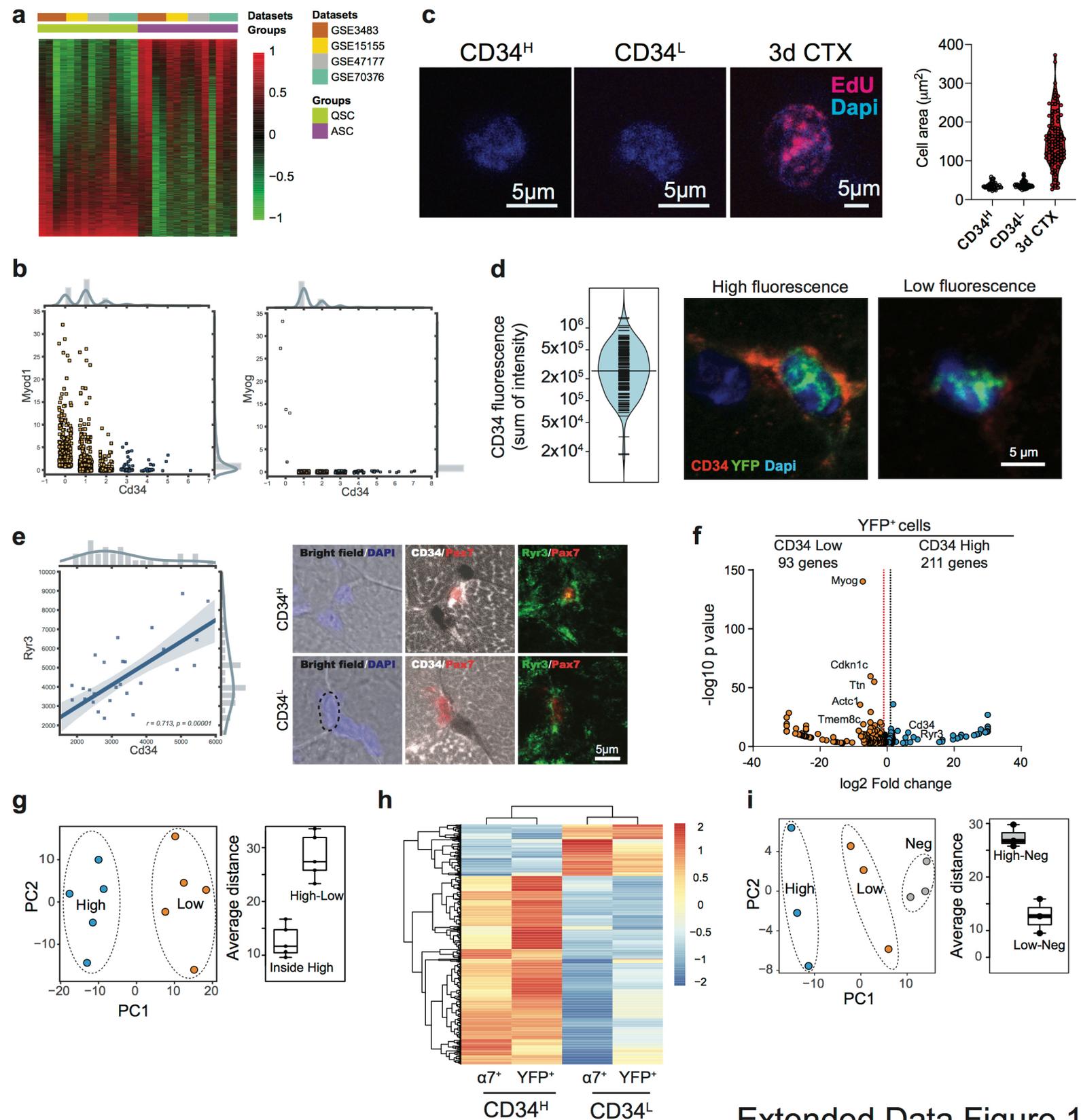
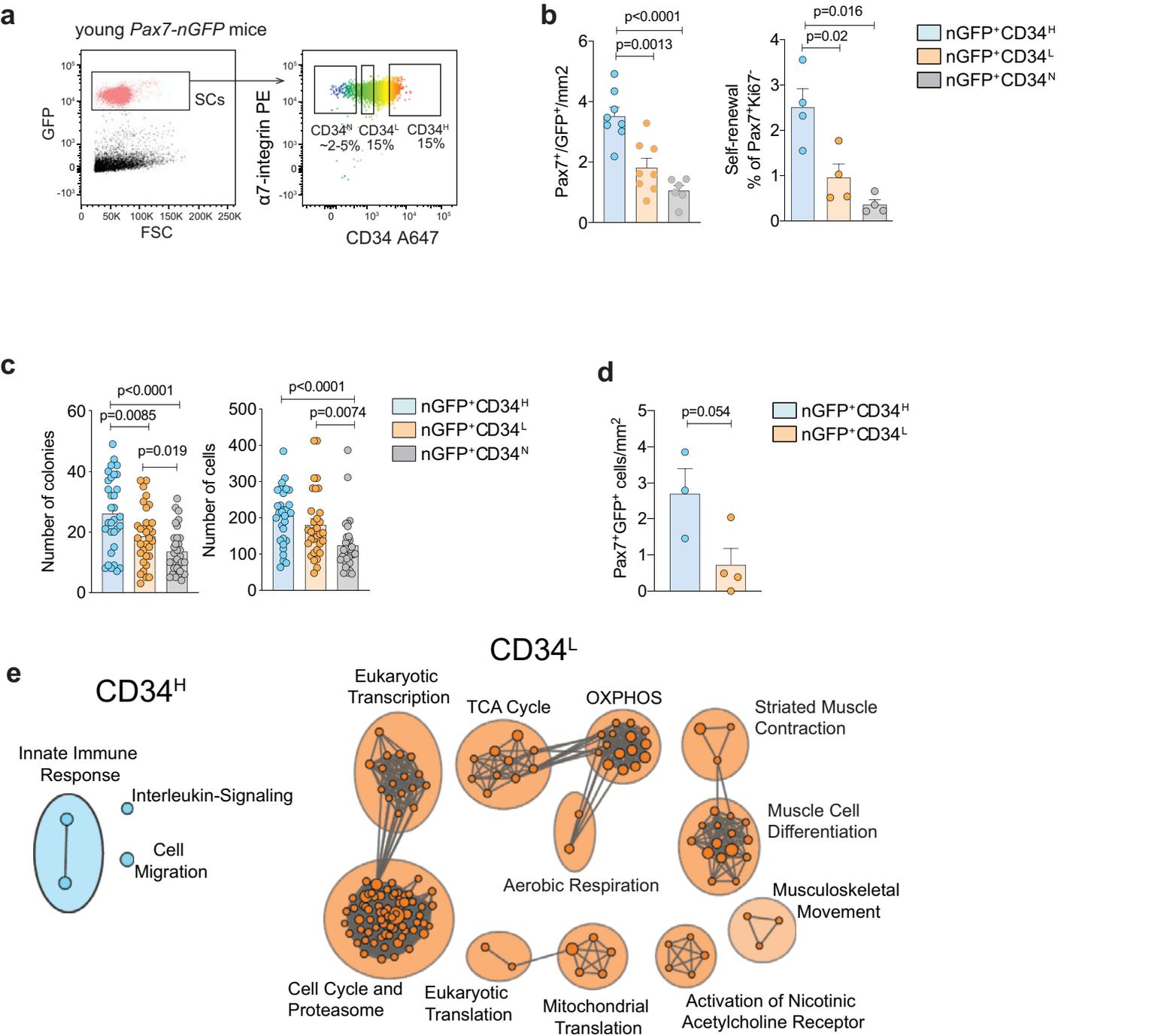
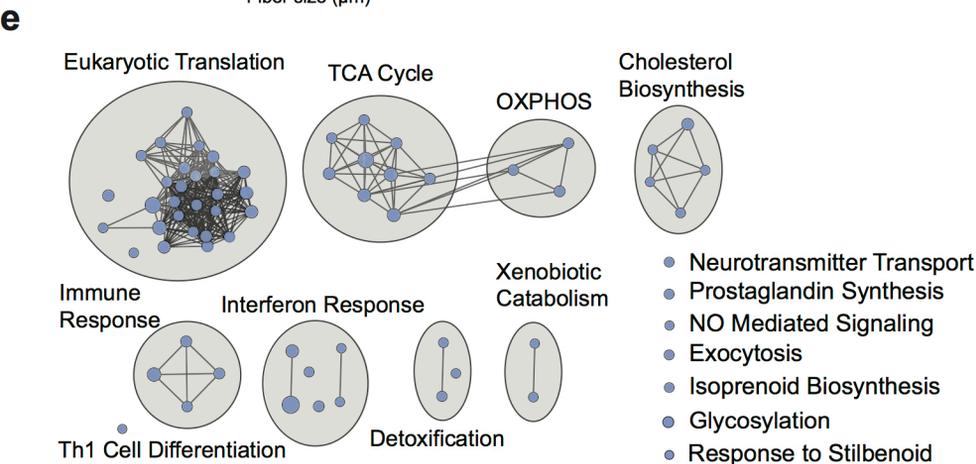
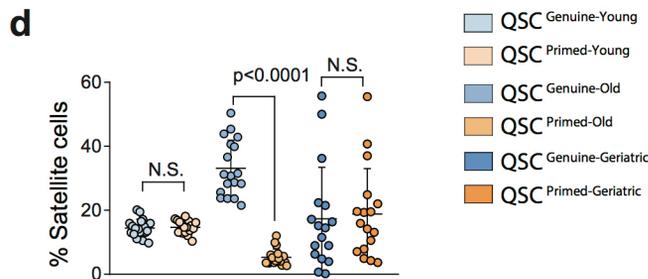
