Extraocular muscle stem cells exhibit distinct cellular properties associated with non-muscle molecular signatures

Daniela Di Girolamo[1,2#], Maria Benavente-Diaz[1,2,3#], Alexandre Grimaldi[1,2,3], Priscilla Thomas Lopes[1,2], Melania Murolo[1,2], Brendan Evano[1,2], Stamatia Gioftsidi[4,5,6], Vincent Laville[1,2], Sebastian Mella[7], Shahragim Tajbakhsh[1,2*], Glenda Comai[1,2*]

[1]Stem Cells & Development Unit, 25 rue du Dr. Roux, Institut Pasteur, 75015 Paris, France
[2]UMR CNRS 3738, Institut Pasteur, Paris, France
[3]Sorbonne Universités, Complexité du Vivant, F-75005, Paris, France
[4]Université Paris-Est, 77420, Champs-sur- Marne, France
[5]Freie Universität Berlin, 14195, Berlin, Germany
[6]Inserm, IMRB U955-E10, 94000, Créteil, France
[7] Hub de Bioinformatique et Biostatistique, Cytométrie et Biomarqueurs, Institut Pasteur, 75015 Paris, France
#Co-first authors
*Co-corresponding authors

Keywords: muscle stem cell diversity, extraocular muscles, mesenchymal stromal cells, MuSC expansion, MuSC sc-RNAseq, matrisome, TF network

Abstract:

The muscle stem cell (MuSC) population is recognized as functionally heterogeneous. Cranial muscle stem cells, which originate from head mesoderm, can have greater proliferative capacity in culture and higher regenerative potential in transplantation assays when compared to those in the limb. The existence of such functional differences in phenotypic outputs remain unresolved as a comprehensive understanding of the underlying mechanisms is lacking. We addressed this issue using a combination of clonal analysis, live imaging, and scRNA-seq, identifying critical biological features that distinguish extraocular (EOM) and limb (Tibialis anterior, TA) MuSC populations. Time-lapse studies using a *Myogenin*^{tdTomato} reporter showed that the increased proliferation capacity of EOM MuSCs is accompanied by a differentiation delay *in vitro*. Unexpectedly, in vitro activated EOM MuSCs expressed a large array of distinct extracellular matrix (ECM) components, growth factors, and signaling molecules that are typically associated with mesenchymal non-muscle cells. These unique features are regulated by a specific set of transcription factors that constitute a coregulating module. This transcription factor network, which includes Foxc1 as one of the major players, appears to be hardwired to EOM identity as it is present in quiescent adult MuSCs, in the activated counterparts during growth and retained upon passages in vitro. These findings provide insights into how high-performing MuSCs regulate myogenic commitment by active remodeling of their local environment.

INTRODUCTION

Skeletal muscle stem cells are arguably the most diverse in terms of molecular properties compared to other adult stem cells that are distributed throughout the body. Genetic and transcriptomic studies have shown that the MuSC population in any particular anatomical location is heterogeneous. Certain subsets are more prone to self-renewal or differentiation, differ in transplantation efficiency, as well as their stem cell-niche interactions, metabolism, and resistance to stress upon activation (Barruet et al., 2020; Chakkalakal et al., 2014; Cho et al., 2020; Dell'Orso et al., 2019; Dumont et al., 2015; Gayraud-Morel et al., 2012; Hernando-Herraez et al., 2019; Micheli et al., 2020; Morree et al., 2019; Ono et al., 2012; Rocheteau et al., 2012; Scaramozza et al., 2019; Tierney et al., 2018; Vartanian et al., 2019; Yartseva et al., 2020; Yennek et al., 2014). Despite this diversity, MuSCs share common functions as they are essential for postnatal growth and repair of the skeletal muscle system (Lepper, 2011; Murphy, 2011; Sambasivan, 2011). In response to muscle tissue damage or growth factors in culture media, MuSCs activate, proliferate and undergo cell fate decisions to either self-renew to replenish the MuSC pool (expressing *Pax7)*, or commit to the myogenic program and differentiate into myoblasts and fusion-competent myocytes that express the commitment and differentiation factors *Myod* and *Myog*, respectively (Evano and Tajbakhsh, 2018; Zammit et al., 2006).

An unexpected finding was the discovery that MuSCs in different anatomical locations are programmed with distinct upstream transcription factors (TF) prior to acquiring myogenic identity (Gopalakrishnan et al., 2015; Harel et al., 2009; Kelly et al., 2004; Sambasivan et al., 2009; Tajbakhsh, 1996). This intriguing observation correlates with reports where subsets of cranial MuSCs are functionally more robust in terms of proliferation, engraftment efficiency, and susceptibility to disease, when compared to those in the limb (Randolph and Pavlath, 2015). However, the intrinsic and extrinsic regulators conferring robust features to cranial muscles have not been characterized.

EOMs are a subset of head muscles that govern eye movements. EOMs are derived from unsegmented cranial mesoderm and are regulated by distinct transcriptional factors and signalling molecules compared to the somite-derived limb and trunk muscle groups (Grimaldi and Tajbakhsh, 2021; Michailovici et al., 2015; Sambasivan et al., 2011). Notably, when the transcription factor *Pax3* is inactivated, no limb muscles develop, yet cranial muscles including the EOMs are present (Sambasivan et al., 2009; Tajbakhsh et al., 1997). Conversely, mice lacking the transcription factor *Pitx2* do not form EOMs whereas other cranial and somite derived muscles are unaffected (Diehl et al., 2006; Gage et al., 1999; Zacharias et al., 2011). Additionally, whole muscle transcriptional profiling revealed an extraordinary diversity amongst adult skeletal muscle groups (Kui et al., 2022; Terry et al., 2018). As some muscle subsets, like the EOMs, are preferentially spared in muscular dystrophies and during ageing (Emery, 2002; Formicola et al., 2014; Verma et al., 2017), it has been proposed that intrinsic properties of MuSCs or the respective myofibers could determine their differentially sensitivity to disease (Randolph and Pavlath, 2015; Terry et al., 2018).

Similar to MuSCs derived from other craniofacial muscles (Ono et al., 2010; Terry et al., 2018), EOM MuSCs have a greater proliferative capacity in culture when compared to those in the limb, with this property being conserved in *Mdx* mice, a model of Duchenne muscular dystrophy (Dmd), and in a preclinical rat model of the disease (Stuelsatz et al., 2015; Taglietti et al., 2023). Additionally, EOM MuSCs have a higher engraftment potential compared to those in the limb (Stuelsatz et al., 2015). Although debated (McLoon and Wirtschafter, 2002; McLoon and Wirtschafter, 2003; Stuelsatz et al., 2015; Verma et al., 2017), EOM MuSCs were reported to continuously contribute to myofibres in homeostatic conditions, suggesting a higher level of basal activation of the MuSC population (Keefe et al., 2015; Pawlikowski et al., 2015; Taglietti et al., 2023). In addition, changes in the transcriptomic profile of EOM MuSCs after transplantation into limb muscle (Evano et al., 2020) showed that despite a large transcriptional reprogramming after transplantation, about 10% of EOM specific genes persisted in the heterotypic niche environment. This finding suggests that cell autonomous regulation of MuSC properties predominates to a certain extent in EOM MuSCs (Evano et al., 2020). Moreover, recent single-cell transcriptomic analysis identified the thyroid hormone signalling pathway as a key factor preventing entry into senescence of EOM MuSCs in Dmd rats (Taglietti et al., 2023). Yet, whether deeply-rooted TF gene regulatory networks exist within MuSCs subsets, and contribute to the maintenance of anatomically-distinct MuSC phenotypes remains largely unexplored.

Here, we used droplet-based scRNA-seq to investigate the transcriptional states that govern the outperformance of MuSC subsets during activation. By integrating genome-wide analysis with cell biology approaches using mouse reporter lines, we identified key regulators that confer distinct mesenchymal-like features to activated EOM MuSCs and expression of extracellular matrix (ECM) components. Computational analyses showed that this pool of EOM cells displays features of a more stem-like state that is actively maintained during *in vitro* proliferation through a specific set of transcription factors forming a co-regulatory module.

RESULTS

Functional differences among MuSCs are accompanied by heterogeneity in myogenic gene expression dynamics

In agreement with previous reports (Stuelsatz et al., 2015), we show that culture of EOM and *Tibialis anterior* (TA) MuSCs isolated by FACS using *Tg:Pax7-nGFP* mice (Sambasivan et al., 2009), revealed a larger fraction of proliferative cells in the EOM from day 3 (D3) onwards, peaking at D4 and remaining higher even when differentiation and fusion starts to take place at D5 (**Figure 1A-C**). Cell proliferation, monitored by 5-Ethynyl-2'-deoxyuridine (EdU) uptake, revealed 45,1% of Pax7+EdU+ EOM cells, compared to 19,8% in the TA at D5 (**Figure 1B**). To obtain insights into the cellular basis for the outperformance of MuSC subsets upon exit from quiescence, we performed clonal and live imaging analysis *in vitro*. First, MuSCs from EOM and TA were plated at clonal density in 96-well plates. In parallel, a fraction of the isolated cells was plated as a bulk culture, allowed to activate for 2 days *in vitro*, re-isolated and then plated at clonal density (**Suppl Figure 1A**). The total cellular output in both

cases was quantified two weeks after seeding. In accordance with a previous study (Stuelsatz et al., 2015), EOM MuSCs displayed a markedly higher clonal capacity than those from the TA muscle, with a mean of 1315 cells/clone and 186 cells/clone respectively (7-fold difference; **Suppl Figure 1B,** T1). Surprisingly, the clonogenic properties of EOM MuSCs continued to be elevated after 2 days of activation *in vitro* where they yielded 3656 cells/clone compared to 297 cells/clone for activated TA MuSCs (12-fold difference; **Suppl Figure 1B**, T2).

We then monitored their differentiation dynamics *in vitro* by live imaging using the *Tg:Pax7-nGFP;Myog*^{ntdTom} mouse lines (Benavente-Diaz et al., 2021; Sambasivan et al., 2009) to isolate MuSCs by FACS and then monitor the onset of *Myog* expression using a nuclear localised tdTOMATO fluorescent reporter. After activation for 3 days *in vitro*, EOM and TA cells were imaged every 12 min for ~48h (**Figure 1D**). Individual cells were tracked and tdTOMATO intensity was scored at selected time points. EOM and TA MuSCs displayed distinct profiles of lineage progression, where TA myogenic cells initiated reporter gene expression earlier than EOM cells (**Figure 1E**). The percentage of *Myog*-expressing cells sharply increased in the TA from 5% at 80h post-plating to 40% at 96h. Significantly, only ~7% of EOM cells were tdTOMATO^{+} by 96h *in vitro* (**Figure 1F**).

Since *Myog* expression is followed by cell-cycle withdrawal and terminal differentiation (Andrés and Walsh, 1996; Benavente-Diaz et al., 2021; Guo et al., 1995), a delay in its expression could allow for sustained proliferation of EOM MuSCs. We performed RT-qPCR at 48h post-plating and assessed the expression of key myogenic transcription factors and the cell cycle inhibitor *p21*, which regulates cell cycle exit during myoblast differentiation (Guo et al., 1995; Zhang et al., 1999). In agreement with the live imaging data, *Myog* mRNA was significantly upregulated in TA activated MuSCs compared to EOM, concomitant with an increase in *p21* mRNA (**Figure 1G**). No significant differences were found in *Pax7* nor *Myod* expression (**Figure 1G**). Moreover, EOM progenitors displayed reduced levels of phosphorylated P38 (**Suppl Fig1C,D**). As p38α/β MAPKs play a central role in promoting MEF2 transcriptional activity and initiating the differentiation program (Rugowska et al., 2021), this data suggests that EOM progenitors are less prone to myogenic commitment. Therefore, our results suggest that a delay in the differentiation of EOM MuSCs might contribute to the higher number of myogenic cells produced compared to TA cells.

Activation of EOM MuSCs include a larger progenitor pool and distinct activation signatures

To gain insights into the potential mechanisms that regulate the phenotypic differences and cell fate decisions between EOM and TA MuSCs, we performed scRNA-seq of *in vitro* activated MuSCs using the 10x Chromium platform. We first isolated EOM and TA MuSCs by FACS from *Tg:Pax7-nGFP* mice and cultured them *in vitro* for 4 days. Cell cultures were subsequently trypsinized, cell-sorted and processed for scRNA-seq (**Figure 2A**). Unsupervised clustering divided cells into 2 clusters per sample (**Figure 2B**) that were annotated as progenitors (Prog) or differentiating (Diff) based on the expression of known myogenic makers, such as *Pax7/Myf5* and *Myod/Myog* (**Figure 2B, C**). Notably, expression

of Myod was lower in EOM progenitors (**Figure 2B**) and lower protein levels were detected at the whole population level in vitro (**Suppl Figure 2A**).

Analysis of differentially expressed genes (DEGs) across these clusters showed that each cell state had a distinct transcriptional pattern (**Figure 2D, Suppl Table 1**). As expected, TA progenitors expressed a defined Hox-signature (Evano et al., 2020) and inhibitors of differentiation (Id1, Id2, (Jen et al., 1992; Kumar et al., 2009)). Instead, EOM progenitors displayed markers that had not been previously described in activated MuSC datasets such as *Mgp*, *Bgn*, *Col1A2* and *Acta2* (smooth muscle actin). EOM and TA differentiating cells shared part of their signature (*Myog, Mymx, Myl4*). However, TA progenitors shared some markers with the TA differentiating cells (*Acta1, Mylpf, Tnnt2*), suggesting that Prog and Diff TA clusters are closer at a transcriptional level than their EOM counterparts (**Figure 2D**).

To get insights into these properties and the prospective functional heterogeneity associated with each transcriptomic signature we performed reactome pathway analysis of the DEGs upon in vitro activation (**Figure 2E, Suppl Figure 2B, Suppl Table 2**). Multiple pathways involved in ECM organisation were characteristic of EOM progenitors including “ECM proteoglycans”, “crosslinking of collagen fibrils” and “integrin cell surface interactions”. On the other hand, EOM and TA differentiating cells were both enriched in pathways involved in striated muscle contraction and creatinine metabolism, which are associated with maturation of skeletal muscle. Also, as suggested by the DEG heatmap, TA progenitors shared common pathways with the differentiating clusters. Analysing the molecular function of the DEGs revealed that EOM progenitors were also specifically enriched in Pdgfrβ, a receptor for platelet-derived growth factor, and proteins related to insulin like growth factor binding and integrin signaling (**Figure 2F, Suppl Figure 2C, Suppl Table 3**). Thus, these analyses uncovered a non-canonical signature in the EOM progenitor fraction upon activation.

EOM and TA MuSCs retain distinct molecular signatures after several days in culture

We then set out to determine to what extent the EOM progenitor fraction that was capable of self-renewal *in vitro*, relates to the quiescent MuSCs present *in vivo*. Thus, we assessed the profiles of quiescent MuSCs by scRNA-seq of from EOM and TA immediately after isolation by FACS (**Figure 3A-B**). Despite the homogeneous expression of MuSC markers *Myf5* and *Pax7* in both populations, *Myod* expression was restricted to the TA MuSCs (**Figure 3C-D**), suggesting an earlier MuSC activation in response to the isolation procedure (Brink et al., 2017; Machado et al., 2017; Velthoven et al., 2017). Differentiation markers like *Myog* were absent in both populations (**Figure 3C-D**). Visualization of the top 25 most variably expressed genes documented distinct transcriptional programs of EOM and TA quiescent clusters (**Figure 3E, Suppl Table 4**). Moreover, microarray analysis of adult quiescent *Tg:Pax7-nGFP* MuSCs revealed that the EOM quiescent signature was also different from that of MuSCs from other cranial-mesoderm derived muscles such as the esophagus and masseter (**Suppl Figure 3A**). Interestingly, we found several conserved genes between the quiescent and activated sc-RNAseq states at each anatomical location, while others were unique to the quiescent or activated cell states (**Figure 3F-I**). In the TA, *Lbx1, Vgll2* and *Hox* genes were found as common signature throughout cell

states (**Figure 3G,I, Suppl Figure 3B, C**). *Lbx1* is a homeobox transcription factor required for the migration of myogenic progenitor cells to the limbs (Brohmann et al., 2000; Gross et al., 2000) and *Vgll2* (previously named Vito-1) is a key cofactor of the myogenic differentiation program (Günther et al., 2004; Maeda et al., 2002). All three genes had been already identified in the limb both at the MuSC levels and in entire muscles in the adult (Evano et al., 2020; Honda et al., 2017; Terry et al., 2018). *Pcdh7*, an integral membrane protein belonging to cadherin superfamily that regulates intercellular adhesion (Wang et al., 2020; Yoshida, 2003) is unique to TA quiescent state and the microtubule protein *Tubb3* (Duly et al., 2022) is a marker of the TA activated state (Figure 3I).

EOM genes unique to the quiescent state included *Tshr*, encoding the thyroid-stimulating hormone receptor, which has been previously identified in bulk RNAseq as "EOM-resistant" kept upon engraftment into the limb (Evano et al., 2020), and shown to control senescence in MuSCs of DMD rats (Taglietti et al., 2023). Matrix Gla protein (Mgp), which is a critical regulator of angiogenesis in multiple organs (Kida and Yamaguchi, 2022), was identified as exclusively upregulated in EOM cells upon activation (Figure 3H). Amongst the conserved EOM genes across cell states we identified *Pitx2*, a well-known major upstream regulator of EOM development (Gage et al., 1999), *Fos* gene family members (Fos and Fosb), with *Fos* being recently identified in a subset of limb MuSCs with enhanced regenerative capacity (Almada et al., 2021), and *Igfbp7*, a specific marker of quiescent MuSCs that is downregulated upon activation (Fukada et al., 2007) but also reported to be upregulated in MuSCs upon exercise (Chen et al., 2020) (**Figure 3F, Suppl Figure 3B, C**). Additionally, the EOM common signature included several ECM components and regulators (e.g. *Bgn*, *Loxl1*, *Col1a2*, *Col6a1*), TFs associated with fibrosis and connective tissue development (e.g. *Foxp1* (Grimaldi et al., 2022; Shao and Wei, 2018), *Egr1* (Havis and Duprez, 2020)) and *Acta2* (alpha smooth muscle actin), a marker of smooth muscle, fibroadipogenic progenitors (Joe et al., 2010; Uezumi et al., 2010) and smooth muscle-mesenchymal cells (SMMCs, (Giordani et al., 2019)) (**Figure 3F,H**). Finally, other top markers for EOM MuSCs identified in the quiescent microarray including *Foxc1,* which is involved in ocular development (Smith et al., 2000) and reported to regulate the balance between myogenic and vascular lineages within somites (Lagha 2009,Mayeuf-Louchart 2016), *Eya2, Islr* and *Igfbp4*, were also enriched in the activated EOM sc-RNAseq dataset (**Suppl Figure 3B**). Altogether, our analysis revealed a closer molecular overlap for EOM MuSCs across cell states, including several TFs and ECM markers that continued to be distinct from TA MuSCs.

EOM MuSC activation involves extensive ECM remodelling and a mesenchymal-like signature

Given the well-established role of ECM synthesis and remodelling on MuSC proliferation and self-renewal (Baghdadi et al., 2018; Rayagiri et al., 2018; Tierney et al., 2016; Yin et al., 2013), we sought to further characterise the EOM and TA signatures using the Matrisome database (Matrisome DB) (Naba et al., 2015), which compiles *in silico* and experimental data on ECM constituents. We first identified the components of the Matrisome DB present in our single-cell dataset and created a global score for ECM component expression (**Figure 4A**) (see Material and Methods). Significantly, the EOM progenitor

cluster had the highest matrisome score, while no major differences where observed within the other clusters (**Figure 4A**). Next, we identified the matrisome components that were differentially expressed in each cluster and found the EOM progenitor cluster also expressed the highest number of genes for each matrisome category when compared to the other clusters (**Figure 4B**).

We then validated some of the matrisome candidate genes at the protein level in activated MuSCs. EOM and TA MuSCs were isolated by FACS, cultured *in vitro* for 4 days and protein expression was assessed by immunofluorescence or Western blot (**Figure 4C-D**). Immunostaining for Fibronectin (FN1) showed a marked difference in fluorescence intensity between EOM and TA cells, confirming that EOM cells express higher levels of this protein (**Figure 4C**). Given that FN1 secretion by MuSCs promotes MuSC expansion in a cell autonomous manner (Bentzinger et al., 2013), the distinct expression of *Fn1* might contribute to maintenance of the progenitor population during EOM MuSC activation. EOM progenitors also had higher levels of Collagen I (COLI) (**Figure 4C**), a major component of the fibrotic ECM (Dulauroy et al., 2012) shown to suppress differentiation of C2C12 cells and display mutually exclusive expression with Myogenin (Alexakis et al., 2007). Similarly, EOM progenitors expressed higher levels of Collagen IV (COLIV) (**Figure 4C**), which was reported to be secreted by MuSCs as well as myoblasts and fibroblasts in culture (Baghdadi et al., 2018; Kühl et al., 1984). In addition, PDGFRβ, a tyrosine-kinase receptor commonly expressed by mesenchymal cells and pericytes (Hellström et al., 1999; Levéen et al., 1994; Soriano, 1994), was also differentially expressed in EOM progenitors by immunofluorescence (**Figure 4C**). Interestingly, Pdgfrβ and Acta2 were described as a markers of Smooth Muscle-Mesenchymal Cells (SMMCs), a novel Itga7+ Vcam- Pdgfrb+ Acta2+ cell subpopulation present in adult muscle that exhibits myogenic potential and promotes MuSC engraftment following transplantation (Giordani et al., 2019).

