Title:

Human skeletal muscle organoids model fetal myogenesis and sustain uncommitted PAX7 myogenic progenitors.

Authors:

Lampros Mavrommatis[1,2*], Hyun-Woo Jeong[3*], Gemma Gomez-Giro[2,4], Martin Stehling[5], Marie-Cécile Kienitz[6], Olympia E. Psathaki[2,7], M. Gabriele Bixel[3], Gabriela Morosan-Puopolo[1], Daniela Gerovska[8], Marcos J. Araúzo-Bravo[8,9], Jens C. Schwamborn[4], Hans R. Schöler [2,10], Ralf H. Adams[3,10], Matthias Vorgerd[11,12], Beate Brand-Saberi[1], Holm Zaehres[1,2*].

[1] Ruhr-University Bochum, Medical Faculty, Department of Anatomy and Molecular Embryology, 44801 Bochum, Germany

[2] Max Planck Institute for Molecular Biomedicine, Department of Cell and Developmental Biology, 48149 Münster, Germany

[3] Max Planck Institute for Molecular Biomedicine, Department of Tissue Morphogenesis, 48149 Münster, Germany

[4] Luxembourg Centre for Systems Biomedicine, LCSB, Developmental and Cellular Biology, University of Luxembourg, 4367 Belvaux, Luxembourg

[5] Max Planck Institute for Molecular Biomedicine, Flow Cytometry Unit, 48149 Münster, Germany

[6] Ruhr-University Bochum, Medical Faculty, Department of Cellular Physiology, 44801 Bochum, Germany

[7] Center for Cellular Nanoanalytics Osnabrück, CellNanOs, University of Osnabrück, 49076 Osnabrück, Germany

[8] Computational Biology and Systems Biomedicine, Biodonostia Health Research Institute, San Sebastián 20014, Spain

[9] IKERBASQUE, Basque Foundation for Science, Bilbao 48013, Spain

[10] Westphalian Wilhelms-University Münster, Medical Faculty, 48149 Münster, Germany

[11] Heimer Institute for Muscle Research, University Hospital Bergmannsheil, 44789 Bochum, Germany

[12] Neurological Clinic Bergmannsheil, Ruhr-University Bochum, 44789 Bochum, Germany

[*] Correspondence: Lampros Mavrommatis (lampros.mavrommatis@rub.de), Hyun-Woo Jeong (hyun-woo.jeong@mpi-muenster.mpg.de), Holm Zaehres (holm.zaehres@rub.de).

Abstract

In vitro **culture systems which structurally recapitulate human myogenesis and promote PAX7^{+} myogenic progenitor maturation have not been established. Here we report human skeletal muscle organoids differentiated from induced pluripotent stem cell lines that contain paraxial mesoderm and neuromesodermal progenitors and develop into organized structures, reassembling neural plate border and dermomyotome formation. Culture conditions cause neural lineage arrest and promote fetal hypaxial myogenesis towards limb axial anatomical identity and generate sustainable uncommitted PAX7 myogenic progenitors, as well as fibroadipogenic (PDGFRa+) progenitor populations equivalent to those from second trimester of human gestation. Single cell comparison to human fetal and adult myogenic progenitors reveals distinct molecular signatures for non-dividing myogenic progenitors in activated (CD44^{High}/CD98^{+}/MYOD1^{+}) and dormant (PAX7^{High}/FBN1^{High}/SPRY1^{High}) states. Our approach, further validated with Duchenne and CRISPR/Cas9 genome-edited Limb-girdle muscular dystrophy (LGMD2A) patient iPSC lines, provides a robust *in vitro* developmental system for investigating muscle tissue morphogenesis and homeostasis.**

Introduction

Most skeletal muscle progenitors of vertebrates develop from the paraxial mesoderm via the transient embryonic developmental structures somites and dermomyotome into the skeletal muscle system that spans the whole body. Researchers facilitated two-dimensional (2D) culture conditions to guide pluripotent stem cell (PSC) differentiation towards skeletal muscle lineage using sequential growth factor applications and/or conditional PAX7 expression (Chal et al., 2015, Xi et al., 2017, Shelton et al., 2014, Borchin et al., 2013, Darabi et al., 2012). Subsequently surface marker combinations could be used to isolate of myogenic progenitors with *in vivo* repopulation potential (Magli et al., 2017, Hicks et al., 2018, Al Tanoury et al., 2020). A few three-dimensional (3D) differentiation approaches provide cohorts of terminally differentiated myofibers and focus their investigation on potential interactions with the vasculature and nervous system without evaluating the developmental identity or sustainability of myogenic progenitors (Faustino Martins et al., 2020, Maffioletti et al., 2018, Rao et al., 2018). Single cell technologies increasingly provide databases for deciphering myogenic trajectories and expression profiles of myogenic stem and progenitor cells (Xi et al., 2020) and PSC differentiation protocols can be evaluated in their potency to mimic human development. The translational perspective to model muscular dystrophies in vitro and provide expandable populations of muscle progenitors requests methods that promote self-renewal and preserve myogenic progenitors in a quiescent, non-diving, state (Quarta et al., 2016, Montarras et al.,2005).

Here, we introduce human skeletal muscle organoids to structurally investigate myogenic differentiation from human induced pluripotent stem cells (iPSCs) in a three-dimensional environment mimicking pathway discovered in chicken and mouse (Buckingham and Rigby, 2014). First, we drive differentiation towards paraxial mesoderm through application of the GSK3 inhibitor CHIR99021, BMP inhibitor LDN193189 and bFGF, while subsequent stimulation with WNT1A, SHH, FGF and HGF promotes derivation of organized structures, reassembling neural plate border and dermomyotome formation. Omitting FGF from culture conditions causes neural lineage arrest, while stimulation with HGF selectively promotes propagation of undifferentiated myogenic progenitors only, concluding into generation of fetal myofibers in the proximity of $PAX3^+$/ $PAX7^+$ cells. Sustainable $PAX7^+$ myogenic progenitors reside predominantly in a non-divided quiescent state at least 14 weeks after differentiation induction. Single cell analysis position them along a quiescent-activation myogenic trajectory discriminating dormant (PAX7+, FBN1+, SPRY11+, CHODL1+) and activated (CD44+, CD98+, MYOD1+, VEGFA+) states. Furthermore, a PDGFRa+ fibroadipogenic progenitor equivalent to those in human fetal muscle tissues emerges in our organoid system. We are providing a novel developmental model for investigating human myogenesis with translational potential to model and cure muscle disorders.

Results

Generation of human fetal skeletal muscle organoids

Organoid cultures from pluripotent stem cells often require pre-patterning before embedding them into Matrigel in order to promote their structural development **(**Lancaster et al., 2013; Spence et al., 2011; Koehler et al., 2017). In our three-dimensional approach, we decided to provide the cells with immediate matrix support upon embryoid body formation with the rationale to preserve cell state transitions. Initial screening before embedding, indicated high expression of the pluripotent markers OCT4, SOX2 and NANOG, moderate of the neural tube marker PAX6 and low expression of mesodermal markers BRACHYURY and MSGN1 **(Figure 1 - figure supplement 1A)**. Upon Matrigel embedding stimulation with CHIR99021, LDN193189 and bFGF promoted paraxial mesoderm formation through derivation of BRACHYURY and TBX6 positive cells **(Figures 1A,C - figure supplement 1A)**. Immunostaining at Day 5 depicted presence of mesodermal (BRACHYURY$^+$), paraxial mesodermal (TBX6$^+$) and neuromesodermal progenitors (SOX2$^+$/ BRACHYURY$^+$) (Gouti et al., 2017; Henrique et al., 2015) **(Figure 1C).** From that stage we attempted to mimic the determination front formation and promote anterior somitic mesoderm (ASM) (Aulehla and Pourquie,2010; Shimozono et al., 2013), by maintaining constant CHIR99021, LDN193189 levels and by reducing bFGF levels to 50% and simultaneously introducing retinoic acid to the culture. Consequently, distinct PAX3$^+$ but SOX2$^-$ cells emerged on the organoid surface **(Figure 1D)**, followed by a significant downregulation of the PSM markers HES7, TBX6 and MSGN1 **(Figures 1E – figure supplement 1A)**. Concomitantly, we observed significant upregulation of the ASM marker MEOX2 and the neural crest marker TFAP2A but not of SOX2 or PAX6, excluding a shift towards neural tube formation at this stage **(Figures 1E,F - figure supplement 1A)**. Until this stage organoid culture exhibits small variability and high reproducibility **(Figure 1B)**. Dermomyotomal fate was promoted by Sonic Hedgehog (SHH) and WNT1A stimulation, while maintaining same levels of BMP inhibition avoided lateral mesoderm formation **(Figure 1A)**. Expression profiling at Day11 depicted presence of neural tube/crest (SOX2, PAX6, TFAP2A, SOX10 upregulation) and mesodermal (UNCX, TBX18, PAX3) lineages and further downregulation of the PSM markers **(Figures 1E,F - figure supplement 1A)**. Notably, at this stage, the dermomyotomal/neural crest marker PAX7 emerged together with markers that define the dorsomedial (EN1) or the ventrolateral portion (SIM1) of the dermomyotome (Cheng et al., 2004) **(Figures 1D,F - figure supplement 1A)**.

Consequently, for favoring myogenesis we stimulated the organoid culture with FGF and HGF (Charge and Rudnicki, 2004) **(Figure 1A)**. Surprisingly, at day 17, organoids constituted a mosaic of neural crest and myogenic progenitor cells. Cells with epithelial morphology were TFAP2A$^+$, SOX2$^+$, PAX3$^+$ and PAX7$^+$, indicating formation of the neural plate border epithelium (Roellig et al., 2017) **(Figure 1G)**. In cell clusters with mesenchymal morphology, we detected

specified migrating PAX3$^+$/SOX10$^+$ neural crest progenitors and cells of myogenic origin by being PAX3$^+$ /SOX2$^-$/ SOX10$^-$ **(Figure 1G)**. At this stage, myogenic lineages seemed to be represented mainly by PAX3 and not PAX7 positive cells (9.35±0.07% PAX3$^+$), as the PAX7 expression pattern (15.3±0.1% PAX3$^+$/PAX7$^+$ and 5.11±0.13% PAX7$^+$) was mainly overlapping with that of SOX2 and TFAP2A **(Figures 1G - figure supplement 1B)**. Morphologically, neural plate boarder and dermomyotomal populations exhibited uneven distribution, and thereby the subsequent neural crest and myogenic migration processes introduced variability in organoid´s size **(Figure 1B,G)**. From 17 d.p.d onwards, we omitted FGF stimulation to cease neural crest development and promote the delamination/migration of the PAX3$^+$/ PAX7$^+$/SOX10$^-$ progenitor cells, we omitted FGF stimulation (Murphy et al., 1994) **(Figure 2A)**. Strikingly, within a short time-period, we could already observe committed myogenic populations by detecting fast MyHC myofibers in the proximity of PAX7 positive but SOX2, TFAP2A, SOX10 negative cells **(Figures 1H - figure supplement 1C to E)**. Consistently, expression profiling indicated downregulation of neural tube/crest lineage markers and significant upregulation of muscle precursor migrating markers (Buckingham and Rigby, 2014), like LBX1, CXCR4 and MEOX2 **(Figures 1E, 2A and Figure2 - figure supplement 1A)**.

At 8 weeks organoids showed profound changes in their transcription profiling **(Figure 2 - figure supplement 1B),** when gene ontology enrichment analysis highlighted an ongoing developmental transition within the organoid culture, with muscle development, muscle contraction, myofibril assembly, as well as neurogenesis and nervous system development being among the top upregulated and downregulated gene ontology terms, respectively **(Figure 2B)**. In addition, we detected downregulation of key markers characterizing neural tube and neural crest development (Soldatov et al., 2019), such as PAK3, DLX5, B3GAT1, FOXB1, HES5 and OLIG3 **(Figure 2 - figure supplement 1C)**. Interestingly, we could spatially visualize this process using immunocytochemistry for SOX2, TFAP2A and SOX10 expressing cells that are restricted to the inner portion of the organoid, and probably not susceptible to culture stimulation at 5 weeks **(Figures 1H, 2C,D and Figure 2 - figure supplement 1D)**. This neural / myogenic lineage spatial orientation could be visualized even at day 84 post differentiation through presence of TUJ1$^+$ neurons only at the inner portion of the organoid and close to SOX2 positive epithelium, while FastMyHC positive myofibers occupied the organoids exterior part **(Figure 2E).** On the other hand, substantial proportions of migratory cells in the proximity of FastMyHC$^+$, MF20$^+$ myofibers were expressing PAX7 (40±0.3%) but not SOX2, TFAP2A, SOX10 or MYOD1 (3.55±0.32%) (**Figures 2F - figure supplement 1D,E**). This behavior is further illustrated by the presence of MYOD1 positive cells only at the organoid´s periphery **(Figure 2D)**.

Culture progression led to significant increase of PAX7$^+$ and MYOD1$^+$ cells at 8 weeks (45,3±3,4% and 49,8±0.62%, respectively) (**Figures 2F - figure supplement 1E**), suggesting an interval in which the organoid culture predominantly commits into myogenic lineage. Expression profiling at this stage, depicted upregulation of markers characteristic for limb migrating myogenic cells e.g LBX1, PAX7, SIX1/4, EYA1/4, PITX2, MYF5, TCF4 and MSX1 **(Figure 2 - figure supplement 1A)**. Consequently, HOX gene cluster profiling at 4th week correlated HOX A9, B9, C9 upregulation with the limb bud site, while upregulation of the HOX 10-13 cluster during 8th,16th weeks, attributed the later organoid development to a more distal limb axial anatomical identity (Shubin et al., 1997; Xu et al., 2011, Raines et al., 2015), **(Figure 2G)**.

Lineage representation and developmental identity for skeletal muscle organoids

Analysis of organoid culture at single cell resolution, indicates predominant presence of skeletal muscle lineage (n= 3945 cells, 91,26% of total population), complemented with two smaller cell clusters of mesenchymal (n=165 cells, 3,82% of total population) , and neural (n=213 cells, 4,93% of total population) origin **(Figures 3A - figure supplement 1A**). Skeletal muscle lineages further divided into distinct sub-clusters: the myogenic muscle progenitors at non dividing (PAX7$^+$/PAX3$^-$) and mitotic states (PAX7$^+$/CDK1$^+$/KI67$^+$), the myoblasts (MYOD1$^+$), the myocytes (MYOG$^+$) and the myotubes (MYH3$^+$) **(Figures 3A - figure supplement 1A,B)**. Further, investigation on the identity of the mesenchymal cluster suggested fibroadipogenic potential of a distinct cell population with PDGFRa among the top upregulated genes (Uezumi et al., 2010,2011; Xi et al., 2020) **(Figure 3A,F - figure supplement 1A,B).** Consistently, we could detect upregulation of fibrotic markers. Moreover, adipogenic potential was highlighted through upregulation of PREF-1 and EBF2 **(Figure 3F)**. Occasionally, we could detect structures resembling adipocytes from 9-10th weeks onwards in the organoid, whereby we verified their adipogenic nature by PRDM16 and Oil red O positive staining **(Figures 3B,D - figure supplement 1C,D)**. A SOX2$^+$/TUBB3$^+$ positive neural cluster potentially derived from the neural plate border epithelium at younger stages, could be located this population towards the interior part of the organoids **(Figure 3C - figure supplement 1E).**

Myogenesis progression based on *t*-SNE feature and Violin plots of key markers from each stage indicated gradual transitions from myogenic progenitor to myotube subclusters **(Figures 3E - figure supplement 1F)**. MyHC expression profiling at single cell resolution demonstrated predominant presence of embryonic (MYH3$^+$) and perinatal (MYH8$^+$) myofibers, co-expressing β-Enolase and M-cadherin in the vicinity of PAX7$^+$ cells **(Figure 3 G - J)**. Consistently, we detected MCAM expression on muscle sub-clusters **(Figue 3G)**, while bulk RNA seq at week 16th highlighted expression of additional MyHC isoforms, e.g. MYH2 (II_a), MYH4(II_b) and of the transcription factor NFIX (Moore et al.,

1993; Barbieri et al., 1990; Alexander et al. 2016; Messina et al., 2010). At more mature stages, we could detect adult MYH1 or slow MYH7 isoforms **(Figures 3 - figure supplement 1G),** presumably due to myofiber stimulation via spontaneous contraction **(Video S1, S2)**. In agreement, differential expression comparison between 8- and 16-weeks organoids indicated less variance in transcription profiling **(Figures 3 - figure supplements 1H, 2A)**. Consistently, sarcomere organization, ion transport, response to stimulus or synapse structure maintenance and mitosis, cell cycle and DNA packaging and nuclear division were among the top upregulated and downregulated gene ontology terms, respectively **(Figures 3 - figure supplement 2B-D)**. Notably, the ongoing maturation process did not affect the pool of progenitor cells, as even at week 14th we could report significant increase in PAX7^{+} (50,16±2.19%) and MYOD1^{+} (58.53±0.92%) cells **(Figures 2F and Figures 3 - figure supplement 2E,F).**

Functionality and maturation of organoid derived myofibers

Regarding localization and functionality of organoid-derived myofibers immunocytochemistry revealed, positive stainings for DYSTROPHIN, and a continuous sheath of laminin around individual muscle fibers. Ultrastructural and 2-photon microscopy analysis depicted well developed sarcomeres **(Figures 3 - figure supplement 3A-D)**. Moreover, patch clamp experiments, upon superfusion of cells with Acetylcholin (ACh), indicated inward current generations that rapidly declines to a low steady state level **(Figures 3K - figure supplement 3E**). The I/V-curve of this current component showed an almost linear behavior and a reversal potential around 0 mV **(Figure 3L)**. These data are in line with previous studies that described currents with analogous properties as characteristic nAChR currents (Jahn et al., 2001; Shao et al., 1998). Application of a fluorescent biosensor for monitoring changes in cytosolic Ca^{2+} (Twitch2B,) revealed efficiency of nAChR in modulating intracellular [Ca^{2+}]. Summarized FRET recordings **(Figures 3M - figure supplement 3F**), following application of ACh (10 μM), illustrated a rapid increase in FRET ratio that gradually declined in the presence of ACh, probably reflecting desensitization of nAChRs. These results demonstrated that nAChR in skeletal muscle cells are functional in terms of inducing Ca^{2+} release from intracellular stores.

Genome editing on LGMD2A patient line using CRISPR/Cas9

Using our organoid protocol, we successfully derived fetal muscle progenitors and electrophysiologically functional myofibers from hiPSC lines with Duchenne Muscular Dystrophy and Limb Girdle Muscular Dystrophy (LGMD2A) genetic backgrounds (**Figure 3 - figure supplement 3E,F**). For the LGMD2A iPSC patient line, being heterozygous for two mutations in neighboring exons **(Figure 4A)**, we developed a new strategy for introducing biallelic modifications within the genome using the CRISPR/Cas9 genome editing technology to rescue the patient mutations. In this approach, we introduce a stringent selection process of successful biallelic events based on fluorescent interrogation

and drug administration. Consequently, the selection cassette system composed of two plasmids, one expressing under the constitutively active CAG promoter the reverse Tetracycline Activator (rtTA) protein and another under the inducible TRE-CMV promoter the tdTomato plus puromycin (**Figure 4B**). This approach generated a depending loop between the expression cassettes, ensuring that only upon integration of both plasmids red fluorescence could be detected. In addition, to ensure that both alleles are selection cassettes free upon PiggyBac mediated excision removal, it was a necessity that both plasmids were expressing the eGFP protein under a constitutively active promoter **(Figure 4B)**. This approach allowed us to simultaneously rescue the LGMD2A patient mutations by deletion (intron skipping), insertion (EXON 4 deletion mutation) and base exchange (Exon3) in a single step **(Figures 4C - figure supplement 1 A-D)**. Upon organoid differentiation we could successfully derive PAX7 myogenic progenitor and FastMyHC myofibers **(Figure 5E)**. Further investigation on LGMD2A lines highlighted differences on mitochondrial proteins, such as TOM20, and on prolonged intracellular Ca2+ concentration upon activation with ACh, as well as reduced SERCA2 expression for the patient LGMD2A line (Kramerova et al., 2009; Toral-Ojeda et al., 2016) **(Figures 4F,G - figure supplement 1E,F).**

Identity and sustainability of organoid derived PAX7 myogenic progenitors

Skeletal muscle organoids remarkably foster the sustainable propagation of PAX7^{+} progenitors **(Figures 2F, 3A)**. Initial screening of PAX7^{+} progenitors indicated expression of several satellite cell markers (Fukada et al., 2007), such as CD82, CAV1, FGFR1, FGFR4, EGFR, M-Cadherin, NCAM, FZD7, CXCR4, ITGβ1, ITGA7, SCD2 and SCD4 **(Figure 5 - figure supplement 1A)** Further investigation verified NOTCH signaling activity in the myogenic subcluster, HES1, HEYL, HEY1 and NRARP*,* while in dormant myogenic progenitors, we could further detect high expression of SPRY1, and of the cell cycle inhibitors p57 (CDKN1C), p21, and PMP22 (Xi et al., 2020; Fukada et al., 2007; Bjornson et al., 2012; Shea et al., 2010) **(Figure 5 - figure supplements 1B, 2A,D)**. Moreover, proliferation assays performed at the same stage demonstrated 4.78±0.28% edU^{+} cells, while substantial proportion of PAX7^{+} myogenic progenitors remained quiescent **(Figure 5 - figure supplements 1C,D)**. t-SNE clustering divided myogenic progenitors into three distinct groups with unique molecular signatures, the "CD44^{high}", the "FBN1^{high}" and the "CDK1^{+}" cluster **(Figures 5A - figure supplements 2A)**. The "CD44^{high}" cluster, further characterized by CD98 upregulation and adopted a molecular signature similar to the activated satellite cell state (Porpiglia et al., 2017) **(Figures 5A,B - figure supplement 2A,D).** At this state, myogenic progenitors express VEGFA **(Figure 5 - figure supplement 2A,D)** as previously been described for murine satellite cells (Verma et al., 2018)**.** Consistently, the CD44^{+} / PAX7^{+} positive 'activated' myogenic population located at sites more accessible to HGF signaling, *e.g* the organoid´s exterior portion and forming bulges **(Figure 5B)**. The "FBN1^{High}" sub-cluster was further characterized by "PAX7^{high} / SPRY1^{high}/CHODL^{high} / FBN1^{high}" expression

(Figures 5A - figure supplement 2A,D). Especially, PAX7^{high} cells within the dormant state co-expressed NOTCH3, JAG1 and LNFG, together with CHODL markers. Further, we could verify presence of FBN1^{+} microfibrils but, compared to CD44^{+} cells, they occupied areas without direct access to activation signals **(Figure 4C)**. Last, the "CDK1^{+}" cluster was the only proliferative population, further expressing markers mediating satellite cell activation and proliferation, such as DEK and EZH2 (Cheung et al., 2012; Juan et al., 2011) **(Figures 5A,D - figure supplements 1A, 2A,B).**

Pseudotemporal ordering of myogenic progenitors, indicated that the "FBN1^{High}" subcluster was the main progenitor population, residing in dormant state, which by downregulating PAX7, FBN1, CHODL, and upregulating CD44, MYOD1, CD98 (SLC3A2) and VEGFA, generated the activated state **(Figures 5D - figure supplement 2B)**. At latter state, myogenic progenitors that upregulate MYOD1 enter mitosis, proliferate and follow a trajectory leading to myogenic commitment **(Figures 5D - figure supplement 2B,D)**. This commitment is further accompanied by PARD3, p38a MAPK and CD9 expression (Porpiglia et al., 2017; Troy et al., 2012) **(Figure 5E).** On the other hand**,** myogenic progenitors that upregulate NOTCH3, SPRY1 and PAX7 follow a loop trajectory that reinstates the dormant stem cell state **(Figures 5D - figure supplement 2B,D)**. Interestingly, "FBN1^{high}" cluster highly expressed extracellular matrix collagens, e.g COL4A1, COL4A2, COL5A1, COL5A2, COL5A3, COL15A1. Notably, their expression declined upon commitment and differentiation. (**Figures 5D,F - figure supplement 2C**).

Organoid derived myogenic progenitor comparison to human fetal and adult progenitor/stem cells

To evaluate developmental identity of the myogenic cluster, we isolated ITGβ1^{+} / CXCR4^{+} organoid-derived myogenic progenitors via FACS and compared them to human fetal muscle progenitors demonstrating high similarity (Pearson´s correlation co-efficiency, rho=0.9), myogenic markers, eg. PAX7, MYF5, MYOG, MYOD1, exhibited similar expression levels **(Figure 5 - figure supplement 3A - F).** Differential expression comparison verified expression of extracellular matrix collagens and proteins, such as COL4A1, COL5A1, COL6A1, COL15A1, FBN1, CHODL also in myogenic progenitors from 17th week human fetal tissue (Pearson´s correlation co-efficiency, rho=0.97) **(Figure 5G)**. Further, to evaluate their myogenic potency *in vitro*, isolated ITGβ1^{+} / CXCR4^{+} organoid-derived myogenic progenitor cells were re-plated and allowed to differentiate under the same organoid culture conditions, proving their capacity to generate FastMyHC+ and PAX7+ populations within 10 days **(Figure 5 - figure supplement 3B,C**). Subsequently, comparison to already available transcriptomic dataset of human adult satellite cells at single cell resolution, divided myogenic progenitors and adult satellite cells into four clusters with distinct molecular signatures **(Figure 5H)**. Interestingly, myogenic progenitors were enriched for extracellular matrix proteins, while satellite cells mainly upregulated genes repressing transcription/translation, such as ZFP36, TSC22D1, or genes related to early activation response, such as

MYF5, JUN, FOS **(Figure 5 - figure supplement 4A,B).** In line, organoid-derived myogenic progenitors exhibited higher NOTCH signaling activity in comparison to satellite cells, with NOTCH3 and RBPJ being among the markers enriched in organoid-derived myogenic progenitors **(Figure 5J**). In contrast to organoid-derived non-dividing myogenic progenitors, adult satellite cells exhibited PAX7^{low} / MYF5^{high} expression profiling, presumably due to tissue dissociation, thereby indicating a tendency for activation rather than preserving or reinstate quiescency (Machado et al., 2017; Seale et al., 2000) **(Figure 5I).** Eventually, pseudotime ordering divided adult satellite cells and organoid-derived myogenic into two distinct clusters with adult satellite cells being downstream of non-dividing myogenic progenitors **(Figure 5 - figure supplement 4C).** Consistently, downregulation of genes promoting quiescence like PAX7, NOTCH3, RBPJ, and upregulation of activation genes like MYF5, JUN, FOS along the trajectory (**Figure 5 - figure supplement 4D)** was a further indication that organoid derived myogenic progenitors reside in dormant non dividing state and that the organoid platform promotes sustainability of myogenic progenitors.