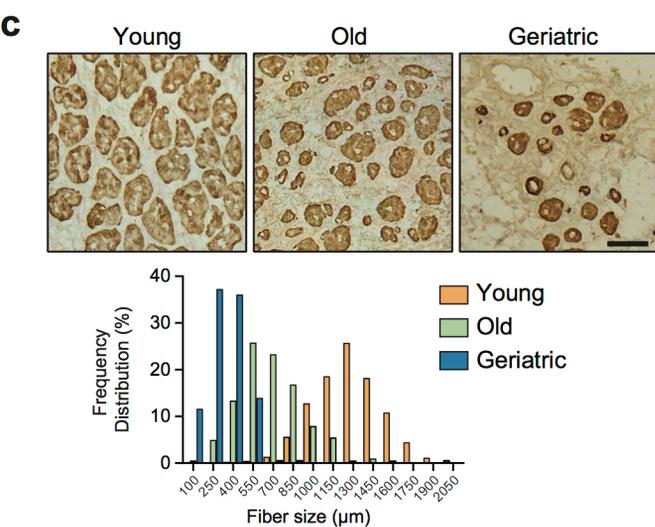
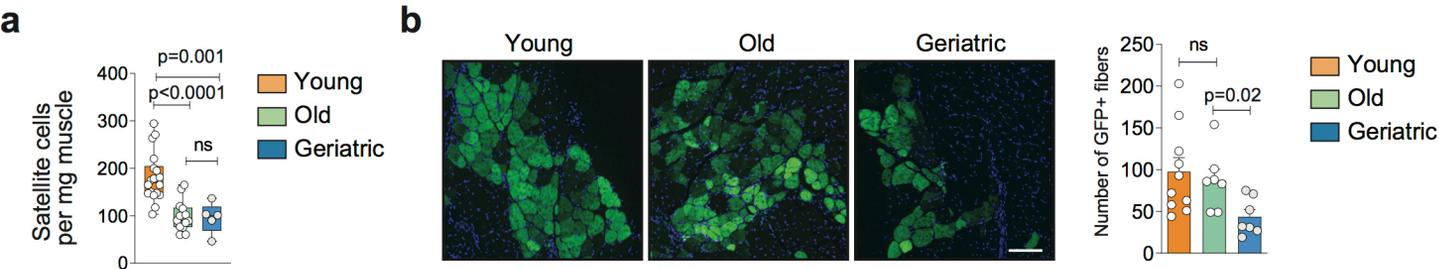
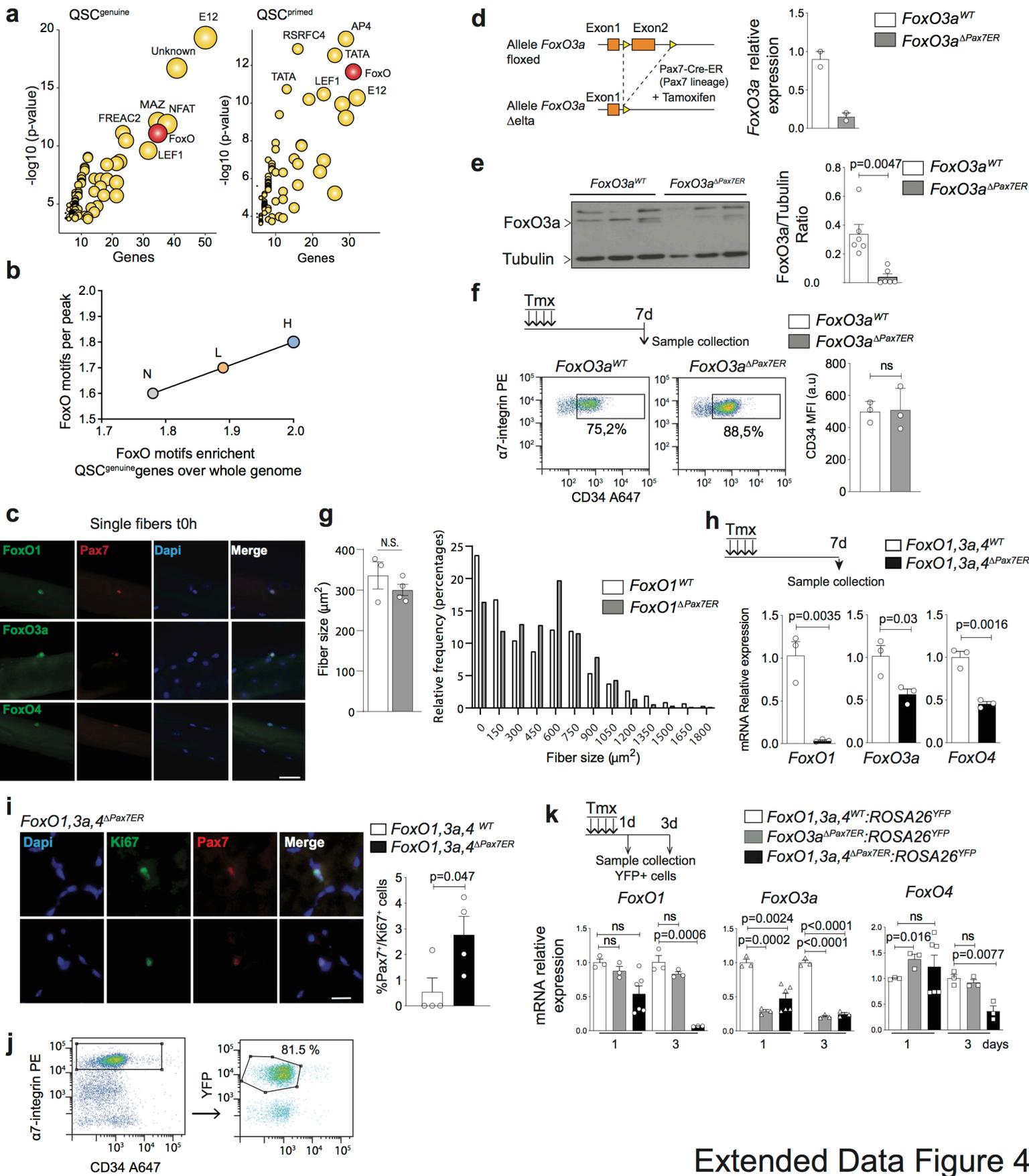