Western blot analysis confirmed that EOM progenitors produce higher levels of Caveolin1 (CAV1) and Cavin1 (**Figure 4D**), which are co-expressed in caveolae and shown to be downregulated upon differentiation of rhabdomyosarcoma cells (Faggi et al., 2015). CAV1 was also shown to be a marker of quiescent and Pax7+ activated mouse MuSCs (Gnocchi et al., 2009) whereas in human, the CAV1+ MuSC subpopulation is associated with ECM organization, expression of quiescent markers and display increased engraftment after transplantation (Barruet et al., 2020). SPARC, MGP and IGFBP7 were also found to be upregulated in EOM activated MuSCs by Western blot (**Figure 4D**). The activity of these proteins seems to be context dependent and they can promote or suppress proliferation in different cell types (Ahmad et al., 2017; Artico et al., 2021; Cho et al., 2000; Jing et al., 2019; Kuronuma et al., 2020; Li et al., 2020; Melouane et al., 2018; Said et al., 2013). Given the large number of matrisome genes differentially expressed in EOM progenitors, we also examined the expression of MMP2, a matrix remodelling protein (Gonçalves et al., 2022). Western Blot analysis revealed an enrichment in the active form of MMP2 in EOM progenitors whereas the TA progenitors displayed an enrichment of the pro-MMP2 or latent form (**Figure 4D)**. This result is in agreement with the observation that activation of MMP2 in vitro and in vivo increases the proportion and mobility of Pax7+ cells (Mu et al., 2010).

In parallel, we generated various gene signature scores to assess to what extent activated EOM MuSCs resemble other cells described in skeletal muscle displaying mesenchymal features (**Suppl Fig 4, Suppl Table 5**). As expected, EOM progenitors displayed a higher score for smooth muscle and mesenchymal cells (SMMCs, (Giordani et al., 2019)), FAPs (Oprescu et al., 2020), myotendinous junction B myonuclei (Kim et al., 2020), Twist2+ population(Liu et al., 2017), fetal MuSCs (Tierney et al., 2016), developing limb connective tissues (Lima et al., 2021) and the skeletal muscle mesenchyme identified in human fetal limb (Xi et al., 2020). Instead, TA progenitors and the respective EOM and TA differentiated fractions displayed a higher score for myogenic commitment and differentiation (**Suppl Fig 4**).

Finally, we focused on Pdgfrβ for further analysis as it is a component of the matrisome (**Figure 4B**) and a gene identified on the EOM molecular functions (**Suppl Figure 2C**) differentially expressed in EOM progenitors (**Figure 4E,F**). Pdgfrβ is a tyrosine-kinase receptor commonly expressed by pericytes (Hellström et al., 1999; Soriano, 1994) and it is involved in multiple signalling pathways that regulate cell migration, proliferation, and differentiation (Heldin and Westermark, 1999). We isolated this subpopulation by FACS from in vitro activated EOM cells using a PDGFRβ antibody. The flow cytometry analysis corroborated the scRNA-seq and immunofluorescence data (**Figure 4C**), where PDGFRβ was enriched in EOM samples (**Suppl Figure 4B,C**). To distinguish the functional properties of PDGFRβ+ and PDGFRβ- EOM cells, we sorted both subpopulations from EOM cultures, which represent respectively 83 and 17% of total nuclei at D4 and cultured them in growth promoting conditions for an additional 24h (**Figure 4G-I**). Cells were treated with EdU for 2h prior to fixation to monitor cell proliferation. PDGFRβ+ cells had a significantly higher proliferative capacity as measured by the total nuclear output and incorporation of EdU compared to PDGFRβ- cells (**Figure 4H,I**). Although the percentage of MYOD-expressing cells was not overtly different between the two populations, MYOG was expressed by over 50% of PDGFRβ- cells, whilst being detected in only 7% of PDGFRβ+ cells (**Figure 4H,I**). Therefore, the PDGFRβ+ myoblast subpopulation is characterised by a higher proliferative potential and decreased differentiation status.

Altogether, this analysis showed that EOM progenitors have an unusual transcriptome profile, express a wide range of ECM-related factors and harbour a mesenchymal signature that may endow them with a higher remodelling capacity of their immediate niche and promotion of an undifferentiated state.

EOM transcriptome profile is associated with a unique transcription factor network

To identify the underlying programs of myoblast heterogeneity between the two muscle groups, we inferred single cell regulatory networks using pySCENIC (Aibar et al., 2017; Sande et al., 2020). SCENIC uses co-expression patterns and transcription factor binding motifs to expose 'regulons" (transcription factors and their putative targets). This pipeline generates a new matrix comprising the activity of each regulon, in each cell (**Figure 5A, Suppl Table 6**). As expected, regulons associated with myoblast differentiation such as *Myod*, *Myog* and *Mef2* family were found to be specifically active in differentiated cells of both EOM and TA (**Figure 5B**). Of note, the top 5 regulons of TA progenitors were also found to be active in EOM progenitors, whereas the top 5 regulons of EOM progenitors were unique to this

cluster (**Figure 5B**). EOM progenitors were characterised by a unique signature of regulons involved in connective tissue/ECM remodelling including *Egr1* (Havis and Duprez, 2020) and Creb3l1, a downstream effector of Thr signaling (García et al., 2017) that plays a critical role during bone development and activates extracellular matrix genes such as Col1a1 and Fn1 (Sampieri et al., 2019). Top regulons of EOM progenitors are also involved in cell proliferation (*Foxc1* (Yang et al., 2017), *Sox4* (Moreno, 2019), *Fos* (Almada et al., 2021), *Klf6* (Dionyssiou et al., 2013), *Ebf1* (Györy et al., 2012)), commitment into endothelial and smooth muscle fates (*Foxc1* (Han et al., 2017; Mayeuf-Louchart et al., 2016; Whitesell et al., 2019; Yang et al., 2017)) as well as differentiation into mesenchymal lineages such as pericytes, adipocytes and chondrocytes (Ebf1, (El-Magd et al., 2021; Jimenez et al., 2007; Pagani et al., 2021). Notably, some of these regulons were also active in the EOM during development and implicated in non-myogenic cell fates from bipotent myogenic cells (Grimaldi et al., 2022). Interestingly also, for some of these EOM progenitor regulons, the genes themselves were already enriched in quiescent EOM MuSCs and their expression was upregulated upon activation (**Suppl Fig 5A, B**).

To visualise the potential interactions between these regulons, we built a network restricted to transcription factors. To do so, target genes that were not at the head of regulons themselves were removed (Grimaldi et al., 2022) (see Methods) (**Figure 5C).** This results in a layout that visually represents the relationships between the nodes, where each node (circle) is an active transcription factor and each edge (distance between nodes) is an inferred regulation between 2 transcription factors. When placed in a force-directed environment (see Methods), these nodes aggregated based on the number of shared edges. Thus, we were able to highlight associated and co-regulating transcription factor modules. Strikingly, the transcription factors of the most specific regulons of each cluster preferentially organised as tightly related modules (**Figure 5C**). In agreement with our previous analyses, the known co-regulating transcription factors found in differentiated cells (*Myod, Myog, Mef2a, Mef2c, Myf6*) formed a tight module in this visualisation (**Fig 5C**, common differentiating category). Interestingly, the TA muscle progenitor module was composed of genes that were reported to be required for limb embryonic development (*Hox* genes, *Lbx1*) (Gross et al., 2000; Pineault and Wellik, 2014; Swinehart et al., 2013). Notably, *Hoxa11*, is a determinant of limb axial identity during embryonic development (Zakany and Duboule, 2007), and HoxA and HoxC clusters were found as signatures of adult TA MuSCs (Evano et al., 2020; Yoshioka et al., 2021); **Suppl Fig5B**).

Within the EOM progenitor module, the network included *Foxc1, Egr1*, *Creb3l1*, *Dmrta2*, *Sox4*, *Fos* and *Egr1* transcription factors together with *Pax7* and *Hes1* which are known to support the maintenance of quiescence (Baghdadi et al., 2018; Mourikis et al., 2012; Olguin and Olwin, 2004; Relaix et al., 2005). In this context though, the network seems to maintain a proliferative progenitor state in activated EOM MuSCs cells as assessed by the higher protein levels of PAX7, CCND1 (Cyclin D1), together with other EOM regulon TFs (FOXC1, EBF1 and CREB3L1) (**Suppl Fig5C, D**). The vicinity of Hes1 to the EOM regulatory module prompted us to explore its direct upstream regulators (one level of ancestry in the network) (**Figure 5D**). Strikingly, 8 potential activators (including *Hes1* itself) were part of the EOM module. By contrast, only 2 upstream regulators were part of the TA regulatory module (**Figure 5D**). Of

note, *Fos* and *Egr1* have also been reported as part of a stress signature following tissue dissociation (Machado et al., 2021). Yet, in our dataset, the expression of these genes (i.e. the StressIndex) was correlated with the expression of *Pax7* and anti-correlated with the expression of *Myod* (**Suppl Fig5E,F**). Moreover, when we applied a regression of the StressIndex or removed these genes from the count matrix, the general aspect of the data did not change (**Suppl Fig5G)**, pointing at a role of these genes in the EOM progenitor maintenance independently of the stress response.

EOM features are present during the growth phase in vivo and retained upon passages in vitro

From the analysis above, it appears clear that EOM MuSCs possess an unusual transcriptomic state that may repress myogenic commitment and maintain a more "stem-like" state. We then wondered whether these EOM features are cell-wired and persist upon passage in vitro. Cells were isolated from *Tg:Pax7-nGFP;Myog*^{ntdTom} mice and analyzed upon in vitro activation or upon one or two passages (**Figure 6A, Suppl Fig 6A, B**). While the total cell number was reduced with passages for both EOM and TA MuSCs (**Suppl Fig 6A, B**), the normalized cellular output was consistently higher for the EOM and correlated with a lower tdT/GFP ratio (**Figure 6B, C**). Real-time qPCR analysis revealed that the expression of Pax7 and Hey1, a bHLH transcription factor that is required in a cell-autonomous manner for maintenance of MuSCs (Noguchi et al., 2019), EOM specific regulon TFs (Foxc1, Sox4, Ebf1, Creb3l1) and genes identified by the matrisome or molecular functions (Bgn, Sparc, Igfbp2, Igfbp7, Pdgfrb) were retained or even increased in EOM cultures after several passages compared to activated EOM MuSCs cultured only for 4 days (**Figure 6D**). These results were confirmed at the protein level by Western blot analysis (**Figure 6E, F**) where the specific EOM signature described above, that seems to maintain a more proliferative state in activated EOM MuSCs cells, persists even after several passages in vitro. Indeed, EOM MuSCs cells retained higher protein levels of PAX7, EOM regulon TFs (**Figure 6E**) as well as EOM matrisome or molecular functional genes (**Figure 6F**) compared to TA MuSCs cells at same passage. These results show that EOM MuSCs cells retain a cell autonomous non-canonical signature that is hard-wired even after extended cell culture.

Next, we asked whether EOM progenitor features are already present in activated fetal development and postnatal stages, where extensive muscle growth and MuSC expansion occurs, or if these features were acquired in adulthood upon reactivation from the quiescent state. To address these possibilities, we first isolated EOM and TA MuSCs from E18.5 and P21 *Tg:Pax7-nGFP* mice and performed qRT-PCR for Myogenin after 48h in culture (**Suppl Figure 7A**). As observed for adult MuSCs (**Figure 1D**), EOM cultures were less differentiated. However, as GFP protein persists during myogenic commitment (Sambasivan et al., 2009) and this can introduce a bias in the initial populations, we isolated the GFP+/tdT- fractions from *Tg:Pax7-nGFP;Myog*^{ntdTom} mice at P7 by FACS (**Figure 7A, Suppl Figure 7B**) and performed RT-qPCR. Similarly, to the *in vitro* sc-RNAseq data, significantly lower levels of *Myogenin* were detected in EOM GFP+/tdT- cells (**Figure 7B**). Finally, we confirmed higher transcript levels for the EOM specific regulon TFs (*Foxc1*, *Ebf1*, *Sox4*, *Creb3l1*) and matrisome or molecular functional components identified by sc-RNAseq of activated MuSCs *in vitro* (**Figure 7B**). Altogether, our results

indicate that EOM MuSCs repress myogenic commitment and maintain a more "stem-like" state upon activation, and this property is conserved from development to adulthood.

Foxc1 marks the EOM MuSC lineage and plays a role in progenitor cell maintenance

We then focused on Foxc1 for further analysis as it is one of the top regulons and DEG of the activated sc-RNAseq dataset (**Figure 5 A,B, Suppl Figure 3B, Suppl Figure 5B, D**) and expressed continuously in the EOM lineage. Notably also, bulk RNAseq from adult muscles (Terry et al., 2018) showed that Foxc1 expression is higher in entire EOMs than in any of the other 10 muscle groups analyzed (**Suppl Fig 7C**). Finally, immunostaining on tissue sections of *Pax7*^{CreERT2}*:R26*^{tdTOM} mice displayed colocalization of FOXC1 with tdTOMATO in EOM MuSCs and expression of FOXC1 in EOM myonuclei at P10 and in adulthood but not in the TA (**Figure 7C, D**). Thus, differences in Foxc1 expression between the EOM and TA seem to arise during development and are kept upon activation, and throughout the myogenic lineage.

As differences in *Myogenin* expression at the mRNA level are already evident at Day2 upon in vitro activation (**Figure 1G**), we decided to use this early timepoint to functionally validate a potential role of Foxc1 in the maintenance of the progenitor state. Thus, we silenced *Foxc1* via a siRNA approach in EOM activated MuSCs (**Suppl Figure 7D-F**). While the total cell number and percentage of Pax7 cells was not changed 2 days after Foxc1 silencing, immunostaining for FOXC1 showed silencing of 82% of the cells at the protein level (**Suppl Figure 7E**) and RT-qPCR showed a 64% reduction at the transcript level (**Suppl Figure 7F**). Moreover, a 2.7 fold increase in *Myogenin* expression was detected by qPCR following silencing of *Foxc1* as early as Day2 (**Suppl Figure 7F**). To look at later timepoints, we transduced EOM activated MuSCs with lentiviruses expressing different short-hairpin RNAs against Foxc1 (**Figure 7E**). Immunofluorescence and EdU uptake at D5 revealed an efficient depletion of Foxc1 protein, and concomitant severe reduction in the total cell number, and number of EdU+ cells (**Figure 7F-H**). We then overexpressed Foxc1 in activated TA MuSCs to assess whether overexpression of a single factor could confer at least a subset of EOM features to TA cells (**Figure 7E, I-L**). This approach resulted in a 3-fold increase in the total cell number (**Figure 7I**), number of FOXC1+ cells (**Figure 7J**) and reduction of Myogenin protein levels (**Figure 7K,L**). As *Foxc1* direct targets (**Suppl Table 7**) include matrisome components (e.g. *Sparc*, *Pdgfrβ*, *Fbn1*, **Figure 4B**) and other EOM specific regulon TFs (e.g. *Ebf1*, *Creb3l1*, *Egr1*, **Figure 5B**), we assessed expression of these genes upon *Foxc1* overexpression (**Figure 7L**). Many of the matrisome genes assessed were upregulated in this context.

Altogether, our findings support the notion that *Foxc1* is a key TF for maintenance of EOM MuSCs in a progenitor like state and matrix deposition. Overexpression of this factor in TA cells seems to be sufficient to recapitulate part of the EOM in vitro activated phenotype.

Transcription dynamics expose EOM and TA disparities in progenitor-state maintenance

We first attempted to correlate the number of regulons responsible for most of the matrisome gene expression differences between EOM and TA upon activation (**Figure 8A-C, Suppl Table 8**). We observed that activated EOM MuSCs consistently regulated a higher number of matrisome genes than TA cells independently of the number of regulons considered (**Figure 8A**). However, the ratio of the number of regulations of matrisome genes between EOM and TA activated MuSCs showed that the difference was maximal when considering the top 5 regulons, *Alx4*, *Dmrta2*, *Foxc1*, *Zmiz1* and *Fos* (**Figure 8B, C**). Notably, Dmrta2, Foxc1 and Fos were also active regulons in quiescence, with slight disparities between EOM and TA (**Figure 8D, Suppl Table 9**).

Second, we set out to determine whether matrix genes underlie the transition towards progenitors and committed cells during activation. To do so, we inferred RNA velocity using scVelo (Bergen et al., 2020). This method computes local changes in the relative amount of unspliced and spliced variants, which depend on the rates of transcription, degradation and splicing (Manno et al., 2018). ScVelo identifies candidate "driver genes", i.e. the most transcriptionally dynamic genes in a given cluster that are responsible for most of the inferred RNA velocity, and thus underlie its transitional trajectory. We applied the scVelo pipeline to EOM and TA datasets independently, and examined the directional trajectories underlying commitment and progenitor maintenance (**Figure 8E,F**). Two distinct velocity streams stood out in both datasets, towards differentiation (*Myog*^{High}) and towards a progenitor-state (*Pax7*^{High}). Strikingly, a larger fraction of EOM cells appeared to transition towards a progenitor state compared with the TA (**Figure 8E,F**). Conversely, most cells in the TA appeared to be directed towards differentiation, consistent with an overlapping transcriptomic profile with differentiating cells, shown as top DEGs (**Figure 2D**) and regulatory networks (**Figure 5C**). These trajectories did not appear to be specifically correlated with cell cycle phases (**Suppl Figure 8 A,B**), which was shown to influence transcriptomic data in some cases (McDavid et al., 2016). Hence, the velocity streams observed are most likely to reflect transitions between distinct cell states instead of the cell cycle progression of a homogeneous cell state.

Using scvelo built-in functions, we extracted the top driver genes underlying the velocity towards the progenitor state in both datasets (**Figure 8G-I, Suppl Table 10, 11**). Out of the top 100 driver genes, 30 were common to both datasets, including *ColV*, which plays a critical role in maintenance of quiescence (Baghdadi et al., 2018) (**Figure 8G,H**). We assessed the GO molecular functions associated with these driver genes to identify the unique and shared pathways employed by both datasets during myogenic progenitor maintenance (**Figure 8J**). In agreement with our previous results, this transition was characterised by the active upregulation of ECM components specifically in the EOM.

In summary, our results establish a new gene regulatory network (GRN) permitting an enhanced expansion capacity and delayed commitment to differentiation of EOM MuSCs upon activation. This regulatory module is hardwired to EOM identity and relies on the active maintenance of progenitor MuSC features.

DISCUSSION

One of the fundamental, yet unanswered questions in stem cell and regenerative biology regards the mechanisms that allow an orchestrated balance between proliferation, differentiation and self-renewal. An unusual feature of skeletal muscle stem cells is their reliance on distinct GRNs in different anatomical locations. How these pathways confer the unique functional properties of each stem cell population in cranial and limb/limb muscles remains poorly understood. Indeed, most studies on adult and developmental myogenesis have focused on trunk and limb muscles and only a handful of transcription factors and signalling pathways have been identified as hallmarks of specific muscle groups. In this study we performed scRNA-seq accompanied by experimental validations to characterise the GRN underlying the functional heterogeneity of MuSCs derived from different muscles, taking as EOM and TA as archetype muscles for high and low performing stem cell populations.

EOM MuSCs are more refractory to *in vitro* differentiation

Although subsets of cranial derived MuSCs were previously described to be more proliferative and to have a higher engraftment capacity than trunk MuSCs (Ono et al., 2010; Randolph et al., 2015; Stuelsatz et al., 2015), the mechanisms responsible for these phenotypic differences remained elusive. As these features depend on the balance between proliferation and differentiation, one possibility is that different upstream regulators define the pace at which MuSCs progress in the lineage in diverse anatomical positions.

Here, by monitoring *Myog* expression, we showed that EOM MuSCs have a lower propensity to differentiate following activation and thus persist as a proliferative population. As foetal and early postnatal EOM MuSCs are also refractory to differentiation, we propose that this property might be hardwired to some extent by unique GRNs that are retained throughout development and adulthood. Notably, foetal MuSCs were shown to be more resistant to myogenic progression upon in vitro expansion and contribute more efficiently upon transplantion than the adult counterparts (Sakai et al., 2013; Tierney et al., 2016). In addition, a recent study showed that limb MuSCs cells isolated at birth displayed prolonged expansion rate and a delay in differentiation and fusion compared with those isolated at later stages of postnatal growth (P7 and P15) and adulthood (Gattazzo et al., 2020). Therefore, is tempting to speculate that EOM MuSCs retain features of the respective foetal and neonatal precursor cells.

EOM MuSCs are strong candidates for cell-based therapies given their robust engraftment capacity in vivo (Stuelsatz et al., 2015). So far, major obstacles in the clinic are the large cell numbers required for transplantation and the fact that ex vivo amplification of trunk MuSCs leads to a drastic decline in regenerative potential due to commitment to differentiation (Briggs and Morgan, 2013; Ikemoto et al., 2007). Substantial advances have been made by modifying the MuSC culture conditions (Charville et al., 2015; L'honoré et al., 2018), or by the use of teratoma-derived skeletal myogenic progenitors (Xie et al., 2021). As such, the identification of factors that regulate cell fate decisions in distinct MuSC populations serves as a resource for advancing knowledge in the context of regenerative medicine.

EOM progenitors exhibit a mesenchymal genetic signature upon activation

Recent scRNA-seq analyses have provided some insights into the transcriptional landscape regulating MuSC quiescence, activation and self-renewal in somite-derived muscles (Dell'Orso et al., 2019; Hernando-Herraez et al., 2019; Machado et al., 2021; Micheli et al., 2020; Yartseva et al., 2020). To characterise these processes in cranial MuSCs, we performed a comparative scRNA-seq analysis of activated EOM and TA MuSCs. Distinct transcriptional profiles divided the myoblast pool into two subpopulations: those that resembled progenitors and those that were differentiating. The EOM progenitor population was characterised by a higher *Pax7* levels, and thus reminiscent of *in vitro* reserve cells (Laumonier et al., 2017; Yoshida et al., 1998; Zammit et al., 2004) and ontology analysis of its DEGs revealed a surprising enrichment in ECM organisation processes and *Pdgf* signalling.