Organoid derived myogenic progenitors exhibited expression profiling similar to that of stage 4 myogenic progenitors, correlating to 17-18th week of human fetal development, recently reported in the human skeletal muscle atlas (Xi et al., 2020). SPRY1, PLAGL1, POU4F1, TCF4, PAX7, CD82, CD44 markers, as well as ECM proteins, nuclear factor I family members (NFIA, NFIB, NFIC) and NOTCH signaling are specifically upregulated markers. Furthermore, in comparison to 2D culture protocols, the organoid approach preserves myogenic progenitors dormant without activating MYOD1 and exhibits higher expression on ECM proteins and key markers characterizing the stage 4 fetal myogenic progenitors **(Figure 5 - figure supplement 5)**. ERBB3 and NGFR markers (Hicks et al., 2017) demarcate populations of earlier developmental stages and are not upregulated either in organoid derived or in stage 4 fetal myogenic progenitors **(Figure 5 - figure supplement 5)**. In addition, organoid myogenic progenitors downregulate cycling genes like MKI67, TOP2A, EZH2, PCNA, like stage 4 fetal myogenic progenitors **(Figure 5 - figure supplement 5)**.

Discussion

Human skeletal muscle organoids introduce a new cell culture system to study human myogenesis, particularly the occurrence of fetal myogenic progenitors. We could demonstrate that modulation of matrigel embedded embryonic bodies with WNT, BMP and FGF signaling at early stages leads to paraxial mesoderm formation **(Figure 1B)**. Further, we show that under guided differentiation we could promote the concomitant development of neural and paraxial mesodermal lineages and derive mesodermal populations with somitic and dermomyotomal origin **(Figure 1C - F).** Especially, upon WNT1A and SHH stimulation, neural lineage is directed towards dorsal neural tube / crest development what benchmarks the structural recapitulation of neural plate border epithelium **(Figure 1G)**. In vitro neural

plate border can be utilized to assess transcriptomic networks and cell fate decision during human neural crest formation.

Delaminating from the dermomyotome, undifferentiated PAX3 progenitor cells reorientate and align in the cranio-caudal axis to form the underlying myotome or completely detach from the dermomyotome and migrate into the surrounding mesenchyme at the limb sites, where they propagate and become committed to skeletal muscle progenitors (Relaix et al., 2005). By stimulating organoid culture at the neural plate border/dermomyotomal stage with bFGF/HGF we could further visualize both the migration of myogenic progenitors and migration/specification of neural crest populations (**Figures 1A,G,H, 2A).** Further, by omitting FGF during organoid development, we could detect a continuous upregulation of genes involved in the myogenic migration process, such as LBX1, PAX3, PAX7 and CXCR4, but not in genes characterizing neural tube or neural crest development, such as SOX10, TFAP2A, PAK3, DLX5, B3GAT1, FOXB1, HES5 and OLIG3. This indicates that organoid culture conditions and especially stimulation with HGF favored skeletal muscle over neural lineage **(Figures 2C-E)**. Interestingly, we could show that by stimulating the organoid culture with SF/HGF, an essential signal for detachment and subsequent migration, but not for specification of the cells at the lateral lip of the dermomyotome (Dietrich et al., 1999), we could preserve the PAX3+/PAX7+ myogenic migratory population in an undifferentiated and not committed state **(Figure 2D,E).** Strikingly, expression profiling based on HOX gene cluster supported this notion, as over time the organoid culture adopted more distal than proximal limb axial anatomical identify **(Figure 2F)**.

Fetal development is characterized by the PAX3+ to PAX7+ myogenic progenitor transition (Seale et al., 2000; Relaix et al., 2005) which we were able to demonstrate in our organoid culture. Our data further supported that organoid culture represented fetal stages as we could detect NFIX upregulation and verify presence of myofibers expressing fetal MyHC isoform as well as NCAM, M-Cad or MCAM in the proximity of PAX7 progenitors **(Figure 3G to J)**. Consequently, transcriptome comparison indicated high similarity to those of human fetal muscle tissue (17-week, Pearson correlation, rho=0.9) **(Figure 5G)**, as well as expression of several satellite cell markers **(Figure 5 - figure supplement 1A).** This was further verified by comparison to a single cell dataset from the human skeletal muscle cell atlas **(Figure 5 - figure supplement 4, 5)**. Interestingly, comparison at single cell resolution clustered portions of adult satellite cells together with organoid derived myogenic progenitors **(Figure 5H**). In addition, pseudotime ordering indicated that organoid derived myogenic progenitors reside developmentally upstream of activated satellite cells, upregulating markers associated with quiescence such as NOTCH signaling and extracellular matrix proteins **(Figures 5E,F,J - figure supplement 4C,D)**. Preservation of myogenic progenitors in a non-dividing state without activating

MYOD1 and self-renewal **(Figures 5D-F - figure supplement 1C,D)** seemed to be responsible for the observed sustainable propagation of PAX7 progenitor cells even 14 week post differentiation **(Figures 2F, 3A),** while in current 2D or 3D protocols there is a constant decline of the myogenic progenitor reservoir (Chal et al., 2015, Faustino-Martins et al., 2020) **(Figure 5 - figure supplement 5)***.* Upregulation of the committed myogenic marker MYOD1 in current protocols could explain the constant progenitor decline over time, as well as their more undifferentiated state.

In this context, we could observe high expression of extracellular matrix proteins and upregulated NOTCH signaling in dormant non-dividing myogenic progenitors **(Figure 5F,J)**. This phenotype is similarly described for human fetal tissue myogenic progenitors (**Figure 5G, Pearson correlation, rho=0.97, figure supplement 5)**. Studies evaluating engraftment potential of satellite cells and fetal muscle progenitors propose that muscle stem cells in a quiescent non-committed state exhibit enhanced engraftment potential (Hicks et al., 2017; Quarta et al., 2016; Montarras et al., 2005; Tierney et al., 2016). Our data demonstrate that upon activation and commitment dormant myogenic progenitors downregulate extracellular matrix proteins and upregulate expression of morphogens / receptors that makes them susceptible to signals, like VEGFA that communicates with vasculature during tissue reconstitution, or CD9, CD44, CD98 participating in cellular activation (Porpiglia et al., 2017) **(Figure 5D - figure supplement 2B,D)**. It would be of great interest for future studies to further investigate whether increased engraftment is attributed to high expression of extracellular matrix proteins and whether their expression affects the engraftment potential of myogenic progenitors. Especially, high expression of Fibrillin1 on dormant non-dividing myogenic progenitors could potentially be a part of a mechanism, in which myogenic progenitors are able to avoid fibrosis by regulating TGF-β signaling levels in their surroundings (Cohn et al., 2007)*.*

We have established a novel 3D *in vitro* developmental system, which resembles distinct stages of human myogenesis and provides myogenic progenitors in dormant and activated states for at least 14 weeks in culture. We demonstrate that in contrast to 2D protocols, organoid culture sustains uncommitted MyoD1-negative myogenic progenitors as well as fibroadipogenic (PDGFRa+) progenitors, both resembling their fetal counterpart. Future work will elucidate signaling transduction pathways during skeletal muscle organoid development to model and understand human myogenesis in more detail.

ACKNOWLEDGEMENTS

We are grateful to Drs. Karl Köhrer, Tobias Lautwein, Patrick Petzsch and Thorsten Wachtmeister, Genomics & Transcriptomics Laboratory, Heinrich-Heine-University Düsseldorf for performing single cell and bulk RNAseq experiments with their Illumina HiSeq platform and data provision. We are further grateful to Dr. Kenjiro Adachi, Max Planck Institute Münster for providing the PiggyBac transposase expression vector, Dr. Oliver Griesbeck for the pcDNA3[Twitch-2B] plasmid and Dr. Stephan Hahn, Ruhr-University Bochum for access to equipment from his lab (CFX 96 Real-Time PCR detection system, BIO-RAD). We further would like to thank Dr. Christoph Clemen, Köln for performing experiments and Ingrid Gelker, Martina Sinn, Max Planck Institute Münster as well as Eva-Maria Konieczny, Rana Houmany, Boris Burr, Ruhr-University Bochum for their technical assistance. Electron microscopy experiments were supported by the Deutsche Forschungsgemeinschaft SFB 944. We thank Drs. George Q. Daley and Thorsten Schlaeger, Boston Children's Hospital for providing the Duchenne Muscular Dystrophy patient-derived iPS cell lines DMD-iPS1 and DMD-iPS2 in the course of our study. Our study was supported by research grants from FoRUM F873-16, Medical Faculty, Ruhr University Bochum, from Deutsche Gesellschaft für Muskelkranke e.V (DGM Foundation), Freiburg, the Georg E. und Marianne Kosing-Stiftung, Deutsches Stiftungszentrum, Essen and Deutsche Duchenne Stiftung der action benni & co e.V., Bochum.

AUTHOR CONTRIBUTIONS

L.M. conceived and designed experiments and the study, performed experiments, analyzed data, and wrote manuscript, H.W.J. performed bulk RNA seq. experiments, analyzed single cell and bulk RNAseq data, and wrote manuscript, G.G.G contributed to experimental design, analyzed data, and wrote manuscript. M.C.K. designed, performed, analyzed electrophysiology experiments and wrote manuscript. M.S performed and analyzed FACS experiments. O.E.P. performed electron microscopy experiments. M.G.B. performed multi-photon microscopy experiments. D.G., M.J.A.B. performed bioinformatic analysis, G.M.P performed experiments, J.C.S. H.R.S., R.H.A., provided materials and supervised experiments, M.V. provided patient and healthy biopsies and supervised study, B.B.S. provided materials, supervised study and edited manuscript, H.Z. supervised study, experimental design, and wrote manuscript.

Materials and Methods

hiPSCs culture

Human induced pluripotent stem cell (hiPSC) lines, Cord Blood iPSC (CB CD34$^+$, passage 15 – 35), LGMD2A patient iPSC (passage 5-25), Isogenic LGMD2A Exon 3-4 edited (passage 2-13), Duchenne Muscle Dystrophy iPSC patient line iPSCORE_65_1 (WiCell, cat.no. WB60393, passage 22-30) and DMD_iPS1 (passage 21-30) & BMD_iPS1 (passage 17-25) (Boston Children's Hospital Stem Cell Core Facility) (Dorn et al., 2015; Park et al., 2008; Panopoulos et al., 2017) were cultured in TESR-E8 (StemCell Technologies) on Matrigel GFR (Corning) coated 6 well plates. CB CD34$^+$ and LGMDA2 hiPSC lines were generated using the SFFV promotor – OSKM vector as described in Dorn et al., 2015. LGMD2A patient material was collected at the University Hospital Bergmannsheil after ethical approval from the ethics commission of the Ruhr-University Bochum, Medical Faculty (15-5401, 08/2015).

Human skeletal muscle organoid differentiation protocol

Prior differentiation, undifferentiated human PSCs, 60-70% confluent, were enzymatically detached and dissociated into single cells using TrypLE Select (ThermoFisher Scientific). Embryoid bodies formed via hanging drop approach with each droplet containing 3-4x10^4 human single PSCs in 20 µl were cultured hanging on TESR-E8 supplemented with Polyvinyl Alcohol (PVA) at 4mg/ml (SigmaAldrich) and rock inhibitor (Y-27632) at 10µM (StemCell Technologies) at the lid of Petri dishes. At the beginning of skeletal muscle organoid differentiation, embryoid bodies at the size of 250-300µm embedded into Matrigel and cultured in DMEM/F12 basal media (ThermoFisher Scientific) supplemented with Glutamine (ThermoFisher Scientific), Non Essential Amino Acids (ThermoFisher Scientific), 100x ITS-G (ThermoFisher Scientific), (Basal Media) 3µM CHIR99021 (SigmaAldrich) and 0.5µM LDN193189 (SigmaAldrich). At Day 3, human recombinant basic Fibroblast Growth Factor (bFGF) (Peprotech) at 10ng/µl final concentration added to the media. Subsequently, at Day 5 the concentration of bFGF reduced at 5ng/µl and the media further supplemented with 10nM Retinoic Acid (SigmaAldrich). The differentiation media, at Day 7, supplemented only with human recombinant Sonic hedgehog (hShh) (Peprotech) at 34ng/µl, human recombinant WNT1A (Peprotech) at 20ng/µl and 0.5µM LDN193189. At Day 11 the cytokine composition of the media changed to 10ng/µl of bFGF and human recombinant Hepatocyte Growth Factor (HGF) at 10ng/µl (Peprotech). From Day 15 onwards, the basal media supplemented with ITS-X (ThermoFisher Scientific) and human recombinant HGF at 10ng/µl. The first 3 days of the differentiation the media changed on a daily basis, from 3rd till 30th every second day, while from Day 30 onwards every third day. We evaluated the organoid approach with 6 hiPSCs lines with independent genetic background, with more than 5 independent derivations per line, especially for the control line (CB CD34$^+$) more than 20 derivation, obtaining always similar results. Per derivation and upon embryoid body matrigel embedding, cultures exhibited high

reproducibility. Upon migration and skeletal muscle formation organoids are occupying whole matrigel droplet and through generation of additional bulge reach size of 4-5 mM. Concomitantly, myogenic progenitors fall off the organoid and generate a sheath of myogenic organoids and muscle fibers at the surface of the culture plate. For all the lines, we could demonstrate functional myofibers and PAX7 positive myogenic populations. Myogenic populations from different lines exhibit high similarity, (Pearson correlation, rho = 0.94-0,95, **Figure 5 - figure supplement 3G**).

CRISPR/Cas9 genome editing of LGMD2A iPSC line

Genome editing approaches using CRISPR/Cas9 are based on distinguishing successful bi-allelic events through detection of fluorescent protein incorporation (GFP and RFP) into the genome (Arias-Fuenzalida et al., 2017). To these approaches, bi-allelic events can be distinguished by the presence of double GFP / RFP positive cells, while subsequent drug administration does not exclusively select for them. For a more stringent selection process, we developed a system that in addition to fluorescent interrogation, drug administration aids in selecting for cells with successful bi-allelic events.

*Selection cassette construction***:** The selection cassette system was incorporated into plasmids, flanked by 1kb in length homology arms. The Upper Homology Arm sequence was corresponding to the wild type sequence of EXON3 {130Trp(TGG)>Cys(TGC)} of the Calpain 3 locus till the point where by a silent mutation (CTG>CTT) a TTAA sequence corresponding to the Inverted Terminal Repeats (ITRs) of the PiggyBac system, could be generated **(Figure 4 - figure supplement 1B,C)**. This TTAA sequence marked the start for the selection cassette. The selection cassette system itself composed of two plasmids, one expressing under the constitutively active CAG promoter the reverse Tetracycline Activator (rtTA) protein and another under the inducible TRE-CMV promoter the tdTomato plus puromycin. This approach generated a depending loop between the expression cassettes, ensuring that only upon integration of both plasmids red fluorescence could be detected. In addition, to ensure that both alleles are “selection cassettes free” upon PiggyBac mediated excision removal, it was a necessity that both plasmids were expressing the eGFP protein under a constitutively active promoter. The sequence downstream the Inverted Terminal Repeats (ITRs) sequence of the PiggyBac system (TTAA) was corresponding to the lower homology arm of the selection system. To generate the new chimeric EXON3/4 exon without introducing any further alterations, the sequence of the lower homology arm was a chimera between the rest of the EXON3 sequence followed direct by the EXON4 wild type sequence, while skipping the sequence of the intermediate intron. Further, to avoid alterations on the selection cassette introduced by CAS9 nuclease, we mutated the PAM sequence of the cassette, by introducing another silent mutation (TTC>TTT) **(Figure 4B,C)**. The gRNA sequence introduced into the Cas9 plasmid (Addgene, 58766) was: 5´-TTTGATCATGGGGTATGACT -3´. For performing genome editing of the CAPN3 locus at EXON3 and EXON4 sites,

constructs containing selection cassettes together with the Cas9 nuclease were delivered to the cells using Lipofectamine Stem Cell transfection reagent (ThermoFisher Scientific). After transfection, the edited population, was enriched first by Puromycin (SigmaAldrich) selection at concentration of 0.5 mg/ml followed by sorting for dtTomato-eGFP positive cells **(Figure 4 - figure supplement 1B)** (BD Biosciences AriaII). At this stage, the genomic integrity of the CAPN3 locus and the presence of the ITRs of the PiggyBac system on genome was evaluated by PCR amplification of the edited areas using GXL Polymerase (Clontech). For the excision of the cassette, cells were transfected with constructs carrying the excision-only mutant (R372A/K375A) of the hyperactive PiggyBac transposase (Li et al., 2013, Yusa et al., 2011) by using Lipofectamine Stem Cell transfection reagent. After transfection, the negative fluorescent cells were selected by FACS sorting followed by genotyping for final validation **(Figure 4 - figure supplement 1C)**.

*sgRNA off-target analysis***:** To evaluate potent mismatches we applied the off-target prediction module from the CCTop-CRISPR/Cas9 target online predictor program (Stemmer et al., 2015) (https://crispr.cos.uni-heidelberg.de), using the default settings for the max. total mismatches, core length, and max. core mismatches. From the output sequences we evaluated the first ten in probability **(Table S1)** where we could not detect any sequence alteration.

Microarray Karyotype

Genomic DNA from the patient and genome edited (isogenic) LGMD2A pluripotent stem cell lines purified using DNeasy Blood and Tissue Kit (Qiagen). Samples were further processed at the University of Bonn Life & Brain Genomics facility using Illumina iScan technology **(Figure 4 - figure supplement 1A)**.

Immunocytochemistry

Cryosection Immunochemistry: Organoids from different stages were fixed on 4% paraformaldehyde overnight at 4°C under shakings conditions, dehydrated (30% sucrose o/n incubation) and embedded in OCT freezing media. Cryosections were acquired on a Leica CM3050s Cryostat. For the immunostaining process, cryosections were rehydrated with PBS and followed by permeabilization once with 0.1% Tween-20 in PBS, (rinsed 3x with PBS), and then with 0.1% Triton-X in PBS (rinsed 3x with PBS). Subsequently, the sections blocked with 1% BSA / 10% NGS in PBS for 1hr at room temperature. Primary antibody incubations were performed from 1 to 2 days at 4°C, where secondary antibody incubations for 2hr at room temperature.

EdU staining: At 12-week post differentiation, control organoids incubated with 2.5μM BrdU final concentration overnight. To detect EdU, the sections were processed with Click-iT EdU Alexa Fluor 488 cell proliferation kit

(Invitrogen) following the manufacturer's instructions. The samples incubated with secondary antibodies after the click reaction for detecting EdU.

Primary Antibodies: anti-Brachyury (R&DSystems, 1:250), anti-TBX6(Abcam, 1:200), anti-PAX3 (DHSB, 1:250), anti-PAX7(DHSB, 1:250), anti-SOX10 (R&DSystems,1:125), anti-KI67 (ThermoFisher Scientific, clone SolA15, 1:100), anti-TITIN (DHSB, 9D-10, 1:300), anti-MyHC (DHSB, MF20, 1:300), anti-MYOD1 (Santa Cruz Biotechnologies, clone 5.8A, 1:200), anti-PRDM16 (Abcam, ab106410, 1:200), anti-TFAP2A (DHSB, 3B5, 1:100), anti-Dystrophin (Novocastra/Leica Biosystems, clone DY4/6D3, 1:200), anti-Laminin (SigmaAldrich, 1:200), anti-FastMyHC (SigmaAldrich, clone MY-32, 1:300), anti-M-Cadherin (Cell Signaling Technology, 1:200) anti-SOX2 (ThermoFisher Scientific, clone Btjce, 1:100), anti-CD44 (eBioscience, clone IM7, 1:100), anti-Fibrillin1 (Invitrogen, clone 11C1.3, 1:100).

Secondary antibodies: Alexa Fluor® 647 AffiniPure Fab Fragment Goat Anti-Mouse IgM, μ Chain Specific (Jackson Immunoresearch Laboratories, 1:100), Rhodamine Red™-X (RRX) AffiniPure Goat Anti-Mouse IgG, Fcγ Subclass 1 Specific (Jackson Immunoresearch Laboratories,1:100), Alexa Fluor® 488 AffiniPure Goat Anti-Mouse IgG, Fcγ subclass 2a specific (Jackson Immunoresearch Laboratories,1:100), Alexa Fluor 488, Goat anti-Rat IgG (H+L) Cross-Adsorbed Secondary Antibody, (ThermoFisher Scientific, 1:500), Alexa Fluor 488, Donkey anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, (ThermoFisher Scientific, 1:500), Alexa Fluor 647, Donkey anti-Goat IgG (H+L) Cross-Adsorbed Secondary Antibody, (ThermoFisher Scientific, 1:500), Alexa Fluor 488, Donkey anti-Goat IgG (H+L) Cross-Adsorbed Secondary Antibody, (ThermoFisher Scientific, 1:500), Alexa Fluor 568, Donkey anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, (ThermoFisher Scientific, 1:500). Images were acquired on a ZEISS LSM780 inverted confocal microscope.

Oil O Red Staining

For histological visualization of adipocytes within the organoids, Oil O Red Stain kit (Abcam, ab150678) applied on frozen sections derived from PFA fixated organoids following manufactures recommended protocol. Organoid sections upon staining with Oil O Red were visualized at an Olympus BX61 upright microscope.

Flow Cytometry

FACS intracellular staining: Organoids during the 2nd,5th,9th and 14th week of differentiation were dissociated into single cells by incubating them till dissociating at 37°C within papain solution under shaking conditions. Then, the cells pelleted at 400xg for 5min and followed incubation with TryplE Select for 10min to ensure dissociation into singe cells. Further, the cells passed through 70μM (2nd week) – 100μM (5th, 9th, $14t^{h}$ week) cell strainer to avoid aggregates. For both digesting steps 10% FBS/DMEM-F12 as digesting deactivation solution applied to the cells. Then, for intracellular flow

cytometric staining analysis the Transcription Factor Buffer set (BD Pharmigen) was applied and the cells were analyzed using flow cytometer (BD Biosciences FACS ARIAII). Primary antibodies used in this study: anti-PAX7, anti-MYOD1, anti-Pax3 in total amount of 400μg per staining, while as secondaries the Rhodamine Red™-X (RRX) AffiniPure Goat Anti-Mouse IgG, Fcγ Subclass 1 Specific (Jackson Immunoresearch Laboratories), Alexa Fluor® 488 AffiniPure Goat Anti-Mouse IgG, Fcγ subclass 2a specific (Jackson Immunoresearch Laboratories) in 1:50 dilution. As isotype controls, Mouse IgG1 kappa Isotype Control (Invitrogen, clone P3.6.2.8.1), Mouse IgG2a kappa Isotype Control, (Invitrogen, clone eBM2a) were used at 400μg total amount per staining.

FACS EdU assay: At 15-week post differentiation, organoids incubated overnight with 5μM EdU final concentration. Next day, organoids were dissociated into single cells by incubation at 37°C within papain solution for 1-2hr, followed by incubation with TryplE Select for 10min to ensure single cell dissociation. Then, the dissociated cells passed through a 70μm cell culture strainer to remove any remaining aggregates. To detect EdU, the cells were processed with Click-iT EdU Alexa Fluor 488 Flow Cytometry Assay Kit (Invitrogen) according to manufacturer instructions and then analyzed using the 488 channel of a BD Biosciences FACSAria Fusion flow cytometer.

FACS isolation of ITGβ1^{+} / CXCR4^{+} myogenic cell population for RNA sequencing: Organoids from Duchenne and Control iPSC lines and during 15th - 16th week post differentiation dissociated into single cells during incubation with Papain solution till we could observe dissociation upon gentle shaking (1-2hr). To acquire singlets, we filter the cells with through 40μm cell strainer and upon washing with 1% BSA solution we proceed with surface antigen staining. For surface antigen staining 20 min incubation with the Alexa Fluor 488 anti-human CD29 (Biolegend, clone TS2/16), PE anti-human CD184[CXCR4] (Biolegend, clone 12G5) were applied together with the corresponding isotype controls: PE Mouse IgG2a, κ Isotype Ctrl antibody (Biolegend, clone MOPC-173), 488 Mouse IgG1, κ Isotype Ctrl antibody (Invitrogen, clone P3.6.2.8.1). For removing residual antibodies, the cells were washed twice with 1% BSA staining solution and processed by BD Biosciences FACSAria Fusion flow cytometer. Briefly before FACS sorting to discriminate between dead and alive cells DAPI was added to the samples and then DAPI^{-} / CD29^{+} / CXCR4^{+} cells populations were collected into tubes containing RLT buffer supplemented with b-mercaptoethanol to avoid RNA degradation. FACS gating strategy is further depicted in **Figure 5 – figure supplement 3A**.