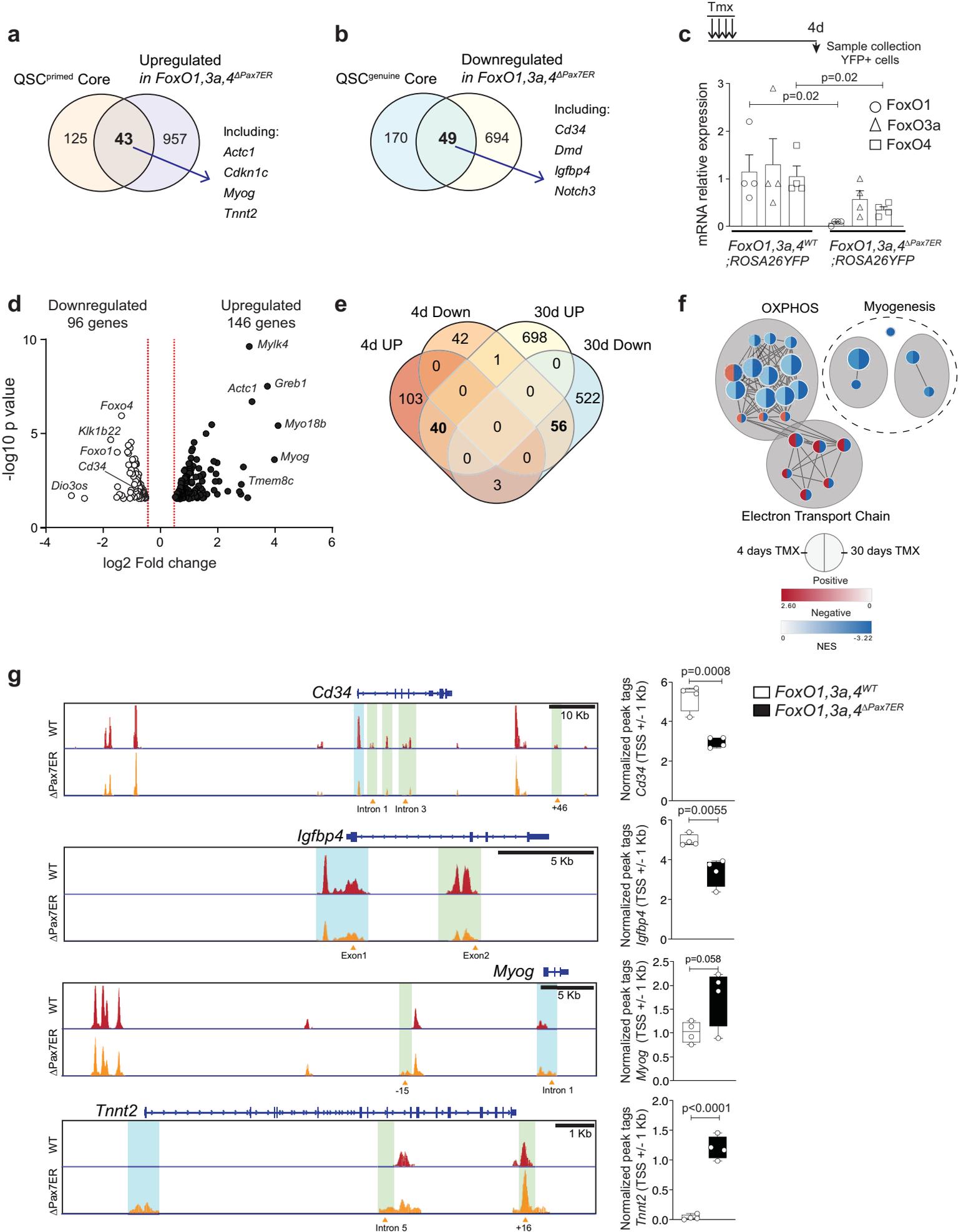
Figure 8

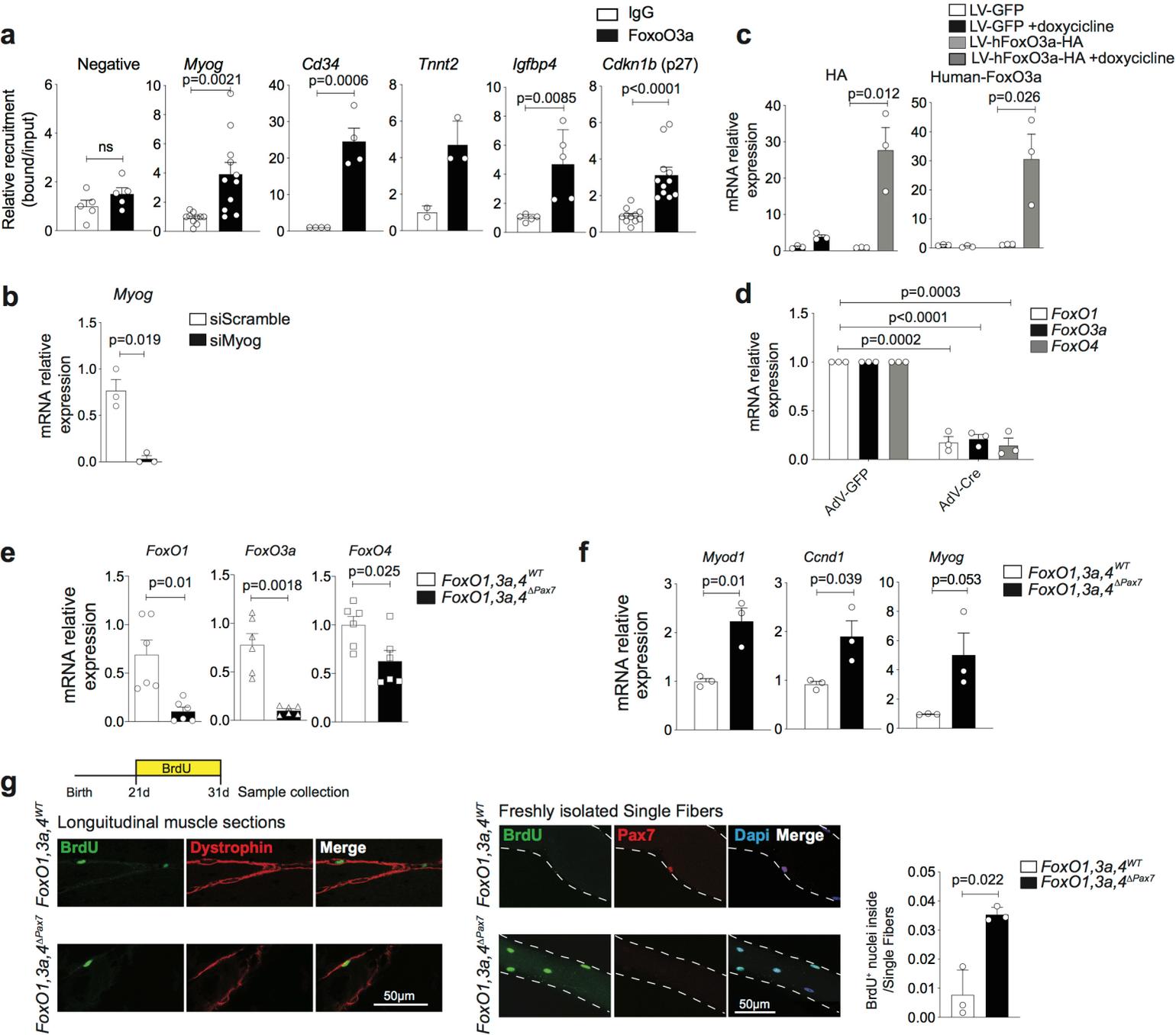


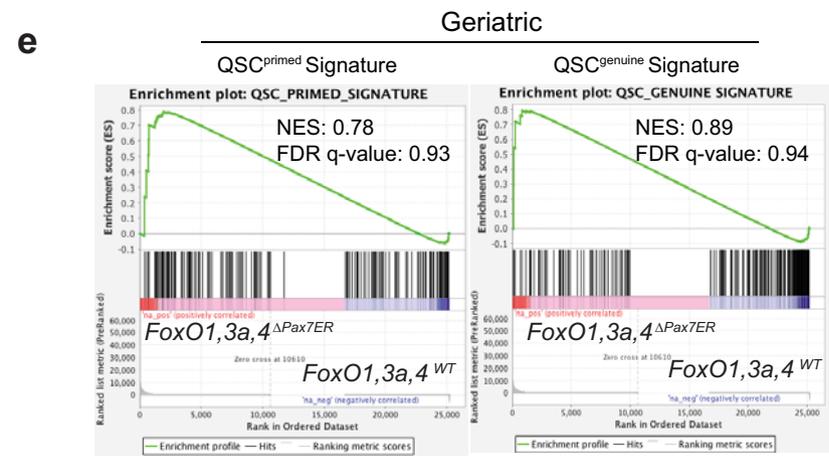