By surveying bulk RNA sequencing and single-cell profiling data of trunk skeletal muscle stem cells, a recent study highlighted a dynamic profile of Pdgf ligands and receptors during myogenesis (Contreras et al., 2021). In addition, treatment with NOTCH and PDGFRβ ligands (DLL4 and PDGF-BB respectively) was shown to enhance migration, expression of stem cell markers and perivascular-like features in MuSCs (Gerli et al., 2019) and embryonic myoblasts (Cappellari et al., 2013). Given that PDGFRβ and several Notch pathway components are co-expressed in EOM progenitors, and at higher levels than the TA, we speculate that cross-talk between these pathways could take place in this subpopulation.

In addition to *Pdgfrβ,* which is a well-known pericyte marker (Lindahl et al., 1997), EOM progenitor cells express *Acta2* which is involved in the contractile apparatus of smooth muscle and also used as a SMMCs (Pdgfrb+ Itga7+ Vcam- Pdgfrb+ Acta2+) and FAPs (Giordani et al., 2019; Joe et al., 2010; Uezumi et al., 2010) marker. However, while the EOM progenitor fraction expressed *Fbn1* and *Loxl1*, which are markers of a new subtype of FAPs identified by scRNA-seq of whole muscle (Rubenstein et al., 2020), they were negative for canonical FAPs and pericyte markers such as *Pdgfrα* (Joe et al., 2010; Uezumi et al., 2010) and *Cspg4* (*Ng2*) (Birbrair et al., 2013), respectively. In addition, the EOM progenitor fraction also related more to fetal MuSCs than the TA counterparts in terms of enrichment in matrisome components such as Fn1, Fbn1, Vcam and collagens (Tierney et al., 2016). It is well established that cell autonomous ECM deposition and remodelling are key instructive cues governing cell fate decisions in MuSCs (Baghdadi et al., 2018; Lukjanenko et al., 2016; Mashinchian et al., 2018; Tierney et al., 2016). Thus, we hypothesize that EOM progenitors might exploit these features to preserve self-renewal potential and mimic their in vivo stem cell niche even after in vitro culture. In the EOM, this plasticity has already been observed in vivo, where EOM myogenic progenitors transition towards non-myogenic cell fates in areas deprived of neural crest cells during mouse embryogenesis (Grimaldi et al., 2022). Additionally, a resident subpopulation expressing reduced levels of canonical myogenic markers but increased levels of mesenchymal markers has been detected in embryonic and foetal human limb muscle (Xi et al., 2020). This plasticity seems to be present also *in vitro*, where directed myogenic differentiation of hiPSCs generates heterogeneous cell types including both myogenic and non-myogenic cells (Xi et al., 2020).

A unique network of transcription factors maintains EOM progenitors

Analysis of single cell regulatory networks (SCENIC) revealed a unique set of regulons whose activity was restricted to EOM progenitors. Of note, the majority of the identified head regulon genes in EOM progenitors are not typical myogenic TFs. One of the most active regulons in this cluster was *Foxc1*, which was described as a pro-mitogenic factor in the cancer field (Yang et al., 2017) and as driver of endothelial/smooth muscle fate alone (Whitesell et al., 2019) or in combination with Foxc2 (Mayeuf-Louchart et al., 2016). Another top regulon was *Egr1*, which was shown to promote the expression of many ECM-related genes in different tissues (Gaut et al., 2016; Guerquin et al., 2013; Milet et al., 2017; Wu et al., 2017). Interestingly, *Foxc1* and *Ebf1* were found as active transcription factors underlying the transition of myogenic towards non-myogenic cell fates in the EOM anlage during embryonic development (Grimaldi et al., 2022). Both of these transcription factors were found in EOM progenitors in this study, suggesting that they might be continuously active from development to adulthood specifically in the EOM. Interestingly, expression of Foxc1 by hair follicle stem cells activates quiescent gene networks upon activation (Wang et al., 2016). As *Foxc1* is expressed both by quiescent and activated EOM MuSCs, this gene might reinforce a progenitor/stem-cell identity throughout the lineage. Notably, *Pitx2* appears as a target of *Foxc1* in our analysis (Suppl Table 7) and its overexpression promotes MuSC proliferation in vitro and enhances the regenerative potential of dystrophic MuSCs in vivo (Vallejo et al., 2018). Other non-canonical myogenic TF identified in EOM progenitors are the Kruppel-like factor family members *Klf4*, *Klf6*, and *Klf9* with *Klf6* identified as top regulons. This gene family participates in the development and homeostasis of numerous tissues (McConnell and Yang, 2010), and *Klf4* is a pioneer factor well-known for its reprogramming capabilities (Schmidt and Plath, 2012)

When assessing the interactions between transcription factors, the EOM Progenitor regulon TFs clustered together in a regulatory module including *Pax7* and the Notch target *Hes1*, supporting our hypothesis that these cells are in a more stem/progenitor state. This module also included the transcription factor Fos/AP-1, whose expression in a subset of freshly isolated trunk MuSCs is related to enhanced regenerative properties such as rapid cell cycle entry, higher clonal capacity and more efficient engraftment (Almada et al., 2021). Further, by means of an inducible FOS expression system, it was reported that elevated FOS activity in recently activated muscle progenitor cells perturbs cellular differentiation by altering the 3D chromosome organization near critical pro-myogenic genes (Barutcu et al., 2022). Altogether, it is tempting to speculate that a common regulatory module allows the establishment and maintenance of a MuSC subpopulation in the EOMs that displays mesenchymal features, partially overlaps with the quiescent state, and is refractory to differentiation.

To investigate the maintenance of this stem state during EOM activation we used RNA velocity to infer the different directional trajectories and their driver genes. The proportion of cells transitioning to a more progenitor state was higher in the EOM when compared to TA, and this progenitor stream was driven by a specific gene set. Amongst these genes there was an enrichment of ECM-related factors, underscoring its active role in the maintenance of a less differentiated state. Interestingly, we identified

Igfbp7 (Insulin growth factor binding protein 7), as unique EOM driver gene, which has been found to be a specific marker of quiescent MuSCs (Fukada et al., 2007). Taken together, our results highlight an active maintenance of a progenitor pool that partially overlaps with the quiescent state during EOM MuSC activation, possibly through the steady induction of inhibitory differentiation signalling cues and promotion of ECM remodelling.

Conclusion

Using *in vitro* and *in silico* approaches, we propose a model where the outperformance of EOM MuSCs depends on the expression of a tightly associated module of transcription factors regulating a characteristic and distinct pattern of ECM-remodelling factors, cell receptors and growth factor binding proteins. These components define the pace at which EOM MuSCs progress through the myogenic lineage and maintenance of a stem-like population. As such, our study lays the groundwork for elucidating the mechanisms of selective sparing of muscle groups in dystrophic disease by providing information on a unique core GRN within this muscle group. In addition, our study could serve as basis for the optimization of in vitro protocols that would exploit cranial MuSC properties in the context of cell-based muscle therapies.

Materials and methods

Animal care

Animals were handled according to national and European Community guidelines and an ethics committee of the Institut Pasteur (CETEA) in France approved protocols. Except when indicated otherwise, males and females of 2-4 months were used. *Tg:Pax7-nGFP* (Sambasivan et al., 2009), *Myog*^{ntdTom} (Benavente-Diaz et al., 2021), *Pax7*^{CreERT2} (Mathew et al., 2011) and *R26*^{tdTom} (Madisen et al., 2009) mouse lines were maintained in a C57Bl/6JRj background. To induce recombination of *Pax7*^{CreERT2}*:R26*^{tdTom} mice a 20 mg/ml stock solution of Tamoxifen was prepared in 5% ethanol and 95% sunflower seed oil by thorough resuspension with rocking at 4 °C. For adult mice, 2 mg of tamoxifen (Sigma #T5648) were administered by gavage during 5 consecutive days and animals sacrificed 5 days later. To induce recombination of pups, the Tamoxifen stock solution was diluted to 15mg/ml with sunflower seed oil and 0.15 mg were administered daily by subcutaneous injection between P4 and P6 daily. Pups were sacrificed at P10 by decapitation and adult mice by cervical dislocation.

MuSC isolation by FACS

Muscles were dissected and minced in ice-cold DMEM. Samples were then incubated in DMEM (Fisher Scientific, 11594446), 0.08% Collagenase D (Sigma, 11088882001), 0.2% Trypsin (ThermoFisher, 15090) and 10 µg/mL of DNAse I (Sigma, 11284932) for 25 minutes at 37°C under gentle agitation for 5 rounds of digestion. After each round, samples were allowed to precipitate for 5 minutes, the supernatant was collected on 4mL of Foetal Bovine Serum (FBS) on ice and fresh digestion buffer was added to the remaining muscle pellet. The collected supernatants were centrifuged for 15 minutes at 550 g at 4°C, resuspended in DMEM 2% FBS and filtered through a 40µm strainer (Corning, 352235) before cell sorting.

Cells were isolated based on size, granularity and GFP fluorescence using an Aria III (BD Biosciences) sorter. Cells were collected directly in MuSC growth media (DMEM: F12 (1:1, GIBCO), 20% FBS (ThermoFisher, #10270), 2% Ultroser (Pall, 15950-017), 1% Penicillin/Streptomycin (GIBCO, 15140-122).

For myoblast isolation based on PDGFRβ expression, dishes were washed twice with PBS prior to trypsinization. Cells were subsequently incubated with α-PDGFRβ antibody in HBSS (Fisher Scientific, 11560616) media with 2% FBS for 30 minutes on ice. After two washes with HBSS 2% FBS, cells were incubated with streptavidin-coupled BV786 and incubated on ice for 25 minutes. After one wash with HBSS 2% FBS, the cell suspension filtered through a 40µm strainer (Corning, 352235) before cell sorting. Cells were isolated based on size, granulosity and GFP or BV786 fluorescence on using an Aria III (BD Biosciences) machine.

Muscle stem cell culture, treatment and transfection

Cells isolated by FACS were plated onto Matrigel® (Corning, 354248) coated dishes and cultured in growth media at 3% O2, 5% CO2, 37°C for the indicated time. To assess proliferation, cells were pulsed with 10^{-6} M of EdU (ThermoFisher, C10640) in cell culture media for 2h prior to fixation.

For loss of function experiments, freshly isolated satellite cells from Tg:Pax7-nGFP mice were transfected in suspension immediately after FACS with the ON-TARGET plus SMARTpool against FOXC1 (Dharmacon, L-047399-01-0005) or Scramble (Dharmacon, ON-TARGETplus Non-targeting Control siRNA, D0018100205) at 200 nM final concentration using Lipofectamine 3000 (ThermoFisher, L3000001) in Opti-MEM (Gibco) as described by the manufacturer. Briefly, a pre-mix of siRNA/Optimem (1.5ul /20ul) and Lipofectamine3000/Optimem (0.3ul / 20ul) were incubated separately 5 min RT, mixed at 1:1 ratio and incubated 15 min more at RT. 2.10^4 cells in 40ul of Optimem were incubated with an equal volume of the transfection mix for 2h at 3% O2, 5% CO2, 37°C in Eppendorf tubes whose caps had had been punctured with a needle to allow gas exchange. Two hours after transfection, three volumes of fresh growth medium were added and cells were plated at 10k/cm2 in Matrigel coated wells containing growth media. Two days upon transfection wells were processed for immunostaining or RNA collected with Tryzol as above.

Lentivirus transduction for gain and loss of function experiments

Control (mCherry), Foxc1-mCherry and Foxc1-shRNA containing vectors were made by VectorBuilder (https://en.vectorbuilder.com/). Freshly isolated satellite cells from *Tg:Pax7-nGFP* mice were plated in matrigel coated wells, cultured overnight and transduced at an MOI (multiplicity of infection) of 100 in 45μl (for 96 well plates) or 125μl (for 48 well plates) of MuSC media containing 5 μg/ml of polybrene. After 4h incubation at 37°C, cells were washed three times with MuSC and cultures for 4 more days prior to fixation for immunostaining, or protein or RNA collection.

RT-qPCR

RNA from in vitro myoblast cultures was extracted using a Trizol-based kit (Zymo Research, R2061) and reverse transcribed using SuperScriptIII (Invitrogen, 18080093). RT-qPCR to assess for mRNA relative expression performed with SYBR green master mix (Roche, 04913914001) in an Applied biosciences machine. Data analysis was performed using the 2-ΔΔCT method (Livak and Schmittgen, 2001) and mRNA expression was normalized with *Rpl13*.

Immunofluorescence

For immunostaining of ECM markers, cells were fixed in warm (37°C) 3% paraformaldehyde (PFA, Electron Microscopy Sciences) 2% sucrose. Upon incubation in the fixative for 15min at RT, cells were washed three times with PBS and incubated with NH4Cl 100mM in PBS to quench aldehyde groups. Cells were washed three times in PBS, permeabilized in 0.5% Triton X-100 (Merck, T8787) for 5 min RT, washed again three times in PBS and blocked with 10% goat serum (GIBCO). Cells were incubated with the indicated primary antibodies overnight at 4°C in PBS with 2% goat serum. Upon PBS washes, cells were incubated with secondary antibodies and Hoechst (ThermoFisher, H1399) for 45 min at RT. EdU was detected after antibody staining if needed according to manufacturer instructions (ThermoFisher, C10640).

Following immunostaining, 96-well plates (Dutscher Dominique, 655090) were imaged on an Opera Phenix high-content microscope (Perkin Elmer) using a 20x or 40x objective. Quantification of acquired

images was performed using Harmony (Perkin Elmer) or Columbus software, using an automated pipeline. First, nuclei were detected based on Hoechst signal. Number of nuclei was assessed and mean and maximum intensity were automatically quantified on the nuclear region for nuclear markers. For acquisition of ECM markers either the Opera Phenix high-content microscope (Perkin Elmer) or Zeiss LSM800 microscope with ZEN software were used.
Immunostaining on tissue sections was performed as (Comai et al., 2020). Immunostaining for FOXC1 on tissue sections required an amplification step using Alexa Fluor 488 Tyramide SuperBoost Kit (IgG rabbit, ThermoFisher, B40943).

Western Blot

Cells in culture were lysed in RIPA buffer (150mM NaCl, 50mM Tris pH8, 1mM EDTA, 1% Triton X100 (Sigma, T8787), 0.5% sodium deoxycholate, 0.1% SDS supplemented with 1X proteases (Sigma, S8820) and phosphatases (Roche, 4906845001) inhibitors) and kept on ice for 30 minutes. The homogenate was then centrifuged for 10 min at 3000g, 4°C. The protein concentration was determined by Bradford assay (Bio-Rad, 5000205) using a BSA standard curve. The protein absorbance was measured at 595 nm by using a microplate reader Infinite M2000 (Tecan). Equal amounts of protein were reconstituted in 4x Laemmli Sample Buffer (Bio-Rad, 1610747) and heated at 95°C for 5 min. 5 µg of protein extracts were run on a 4%–15% polyacrylamide gels (Mini Protein TGX Stain-Free gels, Bio-Rad) and transferred to PVDF (Bio-Rad, 1704156) membrane with Trans-Blot® Turbo™ Transfer system (Bio-Rad). The membrane was blocked with 5% milk or BSA (PAN-Biotech, P06-1391100) in Tris-Buffer Saline 0.2% Tween (Sigma, P9416) (TBS-T) for 1h at room temperature and probed with specific primary antibodies overnight at 4°C. After three washes in TBS-T, the membrane was incubated with HRP or fluorophore-conjugated secondary antibodies and revealed by chemiluminescence (Clarity Max Bio-Rad, 1705062) or fluorescence (Bio-Rad, Chemidoc MP) and detected using the ChemiDoc® Gel Imaging System. Densitometric analysis of the immunoblots was performed using Image Lab Software v.6.1.0 (Bio-Rad).

Time lapse microscopy

MuSC were plated on a microscopy culture chamber (IBIDI, 80826) and cultured in growth media. The plate was incubated at 37°C, 5% CO_2, and 3% O_2 (Zeiss, Pecon). A Zeiss Observer.Z1 connected to a Plan-Apochromat 20x/0.8 M27 objective and Hamamatsu Orca Flash 4 camera piloted with Zen (Zeiss) was used.

Image analysis

Cell tracking was performed using the Manual Tracking feature of the TrackMate plug-in (Tinevez et al., 2017) in Fiji (Schindelin et al., 2012). Fiji and Harmony (Perkin Elmer) were used for image analysis. Figures were assembled in Adobe Photoshop, Illustrator and InDesign (Adobe Systems).

Data analysis and statistics

Data analysis, statistics and visualisations were performed using Prism (Graphpad Software) or using R (Team, 2014) and the package ggplot2 (Wickham, 2009). For comparison between two groups, two tailed unpaired Student's t-tests were performed to calculate p values and to determine statistically significant differences (see Figure Legends).

scRNAseq data generation

MuSCs were isolated on BD FACSAria™ III based on GFP fluorescence and cell viability from *Tg:Pax7-nGFP mice* (Sambasivan et al., 2009). Quiescent MuSCs were manually counted using a hemocytometer and immediately processed for scRNA-seq. For activated samples, MuSCs were cultured *in vitro* as described above for four days. Activated MuSCs were subsequently trypsinized and washed in DMEM/F12 2% FBS. Live cells were re-sorted, manually counted using a hemocytometer and processed for scRNA-seq.

Prior to scRNAseq, RNA integrity was assessed using Agilent Bioanalyzer 2100 to validate the isolation protocol (RIN>8 was considered acceptable). 10X Genomics Chromium microfluidic chips were loaded with around 9000 and cDNA libraries were generated following manufacturer's protocol. Concentrations and fragment sizes were determined using Agilent Bioanalyzer and Invitrogen Qubit. cDNA libraries were sequenced using NextSeq 500 and High Output v2.5 (75 cycles) kits. Count matrices were subsequently generated following 10X Genomics Cell Ranger pipeline.

Following normalisation and quality control, we obtained an average of 5792 ± 1415 cells/condition.

Seurat preprocessing

scRNAseq datasets were processed using Seurat (https://satijalab.org/seurat/) (Butler et al., 2018). Cells with more than 10% of mitochondrial gene fraction were discarded. 4000-5000 genes were detected on average across all 4 datasets. Dimensionality reduction and UMAPs were generated following Seurat workflow. The top 100 DEGs were determined using Seurat "FindAllMarkers" function with default parameters. When processed independently (scvelo), the datasets were first regressed on cell cycle genes, mitochondrial fraction, number of genes, number of UMI following Seurat dedicated vignette, and doublets were removed using DoubletFinder v3 (McGinnis et al., 2019). A "StressIndex" score was generated for each cell based on the list of stress genes previously reported (Machado et al., 2021) using the "AddModule" Seurat function. 94 out of 98 genes were detected in the combined datasets. UMAPs were generated after 1. StressIndex regression, and 2. after complete removal of the detected stress genes from the gene expression matrix before normalization. In both cases, the overall aspect of the UMAP did not change significantly (**Figure S5**). Although immeasurable confounding effects of cell stress following isolation cannot be ruled out, we reasoned that our datasets did not show a significant effect of stress with respect to the conclusions of our study.

Matrisome analysis

After subsetting for the features of the Matrisome database (Naba et al., 2015) present in our single-cell dataset, the matrisome score was calculated by assessing the overall expression of its constituents using the "AddModuleScore" function from Seurat (Butler et al., 2018).

RNA velocity and driver genes

Scvelo was used to calculate RNA velocities (Bergen et al., 2020). Unspliced and spliced transcript matrices were generated using velocyto (Manno et al., 2018) command line function. Seurat-generated filtering, annotations and cell-embeddings (UMAP, tSNE, PCA) were then added to the outputted objects. These datasets were then processed following scvelo online guide and documentation. Velocity was calculated based on the dynamical model (using *scv.tl.recover_dynamics(adata)*, and *scv.tl.velocity(adata, mode='dynamical')*) and differential kinetics calculations were added to the model (using *scv.tl.velocity(adata, diff_kinetics=True)*). Specific driver genes were identified by determining the top likelihood genes in the selected cluster. The lists of top 100 drivers for EOM and TA progenitors are given in Suppl Tables 10 and 11.

Gene regulatory network inference and transcription factor modules

Gene regulatory networks were inferred using pySCENIC (Aibar et al., 2017; Sande et al., 2020). This algorithm regroups sets of correlated genes into regulons (i.e. a transcription factor and its targets) based on binding motifs and co-expression patterns. The top 35 regulons for each cluster was determined using scanpy "scanpy.tl.rank_genes_groups" function (method=t-test). Note that this function can yield less than 35 results depending on the cluster. UMAP and heatmap were generated using regulon AUC matrix (Area Under Curve) which refers to the activity level of each regulon in a given cell. Visualizations were performed using scanpy (Wolf et al., 2018). The outputted list of each regulon and their targets was subsequently used to create a transcription factor network. To do so, only genes that are regulons themselves were kept. This results in a visual representation where each node is an active transcription factor and each edge is an inferred regulation between 2 transcription factors. When placed in a force-directed environment, these nodes aggregate based on the number of shared edges. This operation greatly reduced the number of genes involved, while highlighting co-regulating transcriptional modules. Visualization of this network was performed in a force-directed graph using Gephi "Force-Atlas2" algorithm (https://gephi.org/). Of note, a force-directed graph is a type of visualization technique where nodes are positioned based on the principles of physics that assign forces among the set of edges and the set of nodes. Spring like attractive forces are used to attract pairs of edges towards each other (connected nodes) while repulsive forces, like those of electrically charged particles, are used to separate all pairs of nodes. In the equilibrium state for this system, the edges tend to have uniform length (because of the spring forces), and nodes that are not connected by an edge tend to be drawn further apart (because of the electrical repulsion).