Bulk RNA sequencing

RNA extraction: Total RNA was extracted from single organoids or cultured cells by using the RNAeasy Micro Kit (Qiagen) according to the manufacturer's instructions. Subsequently, before library preparation, the RNA integrity was evaluated on an Agilent 2100 Bioanalyzer by using the RNA 6000 Pico kit (Agilent).

cDNA library preparation: For 4-week and 8-week organoids cDNA library was prepared by using the whole transcriptome Illumina TruSeq Stranded Total RNA Library Prep Kit Gold (Illumina), followed by evaluation on an Agilent 2100 Bioanalyzer by using the DNA 1000 kit. The resulting mRNA library sequenced as 2 X 75 bp paired-end reads on a NextSeq 500 sequencer (Illumina). For 16-weeks and ITGβ1$^+$ / CXCR4$^+$ sorted cells cDNA library was prepared using the whole transcriptome Ovation Solo RNA seq Library Preparation Kit (TECAN, NuGEN), followed by evaluation on an Agilent 2100 Bioanalyzer by using the DNA 1000 Chip. The resulting mRNA library sequenced as 1 X 150 bp single reads on a HiSeq. 3000 sequencer (Illumina).

Bulk RNA seq bioinformatic analysis

Sequenced reads were aligned to the human reference genome (hg38) with TopHat2 (version 2.1.1), and the aligned reads were used to quantify mRNA expression by using HTSeq-count (version 0.11.2). DESeq2 (Love at al., 2014) was used to identify differentially expressed genes across the samples. ITGβ1$^+$ / CXCR4$^+$ organoid derived myogenic cell populations were compared to already available transcriptomic dataset of human fetal muscle progenitors (GSM2328841-2) (Hicks et al., 2018).

scRNA sequencing

Sample and cDNA library preparation: Single cells acquired upon incubation for 1hr with solution containing papain and EDTA. Upon dissociation, we estimated the cell number and viability; then cells resuspended on solution containing 0.5% BSA in PBS to reach concentration of 390 cells per µl. The cDNA library was prepared using the Chromium Single Cell 3′ Reagent Kits (v3): Single Cell 3′ Library & Gel Bead Kit v3 (PN-1000075), Single Cell B Chip Kit (PN-1000073) and i7 Multiplex Kit (PN-120262) (10x Genomics) according to the manufacturer's instructions. Then, the cDNA library was run on an Illumina HiSeq 3000 as 150-bp paired-end reads.

Single cell RNA seq bioinformatic analysis

Sequencing data were processed with UMI-tools (version 1.0.0), aligned to the human reference genome (hg38) with STAR (version 2.7.1a), and quantified with Subread featureCounts (version 1.6.4). Data normalization and further analysis were performed using Seurat (version 3.1.3, Stuart et al., 2019). For initial quality control of the extracted gene-cell matrices, we filtered cells with parameters low threshold = 500, high threshold =6,000 for number of genes per cell (nFeature_RNA), high threshold = 5 for percentage of mitochondrial genes (percent.mito) and genes with parameter min.cells = 3. Filtered matrices were normalized by LogNormalize method with scale factor = 10,000. Variable genes were found with parameters of selection.method = "vst", nfeatures = 2000, trimmed for the genes related to cell cycle

(KEGG cell cycle, hsa04110) and then used for principal component analysis. Statistically significant principal components were determined by JackStraw method and the first 5 principle components were used for non-linear dimensional reduction (tSNE and UMAP) and clustering analysis with resolution=0.2. Monocle3 (version 0.2.0, Cao et al., 2019) was used for pseudotime trajectory analysis. We imported data matrix of Seurat objects (assays[["RNA"]]@counts) to Monocle R package and then performed dimensionality reduction with PCA method with parameters max_components=2 and then cluster_cells, learn_graph and order_cells functions were performed subsequently. Organoid-derived myogenic progenitors were compared to already available transcriptomic dataset of adult satellite cells (GSE130646) (Rubenstein et al., 2020).

qPCR expression analysis

By pooling 3 organoids per sample, total RNA was extracted using the RNAeasy Mini Plus Kit (Qiagen). For first strand cDNA synthesis, the High capacity RNA to cDNA Kit (Applied Biosystems) applied using as template 2µg of total RNA. For setting qPCR reactions, the GoTaq qPCR Master Mix (Promega) was used with input template 4ng cDNA per reaction while the reaction was detected on a CFX 96 Real-Time PCR detection system (BIO-RAD). The relative quantification (ΔCT) method applied for detecting changes in gene expression of pluripotent, neural tube, neural crest, posterior/anterior somitic; and dermomyotomal markers, between different time points along the differentiation. qPCR primers applied for each marker are listed in **Table S2.**

Transmission Electron Microscopy (TEM)

Skeletal muscle organoids were fixed for 4h at room temperature (RT) in 2.5% glutaraldehyde (SigmaAldrich) in 0.1M cacodylate buffer pH 7,4 (Sciences Services, Germany), subsequently washed in 0.1M cacodylate buffer pH 7,4, post-fixed for 2h at RT in 1% osmium tetroxide (Sciences Services, Germany) in 0.1M cacodylate buffer pH 7,4, dehydrated stepwise in a graded ethanol series and embedded in Epon 812 (Fluka, Buchs, Switzerland). Ultrathin sections (70 nm, ultramicrotome EM UC7, Leica, Wetzlar, Germany) were afterwards stained for 30 min in 1% aqueous uranyl acetate (Leica, Germany) and 20 min in 3% lead citrate (Leica, Germany). TEM images were acquired with a 200 kV TEM JEM 2100Plus (Jeol, Japan), transmission electron microscope.

Western blot

Skeletal muscle organoid proteins were extracted using a Urea containing lysis buffer (7 M Urea, 2 M Thiourea, 2% CHAPS and 1 M DTT) containing 1X protease inhibitor cocktail (SigmaAldrich). Protein lysates from individual organoids were loaded in different wells of Bolt™ 4-12% Bis-Tris Plus gels (ThermoFisher Scientific). Samples were

resolved following the manufacturer's instructions. Proteins were then transferred to nitrocellulose membranes employing an iBlot2 device (Thermo). Blocking was performed using 2% milk (Carl Roth) dissolved in PBS for 1 hour at room temperature. Relative amount of outer membrane mitochondrial import receptor TOM20 (Santa Cruz, FL-145 1:1000) and of the sarcoplasmatic / endoplasmatic reticulum ATPase SERCA2 (ThermoFisher Scientific, 2A7-A1, 1:1000) compared to the internal control GAPDH (Abcam, ab9485, 1:1000) was evaluated. Membranes were incubated with SuperSignal West Pico Chemiluminescent Substrate (ThermoFisher Scientific) for visualization.

Second harmonic generation (SHG) imaging using multi-photon microscopy

A TriM Scope II multi photon system from LaVision BioTec was used to visualize skeletal muscle fiber organization inside organoids and distinct sarcomeres. The microscope setup is a single beam instrument with an upright Olympus BX51 WI microscope stand equipped with highly sensitive non-descanned detectors close to the objective lens. The TriM Scope II is fitted with a Coherent Scientific Chameleon Ultra II Ti:Sapphire laser (tuning range 680-1080 nm) and a Coherent Chameleon Compact OPO (automated wavelength extension from 1000 nm to 1600 nm). A 20x IR objective lens (Olympus XLUMPlanFI 20x/1.0W) with a working distance of 2.0 mm was used. Muscle fiber SHG signals were detected in forward direction using TiSa light at 850 nm, a 420/40 band pass filter and a blue-sensitive photomultiplier (Hamamatsu H67080-01). 3D-images were acquired and processed with LaVision BioTec ImSpector Software.

Electrophysiology

Current measurement: Membrane currents were measured at ambient temperature (22-24°C) using standard whole-cell patch clamp software ISO2 (MFK, Niedernhausen, Germany). Cells were voltage-clamped at a holding potential of –90 mV, i.e. negative to EnAChR, resulting in inward Na^+ currents. Every 10 s, voltage ramps (duration 500 ms) from -120 mV to +60 mV were applied to assess stability of the recording conditions and to generate I/V curves (membrane currents in response to depolarizing voltage ramps are shown as downward deflections). Signals were filtered (corner frequency, 1 KHz), digitally sampled at 1 KHz and stored on a computer equipped with the hardware/software package ISO2 for voltage control, data acquisition and data analysis. Rapid exposure to a solution containing acetylcholine was performed by means of a custom-made solenoid-operated flow system permitting a change of solution around an individual cell with a half time of about 100 ms. For measurements cells devoid of contact with neighboring cells were selected.

Fluorescence microscopy and imaging: To monitor changes in $[Ca2^+]i$, skeletal muscle cells were transiently transfected with pcDNA3[Twitch-2B] (Addgene, 49531) (0.25 µg per 35 mm culture-dish). Skeletal muscle cells were transfected using either poly-ethyleneimine (PEI) or Lipofectamine (Invitrogen) according to the manufacturer's instructions. Prior

to experiments, cells were seeded on sterile, poly-L-lysine-coated glass cover slips and analyzed or 48 h after transfections. All experiments were performed using single cells at ambient temperature. Fluorescence was recorded buckusing an inverted microscope (Zeiss Axiovert 200, Carl Zeiss AG, Göttingen, Germany) equipped with a Zeiss oil immersion objective (100x/1.4), a Polychrome V illumination source and a photodiode-based dual emission photometry system suitable for CFP/YFP-FRET (FEI Munich GmbH, Germany). For FRET measurements, single cells were excited at 435 nm wavelength with light pulses of variable duration (20 ms to 50 ms; frequency: 5 Hz) to minimize photo-bleaching. Corresponding emitted fluorescence from CFP (F480 or FCFP) or from YFP (F535 or FYFP) was acquired simultaneously and FRET was defined as ratio FYFP/FCFP. Fluorescent signals were recorded and digitized using a commercial hardware/software package (EPC10 amplifier with an integrated D/A board and Patch-master software, HEKA, HEKA Elektronik, Germany). The individual FRET traces were normalized to the initial ratio value before agonist application (FRET/FRET0).

Solutions and chemicals: For FRET measurements an extracellular solution of the following composition was used (mmol/L): NaCl 137; KCl 5.4; $CaCl_2$ 2; $MgCl_2$ 1.0; Hepes/NaOH 10.0, pH 7.4. For whole cell measurements of membrane currents an extracellular solution of the following composition was used (in mmol/L): NaCl 120, KCl 20, $CaCl_2$ 0.5, $MgCl_2$ 1.0, HEPES/NaOH 10.0 (pH 7.4). The pipette solution contained (in mmol/L): K-aspartate 100, KCl 40, NaCl 5.0, $MgCl_2$ 2.0, Na_2ATP 5.0, BAPTA 5.0, GTP 0.025, and HEPES/KOH 20.0, pH 7.4. Standard chemicals were from Merck. EGTA, HEPES, Na_2ATP, GTP and acetylcholine chloride, were from SigmaAldrich.

Statistics

All statistical analysis was conducted using GraphPad Prism6 software. For qPCR analysis we performed one-way ANOVA with Tukey´s multiple comparisons test for each marker. For the FACS intracellular staining quantification, we performed one-way ANOVA with Sidak´s multiple comparisons test between the different time points. For Western blot quantification we performed Mann-Whitney test between two groups. Significance asterisks represent *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, ns: not significant.

Data and code availability

RNA sequencing datasets produced in this study are deposited in the Gene Expression Omnibus (GEO) under accession code GSE147514. Detailed scripts and parameters used for the study are available from the authors upon reasonable request.

To review GEO accession GSE147514:

Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE147514

REFERENCES

Alexander MS, Rozkalne A, Colletta A, Spinazzola JM, Johnson S, Rahimov F, Meng H, Lawlor MW, Estrella E, Kunkel LM, Gussoni E. 2016. CD82 is a marker for prospective isolation of human muscle satellite cells and is linked to muscular dystrophies. *Cell Stem Cell* **19:**800–807. DOI: https://doi.org/10.1016/j.stem.2016.08.006

Al Tanoury Z, Rao J, Tassy O, Gobert B, Gapon S, Garnier JM, Wagner E, Hick A, Hall A, Gussoni E, & Pourquié O. 2020. Differentiation of the human PAX7-positive myogenic precursors/satellite cell lineage in vitro. Development **147:** dev187344. DOI: https://doi.org/10.1242/dev.187344

Arias-Fuenzalida J, Jarazo J, Qing X, Walter J, Gomez-Giro G, Nickels SL, Zaehres H, Schöler HR, Schwamborn JC. 2017. FACS-Assisted CRISPR-Cas9 Genome Editing Facilitates Parkinson's Disease Modeling. *Stem Cell Reports* **9:** 1423-1431. DOI: https://doi.org/10.1016/j.stemcr.2017.08.026

Aulehla A, Pourquie O. 2010. Signaling Gradients during Paraxial Mesoderm Development. *Cold Spring Harb Perspect Biol*. **2:** a000869. DOI: https://doi.org/10.1101/cshperspect.a000869

Baghdadi MB, Castel D, Machado L, Fukada SI, Birk DE, Relaix F, Tajbakhsh S, Mourikis P. 2018. Reciprocal signalling by Notch–Collagen V–CALCR retains muscle stem cells in their niche. Nature **557:**714–718. DOI: https://doi.org/10.1038/s41586-018-0144-9

Barbieri G, DeAngelis L, Feo S, Cossu G, Giallongo A. 1990. Differential expression of muscle-specific enolase in embryonic and fetal myogenic cells during mouse development *Differentiation* **45:**179-184. DOI: https://doi.org/10.1111/j.1432-0436.1990.tb00471.x

Bjornson CR, Cheung TH, Liu L, Tripathi PV, Steeper KM, Rando TA. 2012. Notch signaling is necessary to maintain quiescence in adult muscle stem cells. *Stem cells* **30:**232–242. DOI: https://doi.org/10.1002/stem.773

Borchin B, Chen J, Barberi, T. 2013. Derivation and FACS-mediated purification of PAX3+/PAX7+ skeletal muscle precursors from human pluripotent stem cells. *Stem Cell Reports* **1:**620-631. DOI: https://doi.org/10.1016/j.stemcr.2013.10.007

Buckingham M, Rigby PWJ. 2014. Gene Regulatory Networks and Transcriptional Mechanism that Control Myogenesis. *Developmental Cell* **28:**225-238. DOI: https://doi.org/10.1016/j.devcel.2013.12.020

Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, Trapnell C, Shendure J. 2019. The single-cell transcriptional landscape of mammalian organogenesis. Nature, **566**: 496–502. DOI: https://doi.org/10.1038/s41586-019-0969-x

Chal J, Oginuma M, Al Tanoury Z, Gobert B, Sumara O, Hick A, Bousson F, Zidouni Y, Mursch C, Moncuquet P, Tassy O, Vincent S, Miyanari A, Bera A, Garnier JM, Guevara G, Hestin M, Kennedy L, Hayashi S, Drayton B, Cherrier T, Gayraud-Morel B, Gussoni E, Relaix F, Tajbakhsh S, Pourquié O. 2015. Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. *Nat Biotechnol.* **33:**962–969. DOI: https://doi.org/10.1038/nbt.3297

Chargé SB, Rudnicki MA. 2004. Cellular and molecular regulation of muscle regeneration. Physiol. Rev. **84:** 209–238. DOI: https://doi.org/10.1152/physrev.00019.2003

Cheng L, Alvares LE, Ahmed MU, El-Hanfy AS, Dietrich S. 2004. The epaxial-hypaxial subdivision of the avian somite. *Dev. Biol.* **274:**348–369. DOI: https://doi.org/10.1016/j.ydbio.2004.07.020

Cheung TH, Quach NL, Charville GW, Liu L, Park L, Edalati A, Yoo B, Hoang P, Rando TA. 2012. Maintenance of muscle stem-cell quiescence by microRNA-489. Nature **482:** 524–528. DOI: https://doi.org/10.1038/nature10834

Cohn RD, van Erp C, Habashi JP, Soleimani AA, Klein EC, Lisi MT, Gamradt M, ap Rhys CM, Holm TM, Loeys BL, Ramirez F, Judge DP, Ward CW, Dietz HC. 2007. Angiotensin II type 1 receptor blockade attenuates TGF-beta-induced failure of muscle regeneration in multiple myopathic states. *Nat Med*, **13**:204–210. DOI: https://doi.org/10.1038/nm1536

Darabi R, Arpke RW, Irion S, Dimos JT, Grskovic M, Kyba M, Perlingeiro RC. 2012. Human ES- and iPS-derived myogenic progenitors restore DYSTROPHIN and improve contractility upon transplantation in dystrophic mice. Cell Stem Cell, **10:**610–619. DOI: https://doi.org/10.1016/j.stem.2012.02.015

Dietrich S, Abou-Rebyeh F, Brohmann H, Bladt F, Sonnenberg-Riethmacher E, Yamaai T, Lumsden A, Brand- Saberi B, Birchmeier C. 1999. The role of SF/HGF and c-Met in the development of skeletal muscle. Development **126**:1621-1629.

Dorn I, Klich K, Arauzo-Bravo MJ, Radstaak M, Santourlidis S, Ghanjati F, Radke TF, Psathaki OE, Hargus G, Kramer J, Einhaus M, Kim JB, Kögler G, Wernet P, Schöler HR, Schlenke P, Zaehres H. 2015. Erythroid differentiation of human induced pluripotent stem cells is independent of donor cell type of origin. *Haematologica* **100:** 32-41. DOI: https://doi.org/10.3324/haematol.2014.108068

Faustino Martins JM, Fischer C, Urzi A, Vidal R, Kunz S, Ruffault PL, Kabuss L, Hube I, Gazzerro E, Birchmeier C, Spuler S, Sauer S, Gouti M. 2020. Self-Organizing 3D Human Trunk Neuromuscular Organoids. *Cell Stem Cell* **26:** 172-186. DOI: https://doi.org/10.1016/j.stem.2019.12.007

Fukada S, Uezumi A, Ikemoto M, Masuda S, Segawa M, Tanimura N, Yamamoto H, Miyagoe-Suzuki Y, Takeda S. 2007. Molecular signature of quiescent satellite cells in adult skeletal muscle. *Stem Cells* **25:** 2448-2459. DOI: https://doi.org/10.1634/stemcells.2007-0019

Gouti M, Delile J, Stamataki D, Wymeersch FJ, Huang Y, Kleinjung J, Wilson V, Briscoe J. 2017. A Gene Regulatory Network Balances Neural and Mesoderm Specification during Vertebrate Trunk Development. Dev. Cell **41**: 243–261. DOI: https://doi.org/10.1016/j.devcel.2017.04.002

Henrique D, Abranches E, Verrier L, Storey KG. 2015. Neuromesodermal progenitors and the making of the spinal cord. Development **142**:2864–2875. DOI: https://doi.org/10.1242/dev.119768

Hicks MR, Hiserodt J, Paras K, Fujiwara W, Eskin A, Jan M, Xi H, Young CS, Evseenko D, Nelson SF, Spencer MJ, Handel BV, Pyle AD. 2018. ERBB3 and NGFR mark a distinct skeletal muscle progenitor cell in human development and hPSCs. *Nat Cell Biol.* **20:**46–57. DOI: https://doi.org/10.1038/s41556-017-0010-2

Jahn K, Mohammadi B, Krampfl K, Abicht A, Lochmüller H, Buffer J. 2001. Deactivation and desensitization of mouse embryonic- and adult- type nicotinic receptor channels currents. *Neuroscience Letters* **307:** 89-92. DOI: https://doi.org/10.1016/s0304-3940(01)01929-2

Juan AH, Derfoul A, Feng X, Ryall JG, Dell'Orso S, Pasut A, Zare H, Simone JM, Rudnicki MA, Sartorelli V. 2011. Polycomb EZH2 controls self-renewal and safeguards the transcriptional identity of skeletal muscle stem cells. Genes & development **25:** 789–794. DOI: https://doi.org/10.1101/gad.2027911

Koehler KR, Nie J, Longworth-Mills E, Liu XP, Lee J, Holt JR, Hashino E. 2017. Generation of inner ear organoids containing functional hair cells from human pluripotent stem cells. *Nat Biotechnol.* **35:**583–589. DOI: https://doi.org/10.1038/nbt.3840

Kramerova I, Kudryashova E, Wu B, Germain S, Vandenborne K, Romain N, Haller RG, Verity MA, Spencer MJ. 2009. Mitochondrial abnormalities, energy deficit and oxidative stress are features of calpain 3 deficiency in skeletal muscle. Human Molecular Genetics **18:**3194–3205. DOI: https://doi.org/10.1093/hmg/ddp257

Lancaster MA, Renner M, Martin CA, Wenzel D, Bicknell LS, Hurles ME, Homfray T, Penninger JM, Jackson AP, Knoblich JA. 2013. Cerebral organoids model human brain development and microcephaly. *Nature* **501:**373–379. DOI: https://doi.org/10.1038/nature12517

Le Douarin NM, Calloni GW, Dupin E. 2008. The stem cells of the neural crest. *Cell Cycle* **7:**1013-1019. DOI: https://doi.org/10.4161/cc.7.8.5641

Li X, Burnight ER, Cooney AL, Malani N, Brady T, Sander JD, Staber J, Wheelan SJ, Joung JK, McCray PBJr, Bushman FD, Sinn PL, Craig NL. 2013. piggyBac transposase tools for genome engineering. *Proc Natl Acad Sci U S A* **110**: e2279-2287. DOI: https://doi.org/10.1073/pnas.1305987110

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15:**550. DOI: https://doi.org/10.1186/s13059-014-0550-8

Machado L, Esteves de Lima J, Fabre O, Proux C, Legendre R, Szegedi A, Varet H, Ingerslev LR, Barrès R, Relaix F, Mourikis P. 2017. In Situ fixation redefines quiescence and early activation of skeletal muscle stem cells *Cell Rep.* **21:** 1982-1993. DOI: https://doi.org/10.1016/j.celrep.2017.10.080

Mafffioletti SM, Sarcar S, Henderson ABH, Mannhardt I, Pinton L, Moyle LA, Steele-Stallard H, Cappellari O, Wells KE, Ferrari G, Mitchell JS, Tyzack GE, Kotiadis VN, Khedr M, Ragazzi M, Wang W, Duchen MR, Patani R, Zammit PS, Wells DJ, Eschenhagen T, Tedesco FS. 2018. Three-dimensional human iPSC-derived artificial skeletal muscles model muscular dystrophies and enable multilineage tissue engineering. *Cell Rep.* **23:**899-908. DOI: https://doi.org/10.1016/j.celrep.2018.03.091

Magli A, Incitti T, Kiley J, Swanson SA, Darabi R, Rinaldi F, Selvaraj S, Yamamoto A, Tolar J, Yuan C, Stewart R, Thomson JA, Perlingeiro R. 2017. PAX7 Targets, CD54, Integrin α9β1, and SDC2, Allow Isolation of Human ESC/iPSC-Derived Myogenic Progenitors. Cell Rep. **19**:2867–2877. DOI: https://doi.org/10.1016/j.celrep.2017.06.005

Messina G, Biressi S, Monteverde S, Magli A, Cassano M, Perani L, Roncaglia E, Tagliafico E, Starnes L, Campbell CE, Grossi M, Goldhamer DJ, Gronostajski RM, Cossu G. 2010. Nfix regulates fetal-specific transcription in developing skeletal muscle. *Cell* **140:**554-566. DOI: https://doi.org/10.1016/j.cell.2010.01.027

Montarras D, Morgan J, Collins C, Relaix F, Zaffran S, Cumano A, Partridge T, Buckingham M. 2005. Direct isolation of satellite cells for skeletal muscle regeneration. *Science* **309**:2064-2067. DOI: https://doi.org/10.1126/science.1114758

Moore R, Walsh FS. 1993. The cell adhesion molecule M-cadherin is specifically expressed in developing and regenerating, but not denervated skeletal muscle. *Development* **117:**1409-1420.

Mourikis P, Sambasivan R, Castel D, Rocheteau P, Bizzarro V, Tajbakhsh S. 2012. A critical requirement for notch signaling in maintenance of the quiescent skeletal muscle stem cell state. *Stem Cells* **30**: 243-252. DOI: https://doi.org/10.1002/stem.775

Murphy M, Reid K, Ford M, Furness JB, Bartlett PF. 1994. FGF2 regulates proliferation of neural crest cells, with subsequent neuronal differentiation regulated by LIF or related factors. Development **120:**3519-3528.