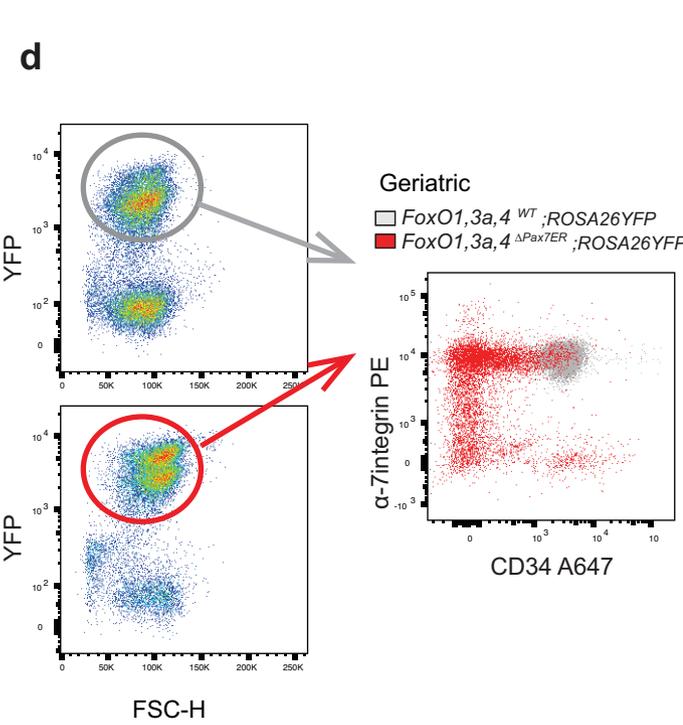
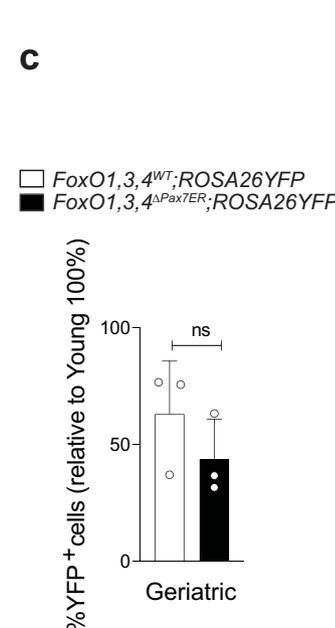
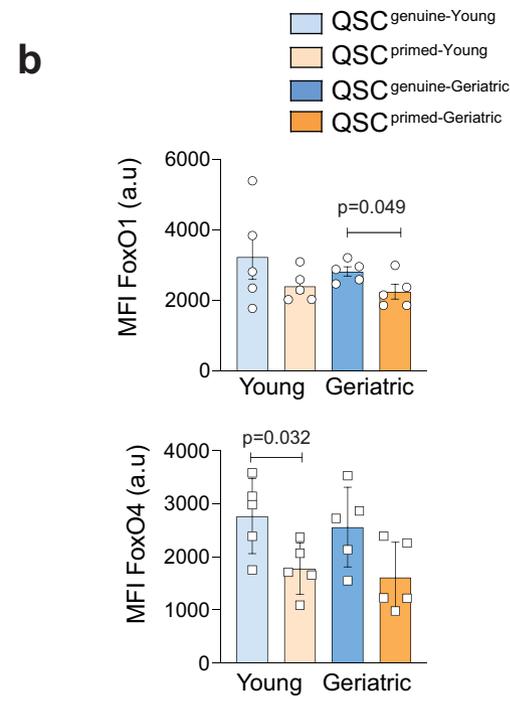
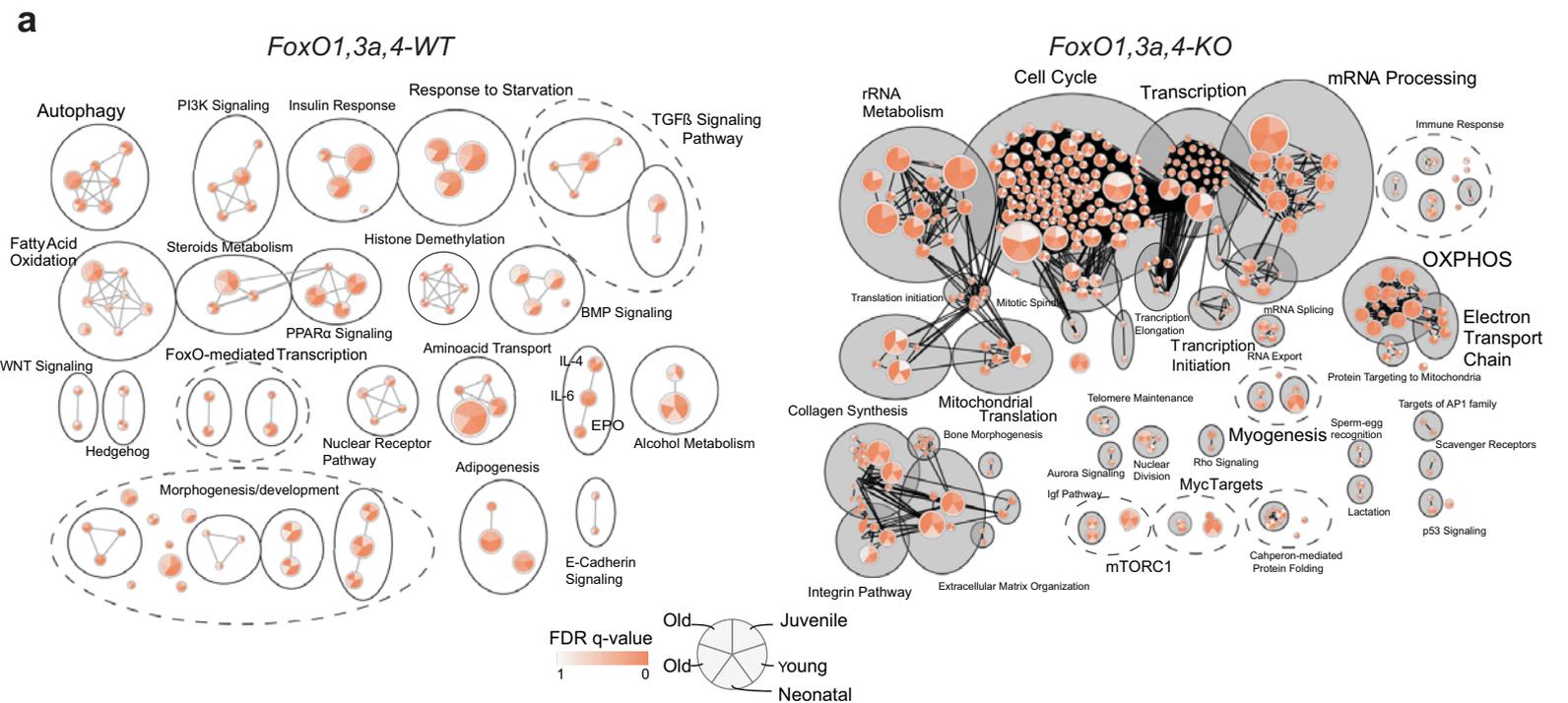