Gene ontology analysis

Gene ontology analyses were performed on top 100 markers (obtained from Seurat function FindAllMarkers) or on top 100 driver genes (obtained from scvelo), using Cluego (Bindea et al., 2009). "GO Molecular Pathway", and "REACTOME pathways" were used independently to identify common and unique pathways involved in each dataset. In all analyses, an enrichment/depletion two-sided

hypergeometric test was performed and p-values were corrected using the Bonferroni step down method. Only the pathways with a p-value lower than 0.05 were displayed.

Literature scores

The scores for SMMCs (Giordani et al., 2019), FAPs (Oprescu et al., 2020), myotendinous junction B myonuclei (Kim et al., 2020), Twist2+ population(Liu et al., 2017), fetal MuSCs (Tierney et al., 2016), developing limb connective tissues (Lima et al., 2021) and the human skeletal muscle mesenchyme (Xi et al., 2020) were calculated by assessing the overall expression of of the markers of each population (Suppl Table 5) using the "AddModuleScore" function from Seurat (Butler et al., 2018).

Microarray data and analysis

EOM, TA, Esophagus, Masseter and Diaphragm muscles from three adult TgPax7nGFP mice were digested and processed for FACs as described before. GFP+ cells were collected and RNA extracted using the RNeasy Micro Kit (Qiagen). After validation of the RNA quality with Bioanalyzer 2100 (using Agilent RNA6000 Pico Chip Kit), 450 pg of total RNA is reverse transcribed following the Ovation Pico WTA System V2 (Tecan). Briefly, the resulting double strand cDNA is used for amplification based on SPIA technology. After purification according to Tecan protocol, 3.6 ug of Sens Target DNA are fragmented and biotin labelled using Encore Biotin Module kit (Tecan). After control of fragmentation using Bioanalyzer 2100, cDNA is then hybridized to GeneChip® MouseGene2.0 (Affymetrix) at 45°C for 17 hours.

After overnight hybridization, chips are washed on the fluidic station FS450 following specific protocols (Affymetrix) and scanned using the GCS3000 7G. The scanned images are then analyzed with Expression Console software (Affymetrix) to obtain raw data (Cel files) and metrics for Quality Controls. Data preprocessing with the oligo R package: Cel files were imported in R with the read.celfiles() function. Background subtraction and quantile normalization were performed using the Robust Multichip Average algorithm (rma() function).The obtained expression matrix was then annotated using the getNetAffx() from the oligo library in combination with the pd.mogene.2.0.st annotation package. Low expressed genes were filtered out using the genefilter R packages functions; first the genes expression standard deviation was computed using the rowSds() function, then the shorth of the calculated standard deviation was computed and all genes with sd above the shorth were kept for the downstream analysis. Differential genes expression analysis was conducted with the Limma bioconductor package. Limma was used to fit a linear model with the pairwise combination of the different anatomical locations as a set of contrasts. This linear model was then used to compute the moderated t-statistics with the eBayes() function. Finally, the multiple test across genes and contrasts was performed using the decideTests() function with default parameters.

Code availability

The code that was used to generate the TF network is available at this address: https://github.com/TajbakhshLab/TFnetwork

Acknowledgements

We acknowledge funding support from the Institut Pasteur, Agence Nationale de la Recerche (Laboratoire d'Exellence Revive, Investissement d'Avenir; ANR-10-LABX-73 to ST and ANR-21-CE13-0005 MUSE to GC), Association Française contre les Myopathies (Grant #20510 to ST and #23201 to GC), and the Centre National de la Recherche Scientifique. M.B.D. was supported by a grant from Laboratoire d'Excellence Revive and La Ligue Contre le Cancer. We gratefully acknowledge the Genomic Patform (Inserm U1016-CNRS UMRS 8104 - Université de Paris Cité) at Institut Cochin, the UtechS Photonic BioImaging (Imagopole), C2RT, Institut Pasteur, supported by the French National Research Agency (France BioImaging; ANR-10–INSB–04; Investments for the Future) and the Center for Translational Science (CRT)-Cytometry and Biomarkers Unit of Technology and Service (CB UTechS) at Institut Pasteur for support in conducting this study.

Author contributions

MBD, GC, and ST conceived the study. MDB, DDG and GC performed most of the experiments. PTL, SG, MM and BE contributed to in vitro experiments and image acquisition. AG, MBD and DDG performed sc-RNAseq. AG and MBD run the bioinformatic analysis with help from SM and VL. AG, MBD and GC wrote the manuscript with input from DDG and ST. GC and ST provided funding.

REFERENCES

Ahmad, S., Jan, A. T., Baig, M. H., Lee, E. J. and Choi, I. (2017). Matrix gla protein: An extracellular matrix protein regulates myostatin expression in the muscle developmental program. *Life Sci* **172**, 55–63.

Aibar, S., González-Blas, C. B., Moerman, T., Huynh-Thu, V. A. A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C. C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. *Nature methods* **14**, 1083–1086.

Alexakis, C., Partridge, T. and Bou-Gharios, G. (2007). Implication of the satellite cell in dystrophic muscle fibrosis: a self-perpetuating mechanism of collagen overproduction. *Am J Physiol-cell Ph* **293**, C661–C669.

Almada, A. E., Horwitz, N., Price, F. D., Gonzalez, A. E., Ko, M., Bolukbasi, O. V., Messemer, K. A., Chen, S., Sinha, M., Rubin, L. L., et al. (2021). FOS licenses early events in stem cell activation driving skeletal muscle regeneration. *Cell Reports* **34**, 108656.

Andrés, V. and Walsh, K. (1996). Myogenin expression, cell cycle withdrawal, and phenotypic differentiation are temporally separable events that precede cell fusion upon myogenesis. *J Cell Biology* **132**, 657–66.

Artico, L. L., Laranjeira, Angelo Brunelli Albertoni, Campos, Livia Weijenborg, Correa, J. R., Zenatti, P. P., Carvalheira, J. B. C., Brambilla, S. R., Nowill, A. E., Brandalise, S. R. and Junes, J. A. (2021). Physiologic IGFBP7 levels prolong IGF1R activation in acute lymphoblastic leukemia. *Blood advances* **18**,.

Baghdadi, M. B., Castel, D., Machado, L., Fukada, S., Birk, D. E., Relaix, F., Tajbakhsh, S. and Mourikis, P. (2018). Reciprocal signalling by Notch–Collagen V–CALCR retains muscle stem cells in their niche. *Nature* **557**, 714–718.

Barruet, E., Garcia, S. M., Striedinger, K., Wu, J., Lee, S., Byrnes, L., Wong, A., Xuefeng, S., Tamaki, S., Brack, A. S., et al. (2020). Functionally heterogeneous human satellite cells identified by single cell RNA sequencing. *Elife* **9**, e51576.

Barutcu, A. R., Elizalde, G., Gonzalez, A. E., Soni, K., Rinn, J. L., Wagers, A. J. and Almada, A. E. (2022). Prolonged FOS activity disrupts a global myogenic transcriptional program by altering 3D chromatin architecture in primary muscle progenitor cells. *Skelet Muscle* **12**, 20.

Benavente-Diaz, M., Comai, G., Girolamo, D. D., Langa, F. and Tajbakhsh, S. (2021). Dynamics of myogenic differentiation using a novel Myogenin knock-in reporter mouse. *Skelet Muscle* **11**, 5.

Bentzinger, C. F., Wang, Y. X., Dumont, N. A. and Rudnicki, M. A. (2013). Cellular dynamics in the muscle satellite cell niche. *Embo Rep* **14**, 1062–72.

Bergen, V., Lange, M., Peidli, S., Wolf, F. A. and Theis, F. J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol* 1–7.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.-H., Pagès, F., Trajanoski, Z. and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093.

Birbrair, A., Zhang, T., Wang, Z.-M., Messi, M. L., Enikolopov, G. N., Mintz, A. and Delbono, O. (2013). Role of Pericytes in Skeletal Muscle Regeneration and Fat Accumulation. *Stem Cells Dev* **22**, 2298–2314.

Briggs, D. and Morgan, J. E. (2013). Recent progress in satellite cell/myoblast engraftment – relevance for therapy. *FEBS Journal* **280**, 4281–4293.

Brink, S. C. van den, Sage, F., Vértesy, Á., Spanjaard, B., Peterson-Maduro, J., Baron, C. S., Robin, C. and Oudenaarden, A. van (2017). Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat Methods* **14**, 935–936.

Brohmann, H., Jagla, K. and Birchmeier, C. (2000). The role of Lbx1 in migration of muscle precursor cells. *Development* **127**, 437–445.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E. and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411–420.

Cappellari, O., Benedetti, S., Innocenzi, A., Tedesco, F. S., Moreno-Fortuny, A., Ugarte, G., Lampugnani, M. G., Messina, G. and Cossu, G. (2013). Dll4 and PDGF-BB Convert Committed Skeletal Myoblasts to Pericytes without Erasing Their Myogenic Memory. *Dev Cell* **24**, 586–599.

Chakkalakal, J. V., Christensen, J., Xiang, W., Tierney, M. T., Boscolo, F. S., Sacco, A. and Brack, A. S. (2014). Early forming label-retaining muscle stem cells require p27kip1 for maintenance of the primitive state. *Development* **141**, 1649–1659.

Charville, G. W., Cheung, T. H., Yoo, B., Santos, P. J., Lee, G. K., Shrager, J. B. and Rando, T. A. (2015). Ex Vivo Expansion and In Vivo Self-Renewal of Human Muscle Stem Cells. *Stem Cell Reports* **5**, 621–632.

Chen, Z., Li, L., Wu, W., Liu, Z., Huang, Y., Yang, L., Luo, Q., Chen, J., Hou, Y. and Song, G. (2020). Exercise protects proliferative muscle satellite cells against exhaustion via the Igfbp7-Akt-mTOR axis. *Theranostics* **10**, 6448–6466.

Cho, W. J., Kim, E. J., Lee, S. J., Kim, H. D., Shin, H. J. and Lim, W. K. (2000). Involvement of SPARC in in Vitro Differentiation of Skeletal Myoblasts. *Biochem Bioph Res Co* **271**, 630–634.

Cho, D. S., Schmitt, R. E., Dasgupta, A., Ducharme, A. M. and Doles, J. D. (2020). Single-cell deconstruction of post-sepsis skeletal muscle and adipose tissue microenvironments. *J Cachexia Sarcopenia Muscle* **11**, 1351–1363.

Comai, G. E., Tesařová, M., Dupé, V., Rhinn, M., Vallecillo-García, P., Silva, F. da, Feret, B., Exelby, K., Dollé, P., Carlsson, L., et al. (2020). Local retinoic acid signaling directs emergence of the extraocular muscle functional unit. *Plos Biol* **18**, e3000902.

Contreras, O., Córdova-Casanova, A. and Brandan, E. (2021). PDGF-PDGFR network differentially regulates the fate, migration, proliferation, and cell cycle progression of myogenic cells. *Cell Signal* **84**, 110036.

Dell'Orso, S., Juan, A. H., Ko, K.-D., Naz, F., Perovanovic, J., Gutierrez-Cruz, G., Feng, X. and Sartorelli, V. (2019). Single cell analysis of adult mouse skeletal muscle stem cells in homeostatic and regenerative conditions. *Development* **146**, dev174177.

Diehl, A. G., Zareparsi, S., Qian, M., Khanna, R., Angeles, R. and Gage, P. J. (2006). Extraocular Muscle Morphogenesis and Gene Expression Are Regulated by Pitx2 Gene Dose. *Investigative Ophthalmology & Visual Science* **47**, 1785–1793.

Dionyssiou, M. G., Salma, J., Bevzyuk, M., Wales, S., Zakharyan, L. and McDermott, J. C. (2013). Krüppel-like factor 6 (KLF6) promotes cell proliferation in skeletal myoblasts in response to TGFβ/Smad3 signaling. *Skelet Muscle* **3**, 7.

Dulauroy, S., Carlo, S. E., Langa, F., Eberl, G. and Peduto, L. (2012). Lineage tracing and genetic ablation of ADAM12+ perivascular cells identify a major source of profibrotic cells during acute tissue injury. *Nature Medicine* **18**, 1262–1270.

Duly, A. M. P., Kao, F. C. L., Teo, W. S. and Kavallaris, M. (2022). βIII-Tubulin Gene Regulation in Health and Disease. *Frontiers Cell Dev Biology* **10**, 851542.

Dumont, N. A., Wang, Y. X., Maltzahn, J. von, Pasut, A., Bentzinger, C. F., Brun, C. E. and Rudnicki, M. A. (2015). Dystrophin expression in muscle stem cells regulates their polarity and asymmetric division. *Nat Med* **21**, 1455–1463.

El-Magd, M. A., Abdelfattah-Hassan, A., Elsisy, R. A., Hawsawi, Y. M., Oyouni, A. A., Al-Amer, O. M. and El-Shetry, E. S. (2021). Expression and function of Ebf1 gene during chondrogenesis in chick embryo limb buds. *Gene* **803**, 145895.

Emery, A. E. (2002). The muscular dystrophies. *Lancet* **359**, 687–695.

Evano, B. and Tajbakhsh, S. (2018). Skeletal muscle stem cells in comfort and stress. *npj Regenerative Medicine* **3**, 24.

Evano, B., Gill, D., Hernando-Herraez, I., Comai, G., Stubbs, T. M., Commere, P.-H., Reik, W. and Tajbakhsh, S. (2020). Transcriptome and epigenome diversity and plasticity of muscle stem cells following transplantation. *Plos Genet* **16**, e1009022.

Faggi, F., Chiarelli, N., Colombi, M., Mitola, S., Ronca, R., Madaro, L., Bouche, M., Poliani, P. L., Vezzoli, M., Longhena, F., et al. (2015). Cavin-1 and Caveolin-1 are both required to support cell proliferation, migration and anchorage-independent cell growth in rhabdomyosarcoma. *Lab Invest* **95**, 585–602.

Formicola, L., Marazzi, G. and Sassoon, D. A. (2014). The extraocular muscle stem cell niche is resistant to ageing and disease. *Front Aging Neurosci* **6**, 328.

Fukada, S., Uezumi, A., Ikemoto, M., Masuda, S., Segawa, M., Tanimura, N., Yamamoto, H., Miyagoe-Suzuki, Y. and Takeda, S. (2007). Molecular Signature of Quiescent Satellite Cells in Adult Skeletal Muscle. *Stem Cells* **25**, 2448–2459.

Gage, P. J., Suh, H. and Camper, S. A. (1999). Dosage requirement of Pitx2 for development of multiple organs. *Development* **126**, 4643–4651.

García, I. A., Demichelis, V. T., Viale, D. L., Giusto, P. D., Ezhova, Y., Polishchuk, R. S., Sampieri, L., Martinez, H., Sztul, E. and Alvarez, C. (2017). CREB3L1-mediated functional and structural adaptation of the secretory pathway in hormone-stimulated thyroid cells. *J Cell Sci* **130**, 4155–4167.

Gattazzo, F., Laurent, B., Relaix, F., Rouard, H. and Didier, N. (2020). Distinct Phases of Postnatal Skeletal Muscle Growth Govern the Progressive Establishment of Muscle Stem Cell Quiescence. *Stem Cell Rep*.

Gaut, L., Robert, N., Delalande, A., Bonnin, M.-A., Pichon, C. and Duprez, D. (2016). EGR1 Regulates Transcription Downstream of Mechanical Signals during Tendon Formation and Healing. *Plos One* **11**, e0166237.

Gayraud-Morel, B., Chrétien, F., Jory, A., Sambasivan, R., Negroni, E., Flamant, P., Soubigou, G., Coppée, J.-Y., Santo, J., Cumano, A., et al. (2012). Myf5 haploinsufficiency reveals distinct cell fate potentials for adult skeletal muscle stem cells. *Journal of Cell Science* **125**, 1738–1749.

Gerli, M., Moyle, L., Benedetti, S., Ferrari, G., Ucuncu, E., Ragazzi, M., Constantinou, C., Louca, I., Sakai, H., Ala, P., et al. (2019). Combined Notch and PDGF Signaling Enhances Migration and Expression of Stem Cell Markers while Inducing Perivascular Cell Features in Muscle Satellite Cells. *Stem Cell Rep* **12**, 461–473.

Giordani, L., He, G. J., Negroni, E., Sakai, H., Law, J., Siu, M. M., Wan, R., Corneau, A., Tajbakhsh, S., Cheung, T. H., et al. (2019). High-Dimensional Single-Cell Cartography Reveals Novel Skeletal Muscle-Resident Cell Populations. *Mol Cell* **74**, 609-621.e6.

Gnocchi, V. F., White, R. B., Ono, Y., Ellis, J. A. and Zammit, P. S. (2009). Further Characterisation of the Molecular Signature of Quiescent and Activated Mouse Muscle Satellite Cells. *Plos One* **4**, e5205.

Gonçalves, P. R., Nascimento, L. D., Gerlach, R. F., Rodrigues, K. E. and Prado, A. F. (2022). Matrix Metalloproteinase 2 as a Pharmacological Target in Heart Failure. *Pharm* **15**, 920.

Gopalakrishnan, S., Comai, G., Sambasivan, R., Francou, A., Kelly, R. G. and Tajbakhsh, S. (2015). A Cranial Mesoderm Origin for Esophagus Striated Muscles. *Developmental Cell* **34**, 694–704.

Grimaldi, A. and Tajbakhsh, S. (2021). Diversity in cranial muscles: Origins and developmental programs. *Curr Opin Cell Biol* **73**, 110–116.

Grimaldi, A., Comai, G., Mella, S. and Tajbakhsh, S. (2022). Identification of bipotent progenitors that give rise to myogenic and connective tissues in mouse. *Elife* **11**, e70235.

Gross, M., Moran-Rivard, L., Velasquez, T., Nakatsu, M., Jagla, K. and Goulding, M. (2000). Lbx1 is required for muscle precursor migration along a lateral pathway into the limb. *Development (Cambridge, England)* **127**, 413–24.

Guerquin, M.-J., Charvet, B., Nourissat, G., Havis, E., Ronsin, O., Bonnin, M.-A., Ruggiu, M., Olivera-Martinez, I., Robert, N., Lu, Y., et al. (2013). Transcription factor EGR1 directs tendon differentiation and promotes tendon repair. *J Clin Invest* **123**, 3564–3576.

Günther, S., Mielcarek, M., Krüger, M. and Braun, T. (2004). VITO-1 is an essential cofactor of TEF1-dependent muscle-specific gene regulation. *Nucleic Acids Res* **32**, 791–802.

Guo, K., Wang, J., Andrés, V., Smith, R. C. and Walsh, K. (1995). MyoD-induced expression of p21 inhibits cyclin-dependent kinase activity upon myocyte terminal differentiation. *Mol Cell Biol* **15**, 3823–9.

Györy, I., Boller, S., Nechanitzky, R., Mandel, E., Pott, S., Liu, E. and Grosschedl, R. (2012). Transcription factor Ebf1 regulates differentiation stage-specific signaling, proliferation, and survival of B cells. *Gene Dev* **26**, 668–682.

Han, B., Bhowmick, N., Qu, Y., Chung, S., Giuliano, A. E. and Cui, X. (2017). FOXC1: an emerging marker and therapeutic target for cancer. *Oncogene* **36**, 3957–3963.

Harel, I., Nathan, E., Tirosh-Finkel, L., Zigdon, H., Guimarães-Camboa, N., Evans, S. M. and Tzahor, E. (2009). Distinct Origins and Genetic Programs of Head Muscle Satellite Cells. *Dev Cell* **16**, 822–832.

Havis, E. and Duprez, D. (2020). EGR1 Transcription Factor is a Multifaceted Regulator of Matrix Production in Tendons and Other Connective Tissues. *Int J Mol Sci* **21**, 1664.

Hellström, M., Kalén, M., Lindahl, P., Abramsson, A. and Betsholtz, C. (1999). Role of PDGF-B and PDGFR-beta in recruitment of vascular smooth muscle cells and pericytes during embryonic blood vessel formation in the mouse. *Dev Camb Engl* **126**, 3047–55.

Hernando-Herraez, I., Evano, B., Stubbs, T., Commere, P.-H., Bonder, M. J., Clark, S., Andrews, S., Tajbakhsh, S. and Reik, W. (2019). Ageing affects DNA methylation drift and transcriptional cell-to-cell variability in mouse muscle stem cells. *Nat Commun* **10**, 4361.

Honda, M., Hidaka, K., Fukada, S., Sugawa, R., Shirai, M., Ikawa, M. and Morisaki, T. (2017). Vestigial-like 2 contributes to normal muscle fiber type distribution in mice. *Sci Rep-uk* **7**, 7168.

Ikemoto, M., Fukada, S., Uezumi, A., Masuda, S., Miyoshi, H., Yamamoto, H., Wada, M. R., Masubuchi, N., Miyagoe-Suzuki, Y. and Takeda, S. (2007). Autologous Transplantation of SM/C-2.6+ Satellite Cells Transduced with Micro-dystrophin CS1 cDNA by Lentiviral Vector into mdx Mice. *Molecular Therapy* **15**, 2178–2185.

Jen, Y., Weintraub, H. and Benezra, R. (1992). Overexpression of Id protein inhibits the muscle differentiation program: in vivo association of Id with E2A proteins. *Gene Dev* **6**, 1466–1479.

Jimenez, M. A., Åkerblad, P., Sigvardsson, M. and Rosen, E. D. (2007). Critical Role for Ebf1 and Ebf2 in the Adipogenic Transcriptional Cascade. *Mol Cell Biol* **27**, 743–757.