Panopoulos AD, D'Antonio M, Benaglio P, *et.al.* 2017. iPSCORE: A Resource of 222 iPSC Lines Enabling Functional Characterization of Genetic Variation across a Variety of Cell Types. *Stem Cell Reports* **8:**1086-1100. DOI: https://doi.org/10.1016/j.stemcr.2017.03.012

Park IH, Arora N, Huo H, Maherali N, Ahfeldt T, Shimamura A, Lensch MW, Cowan C, Hochedlinger K, Daley GQ. 2008. Disease-specific induced pluripotent stem cells. *Cell* **134:** 877-886. DOI: https://doi.org/10.1016/j.cell.2008.07.041

Porpiglia E, Samusik N, Ho ATV, Cosgrove BD, Mai T, Davis KL, Jager A, Nolan GP, Bendall SC, Fantl WJ, Blau HM. 2017. High-resolution myogenic lineage mapping by single-cell mass cytometry. *Nat Cell Biol.* **19**:558–567. DOI: https://doi.org/10.1038/ncb3507

Quarta M, Brett JO, DiMarco R, De Morree A, Boutet SC, Chacon R, Gibbons MC, Garcia VA, Su J, Shrager JB, Heilshorn S, Rando TA. 2016. An artificial niche preserves the quiescence of muscle stem cells and enhances their therapeutic efficacy. *Nat Biotechnol.* **34**:752-759. DOI: https://doi.org/10.1038/nbt.3576

Raines MA, Magella B, Adam M, Potter SS. 2015. Key pathways regulated by HoxA9,10,11/HoxD9,10,11 during limb development. BMC Developmental Biology **15**:28. DOI: https://doi.org/10.1186/s12861-015-0078-5

Rao L, Qian Y, Khodabukus A, Ribar T, Bursac N. 2018. Engineering human pluripotent stem cells into a functional skeletal muscle tissue. Nat Commun **9:**126. DOI: https://doi.org/10.1038/s41467-017-02636-4

Relaix F, Rocancourt D, Mansouri A, Buckingham M. 2005. A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* **435**: 948–953. DOI: https://doi.org/10.1038/nature03594

Roellig D, Tan-Cabugao J, Esaian S, Bronner ME. 2017. Dynamic transcriptional signature and cell fate analysis reveals plasticity of individual neural plate border cells. eLife **6:**e21620. DOI: https://doi.org/10.7554/eLife.21620

Rubenstein AB, Smith GR, Raue U, Begue G, Minchev K, Ruf-Zamojski F, Nair VD, Wang X, Zhou L, Zaslavsky E, Trappe TA, Trappe S, Sealfon SC. 2020. Single-cell transcriptional profiles in human skeletal muscle. *Sci Rep*. **10:**229. DOI: https://doi.org/10.1038/s41598-019-57110-6

Seale P, Sabourin LA, Girgis-Gabardo A, Mansouri A, Gruss P, Rudnick MA. 2000. Pax7 is required for the specification of myogenic satellite cells. *Cell* **102**:777-786. DOI: https://doi.org/10.1016/s0092-8674(00)00066-0

Shao Z, Mellor IR, Brierley JM, Harris J, Usherwood PN. 1998. Potentiation and inhibition of nicotinic acetylcholine receptors by spermine in the TE671 human muscle cell line. *The Journal of Pharmacology and Experimental Therapeutics* **286:**1269-1276.

Shea KL, Xiang W, LaPorta VS, Licht JD, Keller C, Basson MA, Brack AS. 2010. Sprouty1 regulates reversible quiescence of a self-renewing adult muscle stem cell pool during regeneration. *Cell Stem Cell* **6:**117-129. DOI: https://doi.org/10.1016/j.stem.2009.12.015

Shelton M, Metz J, Liu J, Carpenedo RL, Demers SP, Stanford WL, Skerjanc IS. 2014. Derivation and expansion of PAX7-positive muscle progenitors from human and mouse embryonic stem cells. *Stem Cell Reports* **3:**516–529. DOI: https://doi.org/10.1016/j.stemcr.2014.07.001

Sherwood RI, Christensen JL, Conboy IM, Conboy MJ, Rando TA, Weissman IL, Wagers AJ. 2004. Isolation of adult mouse myogenic progenitors: functional heterogeneity of cells within and engrafting skeletal muscle. *Cell* **119:**543-554. DOI: https://doi.org/10.1016/j.cell.2004.10.021

Shimozono S, Iimura T, Kitaguchi T, Higashijima S, Miyawaki A. 2013. Visualization of an endogenous retinoic acid gradient across embryonic development. Nature **496:**363–366. DOI: https://doi.org/10.1038/nature12037

Shubin N, Tabin C, Carroll S. 1997. Fossils, genes and the evolution of animal limbs. *Nature* **388**:639–648. DOI: https://doi.org/10.1038/41710

Soldatov R, Kaucka M, Kastriti ME, Petersen J, Chontorotzea T, Englmaier L, Akkuratova N, Yang Y, Häring M, Dyachuk V, Bock C, Farlik M, Piacentino ML, Boismoreau F, Hilscher MM, Yokota C, Qian X, Nilsson M, Bronner ME, Croci L, Hsiao WY, Guertin DA, Brunet JF, Consalez GG, Ernfors P, Fried K, Kharchenko PV, Adameyko I. 2019. Spatiotemporal structure of cell fate decisions in murine neural crest. *Science* **364:**eaas9536. DOI: https://doi.org/10.1126/science.aas9536

Southard-Smith EM, Kos L, Pavan WJ. 1998. Sox10 mutation disrupts neural crest development in Dom Hirschsprung mouse model. Nat. Genet. **18:**60-64. DOI: https://doi.org/10.1038/ng0198-60

Spence JR, Mayhew CN, Rankin SA, Kuhar MF, Vallance JE, Tolle K, Hoskins EE, Kalinichenko VV, Wells SI, Zorn AM, Shroyer NF, Wells JM. 2011. Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. *Nature* **470:**105–109. DOI: https://doi.org/10.1038/nature09691

Stemmer M, Thumberger T, del Sol Keyer M, Wittbrodt J, Mateo JL. 2015. CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. *PLoS One* **10:**e0124633. DOI: https://doi.org/10.1371/journal.pone.0124633

Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, Satija R. 2019. Comprehensive Integration of Single-Cell Data. *Cell* **177**:1888-1902. DOI: https://doi.org/10.1016/j.cell.2019.05.031

Tierney MT, Gromova A, Sesillo FB, Sala D, Spenlé C, Orend G, Sacco A. 2016. Autonomous Extracellular Matrix Remodeling Controls a Progressive Adaptation in Muscle Stem Cell Regenerative Capacity during Development. *Cell Rep.* **14**:1940-1952. DOI: https://doi.org/10.1016/j.celrep.2016.01.072

Toral-Ojeda I, Aldanondo G, Lasa-Elgarresta J, Lasa-Fernández H, Fernández-Torrón R, López de Munain A, Vallejo-Illarramendi A. 2016. Calpain 3 deficiency affects SERCA expression and function in the skeletal muscle. Expert reviews in molecular medicine, **18:**e7. DOI: https://doi.org/10.1017/erm.2016.9

Troy A, Cadwallader AB, Fedorov Y, Tyner K, Tanaka KK, Olwin BB. 2012. Coordination of satellite cell activation and self-renewal by Par-complex-dependent asymmetric activation of p38α/β MAPK. *Cell Stem Cell* **11:** 541-553. DOI: https://doi.org/10.1016/j.stem.2012.05.025

Uezumi A, Fukada S, Yamamoto N, Takeda S, Tsuchida K. 2010. Mesenchymal progenitors distinct from satellite cells contribute to ectopic fat cell formation in skeletal muscle. Nat. Cell Biol. **12:**143-152. DOI: https://doi.org/10.1038/ncb2014

Uezumi A, Ito T, Morikawa D, Shimizu N, Yoneda T, Segawa M, Yamaguchi M, Ogawa R, Matev MM, Miyagoe-Suzuki Y, Takeda S, Tsujikawa K, Tsuchida K, Yamamoto H, Fukada S. 2011. Fibrosis and adipogenesis originate from a common mesenchymal progenitor in skeletal muscle. *J Cell Sci.* 124:3654-3664. DOI: https://doi.org/10.1242/jcs.086629

Verma M, Asakura Y, Murakonda BSR, Pengo T, Latroche C, Chazaud B, McLoon LK, Asakura A. 2018. Muscle satellite cell cross-talk with a vascular niche maintains quiescence via VEGF and Notch signaling. *Cell Stem Cell* **23:**530-543. DOI: https://doi.org/10.1016/j.stem.2018.09.007

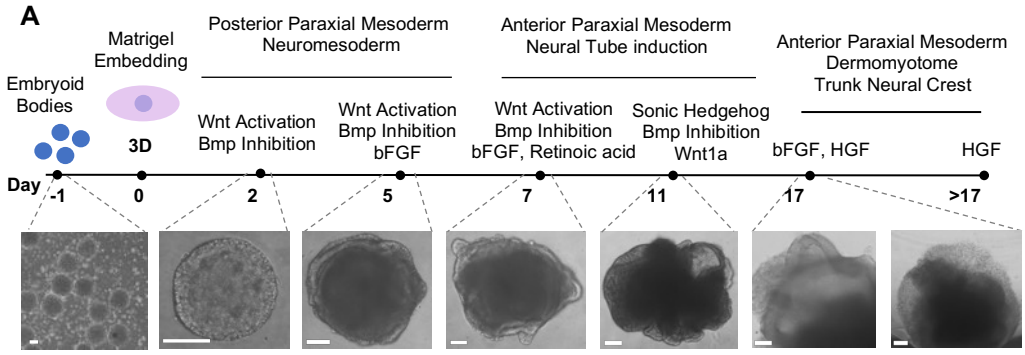
Xi H, Fujiwara W, Gonzalez K, Jan M, Liebscher S, Van Handel B, Schenke-Layland K, Pyle, AD. 2017. In Vivo Human Somitogenesis Guides Somite Development from hPSCs. *Cell Rep.* **18:**1573-1585. DOI: https://doi.org/10.1016/j.celrep.2017.01.040

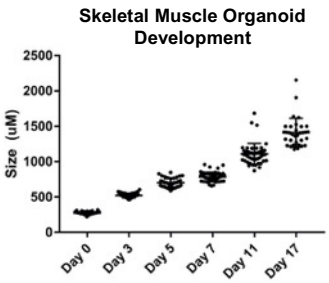
Xi H, Langerman J, Sabri S, Chien P, Young CS, Younesi S, Hicks M, Gonzalez K, Fujiwara W, Marzi J, Liebscher S, Spencer M, Van Handel B, Evseenko D, Schenke-Layland K, Plath K, Pyle AD. 2020. A Human Skeletal Muscle Atlas Identifies the Trajectories of Stem and Progenitor Cells across Development and from Human Pluripotent Stem Cells. Cell Stem Cell, **27**:158–176. DOI: https://doi.org/10.1016/j.stem.2020.04.017

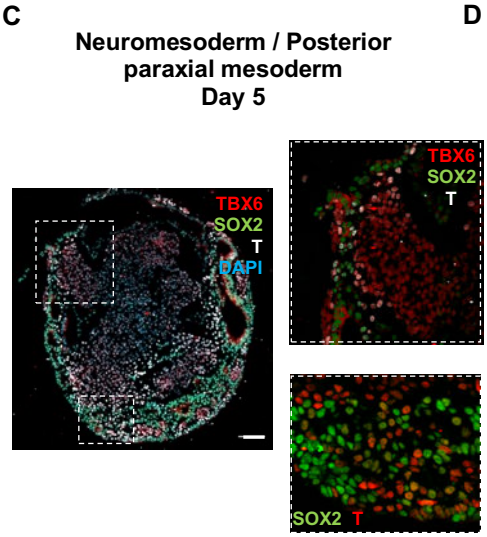
Xu B, Wellik DM. 2011. Axial Hox9 activity establishes the posterior field in the developing forelimb. *Proc Natl Acad Sci U S A*.**108**:4888-4891. DOI: https://doi.org/10.1073/pnas.1018161108

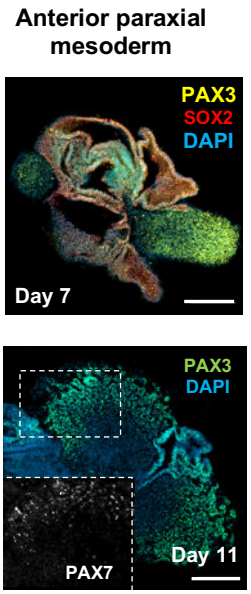
Yusa K, Zhou L, Li MA, Bradley A, Craig NL. 2011. A hyperactive piggyBac transposase for mammalian applications. *Proc Natl Acad Sci U S A* **108:**1531-1536. DOI: https://doi.org/10.1073/pnas.1008322108

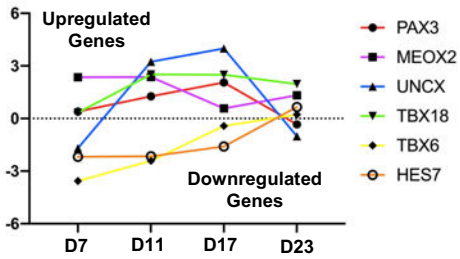
Figures and Tables

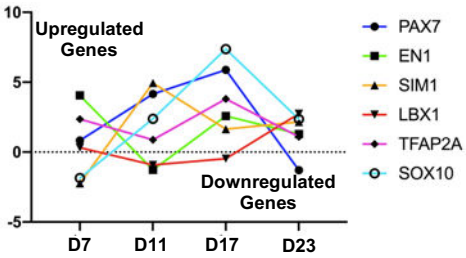

A
Embryoid Bodies
Matrigel Embedding
3D
Posterior Paraxial Mesoderm
Neuromesoderm
Anterior Paraxial Mesoderm
Neural Tube induction
Anterior Paraxial Mesoderm
Dermomyotome
Trunk Neural Crest
Day
-1
0
2
5
7
11
17
>17
Wnt Activation
Bmp Inhibition
Wnt Activation
Bmp Inhibition
bFGF
Wnt Activation
Bmp Inhibition
bFGF, Retinoic acid
Sonic Hedgehog
Bmp Inhibition
Wnt1a
bFGF, HGF
HGF

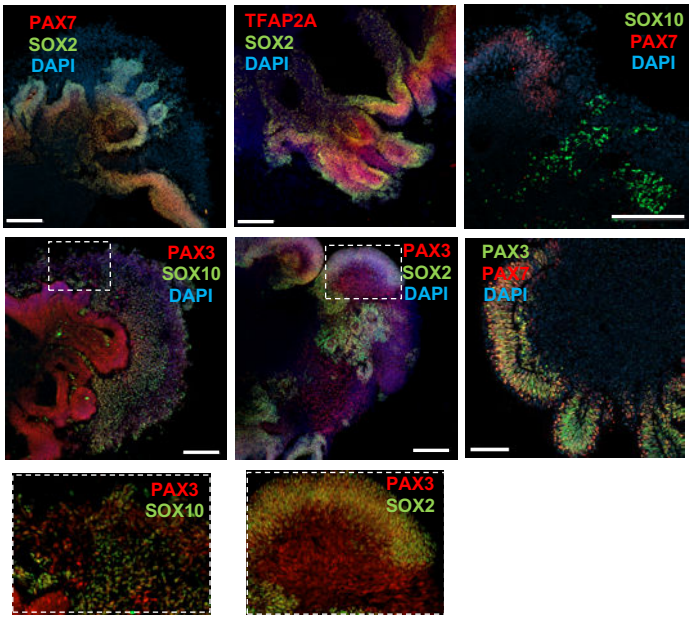

B
Skeletal Muscle Organoid Development
Size (uM)
2500
2000
1500
1000
500
0
Day 0
Day 3
Day 5
Day 7
Day 11
Day 17


C
Neuromesoderm / Posterior paraxial mesoderm
Day 5
TBX6
SOX2
T
DAPI
TBX6
SOX2
T
SOX2 T


D
Anterior paraxial mesoderm
PAX3
SOX2
DAPI
Day 7
PAX3
DAPI
PAX7
Day 11


E
Upregulated Genes
Downregulated Genes
6
3
0
-3
-6
D7
D11
D17
D23
PAX3
MEOX2
UNCX
TBX18
TBX6
HES7


F
Upregulated Genes
Downregulated Genes
10
5
0
-5
D7
D11
D17
D23
PAX7
EN1
SIM1
LBX1
TFAP2A
SOX10


G
Neural Plate Boarder - Specified Neural crest
Dermomyotome - Day 17
PAX7
SOX2
DAPI
TFAP2A
SOX2
DAPI
SOX10
PAX7
DAPI
PAX3
SOX10
DAPI
PAX3
SOX2
DAPI
PAX3
PAX7
DAPI
PAX3
SOX10
PAX3
SOX2

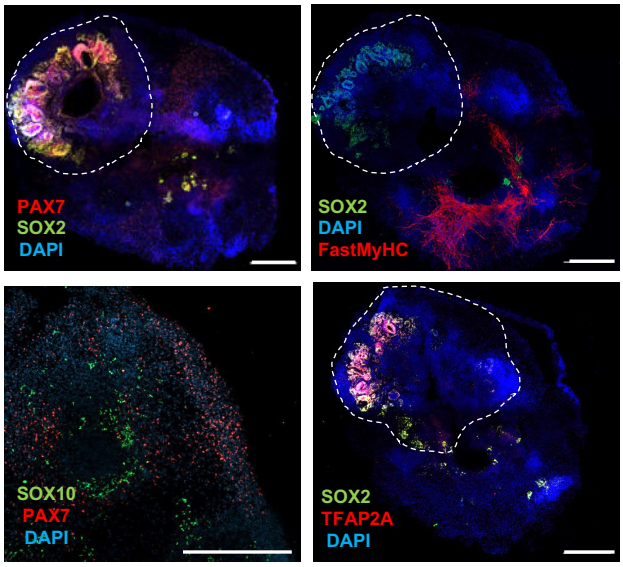

H
Myogenic Progenitor Migration
Day 23
PAX7
SOX2
DAPI
SOX2
DAPI
FastMyHC
SOX10
PAX7
DAPI
SOX2
TFAP2A
DAPI

Figure 1. Skeletal muscle organoid protocol and correlation to fetal development. (A) Stepwise brightfield images emphasizing the crucial stages for the myogenic protocol development, along with the corresponding mixture of the cytokines/growth factors applied in each step **(B)** Graph depicting organoid development in size (Day 0, *n*= 51; Day 3, *n*= 46; Day 5, *n*= 48; Day 7, *n*= 50; Day 11, *n*= 50; Day 17, *n*= 33. For each timepoint we measure organoids from 3 independent derivations) **(C)** Representative expression at Day 4 reveals distinct progenitor populations of mesodermal (T^+), neuromesodermal (T^+, $SOX2^+$) and paraxial mesodermal ($TBX6^+$) origin within the organoid **(D)** Presence of $PAX3^+$ populations of mesodermal origin ($SOX2^-$) and neural origin ($SOX2^+$) from at Day 6 indicate anterior somitic mesoderm formation, while representative organoid overview at Day 10 shows in a mesenchymal state large $PAX3^+$ areas, co-expressing PAX7 **(E)** Graph depicting superimposed qPCR values for anterior somitic mesodermal (ASM) markers, such as PAX3, UNCX, MEOX2, TBX18, and for the posterior somitic mesodermal (PSM) markers, TBX6 and HES7, demonstrate that organoid culture conditions shift the balance towards anterior somitic mesoderm development, as PSM markers are constantly downregulated **(F)** Graph depicting superimposed qPCR values at distinct timepoints during early organoid culture development indicates concomitant upregulation for epaxial and hypaxial dermomyotomal markers, such as PAX7, EN1, SIM1, LBX1, as well as for the neural crest markers, TFAP2A and SOX10 **(G)** Organoid picture at Day 16 indicates formation of $PAX3^+$ / $PAX7^+$ / $SOX2^+$ / $TFAP2A^+$ neural plate border epithelium, and presence of $PAX3^+/SOX2^-$ areas of paraxial mesodermal origin and $PAX3^+$ / $SOX10^+$ areas corresponding to delaminating specified neural crest progenitors **(H)** Representative organoid overview at Day 23, indicates FastMyHC+, PAX7+ areas with myogenic potential towards its exterior portion, while neural lineage is restricted at more inner sites (SOX2, TFAP2A, SOX10) populations where the neural plate border epithelium developed at earlier stages. Dashed line represents the location where the embryoid body was embedded into Matrigel. Statistics: Values at each timepoint represent the difference in mean relative expression for each gene (D7= Day 5 - Day 7, D11= Day 7 - Day 11, D17= Day 11 - Day 17, D23= Day 17 - Day 23) as derived by performing ordinary one-way ANOVA and Tukey´s multiple comparison tests **(E,F)** Scale bars, 200μm **(G)**, 100μm **(A,D,H)**, 50μm **(C)**.

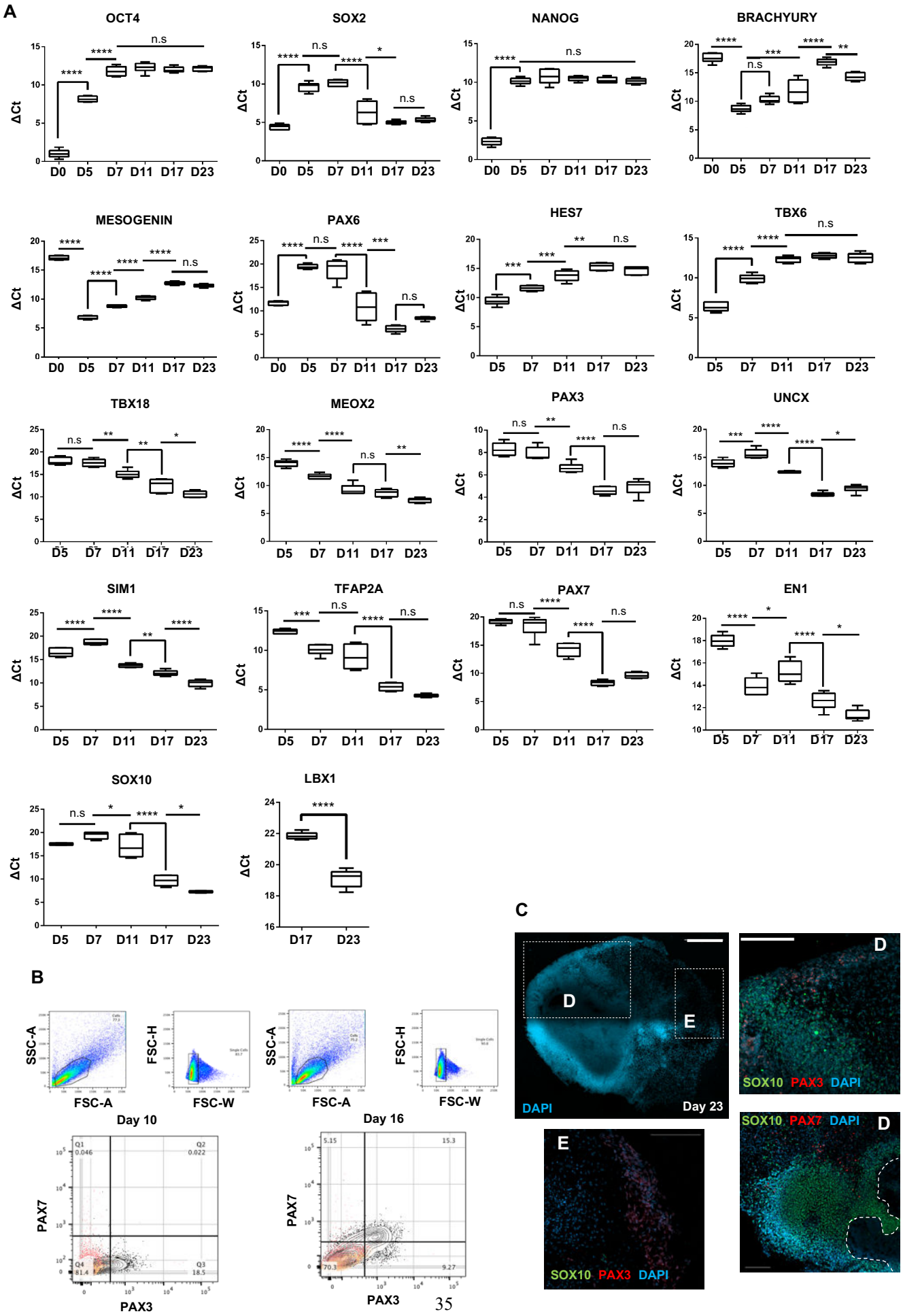

A
OCT4
SOX2
NANOG
BRACHYURY
MESOGENIN
PAX6
HES7
TBX6
TBX18
MEOX2
PAX3
UNCX
SIM1
TFAP2A
PAX7
EN1
SOX10
LBX1
ΔCt
D0 D5 D7 D11 D17 D23
B
SSC-A
FSC-A
FSC-H
FSC-W
Day 10
Day 16
PAX7
PAX3
C
D
E
DAPI
Day 23
SOX10 PAX3 DAPI
SOX10 PAX7 DAPI

Figure 1 – figure supplement 1. Lineage representation and organoid culture progression at early stages. **(A)** qPCR analysis depicting the relative expression of pluripotent (OCT4, SOX2, NANOG), mesodermal / posterior somitic (BRACHYURY, MESOGENIN, TBX6, HES7), anterior somitic (PAX3, UNCX, TBX18, MEOX2), dermomyotomal (PAX7, SIM1, EN1, LBX1) and neural tube / crest markers (SOX2, PAX6, TFAP2A, SOX10) at distinct timepoints of the myogenic differentiation protocol **(B)** Gating strategies and FACS intracellular quantification on PAX3 and PAX7 markers at Day 10 and Day 16 during organoid protocol development **(C)** Tile scan overview at Day 23 highlights the presence of specified neural crest $SOX10^{+}$ populations close to neural plate border epithelium **(D)** while at more outer locations being $SOX10^{-}/PAX3^{+}$ **(E)** For each replicate we pooled 6 organoids, n=6, Each gene expression normalized to housekeeping gene. Statistics: *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, ns: not significant. FACS plots, Yellow histogram: unstained population, Red histogram: Isotype Control, Gray Histogram: $PAX7^{+}$ or $PAX3^{+}$ population**.** Scale bars, 500uM **(C)**, 200uM **(D, E)**.