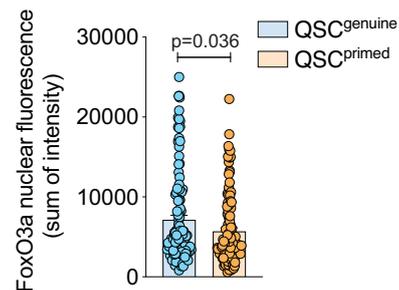
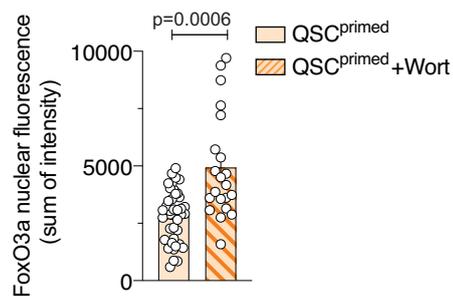
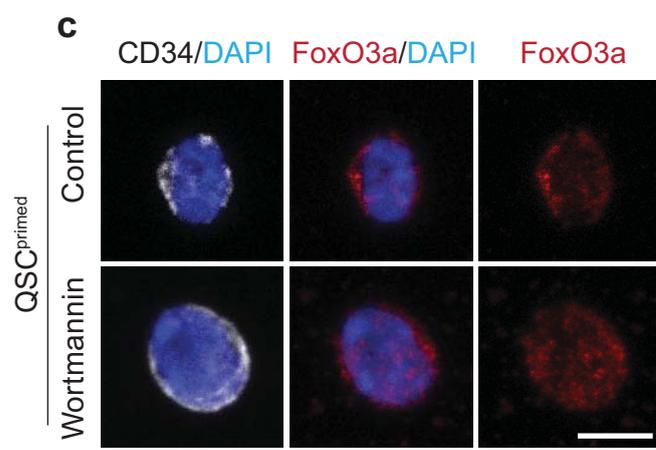
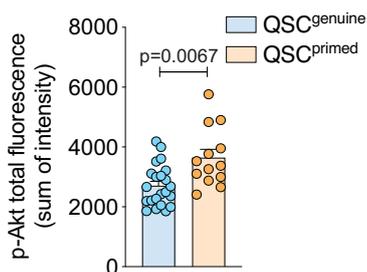