Jing, Y., Jin, Y., Wang, Y., Chen, S., Zhang, X., Song, Y., Wang, Z., Pu, Y., Ni, Y. and Hu, Q. (2019). SPARC promotes the proliferation and metastasis of oral squamous cell carcinoma by PI3K/AKT/PDGFB/PDGFRβ axis. *J Cell Physiol* **234**, 15581–15593.

Joe, A. W. B., Yi, L., Natarajan, A., Grand, F. L., So, L., Wang, J., Rudnicki, M. A. and Rossi, F. M. V. (2010). Muscle injury activates resident fibro/adipogenic progenitors that facilitate myogenesis. *Nat Cell Biol* **12**, 153–63.

Keefe, A. C., Lawson, J. A., Flygare, S. D., Fox, Z. D., Colasanto, M. P., Mathew, S. J., Yandell, M. and Kardon, G. (2015). Muscle stem cells contribute to myofibres in sedentary adult mice. *Nat Commun* **6**, 7087.

Kelly, R. G., Jerome-Majewska, L. A. and Papaioannou, V. E. (2004). The del22q11.2 candidate gene Tbx1 regulates branchiomeric myogenesis. *Human Molecular Genetics* **13**, 2829–2840.

Kida, Y. and Yamaguchi, I. (2022). The vascular protective effect of matrix Gla protein during kidney injury. *Frontiers Mol Medicine* **2**, 970744.

Kim, M., Franke, V., Brandt, B., Lowenstein, E. D., Schöwel, V., Spuler, S., Akalin, A. and Birchmeier, C. (2020). Single-nucleus transcriptomics reveals functional compartmentalization in syncytial skeletal muscle cells. *Nat Commun* **11**, 6375.

Kühl, U., Öcalan, M., Timpl, R., Mayne, R., Hay, E. and Mark, K. von der (1984). Role of muscle fibroblasts in the deposition of type-IV collagen in the basal lamina of myotubes. *Differentiation* **28**, 164–172.

Kui, H., Ran, B., Yang, M., Shi, X., Luo, Y., Wang, Y., Wang, T., Li, D., Shuai, S. and Li, M. (2022). Gene expression profiles of specific chicken skeletal muscles. *Sci Data* **9**, 552.

Kumar, D., Shadrach, J. L., Wagers, A. J. and Lassar, A. B. (2009). Id3 Is a Direct Transcriptional Target of Pax7 in Quiescent Satellite Cells. *Mol Biol Cell* **20**, 3170–3177.

Kuronuma, K., Yokoi, A., Fukuoka, T., Miyata, M., Maekawa, A., Tanaka, S., Matsubara, L., Goto, C., Matsuo, M., Han, H.-W., et al. (2020). Matrix Gla protein maintains normal and malignant hematopoietic progenitor cells by interacting with bone morphogenetic protein-4. *Heliyon* **6**, e03743.

Laumonier, T., Bermont, F., Hoffmeyer, P., Kindler, V. and Menetrey, J. (2017). Human myogenic reserve cells are quiescent stem cells that contribute to muscle regeneration after intramuscular transplantation in immunodeficient mice. *Sci Rep-uk* **7**, 3462.

Levéen, P., Pekny, M., Gebre-Medhin, S., Swolin, B., Larsson, E. and Betsholtz, C. (1994). Mice deficient for PDGF B show renal, cardiovascular, and hematological abnormalities. *Gene Dev* **8**, 1875–1887.

L'honoré, A., Commère, P.-H., Negroni, E., Pallafacchina, G., Friguet, B., Drouin, J., Buckingham, M. and Montarras, D. (2018). The role of Pitx2 and Pitx3 in muscle stem cells gives new insights into P38α MAP kinase and redox regulation of muscle regeneration. *eLife* **7**, e32991.

Li, X., Wei, R., Wang, M., Ma, L., Zhang, Z., Chen, L., Guo, Q., Guo, S., Zhu, S., Zhang, S., et al. (2020). MGP Promotes Colon Cancer Proliferation by Activating the NF-κB Pathway through Upregulation of the Calcium Signaling Pathway. *Mol Ther - Oncolytics* **17**, 371–383.

Lima, J. E. de, Blavet, C., Bonnin, M.-A., Hirsinger, E., Comai, G., Yvernogeau, L., Delfini, M.-C., Bellenger, L., Mella, S., Nassari, S., et al. (2021). Unexpected contribution of fibroblasts to muscle lineage as a mechanism for limb muscle patterning. *Nat Commun* **12**, 3851.

Lindahl, P., Johansson, B. R., Levéen, P. and Betsholtz, C. (1997). Pericyte Loss and Microaneurysm Formation in PDGF-B-Deficient Mice. *Science* **277**, 242–245.

Liu, N., Garry, G. A., Li, S., Bezprozvannaya, S., Sanchez-Ortiz, E., Chen, B., Shelton, J. M., Jaichander, P., Bassel-Duby, R. and Olson, E. N. (2017). A Twist2-dependent progenitor cell contributes to adult skeletal muscle. *Nat Cell Biol* **19**, 202–213.

Livak, K. J. and Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* **25**, 402–408.

Lukjanenko, L., Jung, J. M., Hegde, N., Perruisseau-Carrier, C., Migliavacca, E., Rozo, M., Karaz, S., Jacot, G., Schmidt, M., Li, L., et al. (2016). Loss of fibronectin from the aged stem cell niche affects the regenerative capacity of skeletal muscle in mice. *Nature Medicine* **22**, 897–905.

Machado, L., Lima, J. E. de, Fabre, O., Proux, C., Legendre, R., Szegedi, A., Varet, H., Ingerslev, L. R., Barrès, R., Relaix, F., et al. (2017). In Situ Fixation Redefines Quiescence and Early Activation of Skeletal Muscle Stem Cells. *Cell Reports* **21**, 1982–1993.

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

Madisen, L., Zwingman, T. A., Sunkin, S. M., Oh, S., Zariwala, H. A., Gu, H., Ng, L. L., Palmiter, R. D., Hawrylycz, M. J., Jones, A. R., et al. (2009). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neuroscience* **13**, 133–140.

Maeda, T., Chapman, D. L. and Stewart, A. F. R. (2002). Mammalian Vestigial-like 2, a Cofactor of TEF-1 and MEF2 Transcription Factors That Promotes Skeletal Muscle Differentiation*. *J Biol Chem* **277**, 48889–48898.

Manno, G. L., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M. E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. *Nature* **560**, 494–498.

Mashinchian, O., Pisconti, A., Moal, E. and Bentzinger, F. C. (2018). Chapter Two The Muscle Stem Cell Niche in Health and Disease. *Curr Top Dev Biol* **126**, 23–65.

Mathew, S. J., Hansen, J. M., Merrell, A. J., Murphy, M. M., Lawson, J. A., Hutcheson, D. A., Hansen, M. S., Angus-Hill, M. and Kardon, G. (2011). Connective tissue fibroblasts and Tcf4 regulate myogenesis. *Development* **138**, 371–384.

Mayeuf-Louchart, A., Montarras, D., Bodin, C., Kume, T., Vincent, S. D. and Buckingham, M. (2016). Endothelial cell specification in the somite is compromised in Pax3-positive progenitors of Foxc1/2 conditional mutants, with loss of forelimb myogenesis. *Development* **143**, 872–879.

McConnell, B. B. and Yang, V. W. (2010). Mammalian Krüppel-Like Factors in Health and Diseases. *Physiol Rev* **90**, 1337–1381.

McDavid, A., Finak, G. and Gottardo, R. (2016). The contribution of cell cycle to heterogeneity in single-cell RNA-seq data. *Nat Biotechnol* **34**, 591–593.

McGinnis, C. S., Murrow, L. M. and Gartner, Z. J. (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* **8**, 329-337.e4.

McLoon, L. K. and Wirtschafter, J. (2002). Activated satellite cells are present in uninjured extraocular muscles of mature mice. *T Am Ophthal Soc* **100**, 119–23; discussion 123-4.

McLoon, L. K. and Wirtschafter, J. (2003). Activated Satellite Cells in Extraocular Muscles of Normal Adult Monkeys and Humans. *Investigative Opthalmology Vis Sci* **44**, 1927.

Melouane, A., Carbonell, A., Yoshioka, M., Puymirat, J. and St-Amand, J. (2018). Implication of SPARC in the modulation of the extracellular matrix and mitochondrial function in muscle cells. *Plos One* **13**, e0192714.

Michailovici, I., Eigler, T. and Tzahor, E. (2015). Chapter One Craniofacial Muscle Development. *Curr Top Dev Biol* **115**, 3–30.

Micheli, A. J. D., Laurilliard, E. J., Heinke, C. L., Ravichandran, H., Fraczek, P., Soueid-Baumgarten, S., Vlaminck, I., Elemento, O. and Cosgrove, B. D. (2020). Single-Cell Analysis of the Muscle Stem Cell Hierarchy Identifies Heterotypic Communication Signals Involved in Skeletal Muscle Regeneration. *Cell Reports* **30**, 3583-3595.e5.

Milet, C., Bléher, M., Allbright, K., Orgeur, M., Coulpier, F., Duprez, D. and Havis, E. (2017). Egr1 deficiency induces browning of inguinal subcutaneous white adipose tissue in mice. *Sci Rep-uk* **7**, 16153.

Moreno, C. S. (2019). SOX4: The Unappreciated Oncogene. *Semin Cancer Biol* **67**, 57–64.

Morree, A. de, Klein, J. D. D., Gan, Q., Farup, J., Urtasun, A., Kanugovi, A., Bilen, B., Velthoven, C. T. J. van, Quarta, M. and Rando, T. A. (2019). Alternative polyadenylation of Pax3 controls muscle stem cell fate and muscle function. *Science* **366**, 734–738.

Mourikis, P., Sambasivan, R., Castel, D., Rocheteau, P., Bizzarro, V. and Tajbakhsh, S. (2012). A Critical Requirement for Notch Signaling in Maintenance of the Quiescent Skeletal Muscle Stem Cell State. *Stem Cells* **30**, 243–252.

Mu, X., Urso, M. L., Murray, K., Fu, F. and Li, Y. (2010). Relaxin Regulates MMP Expression and Promotes Satellite Cell Mobilization During Muscle Healing in Both Young and Aged Mice. *Am J Pathology* **177**, 2399–2410.

Naba, A., Clauser, K. R., Ding, H., Whittaker, C. A., Carr, S. A. and Hynes, R. O. (2015). The extracellular matrix: Tools and insights for the "omics" era. *Matrix Biology J Int Soc Matrix Biology* **49**, 10–24.

Noguchi, Y., Nakamura, M., Hino, N., Nogami, J., Tsuji, S., Sato, T., Zhang, L., Tsujikawa, K., Tanaka, T., Izawa, K., et al. (2019). Cell-autonomous and redundant roles of Hey1 and HeyL in muscle stem cells: HeyL requires Hes1 to bind diverse DNA sites. *Development* **146**, dev163618.

Olguin, H. C. and Olwin, B. B. (2004). Pax-7 up-regulation inhibits myogenesis and cell cycle progression in satellite cells: a potential mechanism for self-renewal. *Dev Biol* **275**, 375–388.

Ono, Y., Boldrin, L., Knopp, P., Morgan, J. E. and Zammit, P. S. (2010). Muscle satellite cells are a functionally heterogeneous population in both somite-derived and branchiomeric muscles. *Developmental Biology* **337**, 29–41.

Ono, Y., Masuda, S., Nam, H., Benezra, R., Miyagoe-Suzuki, Y. and Takeda, S. (2012). Slow-dividing satellite cells retain long-term self-renewal ability in adult muscle. *J Cell Sci* **125**, 1309–1317.

Oprescu, S. N., Yue, F., Qiu, J., Brito, L. F. and Kuang, S. (2020). Temporal Dynamics and Heterogeneity of Cell Populations during Skeletal Muscle Regeneration. *Iscience* **23**, 100993.

Pagani, F., Tratta, E., Dell'Era, P., Cominelli, M. and Poliani, P. L. (2021). EBF1 is expressed in pericytes and contributes to pericyte cell commitment. *Histochem Cell Biol* **156**, 333–347.

Pawlikowski, B., Pulliam, C., Betta, N., Kardon, G. and Olwin, B. B. (2015). Pervasive satellite cell contribution to uninjured adult muscle fibers. *Skeletal Muscle* **5**, 42.

Pineault, K. M. and Wellik, D. M. (2014). Hox genes and limb musculoskeletal development. *Curr Osteoporos Rep* **12**, 420–7.

Randolph, M. E. and Pavlath, G. K. (2015). A muscle stem cell for every muscle: variability of satellite cell biology among different muscle groups. *Frontiers in Aging Neuroscience* **7**, 190.

Randolph, M. E., Phillips, B. L., Choo, H., Vest, K. E., Vera, Y. and Pavlath, G. K. (2015). Pharyngeal Satellite Cells Undergo Myogenesis Under Basal Conditions and Are Required for Pharyngeal Muscle Maintenance. *STEM CELLS* **33**, 3581–3595.

Rayagiri, S. S., Ranaldi, D., Raven, A., Azhar, N. I. F. M., Lefebvre, O., Zammit, P. S. and Borycki, A.-G. (2018). Basal lamina remodeling at the skeletal muscle stem cell niche mediates stem cell self-renewal. *Nat Commun* **9**, 1075–1075.

Relaix, F., Rocancourt, D., Mansouri, A. and Buckingham, M. (2005). A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* **435**, 948–953.

Rocheteau, P., Gayraud-Morel, B., Siegl-Cachedenier, I., Blasco, M. A. and Tajbakhsh, S. (2012). A subpopulation of adult skeletal muscle stem cells retains all template DNA strands after cell division. *Cell* **148**, 112–25.

Rubenstein, A. B., Smith, G. R., Raue, U., Begue, G., Minchev, K., Ruf-Zamojski, F., Nair, V. D., Wang, X., Zhou, L., Zaslavsky, E., et al. (2020). Single-cell transcriptional profiles in human skeletal muscle. *Sci Rep-uk* **10**, 229.

Rugowska, A., Starosta, A. and Konieczny, P. (2021). Epigenetic modifications in muscle regeneration and progression of Duchenne muscular dystrophy. *Clin Epigenetics* **13**, 13.

Said, N., Frierson, H. F., Sanchez-Carbayo, M., Brekken, R. A. and Theodorescu, D. (2013). Loss of SPARC in bladder cancer enhances carcinogenesis and progression. *J Clin Invest* **123**, 751–766.

Sakai, H., Sato, T., Sakurai, H., Yamamoto, T., Hanaoka, K., Montarras, D. and Sehara-Fujisawa, A. (2013). Fetal Skeletal Muscle Progenitors Have Regenerative Capacity after Intramuscular Engraftment in Dystrophin Deficient Mice. *Plos One* **8**, e63016.

Sambasivan, R., Gayraud-Morel, B., Dumas, G., Cimper, C., Paisant, S., Kelly, R. G. and Tajbakhsh, S. (2009). Distinct Regulatory Cascades Govern Extraocular and Pharyngeal Arch Muscle Progenitor Cell Fates. *Developmental Cell* **16**, 810–821.

Sambasivan, R., Kuratani, S. and Tajbakhsh, S. (2011). An eye on the head: the development and evolution of craniofacial muscles. *Development* **138**, 2401–15.

Sampieri, L., Giusto, P. D. and Alvarez, C. (2019). CREB3 Transcription Factors: ER-Golgi Stress Transducers as Hubs for Cellular Homeostasis. *Frontiers Cell Dev Biology* **7**, 123.

Sande, B. V. de, Flerin, C., Davie, K., Waegeneer, M. D., Hulselmans, G., Aibar, S., Seurinck, R., Saelens, W., Cannoodt, R., Rouchon, Q., et al. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat Protoc* 1–30.

Scaramozza, A., Park, D., Kollu, S., Beerman, I., Sun, X., Rossi, D. J., Lin, C. P., Scadden, D. T., Crist, C. and Brack, A. S. (2019). Lineage Tracing Reveals a Subset of Reserve Muscle Stem Cells Capable of Clonal Expansion under Stress. *Cell Stem Cell* **24**, 944-957.e5.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682.

Schmidt, R. and Plath, K. (2012). The roles of the reprogramming factors Oct4, Sox2 and Klf4 in resetting the somatic cell epigenome during induced pluripotent stem cell generation. *Genome Biol* **13**, 251.

Shao, X. and Wei, X. (2018). FOXP1 enhances fibrosis via activating Wnt/β-catenin signaling pathway in endometriosis. *Am J Transl Res* **10**, 3610–3618.

Smith, R. S., Zabaleta, A., Kume, T., Savinova, O. V., Kidson, S. H., Martin, J. E., Nishimura, D. Y., Alward, W. L. M., Hogan, B. L. M. and John, S. W. M. (2000). Haploinsufficiency of the transcription factors FOXC1 and FOXC2 results in aberrant ocular development. *Hum Mol Genet* **9**, 1021–1032.

Soriano, P. (1994). Abnormal kidney development and hematological disorders in PDGF beta-receptor mutant mice. *Gene Dev* **8**, 1888–1896.

Stuelsatz, P., Shearer, A., Li, Y., Muir, L. A., Ieronimakis, N., Shen, Q. W., Kirillova, I. and Yablonka-Reuveni, Z. (2015). Extraocular muscle satellite cells are high performance myo-engines retaining efficient regenerative capacity in dystrophin deficiency. *Dev Biol* **397**, 31–44.

Swinehart, I. T., Schlientz, A. J., Quintanilla, C. A., Mortlock, D. P. and Wellik, D. M. (2013). Hox11 genes are required for regional patterning and integration of muscle, tendon and bone. *Development* **140**, 4574–4582.

Taglietti, V., Kefi, K., Rivera, L., Bergiers, O., Cardone, N., Coulpier, F., Gioftsidi, S., Drayton-Libotte, B., Hou, C., Authier, F.-J., et al. (2023). Thyroid-stimulating hormone receptor signaling restores skeletal muscle stem cell regeneration in rats with muscular dystrophy. *Sci Transl Med* **15**, eadd5275.

Tajbakhsh, S. (1996). Redefining the Genetic Hierarchies Controlling Skeletal Myogenesis- Pax-3 and Myf-5 Act Upstream of MyoD.pdf.

Tajbakhsh, S., Rocancourt, D., Cossu, G. and Buckingham, M. (1997). Redefining the Genetic Hierarchies Controlling Skeletal Myogenesis: Pax-3 and Myf-5 Act Upstream of MyoD. *Cell* **89**, 127–138.

Team, R. C. (2014). R: A language and environment for statistical computing. R Foundation for Statistical.

Terry, E. E., Zhang, X., Hoffmann, C., Hughes, L. D., Lewis, S. A., Li, J., Wallace, M. J., Riley, L., Douglas, C. M., Gutierrez-Montreal, M. A., et al. (2018). Transcriptional profiling reveals extraordinary diversity among skeletal muscle tissues. *eLife* **7**, e34613.

Tierney, M., Gromova, A., Sesillo, F., Sala, D., Spenlé, C., Orend, G. and Sacco, A. (2016). Autonomous Extracellular Matrix Remodeling Controls a Progressive Adaptation in Muscle Stem Cell Regenerative Capacity during Development. *Cell Reports* **14**, 1940–1952.

Tierney, M. T., Stec, M. J., Rulands, S., Simons, B. D. and Sacco, A. (2018). Muscle Stem Cells Exhibit Distinct Clonal Dynamics in Response to Tissue Repair and Homeostatic Aging. *Cell Stem Cell* **22**, 119-127.e3.

Tinevez, J.-Y., Perry, N., Schindelin, J., Hoopes, G. M., Reynolds, G. D., Laplantine, E., Bednarek, S. Y., Shorte, S. L. and Eliceiri, K. W. (2017). TrackMate: An open and extensible platform for single-particle tracking. *Methods* **115**, 80–90.

Uezumi, A., Fukada, S., Yamamoto, N., Takeda, S. and Tsuchida, K. (2010). Mesenchymal progenitors distinct from satellite cells contribute to ectopic fat cell formation in skeletal muscle. *Nat Cell Biol* **12**, 143–152.

Vallejo, D., Hernández-Torres, F., Lozano-Velasco, E., Rodriguez-Outeiriño, L., Carvajal, A., Creus, C., Franco, D. and Aránega, A. E. (2018). PITX2 Enhances the Regenerative Potential of Dystrophic Skeletal Muscle Stem Cells. *Stem Cell Rep* **10**, 1398–1411.

Vartanian, A. D., Quétin, M., Michineau, S., Auradé, F., Hayashi, S., Dubois, C., Rocancourt, D., Drayton-Libotte, B., Szegedi, A., Buckingham, M., et al. (2019). PAX3 Confers Functional Heterogeneity in Skeletal Muscle Stem Cell Responses to Environmental Stress. *Cell Stem Cell* **24**, 958-973.e9.

Velthoven, C. T. J. van, Morree, A. de, Egner, I. M., Brett, J. O. and Rando, T. A. (2017). Transcriptional Profiling of Quiescent Muscle Stem Cells In Vivo. *Cell Reports* **21**, 1994–2004.

Verma, M., Fitzpatrick, K. R. and McLoon, L. K. (2017). Extraocular Muscle Repair and Regeneration. *Curr Ophthalmol Reports* **5**, 207–215.

Wang, L., Siegenthaler, J. A., Dowell, R. D. and Yi, R. (2016). Foxc1 reinforces quiescence in self-renewing hair follicle stem cells. *Science* **351**, 613–617.

Wang, C., Chen, A., Ruan, B., Niu, Z., Su, Y., Qin, H., Zheng, Y., Zhang, B., Gao, L., Chen, Z., et al. (2020). PCDH7 Inhibits the Formation of Homotypic Cell-in-Cell Structure. *Frontiers Cell Dev Biology* **8**, 329.