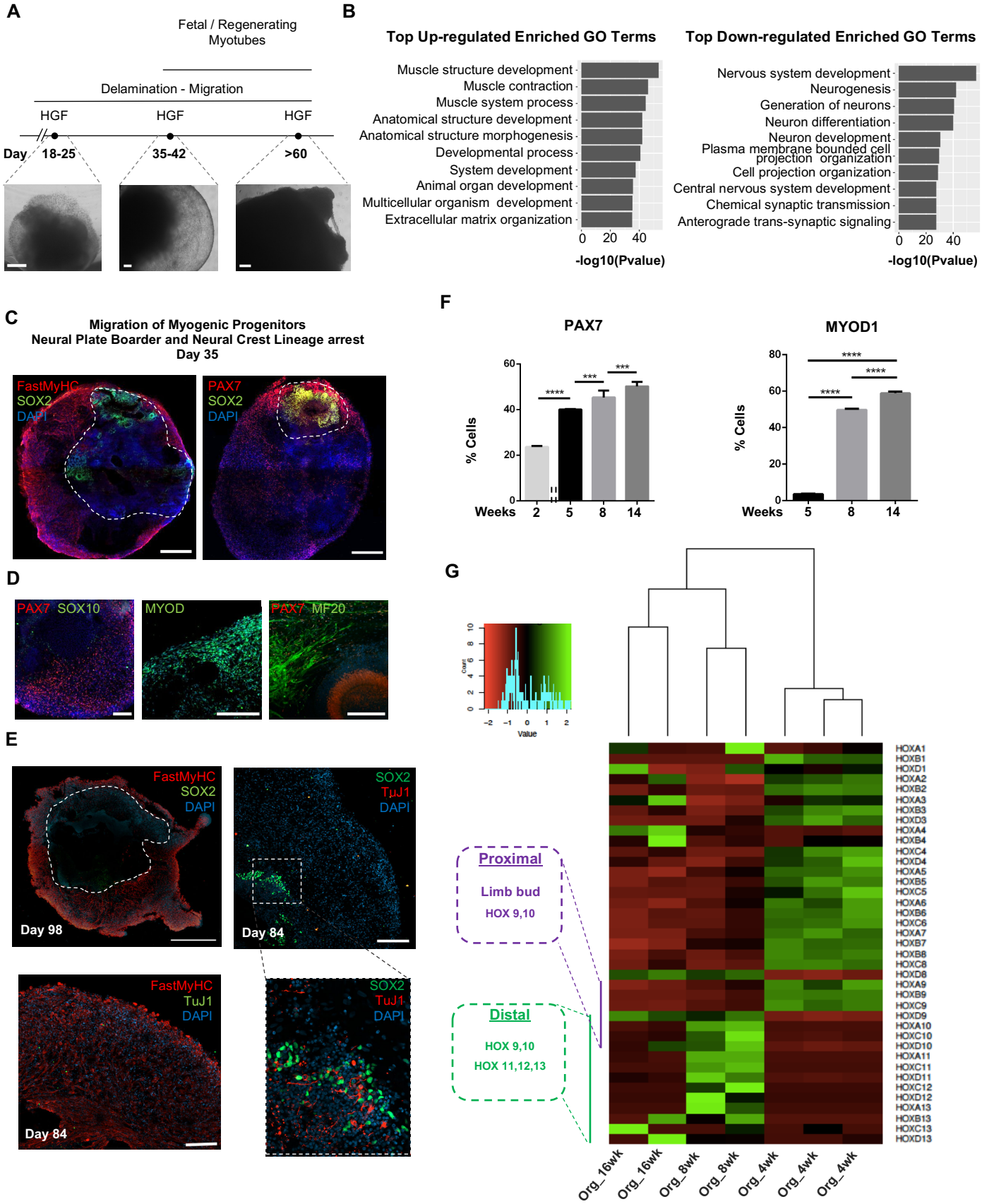
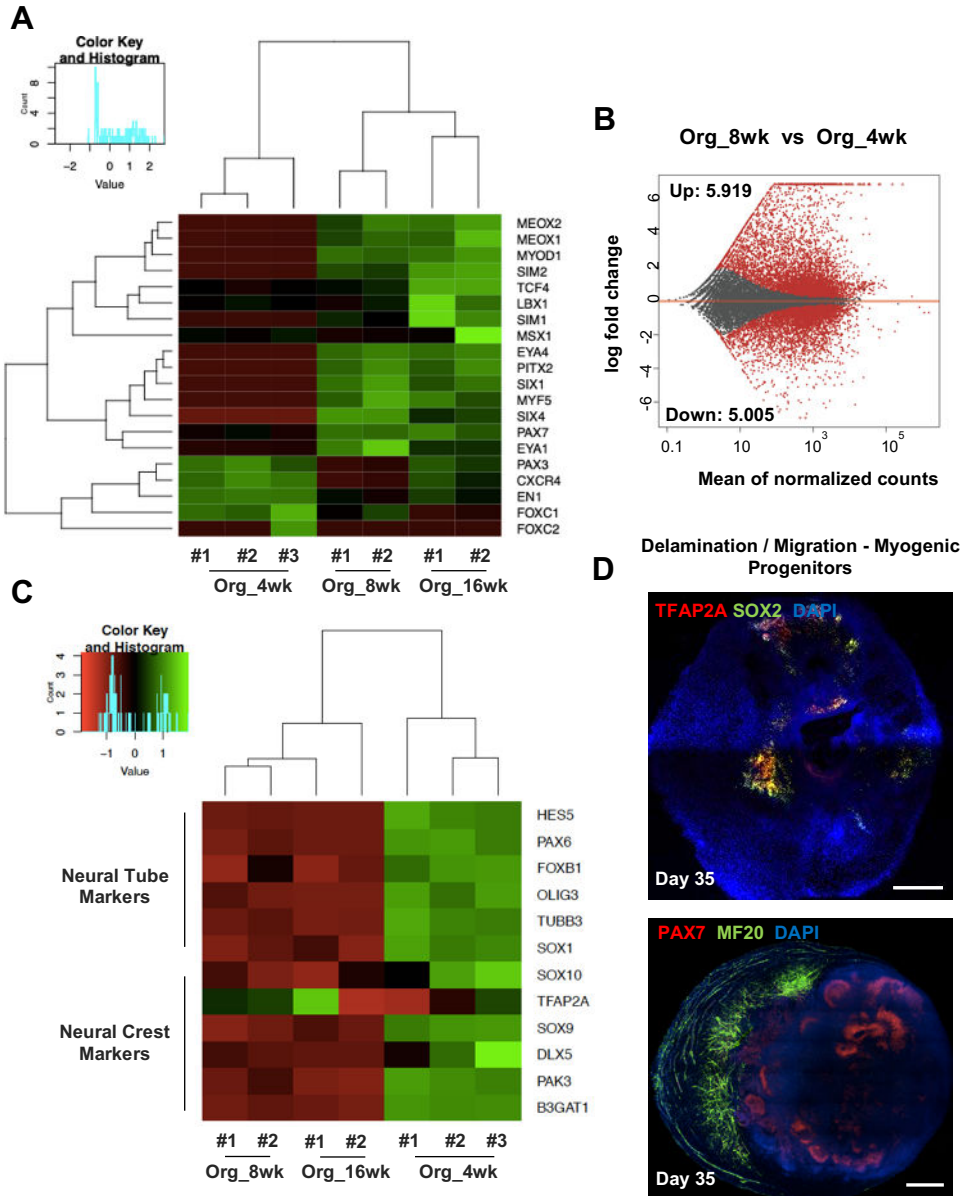

A
Fetal / Regenerating Myotubes
Delamination - Migration
HGF
HGF
HGF
Day
18-25
35-42
>60
B
Top Up-regulated Enriched GO Terms
Muscle structure development
Muscle contraction
Muscle system process
Anatomical structure development
Anatomical structure morphogenesis
Developmental process
System development
Animal organ development
Multicellular organism development
Extracellular matrix organization
0 20 40
-log10(Pvalue)
Top Down-regulated Enriched GO Terms
Nervous system development
Neurogenesis
Generation of neurons
Neuron differentiation
Neuron development
Plasma membrane bounded cell projection organization
Cell projection organization
Central nervous system development
Chemical synaptic transmission
Anterograde trans-synaptic signaling
0 20 40
-log10(Pvalue)
C
Migration of Myogenic Progenitors
Neural Plate Boarder and Neural Crest Lineage arrest
Day 35
FastMyHC
SOX2
DAPI
PAX7
SOX2
DAPI
D
PAX7 SOX10
MYOD
PAX7 MF20
E
FastMyHC
SOX2
DAPI
Day 98
SOX2
TuJ1
DAPI
Day 84
FastMyHC
TuJ1
DAPI
Day 84
SOX2
TuJ1
DAPI
F
PAX7
% Cells
Weeks 2 5 8 14
MYOD1
% Cells
Weeks 5 8 14
G
Count
Value
Proximal
Limb bud
HOX 9,10
Distal
HOX 9,10
HOX 11,12,13
HOXA1
HOXB1
HOXD1
HOXA2
HOXB2
HOXA3
HOXB3
HOXD3
HOXA4
HOXB4
HOXC4
HOXD4
HOXA5
HOXB5
HOXC5
HOXA6
HOXB6
HOXC6
HOXA7
HOXB7
HOXB8
HOXC8
HOXD8
HOXA9
HOXB9
HOXC9
HOXD9
HOXA10
HOXC10
HOXD10
HOXA11
HOXC11
HOXD11
HOXC12
HOXD12
HOXA13
HOXB13
HOXC13
HOXD13
Org_16wk
Org_16wk
Org_8wk
Org_8wk
Org_4wk
Org_4wk
Org_4wk

Figure 2. Neural lineage development during skeletal muscle organoid progression. (A) Stepwise brightfield images depicting delamination / migration of the progenitor population during organoid culture progression and the corresponding myofiber formation **(B)** Gene ontology enrichment analysis between 8w and 4w organoids attributes muscle identity at organoid culture at 8w post differentiation, and highlights muscle system development and neural lineage arrest among the top upregulated and downregulated gene ontology terms respectively **(C)** Organoid overview at Day 35 indicates predominant expression of FastMyHC$^+$ and PAX7$^+$ myogenic populations, while SOX2$^+$ neural populations at this stage are demarcating the position of the SOX2 neural plate border epithelium observed at earlier stages (Day 16) **(D)** At this stage, PAX7 cells are of myogenic origin (PAX7$^+$/SOX10$^-$), while in their proximity we detect MF20$^+$ myotubes. Further, at this stage MYOD1$^+$ cells appear at the organoid periphery **(E)** TUJ1$^+$ neurons are restricted to the inner portion of the organoid and close to SOX2 positive epithelium, while FastMyHC positive myofibers occupy organoids exterior portion **(F)** Histographs based on FACS intracellular quantification depicting the percentage of cells being PAX7$^+$ or MYOD1$^+$ across different time points of differentiation. For each replicate we pooled 10 organoids, n=3, Statistics: *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, ns: not significant **(G)** Heatmap on HOX gene cluster emphasizes organoid culture´s limb axial anatomical identity, by depicting the transition from an initial limb bud (HOX 9-10) towards a more distal identity (HOX 11-13) at 8w and 16w post differentiation. Scale bars, 500μm **(C)**, 200μm **(A, D, E)**.


A
Color Key and Histogram
Count
Value
MEOX2
MEOX1
MYOD1
SIM2
TCF4
LBX1
SIM1
MSX1
EYA4
PITX2
SIX1
MYF5
SIX4
PAX7
EYA1
PAX3
CXCR4
EN1
FOXC1
FOXC2
#1 #2 #3 #1 #2 #1 #2
Org_4wk Org_8wk Org_16wk
B
Org_8wk vs Org_4wk
Up: 5.919
Down: 5.005
log fold change
Mean of normalized counts
C
Color Key and Histogram
Count
Value
Neural Tube Markers
Neural Crest Markers
HES5
PAX6
FOXB1
OLIG3
TUBB3
SOX1
SOX10
TFAP2A
SOX9
DLX5
PAK3
B3GAT1
#1 #2 #1 #2 #1 #2 #3
Org_8wk Org_16wk Org_4wk
D
Delamination / Migration - Myogenic Progenitors
TFAP2A SOX2 DAPI
Day 35
PAX7 MF20 DAPI
Day 35

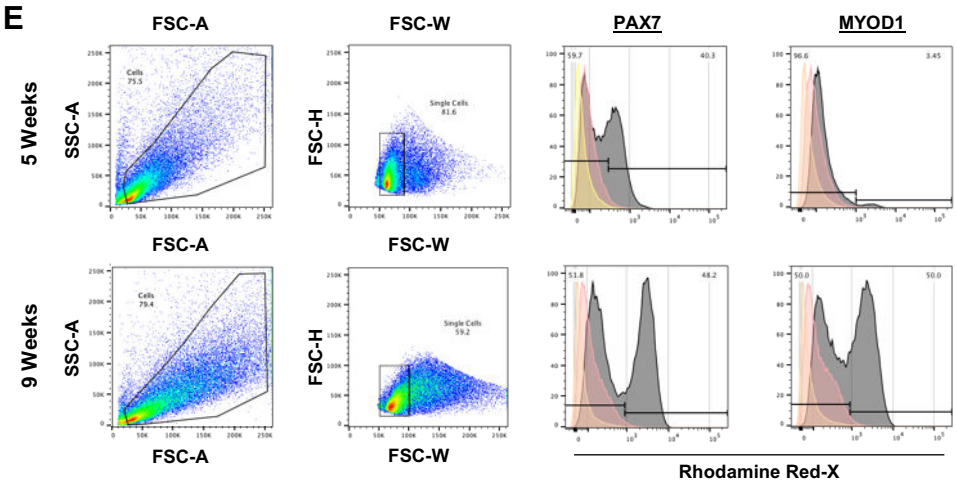

E
FSC-A
FSC-W
PAX7
MYOD1
5 Weeks
9 Weeks
SSC-A
FSC-H
Rhodamine Red-X

Figure 2 – figure supplement 1. Myogenic versus neural fate during organoid development. (A) Heatmap on limb myogenic progenitor markers emphasizes upregulation of myogenic marker along organoid protocol development (**B)** Differential expression comparison highlights upregulated (*n*=5.919) and downregulated (*n*=5.005) genes between 8w vs 4w organoids. (Padj< 0.001) **(C)** Heatmap on neural tube and neural crest markers during the 4th, 8th and 16th week post differentiation highlights the neural lineage arrest from 4th onwards **(D)** Organoid overviews at Day 35 indicates presence of $MF20^{+}$ myofibers in the proximity of $PAX7^{+}$ cells and that the delamination migration process is mainly of myogenic origin since SOX2 and TFAP2A expression is restricted towards the inner portion of the organoid at that stage **(E)** Gating strategy applied to quantify $MYOD1^{+}$ (committed state) and $PAX7^{+}$ (progenitor state) population at five- and eight- weeks post differentiation. Yellow histogram: unstained population, Red histogram: Isotype Control, Gray Histogram: $PAX7^{+}$ or $MYOD^{+}$ population Scale bars, 500uM **(D)**.

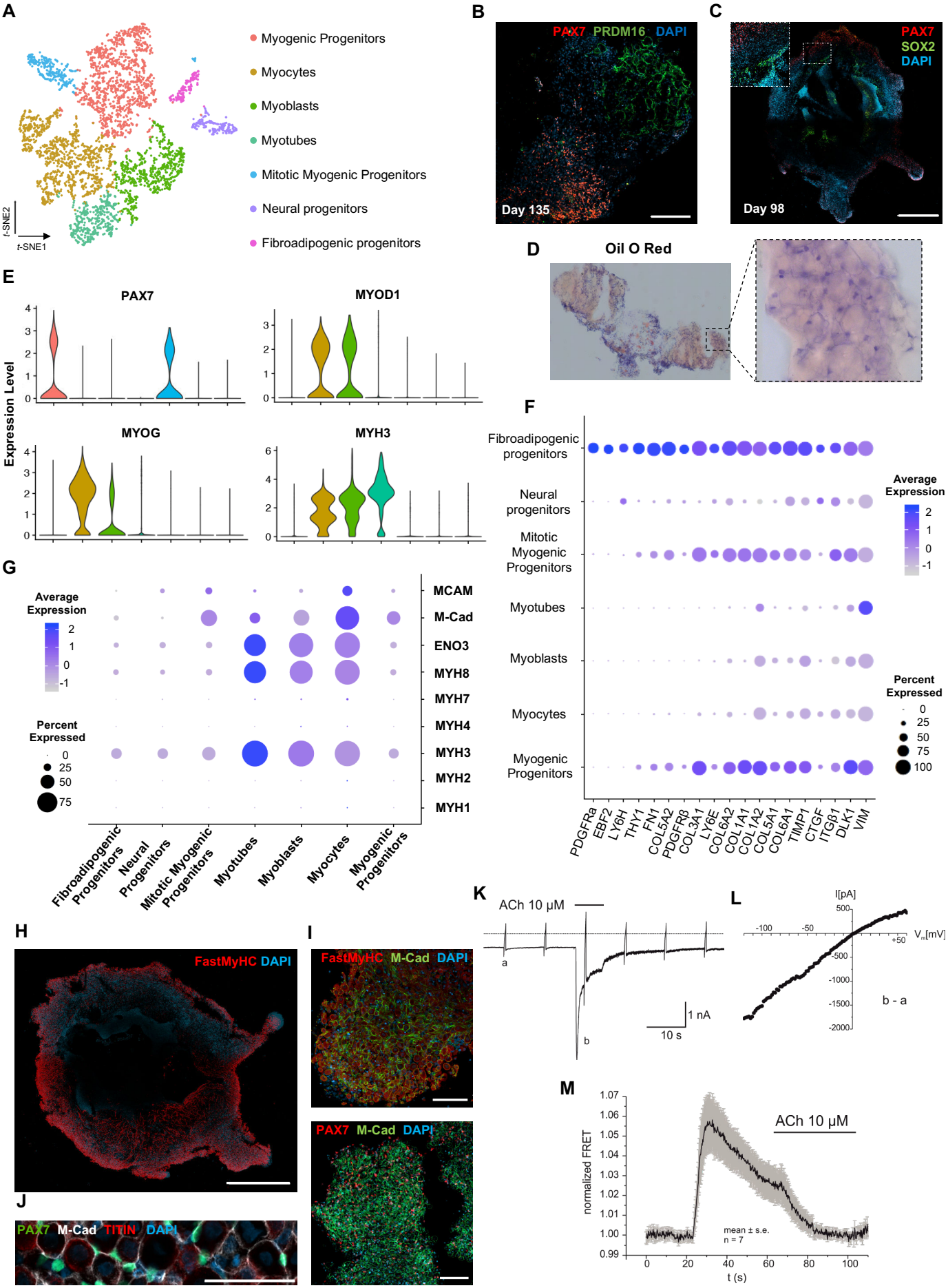

A
Myogenic Progenitors
Myocytes
Myoblasts
Myotubes
Mitotic Myogenic Progenitors
Neural progenitors
Fibroadipogenic progenitors
t-SNE2
t-SNE1
B
PAX7 PRDM16 DAPI
Day 135
C
PAX7
SOX2
DAPI
Day 98
D
Oil O Red
E
PAX7
MYOD1
MYOG
MYH3
Expression Level
F
Fibroadipogenic progenitors
Neural progenitors
Mitotic Myogenic Progenitors
Myotubes
Myoblasts
Myocytes
Myogenic Progenitors
Average Expression
Percent Expressed
PDGFRa
EBF2
LY6H
THY1
FN1
COL5A2
PDGFRβ
COL3A1
LY6E
COL6A2
COL1A1
COL1A2
COL5A1
COL6A1
TIMP1
CTGF
ITGB1
DLK1
VIM
G
MCAM
M-Cad
ENO3
MYH8
MYH7
MYH4
MYH3
MYH2
MYH1
Fibroadipogenic Progenitors
Neural Progenitors
Mitotic Myogenic Progenitors
Myotubes
Myoblasts
Myocytes
Myogenic Progenitors
H
FastMyHC DAPI
I
FastMyHC M-Cad DAPI
PAX7 M-Cad DAPI
J
PAX7 M-Cad TITIN DAPI
K
ACh 10 μM
a
b
1 nA
10 s
L
I[pA]
V_m[mV]
b - a
M
ACh 10 μM
normalized FRET
mean ± s.e.
n = 7
t (s)

Figure 3. Skeletal muscle organoid characterization at single cell resolution. (A) t-SNE visualization of color-coded clustering (n= 4.323 cells) from organoid culture at 12w post differentiation highlights predominant presence of skeletal muscle lineage, represented by clusters corresponding to myogenic progenitors (n=1625 cells, 37% of total population) in non-dining (n=1317 cells) and mitotic (n=308 cells) state, myoblasts (n=731 cells), myocytes (n=1147 cells) and myotubes (n=442). Additional, mesenchymal and neural lineages represented by two smaller cluster of fibroadipogenic (n=165 cells) and neural (n=213 cells) progenitors, respectively **(B)** Immunocytochemistry picture at Day 135 indicates derivation of PRDM16+ adipocytes clustered together at a separate location from PAX7+ myogenic progenitors **(C)** Organoid overview at Day 98 depicts expression of PAX7+ myogenic populations and highlights that SOX2+ neural populations are located towards the interior **(D)** Positive areas with Oil O Red staining indicate derivation of adipocytes in the organoid culture at Day 135 **(E)** Violin plots of key markers from each stage, PAX7, MYOD1, MYOG, MYH3, depict their relative expression levels and emphasize the gradual transition from the myogenic progenitor to myotube subcluster **(F)** Dot plot showing the expression of representative genes related to adipogenesis and fibrogenesis across the 7 main clusters. The size of each circle reflects the percentage of cells in a cluster where the gene is detected, and the color reflects the average expression level within each cluster (gray, low expression; blue, high expression) **(G)** Dot plot showing the expression of representative genes related to fetal myogenesis across the 7 main clusters. The size of each circle reflects the percentage of cells in a cluster where the gene is detected, and the color reflects the average expression level within each cluster (gray, low expression; blue, high expression). **(H-J)** Representative organoid overview at Day 98 indicates predominant expression of Fast MyHC fetal myofibers **(H,J),** positive for M-Cadherin and in the proximity of PAX7+ cells **(I,J) (K)** Representative recording (n=6) of acetylcholine-induced changes in holding current in a single skeletal muscle cell. Acetylcholine (ACh, 10 μM) was applied as indicated by the bar. Holding potential -90 mV. Downward deflections represent membrane currents in response to depolarizing voltage ramps (duration 500 ms) from -120 mV to +60 mV. Dashed line indicates zero current level. **(L)** I/V curve of nAChR currents obtained by subtraction of voltage ramp-induced changes of current in the presence and absence of ACh (10-5 mol/L), corresponding lower-case letters b - a in (right) **(M)** Summarized FRET-recordings (left) from skeletal muscle cells transfected with Twitch2B to monitor the increase in [Ca2+]i during application of ACh (h) Scale bars, 1mm (**C,H)**, 200μm **(B)**, 100μm **(I,J)**.