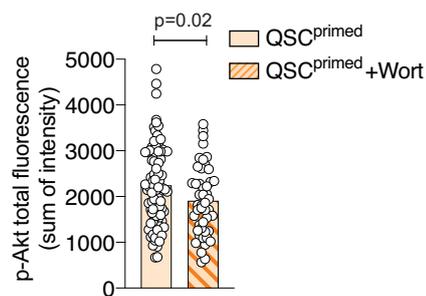
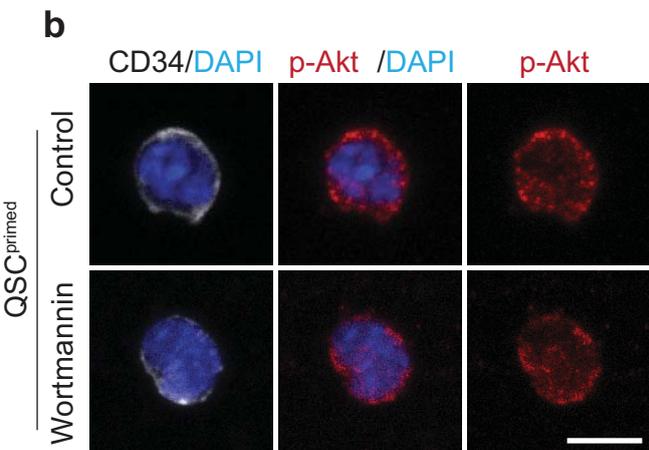
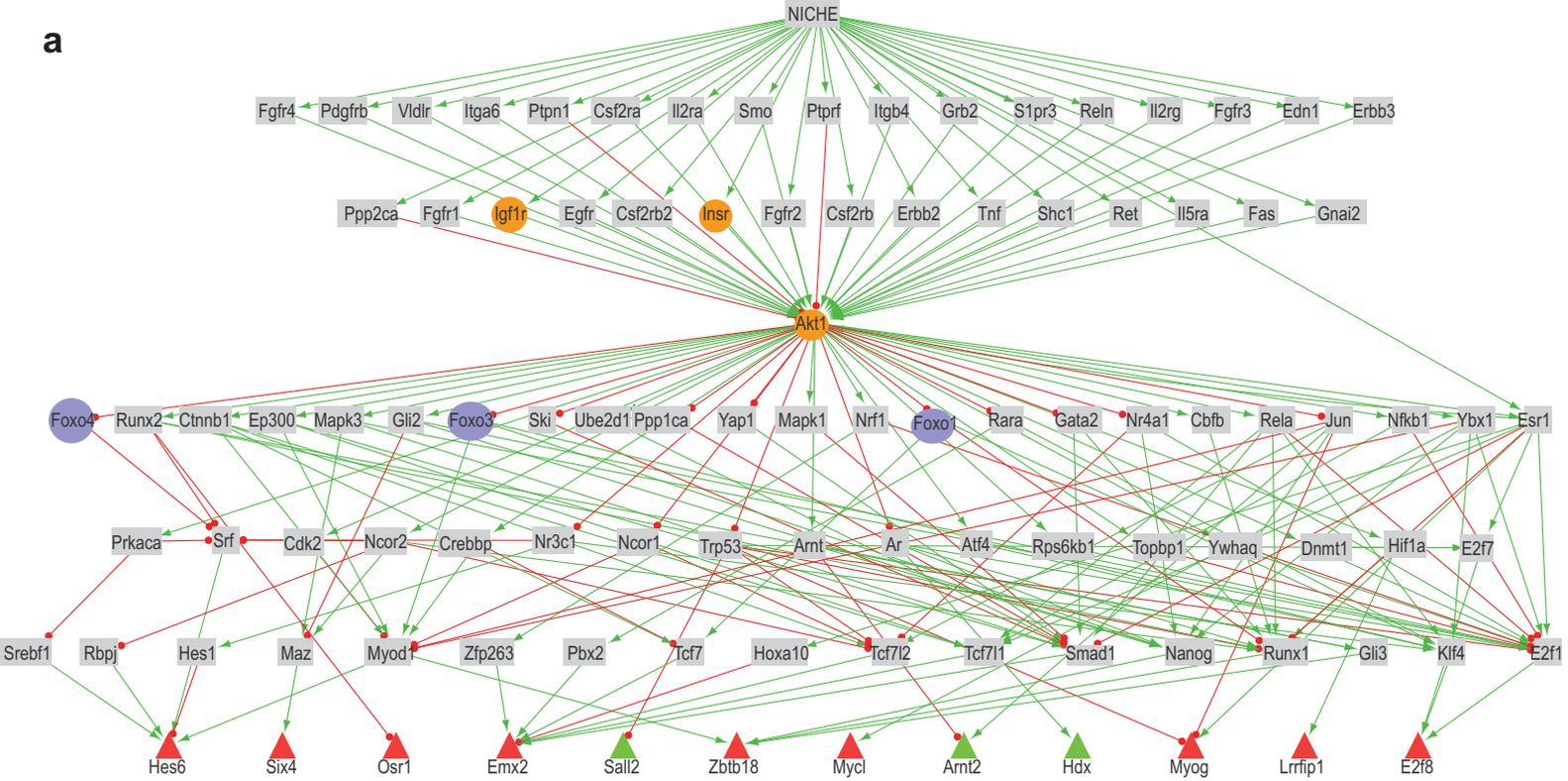