Whitesell, T. R., Chrystal, P. W., Ryu, J.-R., Munsie, N., Grosse, A., French, C. R., Workentine, M. L., Li, R., Zhu, L. J., Waskiewicz, A., et al. (2019). foxc1 is required for embryonic head vascular smooth muscle differentiation in zebrafish. *Dev Biol* **453**, 34–47.

Wickham, H. (2009). ggplot2, Elegant Graphics for Data Analysis.

Wolf, F. A., Angerer, P. and Theis, F. J. (2018). SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* **19**, 15.

Wu, Y.-F., Wang, H.-K., Chang, H.-W., Sun, J., Sun, J.-S. and Chao, Y.-H. (2017). High glucose alters tendon homeostasis through downregulation of the AMPK/Egr1 pathway. *Sci Rep-uk* **7**, 44199.

Xi, H., Langerman, J., Sabri, S., Chien, P., Young, C. S., Younesi, S., Hicks, M., Gonzalez, K., Fujiwara, W., Marzi, J., et al. (2020). A Human Skeletal Muscle Atlas Identifies the Trajectories of Stem and Progenitor Cells across Development and from Human Pluripotent Stem Cells. *Cell Stem Cell* **27**, 158-176.e10.

Xie, N., Chu, S. N., Azzag, K., Schultz, C. B., Peifer, L. N., Kyba, M., Perlingeiro, R. C. R. and Chan, S. S. K. (2021). In vitro expanded skeletal myogenic progenitors from pluripotent stem cell-derived teratomas have high engraftment capacity. *Stem Cell Rep*.

Yang, Z., Jiang, S., Cheng, Y., Li, T., Hu, W., Ma, Z., Chen, F. and Yang, Y. (2017). FOXC1 in cancer development and therapy: deciphering its emerging and divergent roles. *Ther Adv Med Oncol* **9**, 797–816.

Yartseva, V., Goldstein, L. D., Rodman, J., Kates, L., Chen, M. Z., Chen, Y.-J. J., Foreman, O., Siebel, C. W., Modrusan, Z., Peterson, A. S., et al. (2020). Heterogeneity of Satellite Cells Implicates DELTA1/NOTCH2 Signaling in Self-Renewal. *Cell Reports* **30**, 1491-1503.e6.

Yennek, S., Burute, M., Théry, M. and Tajbakhsh, S. (2014). Cell Adhesion Geometry Regulates Non-Random DNA Segregation and Asymmetric Cell Fates in Mouse Skeletal Muscle Stem Cells. *Cell Reports* **7**, 961–970.

Yin, H., Price, F. and Rudnicki, M. A. (2013). Satellite cells and the muscle stem cell niche. *Physiol Rev* **93**, 23–67.

Yoshida, K. (2003). Fibroblast cell shape and adhesion in vitro is altered by overexpression of the 7a and 7b isoforms of protocadherin 7, but not the 7c isoform. *Cell Mol Biol Lett* **8**, 735–41.

Yoshida, N., Yoshida, S., Koishi, K., Masuda, K. and Nabeshima, Y. (1998). Cell heterogeneity upon myogenic differentiation: down-regulation of MyoD and Myf-5 generates “reserve cells”. *J Cell Sci* **111 (Pt 6)**, 769–79.

Yoshioka, K., Nagahisa, H., Miura, F., Araki, H., Kamei, Y., Kitajima, Y., Seko, D., Nogami, J., Tsuchiya, Y., Okazaki, N., et al. (2021). Hoxa10 mediates positional memory to govern stem cell function in adult skeletal muscle. *Sci Adv* **7**, eabd7924.

Zacharias, A. L., Lewandoski, M., Rudnicki, M. A. and Gage, P. J. (2011). Pitx2 is an upstream activator of extraocular myogenesis and survival. *Dev Biol* **349**, 395–405.

Zakany, J. and Duboule, D. (2007). The role of Hox genes during vertebrate limb development. *Curr Opin Genet Dev* **17**, 359–366.

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

Zammit, P. S., Partridge, T. A. and Yablonka-Reuveni, Z. (2006). The Skeletal Muscle Satellite Cell: The Stem Cell That Came in From the Cold. *J Histochem Cytochem* **54**, 1177–1191.

Zhang, P., Wong, C., Liu, D., Finegold, M., Harper, J. W. and Elledge, S. J. (1999). p21CIP1 and p57KIP2 control muscle differentiation at the myogenin step. *Gene Dev* **13**, 213–224.

Figure Legends

Figure 1. EOM and TA MuSCs exhibit functional differences following activation

A. Experimental scheme. MuSCs were isolated by FACS based on GFP fluorescence from *Tg:Pax7-nGFP* mice, plated and pulsed with EdU before fixation and immunostaining at Day (D) 3, D4 and D5.

B. Quantification of the percentage of EdU+ cells at D5. Two-tailed unpaired Student's t-test. *** p-value<0.005. n=5 mice.

C. Immunofluorescence for MF20 and EdU detection at different time points of the experiment in A.

D. Experimental scheme. EOM and TA MuSCs from *Tg:Pax7-nGFP;Myog*^{tdTom} mice were cultured for 72h and imaged for 48h.

E. Representative overlayed DIC and red fluorescence channel images at selected time points showing a temporal increase of tdTOMATO fluorescence. Scale bar 25 µm.

F. Quantification of selected timepoints in (E). Graph represents mean percentage of tdTOMATO^{+} cells over total number of cells at each timepoint for each muscle. Percentage of tdTOMATO^{+} cells was compared between EOM and TA. Two-tailed unpaired Student's t-test; * p-value<0.05, *** p-value<0.005. n=4 mice; > 100 cells counted.

G. EOM and TA MuSCs were cultured for 48h and then processed for RNA extraction. RT-qPCR for *Pax7*, *Myod*, *Myogenin* and *p21* normalized to *Rpl13*. Two-tailed unpaired Student's t-test; *p-value<0.05, *** p-value<0.01. n=8 mice.

Figure 2. Single cell transcriptome signatures of activated EOM and TA MuSCs

A. Experimental scheme of sc-RNAseq pipeline for in vitro activated EOM and TA MuSCs.

B. UMAP visualization of EOM progenitors, EOM differentiating, TA progenitors, and TA differentiating cell clusters (left). Violin plots of the expression of myogenic markers *Pax7* (progenitor), *Myod1* (committed) and *Myog* (differentiating) in each cluster (right).

C. Expression plots of myogenic markers for progenitors (*Myf5*, *Pax7*) and committed/differentiating cells (*Myod*, *Myog*).

D. Heatmap representing top differentially expressed genes in each cluster and expression levels across all cells.

E-F. Reactome pathway (E) and GO Molecular function (F) network analysis on top 100 DEGs of each cluster. Pie charts represent relative contribution of each cluster to this ontology term. EOM-specific selected terms are highlighted.

Figure 3. Molecular signature of EOM and TA quiescent MuSCs

A. Experimental scheme of sc-RNAseq pipeline for quiescent EOM and TA MuSCs.

B. UMAP visualization of EOM and TA quiescent clusters.

C-D. Violin plot (C) and expression plot (D) of *Myf5*, *Pax7*, *Myod* and *Myog*.

E. Heatmap representing top differentially expressed genes in EOM and TA quiescent clusters and their expression levels across all cells.

F-G. Venn diagrams of the overlapping differentially expressed genes between quiescent and global in vitro activated datasets for EOM (F) and TA (G).

H. Expression plots of selected EOM markers exclusive to the quiescence state (Unique Q), common to quiescent and activated states (conserved) and exclusive to activation (Unique Act) from the analysis in F.

I. Expression plots of selected TA markers exclusive to the quiescence state (Unique Q), common to quiescent and activated states (conserved) and exclusive to activation (Unique Act) from the analysis in G.

Figure 4. EOM MuSC activation is accompanied by ECM remodelling.

A. Violin plots of matrisome scores for each cluster of the activated sc-RNAseq dataset.

B. Dot plot visualisation of differentially expressed matrisome components of the activated sc-RNAseq dataset. ECM, extracellular matrix; AP, affiliated proteins; PG, proteoglycans.

C. Immunofluorescence for Fibronectin (FN1), Collagen I (COLI), Collagen IV (COLIV) and PDGFRβ.

D. Western blot showing expression of CAV1, CAVIN1, MGP, MMP2, SPARC, IGFBP7 and H3 for normalization. Cells coming from n=3 mice pooled per lane.

E. Violin plot of *Pdgfrβ* expression on in vitro activated sc-RNAseq dataset.

F. UMAP of *Pdgfrβ* expression on in vitro activated sc-RNAseq dataset.

G. Scheme of isolation and culture of MuSCs isolated from adult *Tg:Pax7-nGFP* mice. EOM cells were resorted based on their PDGFRβ expression, plated for 24h and pulsed with EdU before fixation for immunostaining.

H. Representative images of PDGFRβ positive and negative EOM MuSC fractions stained for MYOD, MYOGENIN and EdU as described in J.

I. Bar plots of quantification of number of cells/cm^2 and percentage of EdU$^+$, MYOD$^+$ and MYOGENIN$^+$ cells. n=3 mice, >100 cells counted. Scale bar 50 µm. Two-tailed unpaired Student's t-test; ** p-value<0.01, *** p-value<0.005.

Figure 5. Distinct gene regulatory networks underlie EOM and TA activation dynamics

A. Top 4 regulon activity for each cluster of in vitro activated sc-RNAseq dataset overlaid onto UMAP representation.

B. Heatmap of Top 7 regulons in each cluster of the activated sc-RNAseq dataset with activity level in each cell

C. Transcription factor network highlighting top regulons of each cluster as well as common regulons (35 regulons maximum) of activated sc-RNAseq dataset. Proximity of the nodes in the network indicate a higher number of shared edges, highlighting core modules.

D. Filtered network showing Hes1 direct regulators (nodes with 1 level ancestry).

Figure 6. EOM properties are retained despite several passages in vitro.

A. Scheme of isolation and passages of EOM and TA MuSCs. Cells were isolated by FACS based on GFP fluorescence from *Tg:Pax7-nGFP;Myog*^{ntdTom} and cultured for 3 days (Act). The entire well was passaged (P1, P2) every 3 days. Following each passage, a fraction of cells was analysed by FACS to assess GFP and tdTom ratios.

B. EOM/TA normalized cell number ratios. Total cell numbers per well were counted upon trypsinization at day 3 (D3), days 5 (D5) and at passages (P)1 and P2. EOM/TA normalized cell number ratio was generated for each time point analysed (n=3 mice). For normalization, the ratio between cell numbers counted per well and averaged EOM cell number in all wells at each time point was calculated.

C. FACS analysis of *Tg:Pax7-nGFP;Myog*^{ntdTom} EOM and TA MuSCs upon 5 days in culture (D5), one or two passages (P1, P2) ($n \geq 3$ mice). Graph represents ratio of TdT^+/GFP^+ cells for each time point. Two-tailed unpaired Student's t-test; *p-value<0.05, ** p-value<0.01.

D. qRT-PCR at D3 (Act, Activated), P1, P2 on whole populations present in dish for key myogenic markers, regulon TFs and Matrisome/Molecular function genes identified in EOM progenitors (n=4 mice).

E. Western blot of total P2 population of EOM and TA MuSCs (n=3 mice; one mouse per lane). Note higher expression of Pax7 and EOM progenitor specific TF regulons (FOXC1, CREB3L1, EBF1) identified in activated EOM MuSCs by sc-RNAseq. ACTIN was used for normalization of protein loading.

F. Western blot of the total P2 population of EOM and TA MuSCs (n=3 mice; one mouse per lane) showing higher expression of matrisome genes (CAV1, COL1A2, COL6, IGFBP7, NRP1, SPARC) identified in activated EOM MuSCs by sc-RNAseq. ACTIN was used for normalization of protein loading.

Figure 7. EOM activated features are present during postnatal growth and can be modulated by Foxc1

A. Scheme of isolation of EOM and TA MuSCs from *Tg:Pax7-nGFP;Myog*^{ntdTom} from postnatal day (P) 7 mice.

B. qRT-PCR for key myogenic markers, regulon TF and Matrisome/Molecular function genes identified in EOM progenitors by scRNAseq analysis. GFP^+/TdT^- Cells were isolated from *Tg:Pax7-nGFP;Myog*^{ntdTom} EOM and TA MuSCs at postnatal day (P) 7 (n=3-6).

C-D. Immunostaining for FOXC1, tdTOMATO (TOM) and Dystrophin on cryosections from EOM and TA muscles isolated from P10 (C) and adult (D) *Pax7*^{CreERT2}*:R26*^{tdTom} mice (n=2 per stage). White arrowheads point to PAX7+FOXC1+ cells. Black arrowheads point to PAX7+FOXC1- cells.

E. Scheme of lentivirus transduction for loss of function (LOF) of EOM and gain of function (GOF) of TA MuSCs. Cells from *Tg:Pax7-nGFP* mice were isolated by FACs, plated and transduced the next day (MOI 100). Cells were analysed at D5.

F-H. Bar plots of quantification of number nuclei, FOXC1+ nuclei and EdU+ nuclei of EOM MuSCs transduced with scramble or shRNA against Foxc1 containing viruses.

I-J. Bar plots of quantification of number nuclei and FOXC1+ nuclei of TA MuSCs transduced with control or Foxc1 expressing viruses. F-L, Two-tailed unpaired Student's t-test; *** p-value<0.005, **** p-value<0.0001.

K. Western blot of TA MuSCs transduced with control or Foxc1 expressing viruses for mCherry, Foxc1 and Myogenin. Vinculin was used for normalization of protein loading.

L. qRT-PCR of TA MuSCs transduced with control or Foxc1 expressing viruses for Myogenin and key ECM proteins and regulators (Creb3l1) identified in EOM progenitors.

Figure 8. RNA velocity reveals distinct population kinetics and potential key regulators of progenitor maintenance

A. Number of regulatory links between regulons and matrisome genes depending on the number of top regulons in EOM and TA in vitro activated sc-RNAseq datasets.

B. Ratio of number of regulations of matrisome genes between EOM and TA MuSCs on in vitro activated sc-RNAseq dataset. Note difference is maximal for top 5 first regulons.

C. Heatmap of top regulons in global in vitro activated sc-RNAseq dataset with activity level in each cell.

D. Heatmap of same top 5 regulons in quiescent sc-RNAseq dataset. Note activities in quiescence is similar between EOM and TA, but seem to diverge during activation.

E. EOM Velocity streams overlaid onto a UMAP representation, along with expression patterns of *Myog* and *Pax7* on in vitro activated sc-RNAseq dataset.

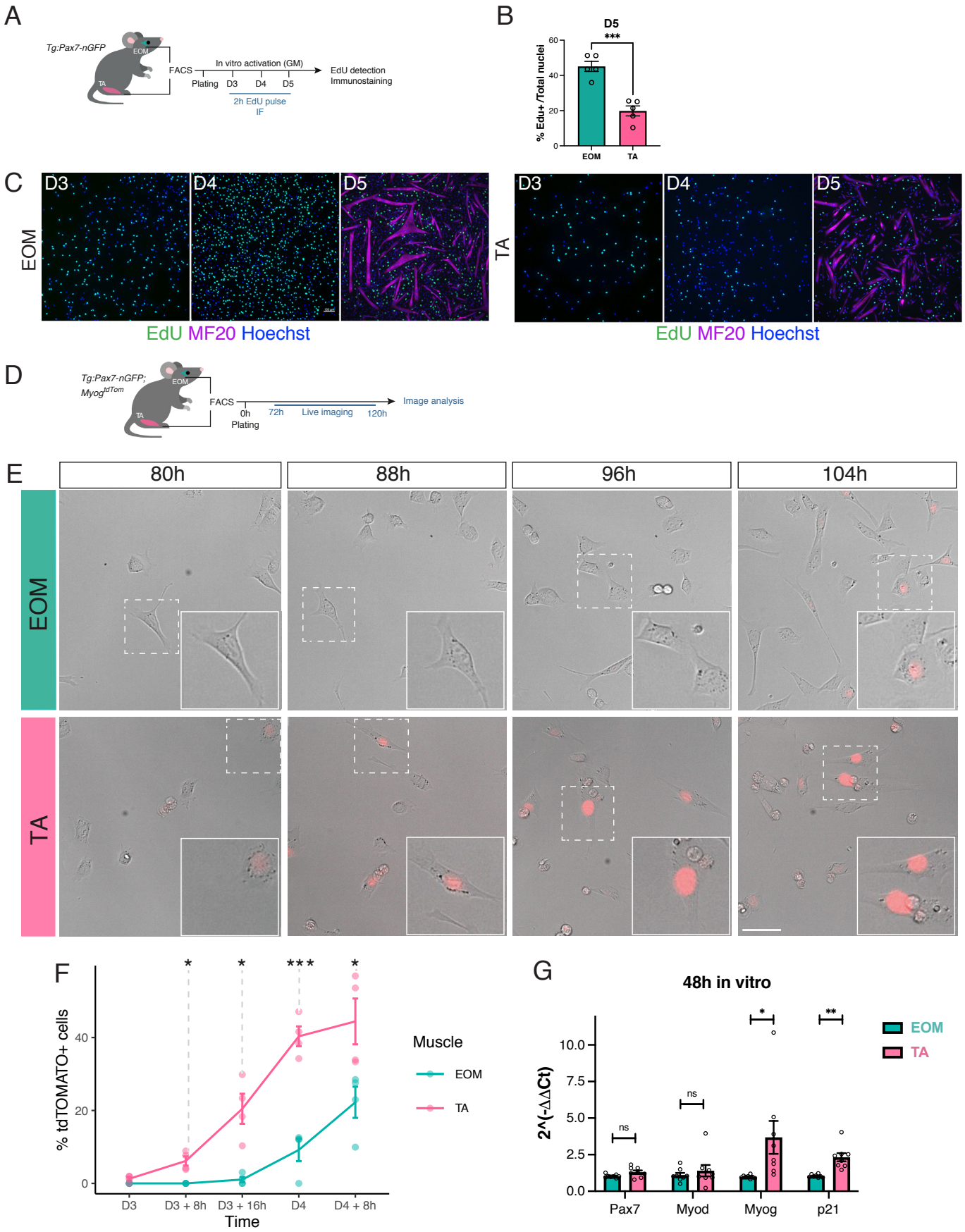
F. TA Velocity streams overlaid onto a UMAP representation, along with expression patterns of *Myog* and *Pax7* on in vitro activated sc-RNAseq dataset.

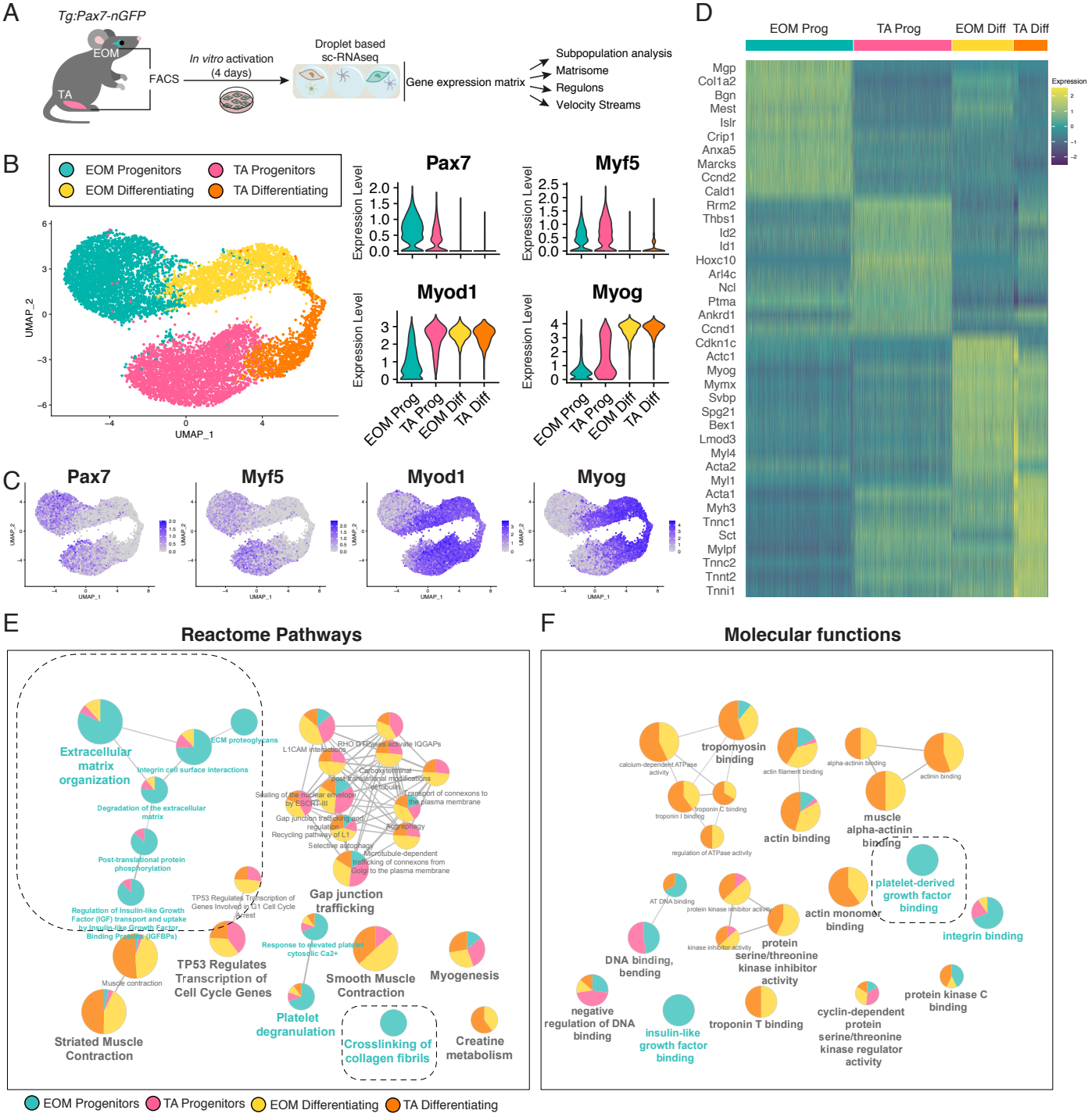
G. Heatmap of driver genes expression, obtained from EOM progenitor velocity stream.