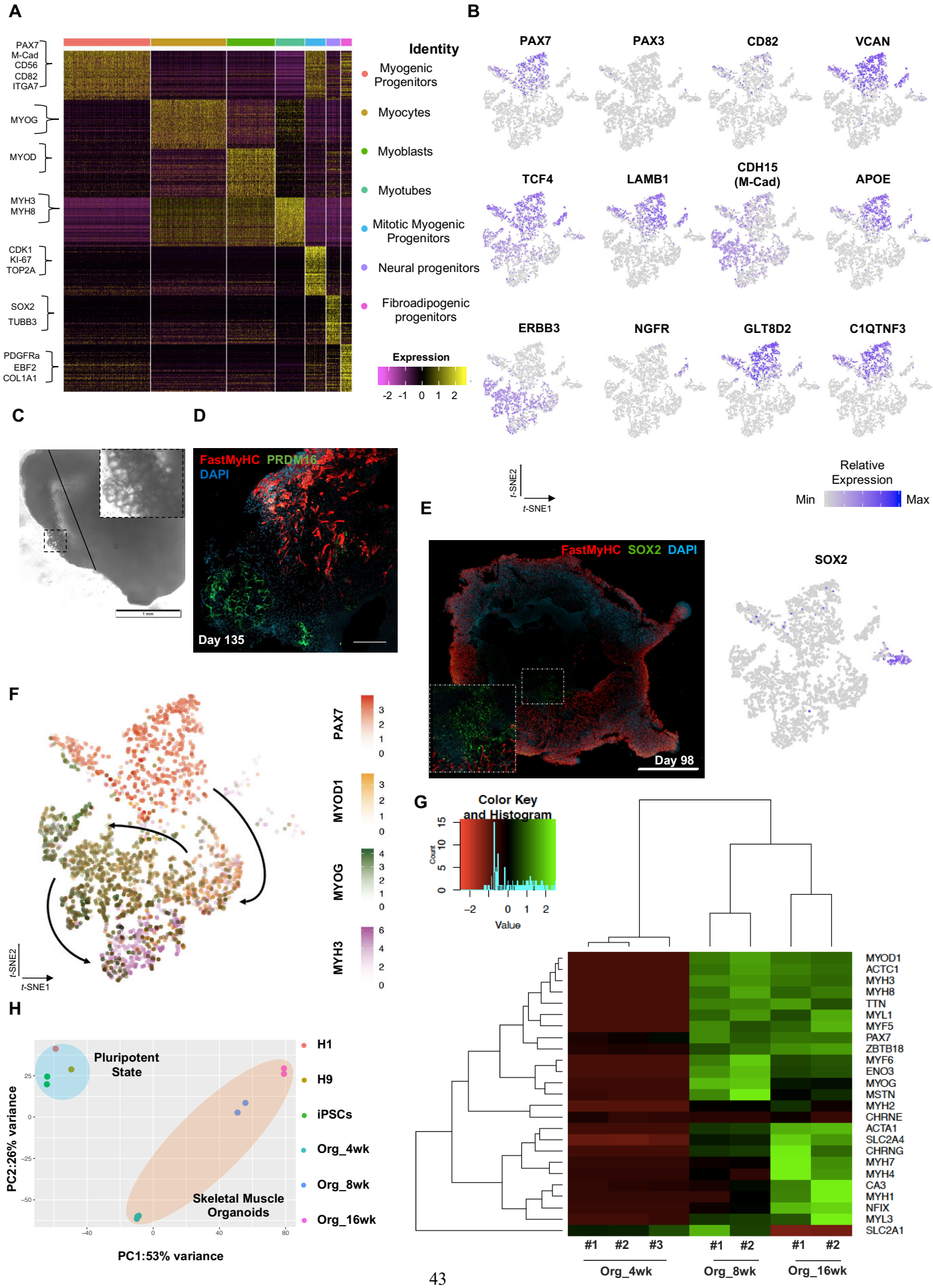

A
PAX7
M-Cad
CD56
CD82
ITGA7
MYOG
MYOD
MYH3
MYH8
CDK1
KI-67
TOP2A
SOX2
TUBB3
PDGFRa
EBF2
COL1A1
Identity
Myogenic Progenitors
Myocytes
Myoblasts
Myotubes
Mitotic Myogenic Progenitors
Neural progenitors
Fibroadipogenic progenitors
Expression
-2 -1 0 1 2
B
PAX7
PAX3
CD82
VCAN
TCF4
LAMB1
CDH15 (M-Cad)
APOE
ERBB3
NGFR
GLT8D2
C1QTNF3
t-SNE2
t-SNE1
Relative Expression
Min
Max
C
D
FastMyHC PRDM16
DAPI
Day 135
E
FastMyHC SOX2 DAPI
Day 98
SOX2
F
PAX7
MYOD1
MYOG
MYH3
t-SNE2
t-SNE1
G
Color Key and Histogram
Count
Value
MYOD1
ACTC1
MYH3
MYH8
TTN
MYL1
MYF5
PAX7
ZBTB18
MYF6
ENO3
MYOG
MSTN
MYH2
CHRNE
ACTA1
SLC2A4
CHRNG
MYH7
MYH4
CA3
MYH1
NFIX
MYL3
SLC2A1
#1 #2 #3
Org_4wk
#1 #2
Org_8wk
#1 #2
Org_16wk
H
Pluripotent State
Skeletal Muscle Organoids
PC2:26% variance
PC1:53% variance
H1
H9
iPSCs
Org_4wk
Org_8wk
Org_16wk

Figure 3 – figure supplement 1. Single cell RNA seq expression profiling and lineage representation in organoid culture at 12th week. (A) Gene signatures of the *t*-SNE described clusters based on the relative expression levels of the 50 most-significant markers for each of the 7 clusters **(B)** *t*-SNE plot of myogenic and fibroadipogenic markers **(C)** Brightfield picture at Day 135 highlights the presence of structures reassembling adipocytes at the organoid culture **(D)** Immunocytochemistry picture from that stage emphasizes the derivation of PRDM16 positive cells in the proximity of FastMyHC positive myofibers **(E)** Organoid overview at Day 98 highlights muscle system development at its outer portion, and neural lineage representation at more inner sites **(F)** *t*-SNE plot of key markers from each stage, PAX7, MYOD1, MYOG, MYH, depict the relative expression levels and emphasize the gradual transition from the myogenic progenitor to myotube subcluster **(G)** Heatmap on embryonic or fetal myogenic markers depicts a fetal environment during 8th or 16th week post differentiation **(H)** Two-dimensional principal component analysis (PCA) based on gene expression profiling between the samples, separates the skeletal muscle organoid cluster from the pluripotent one and highlights the greater difference in variance between the 4w and 8-,16w subgroups in comparison to 8w and 16w subgroup. w.p.d.: weeks post differentiation. Scale bars, 1mM **(D),** 200uM **(E, F)**.

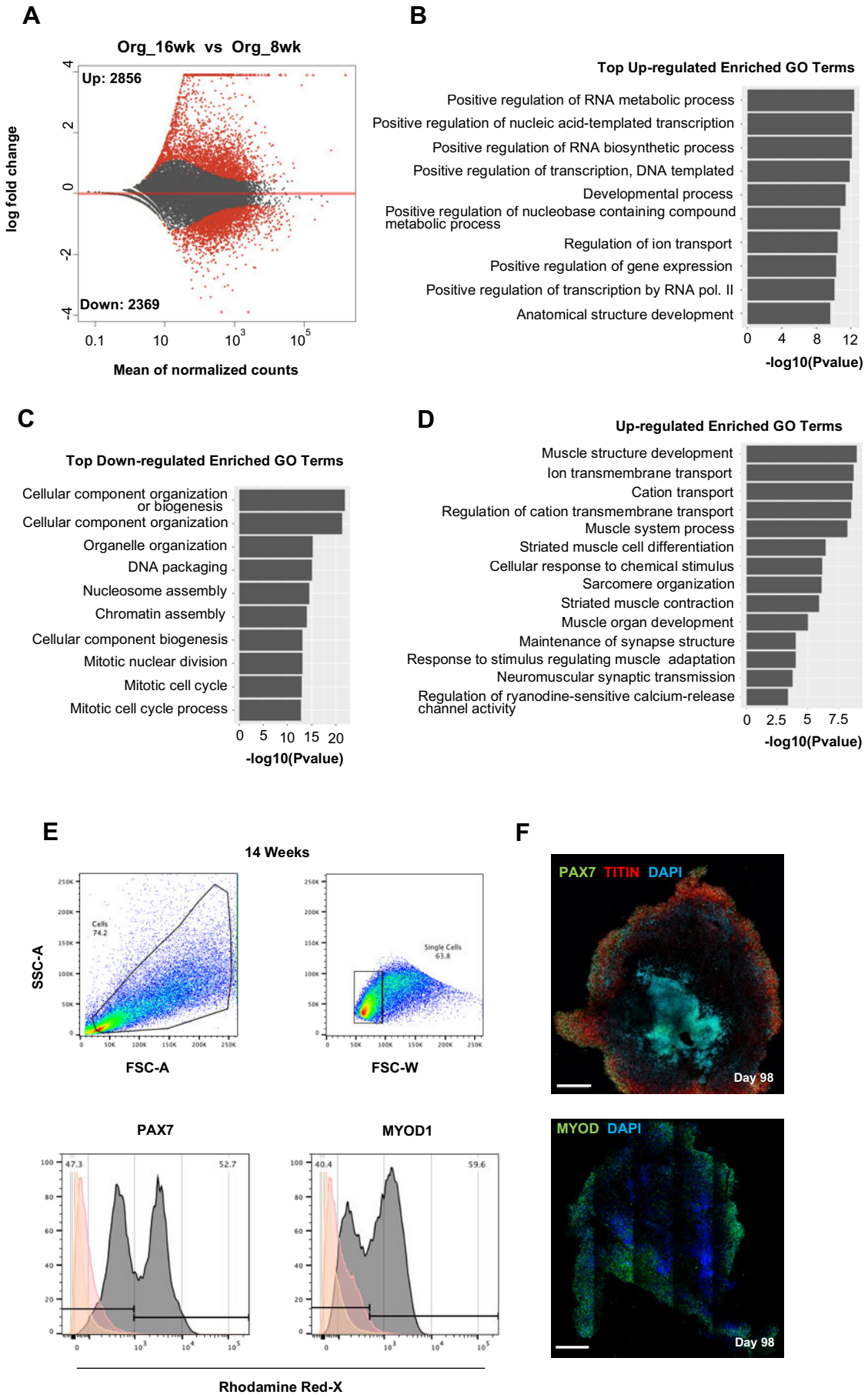

A
Org_16wk vs Org_8wk
Up: 2856
Down: 2369
log fold change
Mean of normalized counts
B
Top Up-regulated Enriched GO Terms
Positive regulation of RNA metabolic process
Positive regulation of nucleic acid-templated transcription
Positive regulation of RNA biosynthetic process
Positive regulation of transcription, DNA templated
Developmental process
Positive regulation of nucleobase containing compound metabolic process
Regulation of ion transport
Positive regulation of gene expression
Positive regulation of transcription by RNA pol. II
Anatomical structure development
-log10(Pvalue)
C
Top Down-regulated Enriched GO Terms
Cellular component organization or biogenesis
Cellular component organization
Organelle organization
DNA packaging
Nucleosome assembly
Chromatin assembly
Cellular component biogenesis
Mitotic nuclear division
Mitotic cell cycle
Mitotic cell cycle process
-log10(Pvalue)
D
Up-regulated Enriched GO Terms
Muscle structure development
Ion transmembrane transport
Cation transport
Regulation of cation transmembrane transport
Muscle system process
Striated muscle cell differentiation
Cellular response to chemical stimulus
Sarcomere organization
Striated muscle contraction
Muscle organ development
Maintenance of synapse structure
Response to stimulus regulating muscle adaptation
Neuromuscular synaptic transmission
Regulation of ryanodine-sensitive calcium-release channel activity
-log10(Pvalue)
E
14 Weeks
SSC-A
FSC-A
FSC-W
Cells 74.2
Single Cells 63.8
PAX7
MYOD1
47.3
52.7
40.4
59.6
Rhodamine Red-X
F
PAX7 TITIN DAPI
Day 98
MYOD DAPI
Day 98

Figure 3 – figure supplement 2. Skeletal muscle organoid culture over time maturation and identity. (A) Differential expression comparison highlights upregulated (*n*=2.856) and downregulated (*n*=2.369) genes between 16w vs 8w organoids (Padj<0.001) **(B-C)** Gene ontology enrichment analysis depicts the top statistically significant upregulated **(b)** and downregulated **(C)** gene ontology (G.O) terms **(D)** Statistically significant upregulated G.O terms at 16th week organoids further highlight skeletal muscle maturation **(E)** Gating strategy applied to quantify MYOD1^{+} (committed state) and PAX7^{+} (progenitor state) population at 14th week post differentiation during skeletal muscle organoid development **(F)** Organoid overview at Day 98 depicts high proportion of cells being PAX7 and MYOD positive towards its outer portion. Yellow histogram: unstained population, Red histogram: Isotype Control, Gray Histogram: PAX7^{+} or MYOD^{+} population**.** Scale bars, 500uM **(F)**

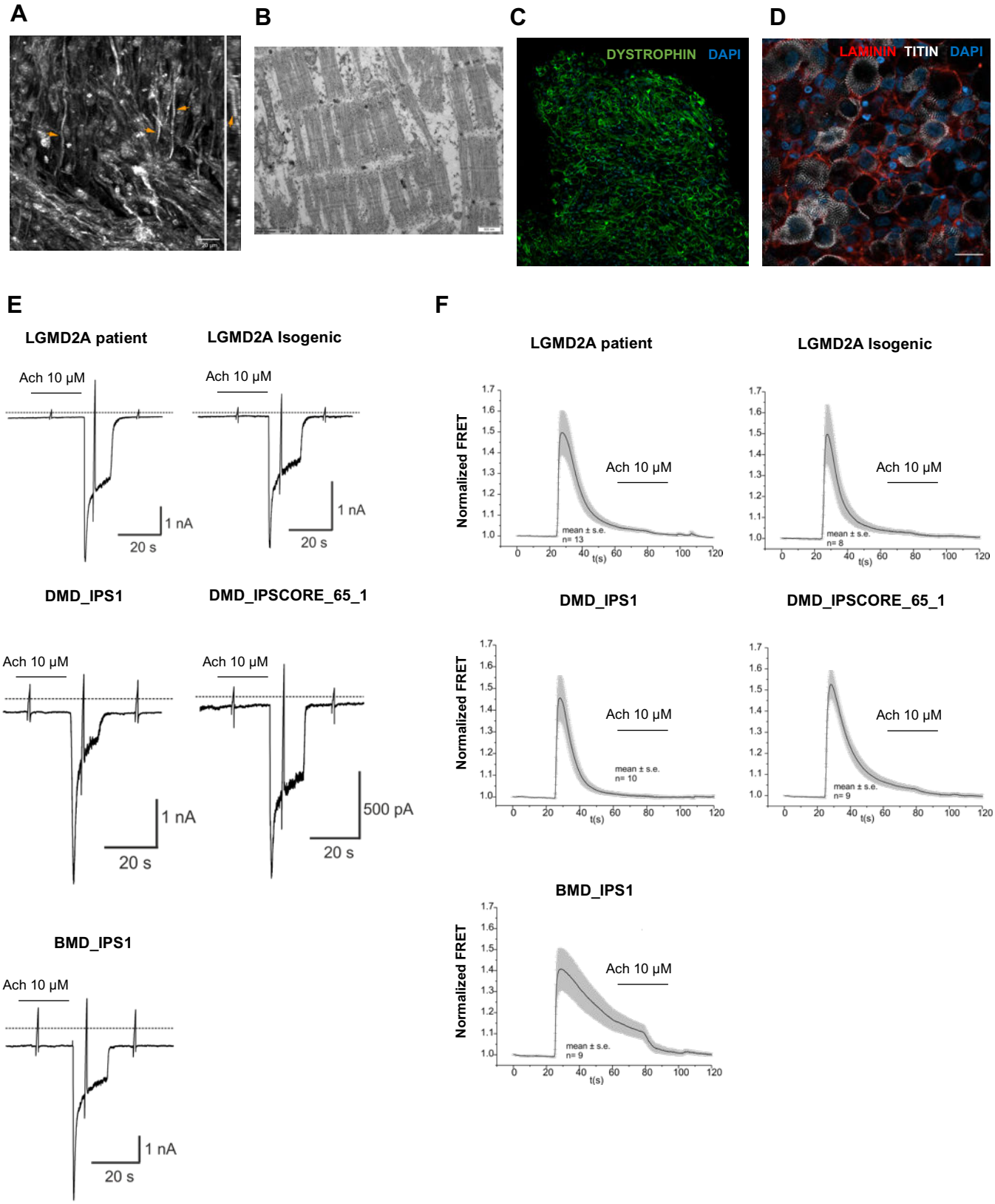

A
B
C
DYSTROPHIN DAPI
D
LAMININ TITIN DAPI
E
LGMD2A patient
LGMD2A Isogenic
Ach 10 μM
1 nA
20 s
DMD_IPS1
DMD_IPSCORE_65_1
500 pA
BMD_IPS1
F
Normalized FRET
t(s)
mean ± s.e.
n= 13
n= 8
n= 10
n= 9

Figure 3 – figure supplement 3. Functional properties of organoid derived skeletal muscle myofibers. (A) Second Harmonic Generation imaging (SHG) from areas with dense fiber network reveals myofibril formation with distinct sarcomeres. (arrows) z plane of stack (right) **(B)** Electron microscopy pictures depicts myofibrils with well-developed sarcomeres **(C,D)** Skeletal muscle organoid derived myofibers expressed LAMININ and DYSTROPHIN **(E)** Representative recording of acetylcholine-induced changes in holding current in a single skeletal muscle cell from organoid-derived skeletal muscle cells from iPSC lines with Duchenne (DMD_IPS1, DMD_IPSCORE_65_1), Becker (BMD_IPS1) or LGMD2A genetic background. ACh (10 μM) was applied as indicated by the bar. Holding potential -90 mV. Downward deflections represent membrane currents in response to depolarizing voltage ramps (duration 500 ms) from -120 mV to +60 mV. Dashed line indicates zero current level. **(F)** Summarized FRET-recordings from organoid-derived skeletal muscle cells from iPSC lines with Duchenne (DMD_IPS1, DMD_IPSCORE_65_1), Becker (BMD_IPS1) or LGMD2A genetic background; transfected with Twitch2B to monitor the increase in [Ca2+]i during application of Acetylcholine (Ach). Scale Bars, 100uM **(C,D)**, 20uM in **(A)**, 500nM **(B)**

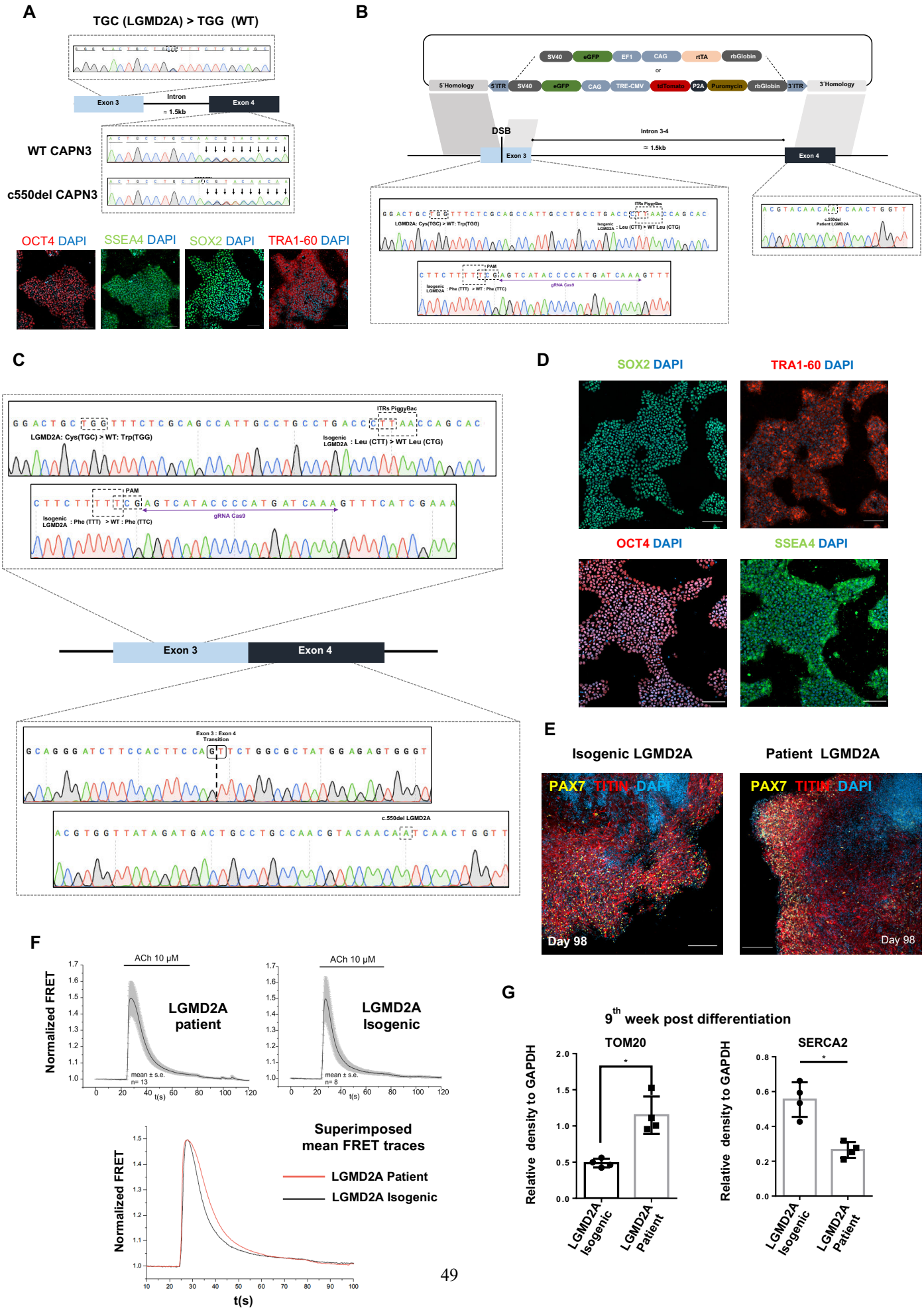
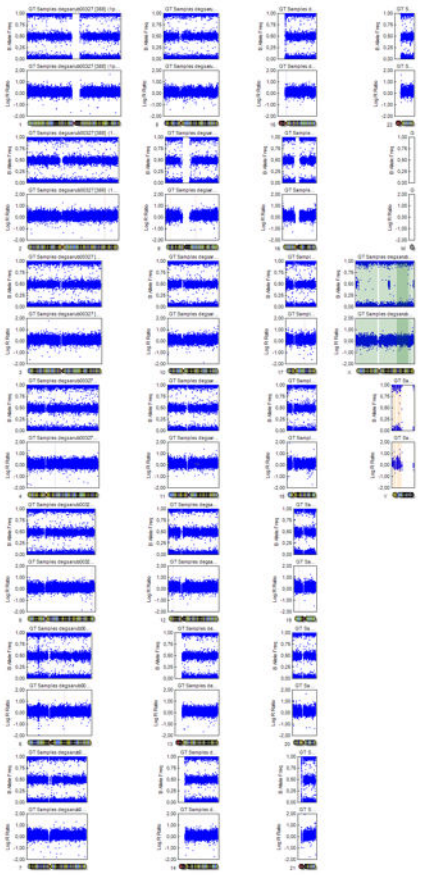

A
TGC (LGMD2A) > TGG (WT)
Exon 3
Intron
≈ 1.5kb
Exon 4
WT CAPN3
c550del CAPN3
OCT4 DAPI
SSEA4 DAPI
SOX2 DAPI
TRA1-60 DAPI
B
SV40
eGFP
EF1
CAG
rtTA
rbGlobin
or
5'Homology
5'ITR
SV40
eGFP
CAG
TRE-CMV
tdTomato
P2A
Puromycin
rbGlobin
3'ITR
3'Homology
DSB
Intron 3-4
≈ 1.5kb
Exon 3
Exon 4
ITRs PiggyBac
LGMD2A: Cys(TGC) > WT: Trp(TGG)
Isogenic LGMD2A : Leu (CTT) > WT Leu (CTG)
PAM
gRNA Cas9
Isogenic LGMD2A : Phe (TTT) > WT : Phe (TTC)
c.550del Patient LGMD2A
C
ITRs PiggyBac
LGMD2A: Cys(TGC) > WT: Trp(TGG)
Isogenic LGMD2A : Leu (CTT) > WT Leu (CTG)
PAM
gRNA Cas9
Isogenic LGMD2A : Phe (TTT) > WT : Phe (TTC)
Exon 3
Exon 4
Exon 3 : Exon 4 Transition
c.550del LGMD2A
D
SOX2 DAPI
TRA1-60 DAPI
OCT4 DAPI
SSEA4 DAPI
E
Isogenic LGMD2A
Patient LGMD2A
PAX7 TITIN DAPI
Day 98
PAX7 TITIN DAPI
Day 98
F
ACh 10 μM
LGMD2A patient
Normalized FRET
mean ± s.e.
n= 13
t(s)
ACh 10 μM
LGMD2A Isogenic
mean ± s.e.
n= 8
t(s)
Superimposed mean FRET traces
LGMD2A Patient
LGMD2A Isogenic
Normalized FRET
t(s)
G
9th week post differentiation
TOM20
Relative density to GAPDH
LGMD2A Isogenic
LGMD2A Patient
SERCA2
Relative density to GAPDH
LGMD2A Isogenic
LGMD2A Patient

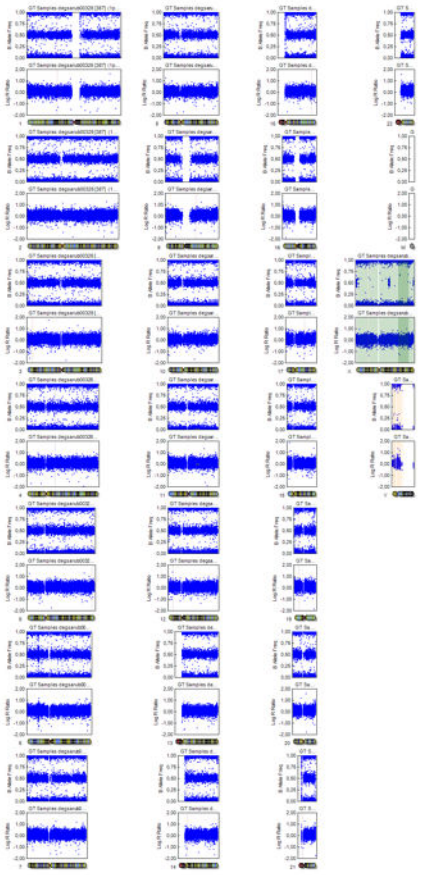
Figure 4. Generation of functional Isogenic LGMD2A hiPSC line using CRISPR/Cas9 genome editing technology. (A) Graph illustrating the genetic alterations on EXON3 and EXON4 of the Calpain 3 locus for the patient LGMD2A line together with immunocytochemistry pictures displaying the pluripotent state of LGMD2A patient line before genome editing **(B)** Graph depicting the genome editing strategy followed for introducing controlled biallelic modification **(C)** Scheme displaying the new EXON3/4 CAPN3 conformation together with pyrograms emphasizing all the silent introduced mutations TTT:TTC (Phe), CTG:CTT (Leu), the EXON3 to EXON4 transition as well as the rescues in EXON3: 130Cys(TGC-LGMD2A)>Trp(TGG-WT) and EXON4: c.550delA **(D)** Immunocytochemistry pictures displaying the pluripotent state of Isogenic LGMD2A line after genome editing **(E)** Immunocytochemistry pictures highlighting derivation of $PAX7^+$ fetal myogenic progenitor and $TITIN^+$ myofiber population during skeletal muscle organoid development for patient and isogenic LGMD2A hiPSC lines **(F)** Summarized FRET-recordings from organoid-derived skeletal myotubes from iPSC lines with LGMD2A (Patient, and Isogenic) genetic background; transfected with Twitch2B to monitor the increase in intracellular [Ca^{2+}] during application of Acetylcholine (Ach). Superimposed FRET traces for Patient and Isogenic LGMD2A derived skeletal myotubes indicated prolonged intracellular Ca^{2+} concentration upon activation with ACh for the patient LGMD2A line **(G)** Relative quantification of the TOM20 and SERCA2 levels from western blots. Data points represent individual organoids (n=4 organoids from independent derivations). Comparison between patient and Isogenic LGMD2A cell lines was performed with a Mann-Whitney test (*p= 0.0286). Scale bars, 200uM **(E),** 100uM **(A,D).**