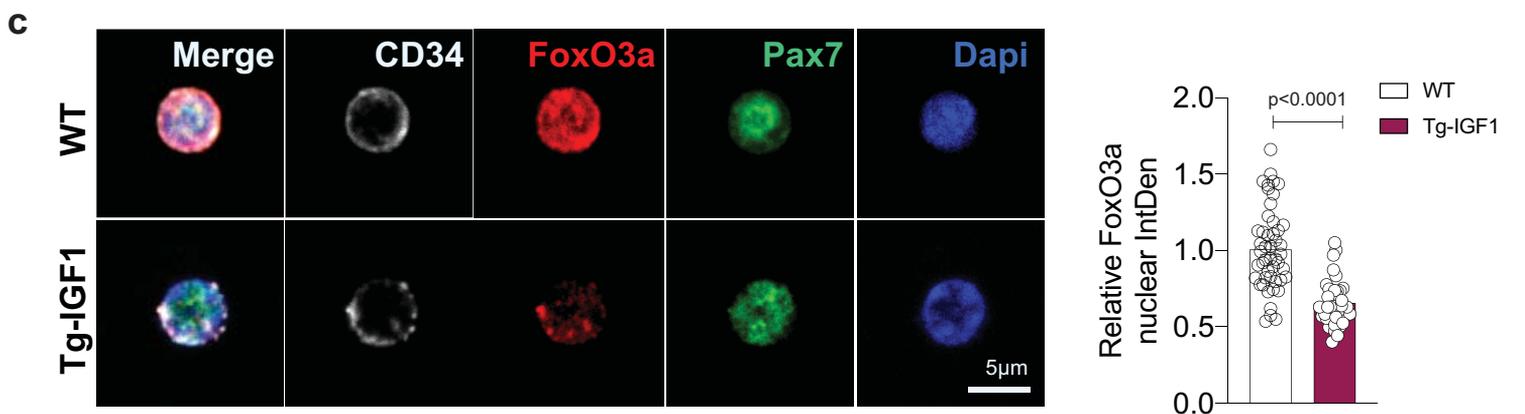
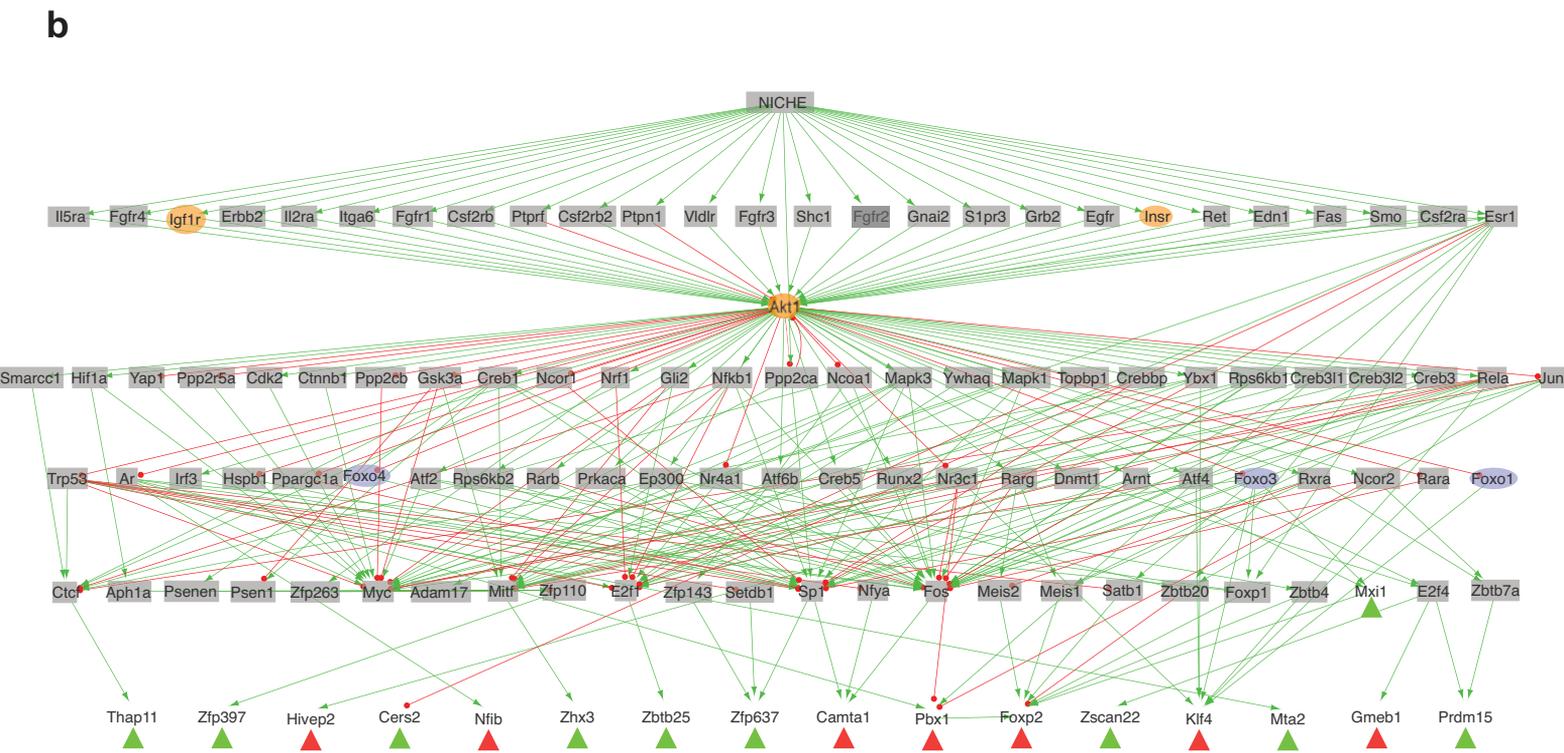
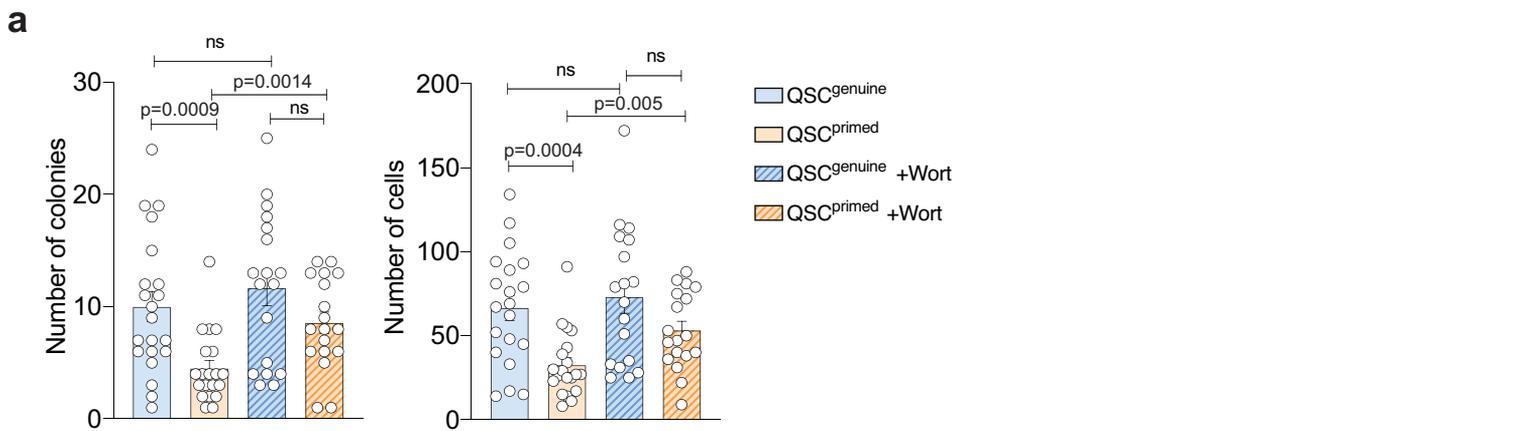


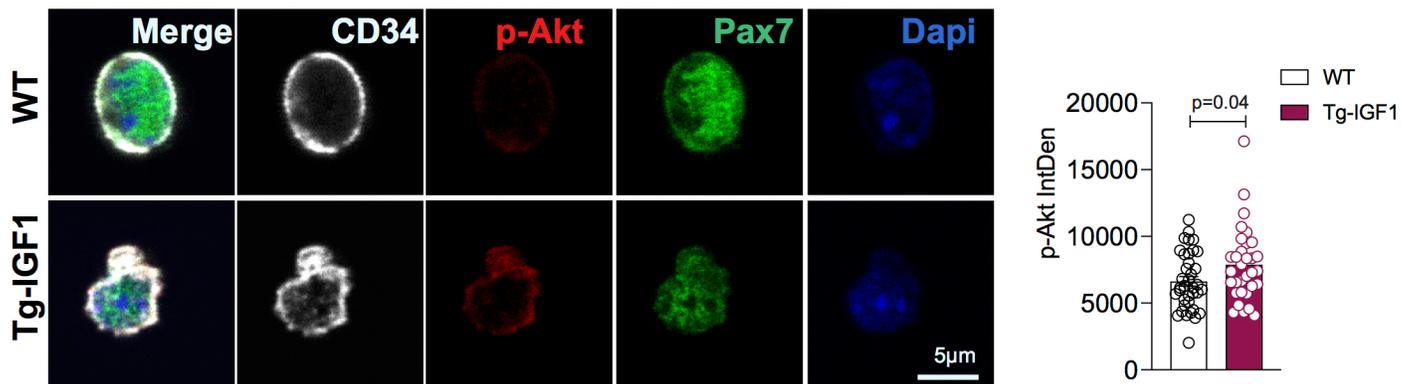