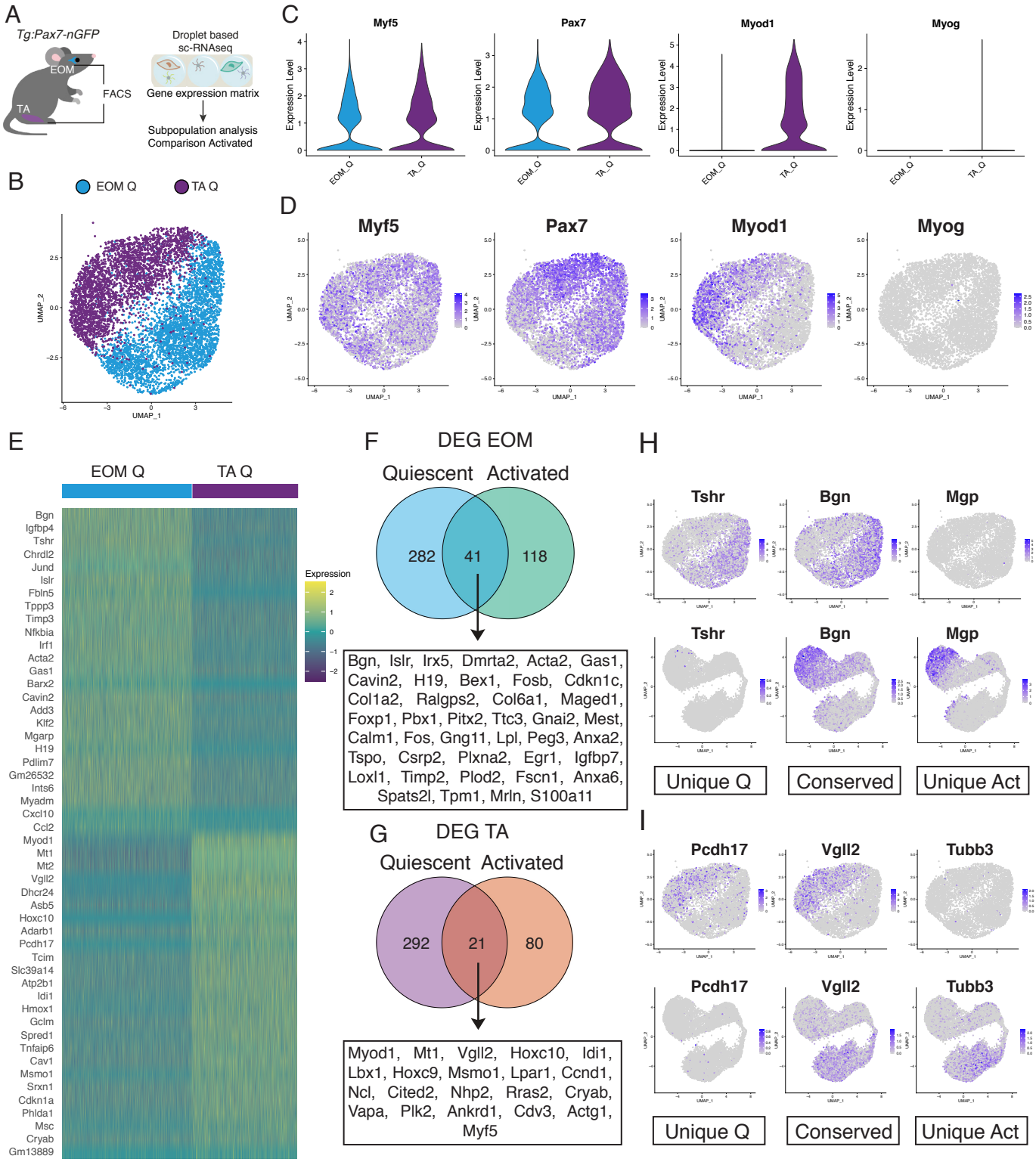
H. Heatmap of driver genes expression, obtained from TA progenitor velocity stream.

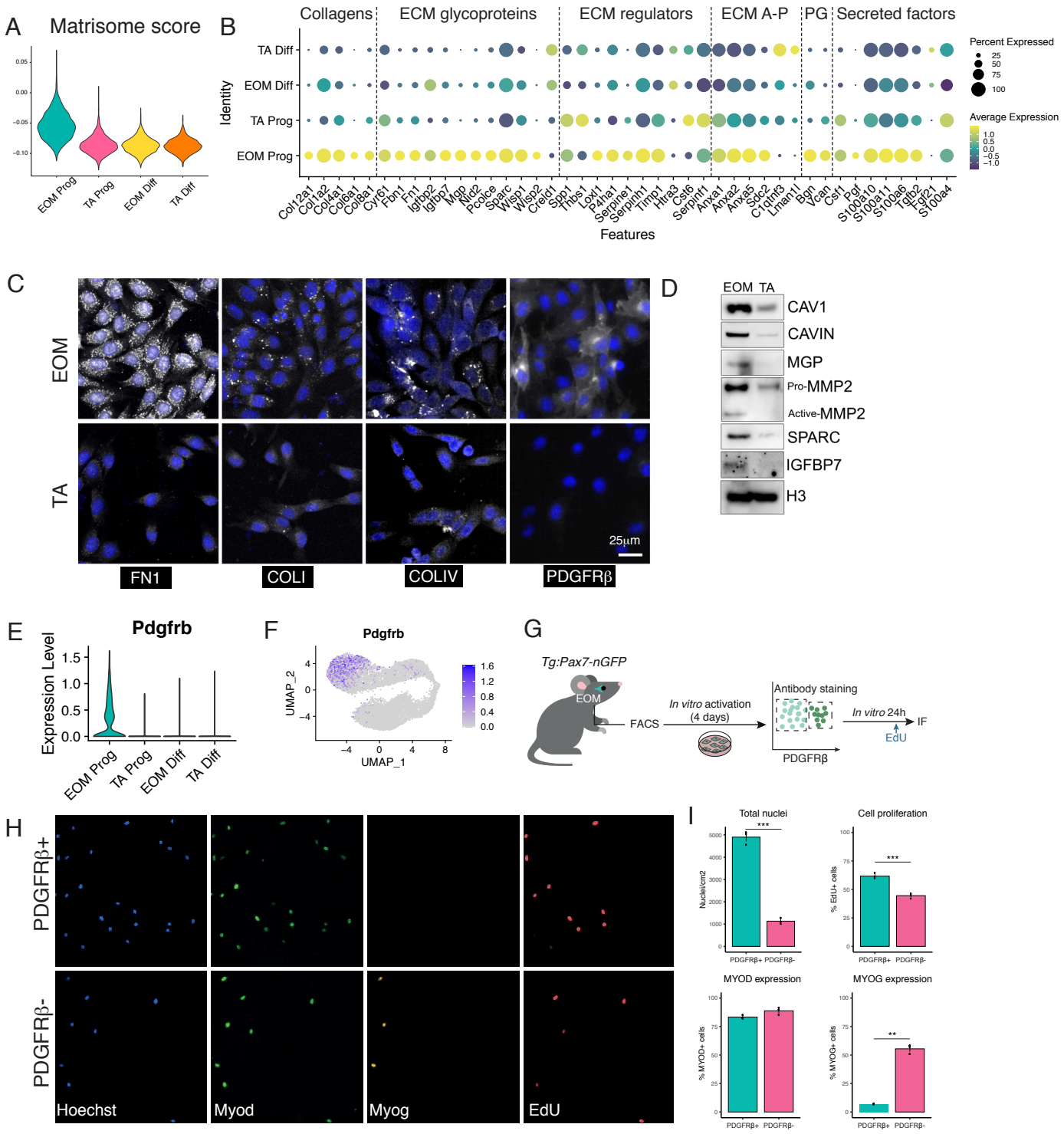
I. Venn diagram of top 100 drivers for EOM and TA progenitor streams.

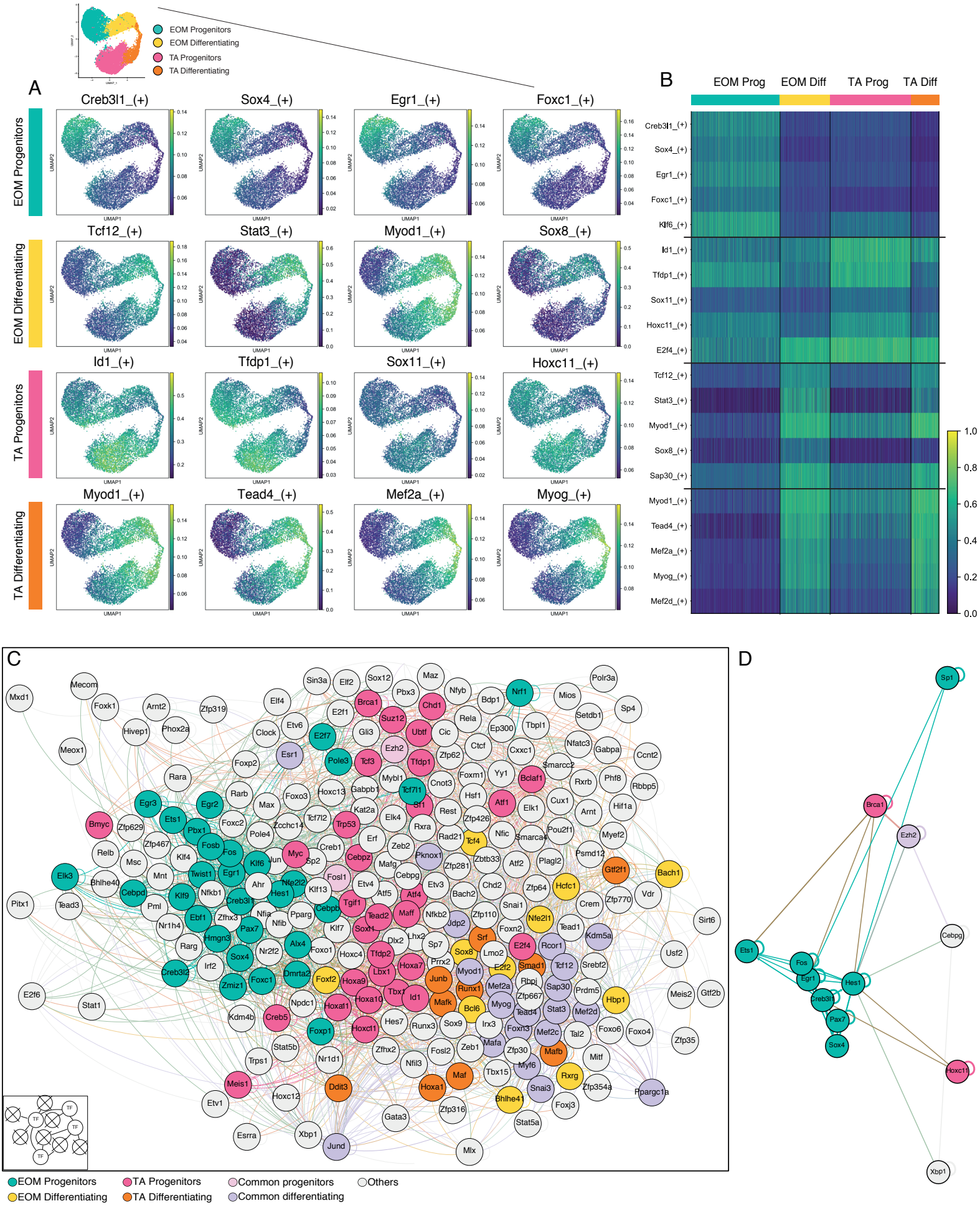
J. GO Molecular function ontology network and associated driver genes exposing unique and common pathways dynamically regulated during EOM and TA progenitor maintenance.

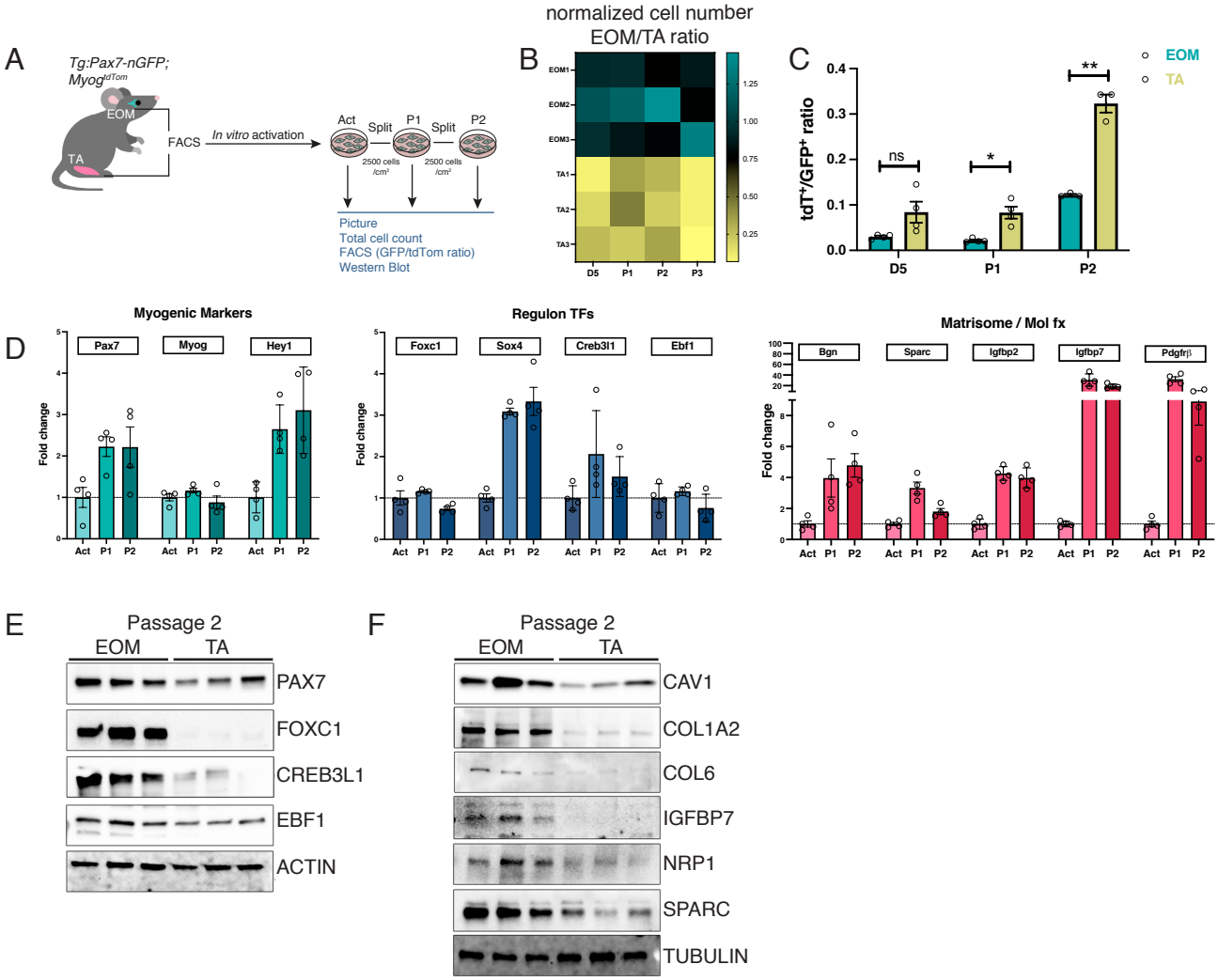

A
Tg:Pax7-nGFP
EOM
TA
FACS
In vitro activation (GM)
Plating D3 D4 D5
2h EdU pulse
IF
EdU detection
Immunostaining
B
D5
***
% Edu+ /Total nuclei
60
40
20
0
EOM
TA
C
D3
D4
D5
EOM
TA
EdU MF20 Hoechst
D
Tg:Pax7-nGFP;
Myog tdTom
FACS
0h
Plating
72h
Live imaging
120h
Image analysis
E
80h
88h
96h
104h
EOM
TA
F
% tdTOMATO+ cells
40
20
0
D3
D3 + 8h
D3 + 16h
D4
D4 + 8h
Time
Muscle
EOM
TA
G
48h in vitro
2^(-ΔΔCt)
10.0
7.5
5.0
2.5
0.0
ns
ns
*
**
Pax7
Myod
Myog
p21
EOM
TA


A
Tg:Pax7-nGFP
EOM
TA
FACS
In vitro activation
(4 days)
Droplet based
sc-RNAseq
Gene expression matrix
Subpopulation analysis
Matrisome
Regulons
Velocity Streams
B
EOM Progenitors
TA Progenitors
EOM Differentiating
TA Differentiating
UMAP_1
UMAP_2
Pax7
Myf5
Myod1
Myog
Expression Level
EOM Prog
TA Prog
EOM Diff
TA Diff
C
Pax7
Myf5
Myod1
Myog
D
EOM Prog
TA Prog
EOM Diff
TA Diff
Expression
Mgp
Col1a2
Bgn
Mest
Islr
Crip1
Anxa5
Marcks
Ccnd2
Cald1
Rrm2
Thbs1
Id2
Id1
Hoxc10
Arl4c
Ncl
Ptma
Ankrd1
Ccnd1
Cdkn1c
Actc1
Myog
Mymx
Svbp
Spg21
Bex1
Lmod3
Myl4
Acta2
Myl1
Acta1
Myh3
Tnnc1
Sct
Mylpf
Tnnc2
Tnnt2
Tnni1
E
Reactome Pathways
Extracellular matrix organization
ECM proteoglycans
Integrin cell surface interactions
Degradation of the extracellular matrix
Post-translational protein phosphorylation
Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Binding Proteins (IGFBPs)
Muscle contraction
Striated Muscle Contraction
TP53 Regulates Transcription of Genes Involved in G1 Cell Cycle Arrest
TP53 Regulates Transcription of Cell Cycle Genes
L1CAM interactions
RHO GTPases activate IQGAPs
Post-translational modifications: tubulin
Sealing of the nuclear envelope by ESCRT-III
Transport of connexons to the plasma membrane
Gap junction trafficking and regulation
Recycling pathway of L1
Aggrephagy
Selective autophagy
Microtubule-dependent trafficking of connexons from Golgi to the plasma membrane
Gap junction trafficking
Response to elevated platelet cytosolic Ca2+
Platelet degranulation
Smooth Muscle Contraction
Myogenesis
Crosslinking of collagen fibrils
Creatine metabolism
F
Molecular functions
tropomyosin binding
calcium-dependent ATPase activity
actin filament binding
alpha-actinin binding
actinin binding
troponin C binding
troponin I binding
regulation of ATPase activity
actin binding
muscle alpha-actinin binding
platelet-derived growth factor binding
AT DNA binding
protein kinase inhibitor activity
kinase inhibitor activity
actin monomer binding
integrin binding
DNA binding, bending
protein serine/threonine kinase inhibitor activity
negative regulation of DNA binding
insulin-like growth factor binding
troponin T binding
cyclin-dependent protein serine/threonine kinase regulator activity
protein kinase C binding
EOM Progenitors
TA Progenitors
EOM Differentiating
TA Differentiating