A

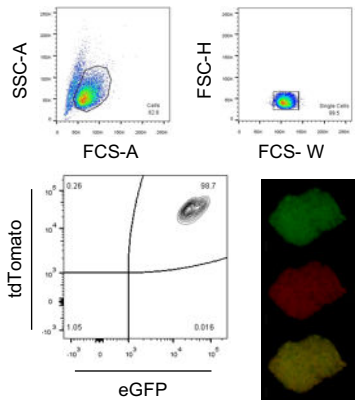
LGMD2A Patient



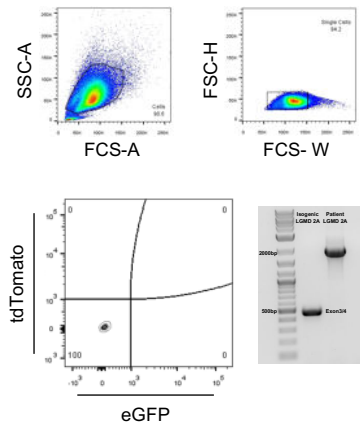
Isogenic LGMD2A



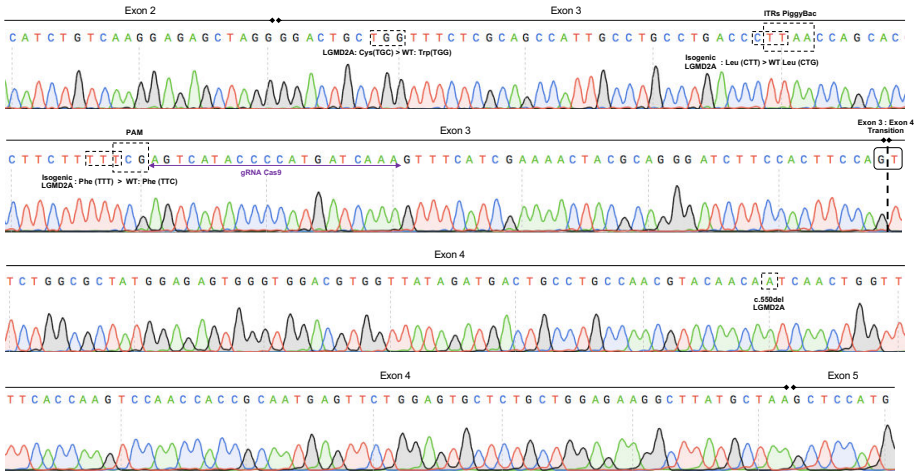
B



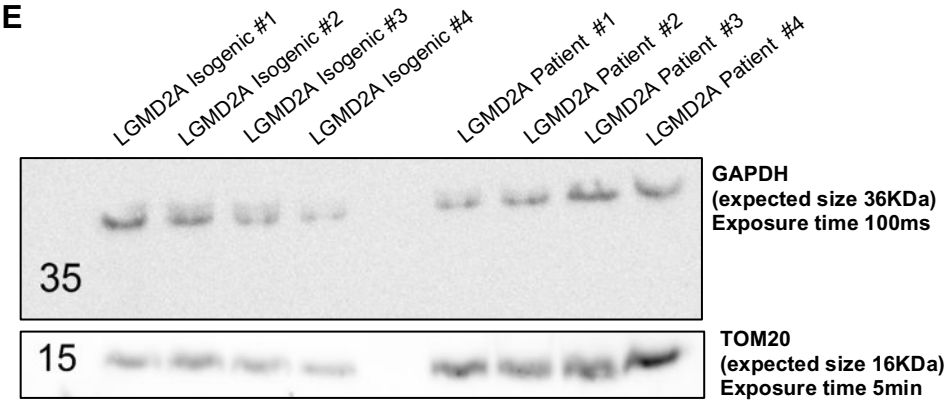
C



D



E



F

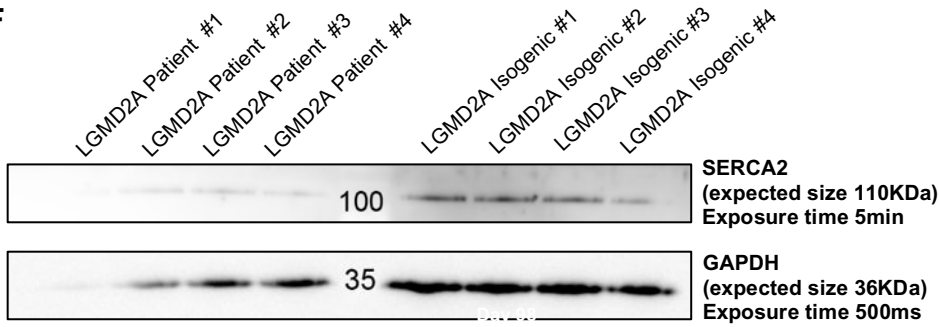


Figure 4 – figure supplement 1. Evaluation on Isogenic LGMD2A hiPSC line integrity and functionality. (A) Karyograms from the patient and isogenic LGMD2A line indicating that no genomic alterations occurred during genome editing **(B)** FACS plots displaying the gating strategy for enrichment of cells incorporated the selection cassettes, together with an immunofluorescence picture of an iPSCs colony after sorting for tdTomato^{+} / eGFP^{+} population **(C)** Gating strategy to distinguish / acquire cells with traceless removal of the selection cassette. (tdTomato^{-} / eGFP^{-} population) and PCR verification of traceless cassette excision at CAPN3 locus, and chimeric Exon3/4 generation **(D)** cDNA from isogenic LGMD2A skeletal muscle organoids highlights that the intron skipping had no effect at the CAPN3 mRNA sequence **(E-F)** Representative western blot staining for TOM20 **(E)** and SERCA2 **(F)** showing an increase in LGMD2A patient derived organoids.

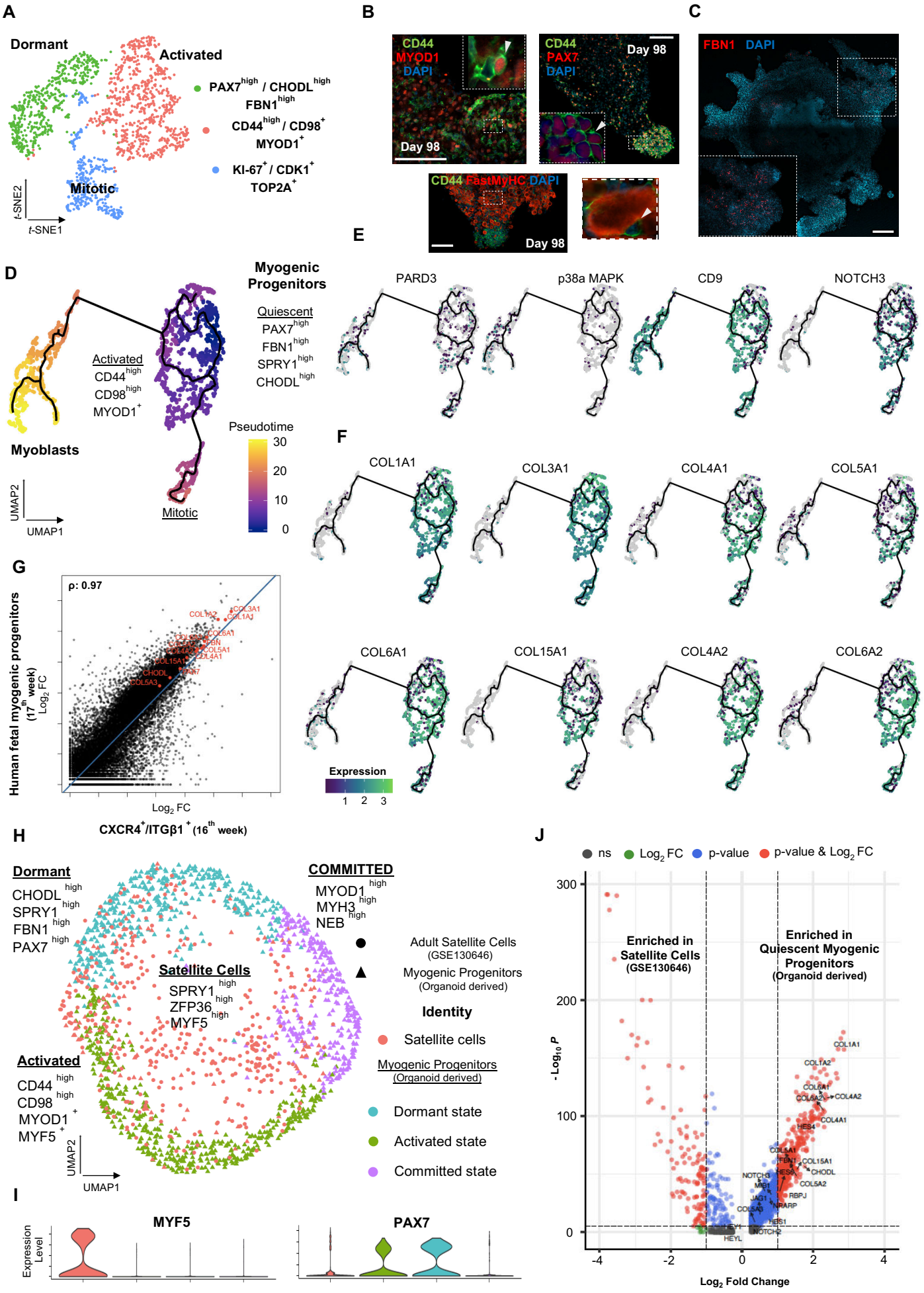

A
Dormant
Activated
Mitotic
PAX7high / CHODLhigh
FBN1high
CD44high / CD98+
MYOD1+
KI-67+ / CDK1+
TOP2A+
t-SNE2
t-SNE1
B
CD44
MYOD1
DAPI
Day 98
CD44
PAX7
DAPI
Day 98
CD44 FastMyHC DAPI
Day 98
C
FBN1 DAPI
D
Myogenic Progenitors
Quiescent
PAX7high
FBN1high
SPRY1high
CHODLhigh
Activated
CD44high
CD98high
MYOD1+
Myoblasts
Mitotic
Pseudotime
30
20
10
0
UMAP2
UMAP1
E
PARD3
p38a MAPK
CD9
NOTCH3
F
COL1A1
COL3A1
COL4A1
COL5A1
COL6A1
COL15A1
COL4A2
COL6A2
Expression
1 2 3
G
ρ: 0.97
Human fetal myogenic progenitors (17th week) Log2 FC
Log2 FC
CXCR4+/ITGβ1+ (16th week)
H
Dormant
CHODLhigh
SPRY1high
FBN1high
PAX7high
COMMITTED
MYOD1high
MYH3high
NEB high
Satellite Cells
SPRY1high
ZFP36high
MYF5high
Activated
CD44high
CD98high
MYOD1+
MYF5+
UMAP2
UMAP1
Adult Satellite Cells (GSE130646)
Myogenic Progenitors (Organoid derived)
Identity
Satellite cells
Myogenic Progenitors (Organoid derived)
Dormant state
Activated state
Committed state
I
MYF5
PAX7
Expression Level
J
ns
Log2 FC
p-value
p-value & Log2 FC
Enriched in Satellite Cells (GSE130646)
Enriched in Quiescent Myogenic Progenitors (Organoid derived)
- Log10 P
Log2 Fold Change
300
200
100
0
-4
-2
0
2
4
COL1A1
COL1A2
COL6A1
COL5A3
COL4A2
COL4A1
HES4
COL5A1
FBN1
COL15A1
CHODL
HES6
COL5A2
NOTCH3
MGP
JAG1
RBPJ
NRARP
COL5A3
HES1
HEYL
NOTCH2

Figure 5. Myogenic progenitor´s identity and comparison to progenitors derived from fetal and adult muscle tissue. (A) t-SNE plot visualization of color-coded clustering indicates myogenic progenitor sub-cluster with distinct molecular signatures, the "dormant": PAX7^{high}/ CHODL^{high}/FBN1^{high}, the "activated": CD44^{high} / CD98^{+} / MYOD1^{+} and the "mitotic": KI-67^{+}/CDK1^{+}/TOP2A **(B)** Organoid overviews illustrate CD44 and PAX7 expressing cells at developing sites, more accessible to HGF activation signal, picture highlights the specificity of CD44 on MYOD, PAX7 positive progenitor expressing cells (arrows) and absence from FastMyHC positive myofibers **(C)** FBN1 positive microfibrils are located towards the organoids interior **(D)** Pseudotime ordering for myogenic progenitors and myoblast corresponding clusters, highlights distinct developmental trajectories promoting myogenic commitment and self-renewal (**E)** UMAP feature plots depicting the relative expression of genes regulating asymmetric divisions and self-renewal, e.g PARD3, p38a/b MAPK, NOTCH3 and, (**F)** the relative expression of extracellular matrix collagens, such as COL3A1, COL4A1, COL5A1, COL5A1, COL6A1, COL15A1, along pseudotime ordering **(G)** The correlation coefficient between the Log_2 fold change (Log_2 FC) values for isolated myogenic progenitors from human fetal tissue (17th week) and for FACS sorted CXCR4^{+} / ITGβ1^{+} organoid derived myogenic progenitors (16th week). PAX7, COL1A1, COL3A1, COL4A1, COL5A1, COL15A1, FBN1 and CHODL and further extracellular matrix related genes are highlighted on the plot. Pearson's correlation coefficient, rho=0.9 for global expression comparison and rho=0.97 for the selected genes **(H)** UMAP color-based clustering, divides non-dividing myogenic progenitors and adult satellite cells into four clusters with distinct molecular signatures. Satellite cells characterized by SPRY1^{high} /ZFP36^{high} /MYF5^{high} expression, while fraction of them co-clustered with dormant, SPRY1^{high} / FBN1^{high} / CHODL^{high} / PAX7^{high}, activated, CD44^{high} / CD98^{high} / MYOD1^{+}, and committed, NEB ^{high}/ MYH3^{high} / MYOD1^{high} organoid derived myogenic progenitors**.** Round dots correspond to adult Satellite cells from GSE130646 database, while triangle dots correspond to organoid derived myogenic progenitors **(I)** Violin plots depicting the relative expression levels of markers, specific for the quiescent, PAX7, or the activated, MYF5, muscle stem cell state across adult satellite cells (GSE130646) and organoid derived myogenic progenitors subclusters **(J)** Enhanced Volcano plot from comparing transcript levels between all cells from adult satellite cells (GSE130646) and organoid derived myogenic progenitors. Log2 fold-change in normalized gene expression versus -Log_{10} adjusted p-value is plotted. Differentially expressed genes with adjusted p-value < 0.001 are colored blue and genes with adjusted p-value > 0.001 and log2 fold-change > 1 are colored green. Genes with log2 fold-change > 1 and adjusted p value < 0.001 are labeled red. No significant genes (ns) are labeled with gray color. Scale Bars, 500uM **(C)**, 100uM in **(B).**

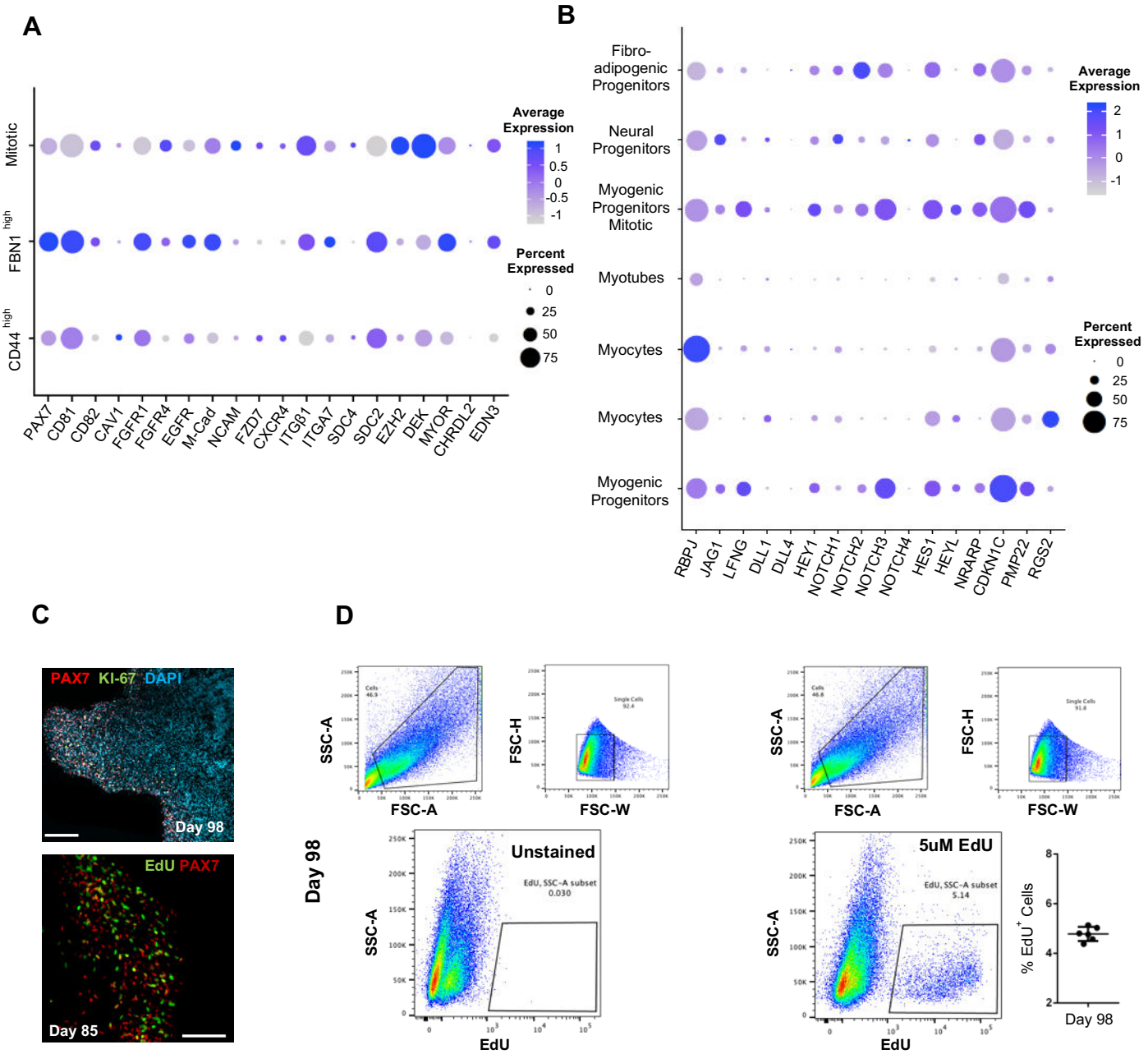

A
Mitotic
FBN1 high
CD44 high
PAX7
CD81
CD82
CAV1
FGFR1
FGFR4
EGFR
M-Cad
NCAM
FZD7
CXCR4
ITGβ1
ITGA7
SDC4
SDC2
EZH2
DEK
MYOR
CHRDL2
EDN3
Average Expression
Percent Expressed
B
Fibro-adipogenic Progenitors
Neural Progenitors
Myogenic Progenitors Mitotic
Myotubes
Myocytes
Myocytes
Myogenic Progenitors
RBPJ
JAG1
LFNG
DLL1
DLL4
HEY1
NOTCH1
NOTCH2
NOTCH3
NOTCH4
HES1
HEYL
NRARP
CDKN1C
PMP22
RGS2
C
PAX7 KI-67 DAPI
Day 98
EdU PAX7
Day 85
D
Day 98
SSC-A
FSC-A
FSC-H
FSC-W
Unstained
EdU, SSC-A subset 0.030
5uM EdU
EdU, SSC-A subset 5.14
EdU
% EdU+ Cells
Day 98

Figure 5 – figure supplement 1. Sub-clustering of myogenic progenitors and NOTCH signaling. (A) Dot plot showing the expression of representative genes related satellite cell identity across the 7 main clusters. **(B)** Dot plot showing the expression of representative genes related to NOTCH signaling and are identified as regulators of satellite cell quiescence across the 7 main clusters **(C)** Illustration at Day98 indicate proportion of PAX7 myogenic progenitors being Ki-67 and EdU negative **(D)** Gating strategy applied to quantify EdU+ proliferating cells in an organoids culture at Day98 (14 weeks). Histograph depicting the percentage of EdU^+ cells at the same stage. Dot Plot: The size of each circle reflects the percentage of cells in a cluster where the gene is detected, and the color reflects the average expression level within each cluster (gray, low expression; blue, high expression), Scale Bars, 100uM in **(C)**.

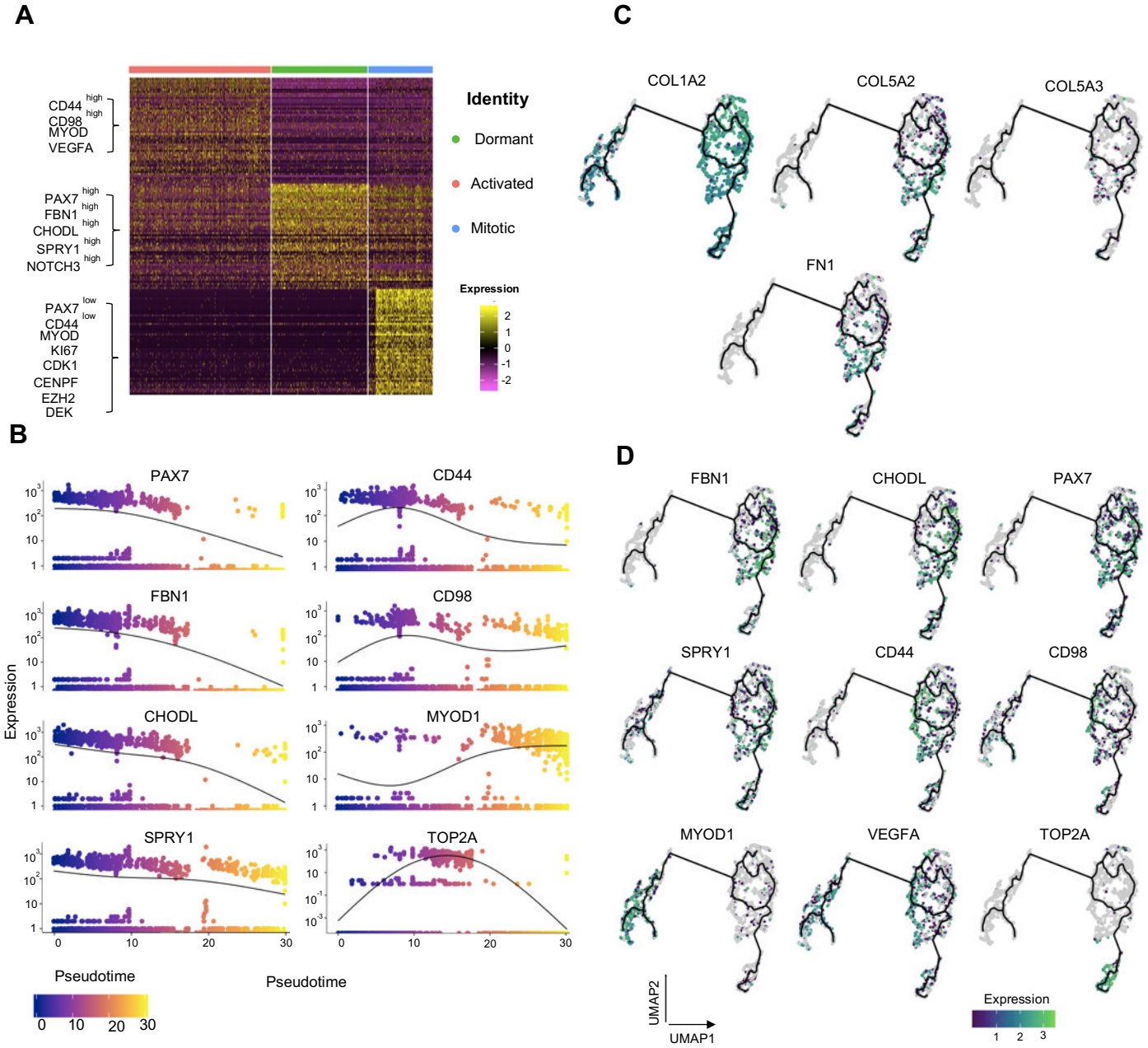
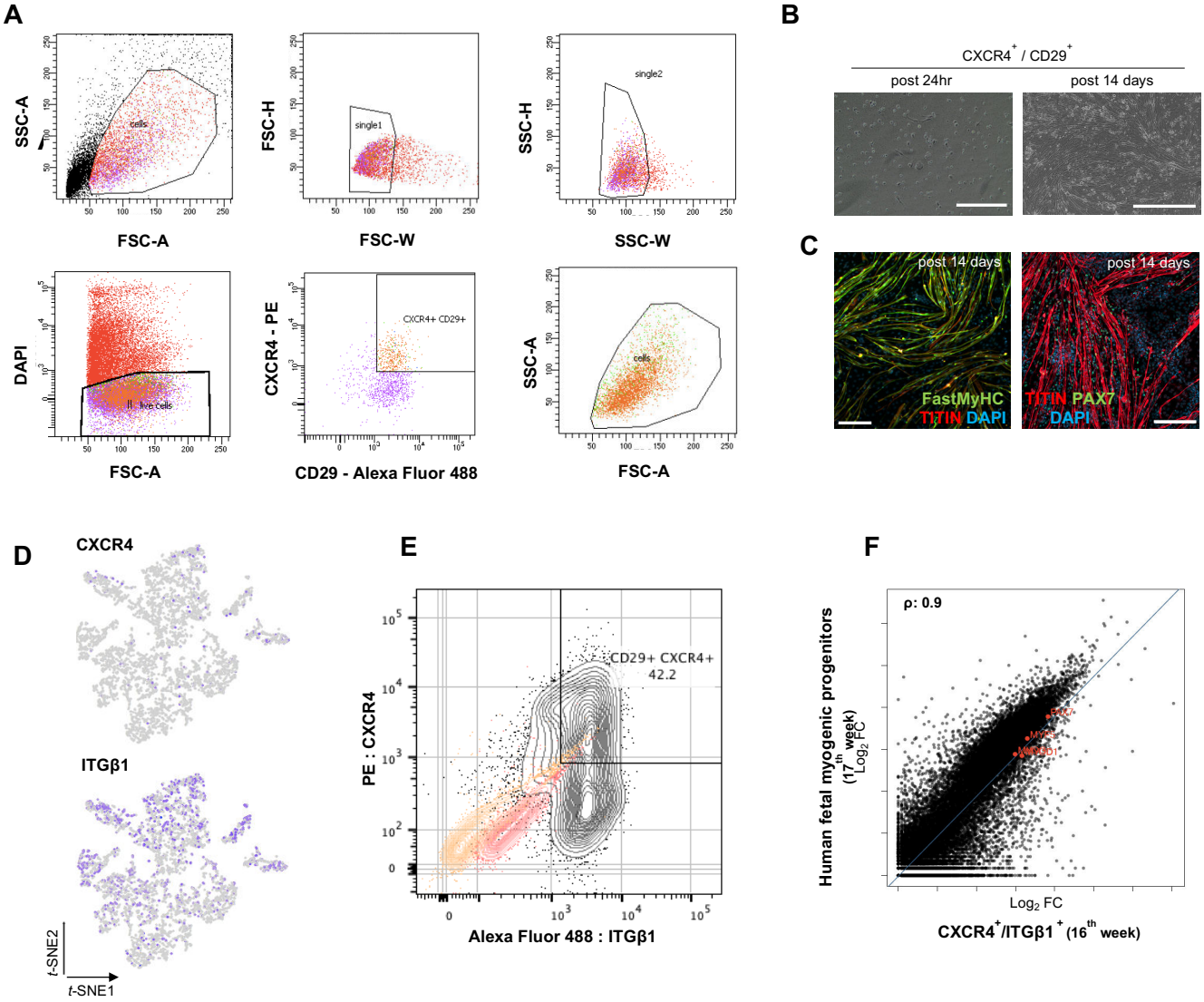