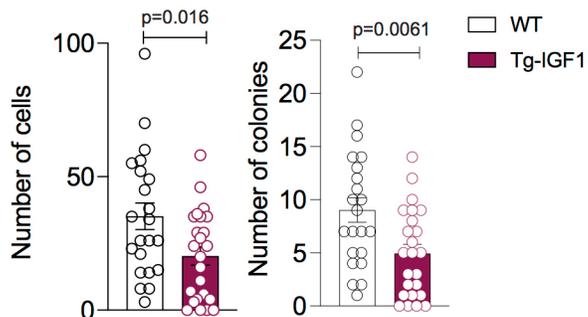
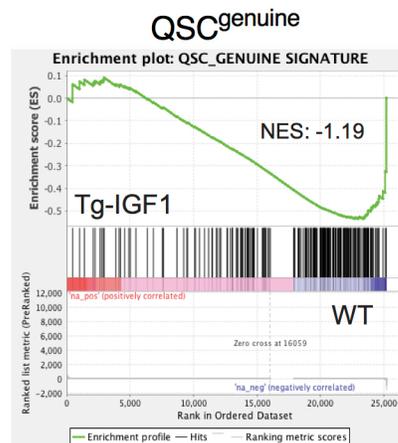




Extended Data Figure 7





**a****b****c****d**