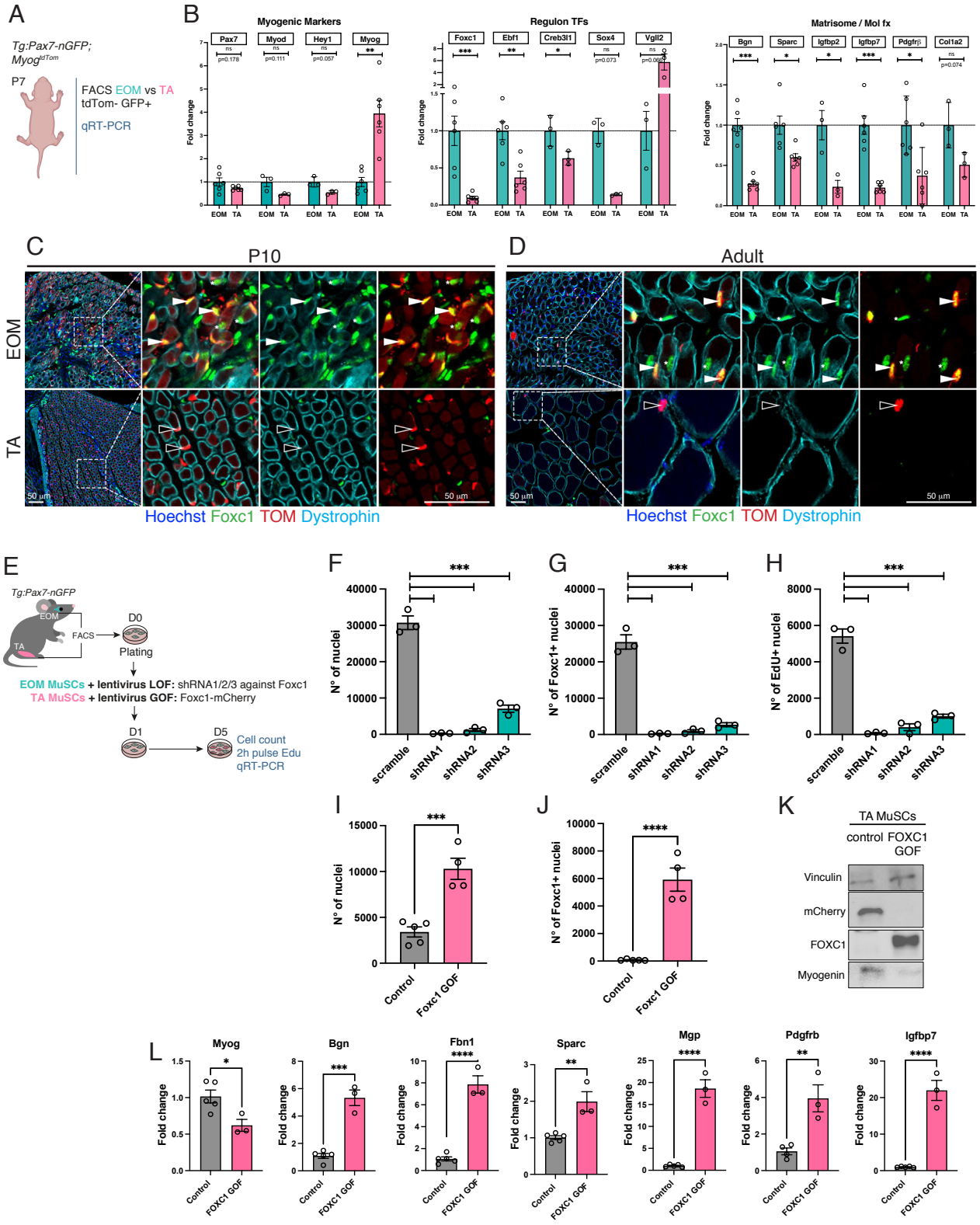

A
Tg:Pax7-nGFP
EOM
TA
FACS
Droplet based sc-RNAseq
Gene expression matrix
Subpopulation analysis
Comparison Activated
B
EOM Q
TA Q
UMAP_1
UMAP_2
C
Myf5
Pax7
Myod1
Myog
Expression Level
EOM_Q
TA_Q
D
Myf5
Pax7
Myod1
Myog
E
EOM Q
TA Q
Expression
Bgn
Igfbp4
Tshr
Chrdl2
Jund
Islr
Fbln5
Tppp3
Timp3
Nfkbia
Irf1
Acta2
Gas1
Barx2
Cavin2
Add3
Klf2
Mgarp
H19
Pdlim7
Gm26532
Ints6
Myadm
Cxcl10
Ccl2
Myod1
Mt1
Mt2
Vgll2
Dhcr24
Asb5
Hoxc10
Adarb1
Pcdh17
Tcim
Slc39a14
Atp2b1
Idi1
Hmox1
Gclm
Spred1
Tnfaip6
Cav1
Msmo1
Srxn1
Cdkn1a
Phlda1
Msc
Cryab
Gm13889
F
DEG EOM
Quiescent
Activated
282
41
118
Bgn, Islr, Irx5, Dmrta2, Acta2, Gas1, Cavin2, H19, Bex1, Fosb, Cdkn1c, Col1a2, Ralgps2, Col6a1, Maged1, Foxp1, Pbx1, Pitx2, Ttc3, Gnai2, Mest, Calm1, Fos, Gng11, Lpl, Peg3, Anxa2, Tspo, Csrp2, Plxna2, Egr1, Igfbp7, Loxl1, Timp2, Plod2, Fscn1, Anxa6, Spats2l, Tpm1, Mrln, S100a11
G
DEG TA
Quiescent
Activated
292
21
80
Myod1, Mt1, Vgll2, Hoxc10, Idi1, Lbx1, Hoxc9, Msmo1, Lpar1, Ccnd1, Ncl, Cited2, Nhp2, Rras2, Cryab, Vapa, Plk2, Ankrd1, Cdv3, Actg1, Myf5
H
Tshr
Bgn
Mgp
Unique Q
Conserved
Unique Act
I
Pcdh17
Vgll2
Tubb3
Unique Q
Conserved
Unique Act

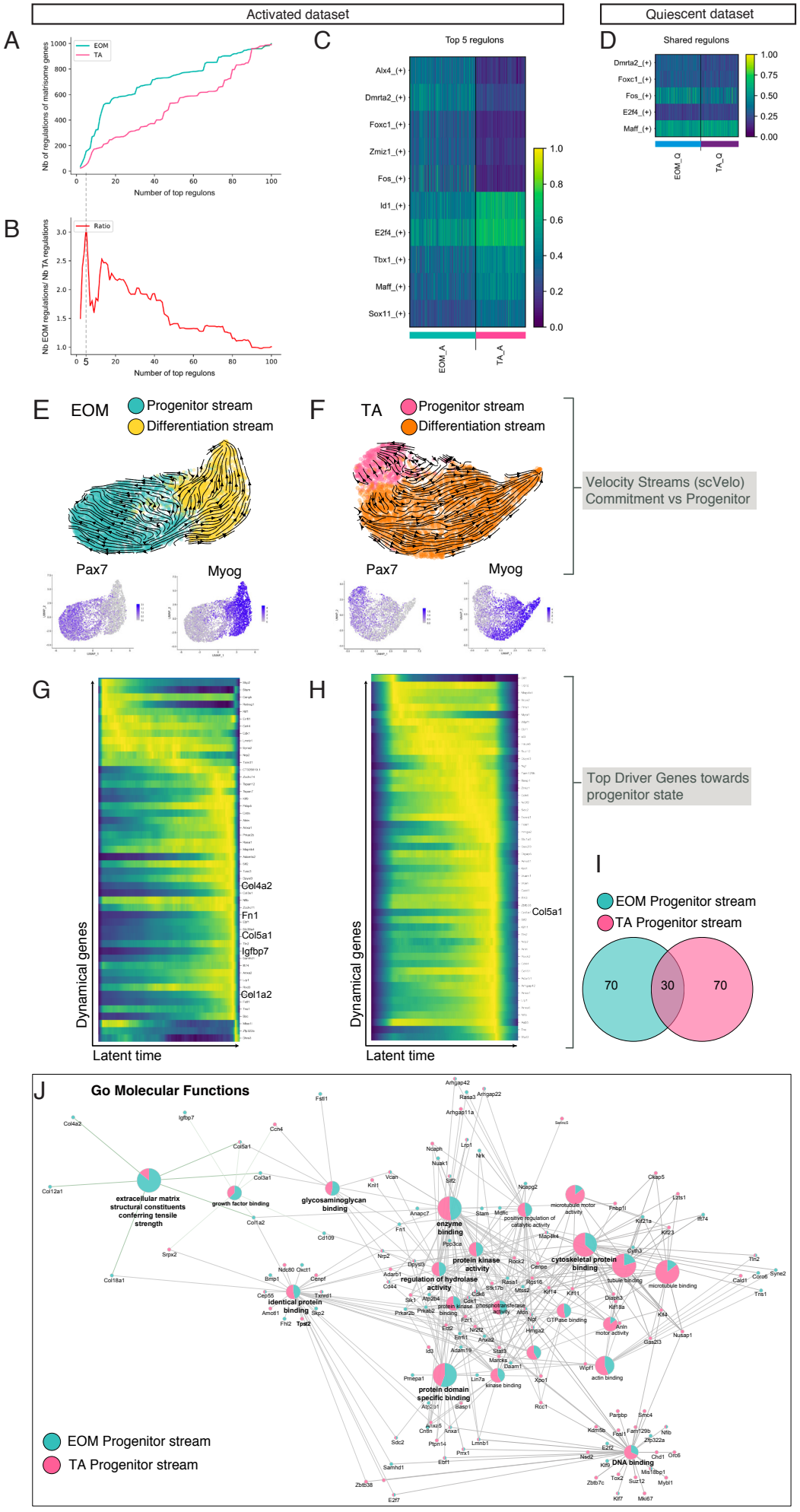

A
Matrisome score
EOM Prog
TA Prog
EOM Diff
TA Diff
B
Collagens
ECM glycoproteins
ECM regulators
ECM A-P
PG
Secreted factors
Identity
Features
Percent Expressed
Average Expression
C
EOM
TA
FN1
COLI
COLIV
PDGFRβ
25µm
D
EOM TA
CAV1
CAVIN
MGP
Pro-MMP2
Active-MMP2
SPARC
IGFBP7
H3
E
Pdgfrb
Expression Level
F
Pdgfrb
UMAP_1
UMAP_2
G
Tg:Pax7-nGFP
EOM
FACS
In vitro activation (4 days)
Antibody staining
In vitro 24h
EdU
IF
PDGFRβ
H
PDGFRβ+
PDGFRβ-
Hoechst
Myod
Myog
EdU
I
Total nuclei
Cell proliferation
MYOD expression
MYOG expression
Nuclei/cm2
% EdU+ cells
% MYOD+ cells
% MYOG+ cells
PDGFRβ+
PDGFRβ-


EOM Progenitors
EOM Differentiating
TA Progenitors
TA Differentiating
A
Creb3l1_(+)
Sox4_(+)
Egr1_(+)
Foxc1_(+)
Tcf12_(+)
Stat3_(+)
Myod1_(+)
Sox8_(+)
Id1_(+)
Tfdp1_(+)
Sox11_(+)
Hoxc11_(+)
Myod1_(+)
Tead4_(+)
Mef2a_(+)
Myog_(+)
UMAP1
UMAP2
B
EOM Prog
EOM Diff
TA Prog
TA Diff
Creb3l1_(+)
Sox4_(+)
Egr1_(+)
Foxc1_(+)
Klf6_(+)
Id1_(+)
Tfdp1_(+)
Sox11_(+)
Hoxc11_(+)
E2f4_(+)
Tcf12_(+)
Stat3_(+)
Myod1_(+)
Sox8_(+)
Sap30_(+)
Myod1_(+)
Tead4_(+)
Mef2a_(+)
Myog_(+)
Mef2d_(+)
1.0
0.8
0.6
0.4
0.2
0.0
C
D
EOM Progenitors
TA Progenitors
Common progenitors
Others
EOM Differentiating
TA Differentiating
Common differentiating


A
Tg:Pax7-nGFP;
Myog tdTom
EOM
TA
FACS
In vitro activation
Act
Split
P1
Split
P2
2500 cells /cm²
2500 cells /cm²
Picture
Total cell count
FACS (GFP/tdTom ratio)
Western Blot
B
normalized cell number
EOM/TA ratio
EOM1
EOM2
EOM3
TA1
TA2
TA3
D5
P1
P2
P3
1.25
1.00
0.75
0.50
0.25
C
EOM
TA
tdT+/GFP+ ratio
ns
*
**
D5
P1
P2
D
Myogenic Markers
Pax7
Myog
Hey1
Fold change
Act
P1
P2
Regulon TFs
Foxc1
Sox4
Creb3l1
Ebf1
Matrisome / Mol fx
Bgn
Sparc
Igfbp2
Igfbp7
Pdgfrβ

E
Passage 2
EOM
TA
PAX7
FOXC1
CREB3L1
EBF1
ACTIN
F
Passage 2
EOM
TA
CAV1
COL1A2
COL6
IGFBP7
NRP1
SPARC
TUBULIN


A
Tg:Pax7-nGFP; Myog tdTom
P7
FACS EOM vs TA
tdTom- GFP+
qRT-PCR
B
Myogenic Markers
Regulon TFs
Matrisome / Mol fx
Fold change
EOM
TA
C
P10
D
Adult
EOM
TA
50 µm
Hoechst Foxc1 TOM Dystrophin
E
Tg:Pax7-nGFP
FACS
D0
Plating
EOM MuSCs + lentivirus LOF: shRNA1/2/3 against Foxc1
TA MuSCs + lentivirus GOF: Foxc1-mCherry
D1
D5
Cell count
2h pulse Edu
qRT-PCR
F
N° of nuclei
scramble
shRNA1
shRNA2
shRNA3
G
N° of Foxc1+ nuclei
H
N° of EdU+ nuclei
I
N° of nuclei
Control
Foxc1 GOF
J
N° of Foxc1+ nuclei
K
TA MuSCs
control FOXC1 GOF
Vinculin
mCherry
FOXC1
Myogenin
L
Myog
Bgn
Fbn1
Sparc
Mgp
Pdgfrb
Igfbp7
Fold change
Control
FOXC1 GOF


Activated dataset
Quiescent dataset
A
EOM
TA
Nb of regulations of matrisome genes
Number of top regulons
B
Ratio
Nb EOM regulations/ Nb TA regulations
Number of top regulons
C
Top 5 regulons
Alx4_(+)
Dmrta2_(+)
Foxc1_(+)
Zmiz1_(+)
Fos_(+)
Id1_(+)
E2f4_(+)
Tbx1_(+)
Maff_(+)
Sox11_(+)
EOM_A
TA_A
D
Shared regulons
Dmrta2_(+)
Foxc1_(+)
Fos_(+)
E2f4_(+)
Maff_(+)
EOM_Q
TA_Q
E
EOM
Progenitor stream
Differentiation stream
Pax7
Myog
F
TA
Progenitor stream
Differentiation stream
Pax7
Myog
Velocity Streams (scVelo)
Commitment vs Progenitor
G
Dynamical genes
Latent time
Col4a2
Fn1
Col5a1
Igfbp7
Col1a2
H
Dynamical genes
Latent time
Col5a1
Top Driver Genes towards
progenitor state
I
EOM Progenitor stream
TA Progenitor stream
70
30
70
J
Go Molecular Functions
extracellular matrix structural constituents conferring tensile strength
growth factor binding
glycosaminoglycan binding
enzyme binding
protein kinase activity
cytoskeletal protein binding
regulation of hydrolase activity
identical protein binding
protein domain specific binding
DNA binding
EOM Progenitor stream
TA Progenitor stream

Supplementary Information

Supplementary Tables

Table 1. Top 100 marker genes of EOM and TA of in vitro activated MuSCs

Table 2. Reactome Pathway Analysis for EOM and TA in vitro activated MuSCs

Table 3. Molecular Function Analysis for EOM and TA in vitro activated MuSCs

Table 4. List of top 100 marker genes of EOM and TA of quiescent MuSCs

Table 5. Genes used to generate signature scores for EOM and TA in vitro activated MuSCs

Table 6. List of top 100 regulons of Prog and Diff clusters of EOM and TA activated MuSCs

Table 7. Foxc1 targets identified by pyScenic

Table 8. List of top 100 regulons of the global EOM and TA activated MuSCs populations

Table 9. List of top 100 regulons of the global EOM and TA quiescent MuSCs populations

Table 10. List of top 100 drivers for EOM progenitors

Table 11. List of top 100 drivers for TA progenitors

Table 12. Antibodies and resources used in this study

Table 13. qRT-PCR primer list

Supplementary Figure Legends

Suppl Figure 1. EOM and TA MuSCs exhibit functional differences following activation

A. Experimental scheme. MuSCs were isolated by FACS based on GFP fluorescence from *Tg:Pax7-nGFP* mice. A fraction of these cells was individually plated on a 96-well plate (T1) and the remainder were activated for 2 days *in vitro* before re-sorting and clonal plating on a 96-well plate (T2). Cells were allowed to proliferate for 14 days and assessed for total cell number per clone.

B. Quantification of cell numbers per clone at both time points. Each dot represents a clone. Representative of 2 independent experiments. The mean of the clone size is indicated.

C. Western blots for AKT/Phospho-AKT, p38/Phospho-p38 and VINCULIN for normalization of protein loading. n =3 with cells coming from two mice pooled per lane.

D. Densitometric analysis of the Western Blots in D. Two-tailed unpaired Student's t-test; *p-value<0.05.

Suppl Figure 2. Single cell transcriptome signatures of activated EOM and TA MuSCs

A. Quantification of Myod staining intensity. Violin plots show the single cell distribution Myod staining intensity of activated EOM and TA MuSCs from *Pax7*^{CreERT2}*:R26*^{mTmG}. Dots superimposed on violin plots correspond to the average value in the independent wells (n = 3 wells, each replicate encompasses measurements from >1000 cells). Grey bars represent mean ± SD. P values were calculated using a two-tailed unpaired Student's *t* test (p-value < 2.2e-16).

B-C. Reactome pathway (A) and GO Molecular function (B) network analysis on top 100 DEGs of each cluster. Pie charts represent relative contribution of each cluster to this ontology term. Genes belonging to each term are highlighted in red.

Suppl Figure 3. Molecular overlap of EOM and TA MuSCs across cell states

A. Heatmap of adult quiescent MuSCs isolated from cranial mesoderm-derived (extraocular muscle (EOM), esophagus (ESO), masseter (MASS)) and somitic mesoderm derived (Tibialis anterior (TA), diaphragm (DIA)) muscles from *Tg:Pax7-nGFP* mice.
B. Heatmap from in vitro activated sc-RNAseq datasets of selected signature genes of EOM and TA MuSCs identified in Figure 3F,G.
C. Heatmap on quiescent sc-RNAseq datasets of selected signature genes of EOM and TA MuSCs identified in Figure 3F,G.

Suppl Figure 4. EOM progenitors exhibit a distinctive mesenchymal cell signature

A. Scores of EOM Prog, TA Prog, EOM Diff and TA Diff fractions for the signature of SMMCs (Giordani et al. Mol Cell 2019), FAPs and Fibroblasts (Oprescu et al. iScience 2020), myotendinous junction B myonuclei (Kim et al. Nat Com 2020), Twist2+ population (Liu et al. Nat Cell Biol 2017), fetal MuSCs (Tierney et al. Cell Reports 2016), developing limb connective tissues (Esteves-Lima et al. Nat Com 2021), Skeletal Muscle Mesenchyme (Xi et al. Cell Stem Cell 2020) and Global Myogenic markers.
B. Scheme of isolation of EOM and TA MuSCs from *Tg:Pax7-nGFP* mice, culture and staining for PDGFRβ-BV786 before flow cytometry analysis.
C. Representative flow cytometry profiles of EOM and TA activated MuSCs as per in E stained with PDGFRβ-BV786.

Suppl Figure 5. Distinct gene regulatory networks underlie EOM and TA activation dynamics

A. UMAP visualization of EOM and TA quiescent (top) and in vitro activated (bottom) clusters.
B. Expression plots of Foxc1, Creb3l1, Ebf1 and Hoxa10 in both datasets.
C. Scheme of isolation procedure, culture and Western Blot analysis (D-F) of EOM and TA MuSCs from *Tg:Pax7-nGFP* mice.
D. Western blots of EOM and TA MuSCs cultured for 4 days showing expression of FOXC1, EBF1, CREB3L1, PAX7, CYCLIND1, H3/TUBULIN for normalization of protein loading. Cells coming from n=3 mice pooled per lane.
E. Stress Index score across the merged activated sc-RNAseq dataset, overlaid onto UMAP representation.
F. Plot highlighting the correlation between Stress Index and the expression of myogenic markers *Pax7* (progenitor) and *Myod1* (committed/differentiating) on the activated sc-RNAseq dataset.
G. UMAP representations of different clusters and the expression of *Myod1* and *Pax7* before and after Stress Index regression, or removal of Stress Index genes from the count matrix. Upon regression the shape of the UMAPs is similar, although it comes out as a mirror image of the control.

Suppl Figure 6. Increased expansion capacity of EOM MuSCs is retained after several passages in vitro

A. FACS plots of *Tg:Pax7-nGFP;Myog*^{ntdTom} EOM and TA MuSCs following 3 days in culture (Act), and one or two passages (P1, P2). Cells were isolated by FACS based on GFP fluorescence from *Tg:Pax7-nGFP;Myog*^{ntdTom} and cultured *in vitro* for 5 days (Act) (n=4). A faction of cells was analysed by Flow

cytometry while the rest of cells were plated for the next time point. This procedure was repeated for all time points (passages P1 and P2, n=3). Cells were split every 3 days.

B. Brightfield images of EOM and TA MuSCs (n=3) at indicated time points (D3, day3; D5, day5; passages P1 and P2). Cells were split every 3 days.

Suppl Figure 7. EOM features emerge prenatally and are kept with passages in vitro

A. Foetal (E18.5) and postnatal (P21) EOM and TA MuSCs from *Tg:Pax7-nGFP* mice were cultured for 48h then processed for RNA extraction. Graph shows RT-qPCR for *Myogenin* normalized to *Rpl13*. Two-tailed unpaired Student's t-test; *p-value<0.05. n=4 mice.

B. FACS plots of *Tg:Pax7-nGFP;Myog*^{ntdTom} EOM and TA MuSCs isolated from P7 mice.

C. Expression table for Foxc1 extracted from MuscleDB dataset (Terry et al. eLife 2018) showing higher expression of Foxc1 in EOM (depicted here as eye muscles) compared to 10 other muscle groups.

D. Scheme of siRNA treatment of EOM MuSCs isolated from adult *Tg:Pax7-nGFP* mice.

E. Quantification of experiment in E. At Day2, total cell numbers were counted and percentage of PAX7+ and FOXC1 + cells out of total number of cells in the well was assessed by immunostaining in scramble control (ctrl) and Foxc1 siRNA treated cells (n=3 mice).

F. qRT-PCR for *Myogenin* on control and Foxc1 siRNA treated EOM cells as per in E. Graph shows qRT-PCR results for *Myogenin* normalized to *Rpl13* on control and Foxc1 siRNA treated EOM cells as described in E. (n=3 mice).

Suppl Figure 8. Cell cycle does not explain the velocity streams of in vitro activated EOM and TA MuSCs

A. UMAP representation of EOM velocity streams overlaid with cluster and cell cycle information.

B. UMAP representation of TA velocity streams overlaid with cluster and cell cycle information.