A
CD44 high
CD98 high
MYOD
VEGFA
PAX7 high
FBN1 high
CHODL high
SPRY1 high
NOTCH3
PAX7 low
CD44 low
MYOD
KI67
CDK1
CENPF
EZH2
DEK
Identity
Dormant
Activated
Mitotic
Expression
2
1
0
-1
-2
B
PAX7
CD44
FBN1
CD98
CHODL
MYOD1
SPRY1
TOP2A
Expression
Pseudotime
0
10
20
30
C
COL1A2
COL5A2
COL5A3
FN1
D
FBN1
CHODL
PAX7
SPRY1
CD44
CD98
MYOD1
VEGFA
TOP2A
UMAP2
UMAP1
Expression
1
2
3

Figure 5 – figure supplement 2. Pseudotime ordering of myogenic progenitor reveals distinct states and cell fate decisions (A) Gene signatures of the *t*-SNE described clusters based on the relative expression levels of the 50 most-significant markers for each of the 3 clusters **(B)** Expression of selected genes along pseudotime. Group of genes selected for myogenic progenitors: PAX7, SPRY1, CHODL (dormant state), CD44, CD98, MYOD1 (activated state), TOP2A (mitotic state), and for myoblasts: MYOD1 **(C)** Cells along the myogenic progenitors trajectory, coloured by their expression of extracellular matrix proteins, such as COL1A2, COL4A2, COL5A2, COL5A3, COL6A2, FN1 **(D)** Cells along the myogenic progenitors trajectory, coloured by their expression of selected transcriptional regulators that define the dormant (PAX7, FBN1, CHODL, SPRY1), activated (CD44, CD98, MYOD1, VEGFA) and mitotic (TOP2A) state of myogenic progenitor and myoblast (MYOD1) clusters.


A
SSC-A
FSC-A
cells
FSC-H
FSC-W
single1
SSC-H
SSC-W
single2
DAPI
live cells
CXCR4 - PE
CD29 - Alexa Fluor 488
CXCR4+ CD29+
B
CXCR4+ / CD29+
post 24hr
post 14 days
C
post 14 days
FastMyHC
TITIN DAPI
TITIN PAX7
DAPI
D
CXCR4
ITGβ1
t-SNE2
t-SNE1
E
PE : CXCR4
Alexa Fluor 488 : ITGβ1
CD29+ CXCR4+
42.2
F
ρ: 0.9
Human fetal myogenic progenitors
(17th week)
Log2 FC
CXCR4+/ITGβ1+ (16th week)

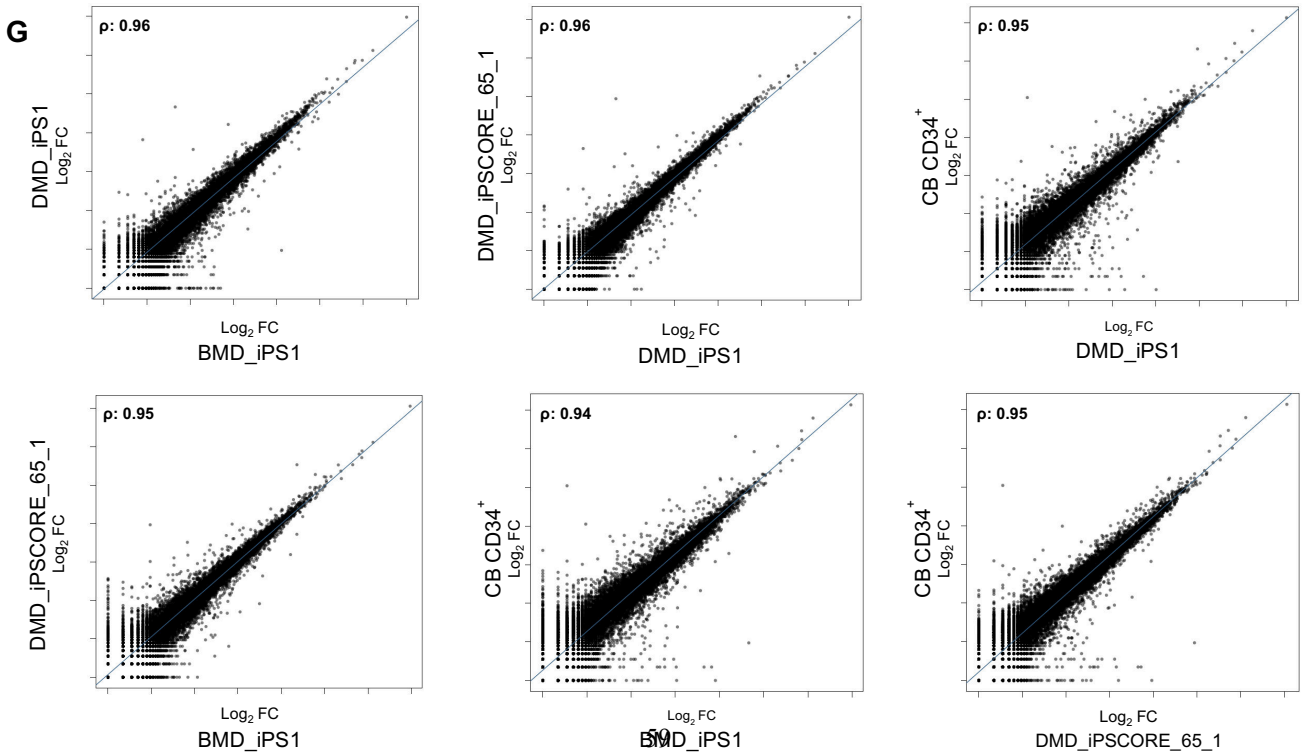

G
ρ: 0.96
DMD_iPS1
Log2 FC
BMD_iPS1
ρ: 0.96
DMD_iPSCORE_65_1
DMD_iPS1
ρ: 0.95
CB CD34+
DMD_iPS1
ρ: 0.95
DMD_iPSCORE_65_1
BMD_iPS1
ρ: 0.94
CB CD34+
BMD_iPS1
ρ: 0.95
CB CD34+
DMD_iPSCORE_65_1

Figure 5 – figure supplement 3. Organoid derived myogenic progenitors and correlation to fetal muscle progenitors. (A) Gating strategy applied for the FACS sorting $CD29^+$ / $CXCR4^+$ cells from 15th -16th week skeletal muscle organoids **(B,C)** Re-plating $CD29^+$ / $CXCR4^+$ cells and culturing for 14 days highlights their fetal myogenic potential, illustrated by brightfield and immunocytochemistry pictures for Fast $MyHC^+$, $TITIN^+$ and $PAX7^+$ populations **(D)** *t*-SNE feature plots on ITGβ1 and CXCR4 demarcate the expression of FACS isolated calls into the myogenic progenitor subcluster **(E)** Gating strategy from **(A)** together with the unstained population (Yellow), Isotype Control (Red) and $CD29^+$ / $CXCR4^+$ (Gray) population **(F)** The correlation coefficient between the Log_2 fold change (Log_2 FC) values for isolated myogenic progenitors from human fetal tissue (17th week) and for FACS sorted $CXCR4^+$ / $ITGβ1^+$ organoid derived myogenic progenitors (16th week). PAX7, MYF5, MYOD1, MYOG are highlighted on the plot **(G)** The correlation coefficient between the Log_2 fold change (Log_2 FC) values for FACS sorted $CXCR4^+$ / $ITGβ1^+$ organoid derived myogenic progenitors (16th week) from CB $CD34^+$, DMD_IPS1, BMD_iPS1 and iPSCORE_65_1 pluripotent lines. Scale bars, 200uM **(B)**, 100uM **(C).**

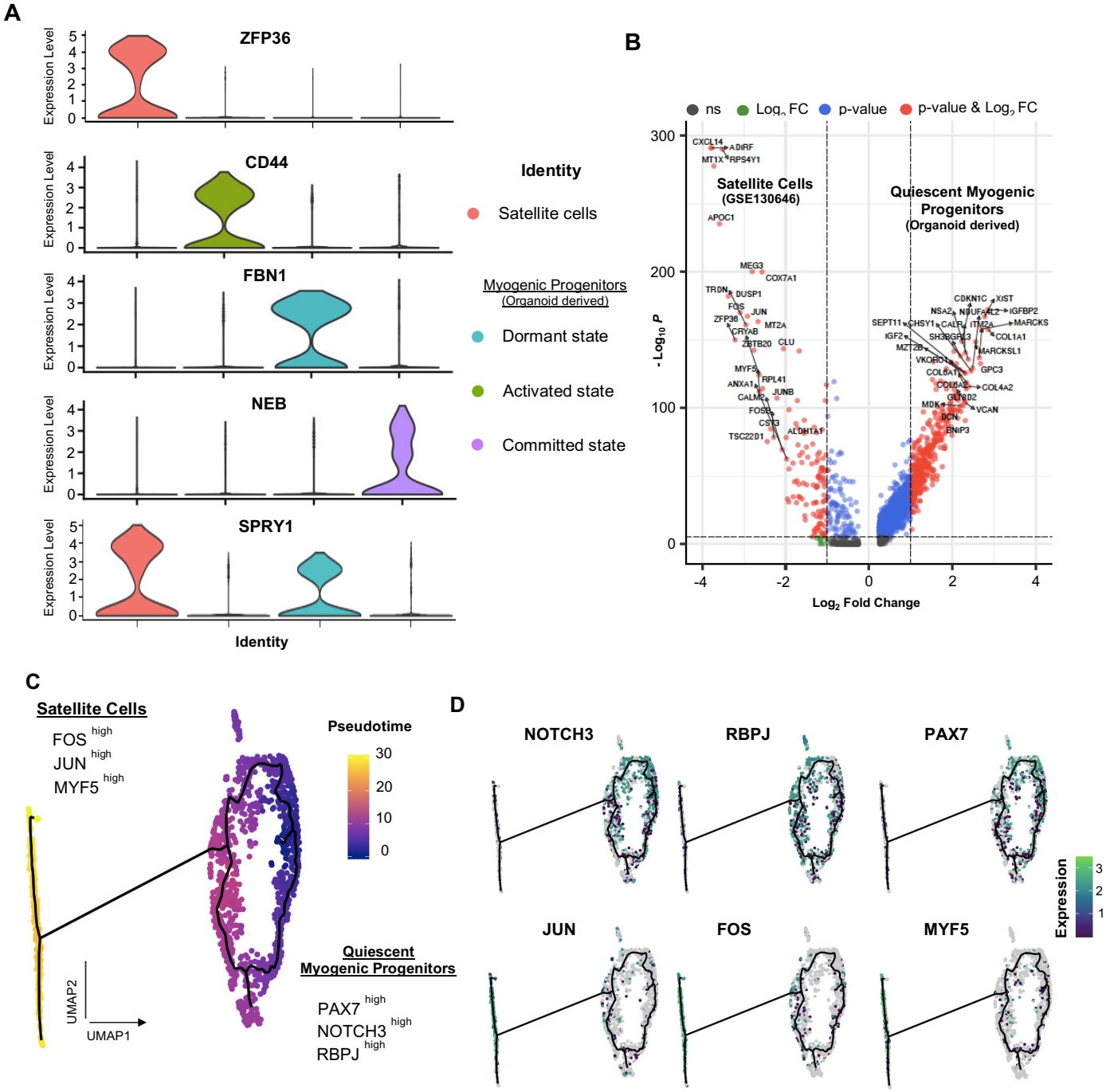
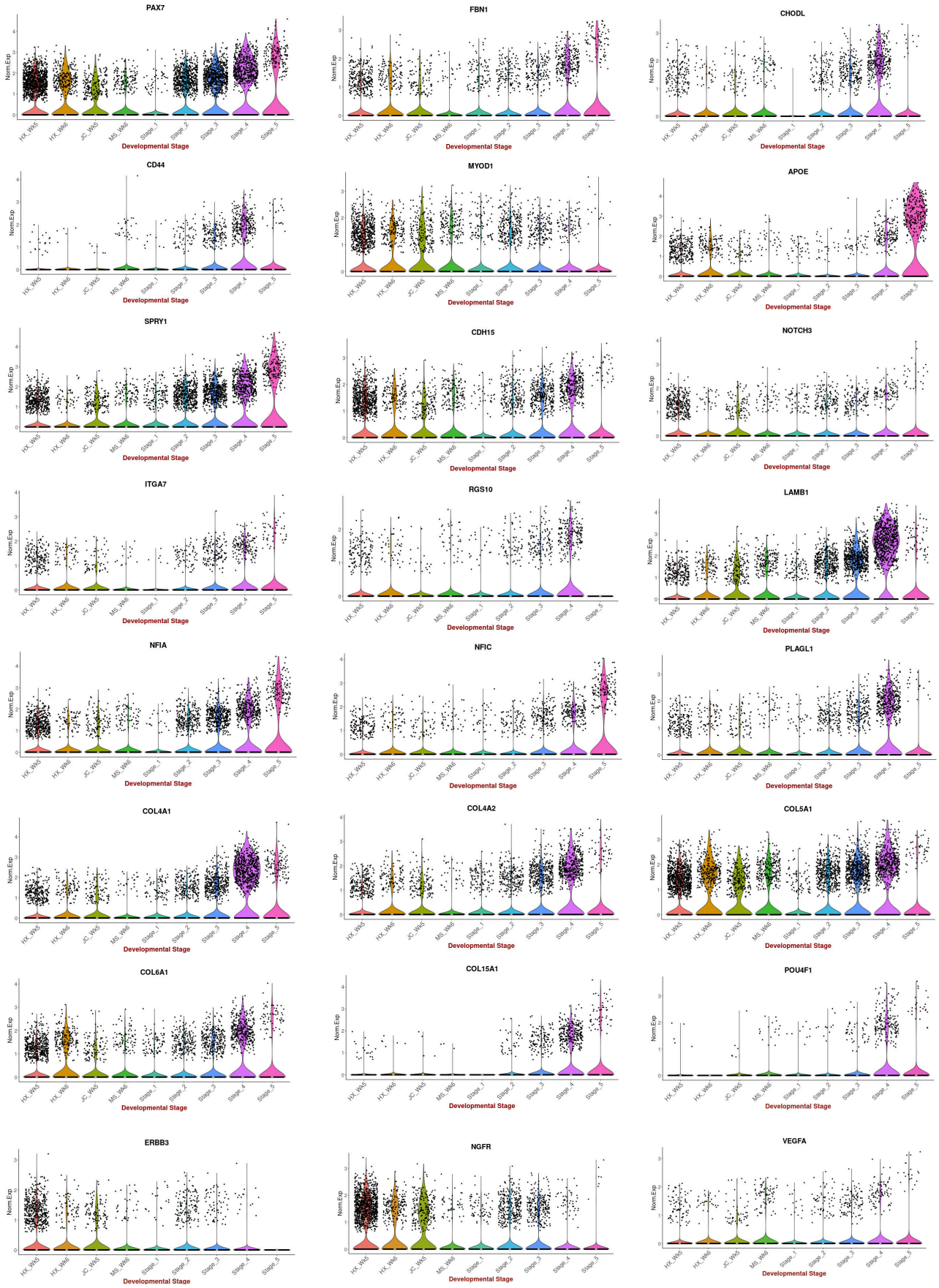

A
ZFP36
CD44
FBN1
NEB
SPRY1
Expression Level
Identity
Identity
Satellite cells
Myogenic Progenitors
(Organoid derived)
Dormant state
Activated state
Committed state
B
ns
Log2 FC
p-value
p-value & Log2 FC
Satellite Cells
(GSE130646)
Quiescent Myogenic
Progenitors
(Organoid derived)
- Log10 P
Log2 Fold Change
C
Satellite Cells
FOS high
JUN high
MYF5 high
Pseudotime
UMAP2
UMAP1
Quiescent
Myogenic Progenitors
PAX7 high
NOTCH3 high
RBPJ high
D
NOTCH3
RBPJ
PAX7
JUN
FOS
MYF5
Expression

Figure 5 – figure supplement 4. Organoid derived myogenic progenitors and correlation to adult human satellite cells. (A) Violin plots highlighting the developmental status for organoid derived myogenic progenitors and satellite cells by illustrating the relative expression of key signature markers for each subcluster such as ZFP36, CD44, FBN1 and NEB **(B)** Enhanced Volcano plot from comparing transcript levels between all cells from adult satellite cells (GSE130646) and organoid derived myogenic progenitors. Log2 fold-change in normalized gene expression versus -Log_{10} adjusted p-value is plotted. Differentially expressed genes with adjusted p-value < 0.05 are colored blue and genes with adjusted p-value > 0.05 and log2 fold-change > 1 are colored green. Genes with log2 fold-change > 1 and adjusted p value < 0.05 are labeled red. No significant genes (ns) are labeled with gray color **(C)** Pseudotime ordering for organoid derived myogenic progenitors and adult satellite cells together with expression of selected genes along pseudotime **(D)** Group of genes selected for dormant myogenic progenitors, PAX7, NOTCH3, RBPJ, and for activated satellite cells MYF5, JUN, FOS.


PAX7
FBN1
CHODL
CD44
MYOD1
APOE
SPRY1
CDH15
NOTCH3
ITGA7
RGS10
LAMB1
NFIA
NFIC
PLAGL1
COL4A1
COL4A2
COL5A1
COL6A1
COL15A1
POU4F1
ERBB3
NGFR
VEGFA
Norm.Exp
Developmental Stage
HX_Wk5
HX_Wk6
JC_Wk5
MS_Wk6
Stage_1
Stage_2
Stage_3
Stage_4
Stage_5

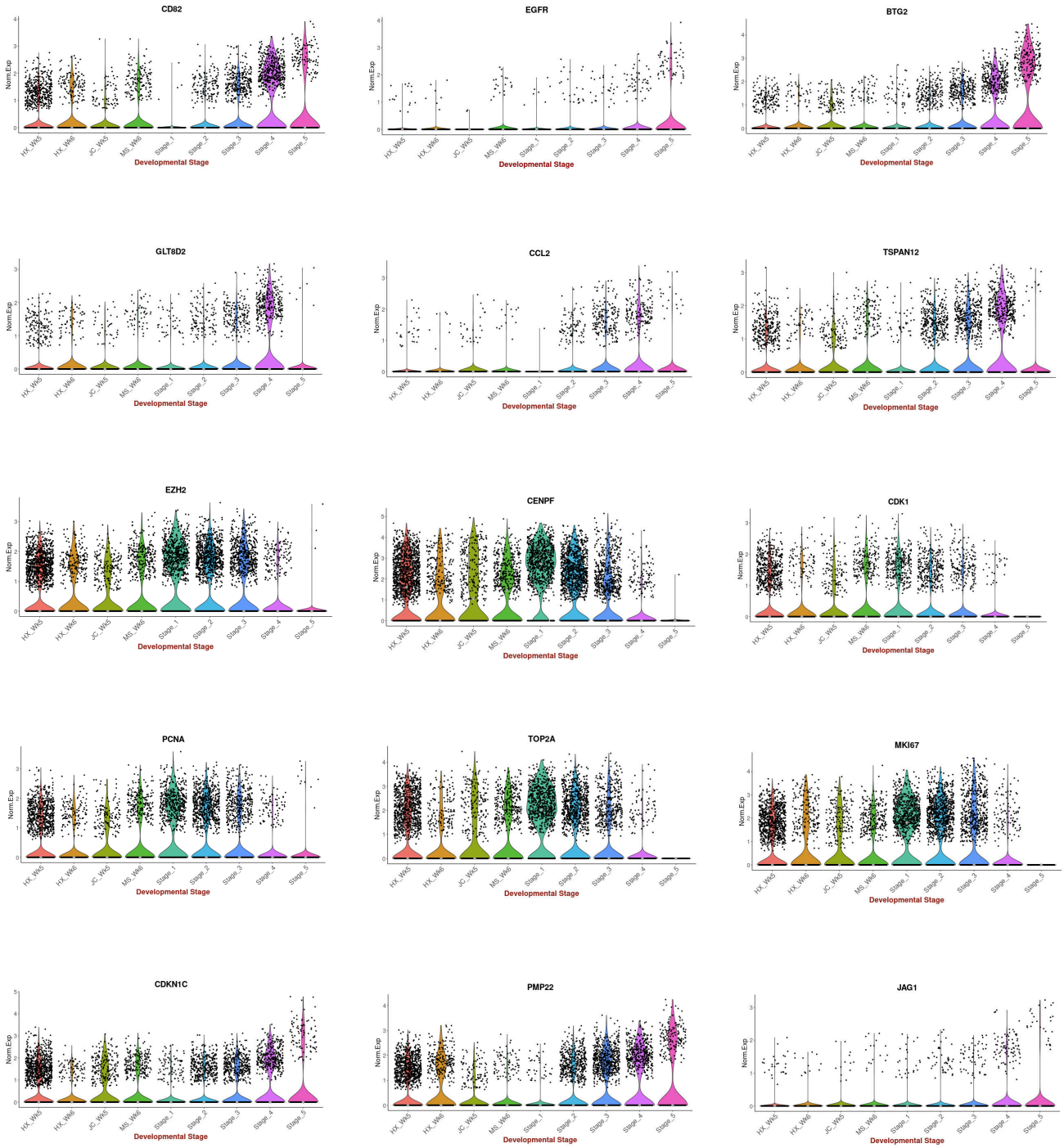

CD82
EGFR
BTG2
GLT8D2
CCL2
TSPAN12
EZH2
CENPF
CDK1
PCNA
TOP2A
MKI67
CDKN1C
PMP22
JAG1
Norm.Exp
HX_Wk5
HX_Wk6
JC_Wk5
MS_Wk6
Stage_1
Stage_2
Stage_3
Stage_4
Stage_5
Developmental Stage

Figure 5 – figure supplement 5. Comparison between In Vivo and In Vitro Multiple Protocols SMPCs and SCs. Violin plots highlighting expression of key markers associated with fetal development and are upregulated in organoid derived myogenic progenitors between *in vivo* myogenic progenitors and those derived from 2D differentiation protocols ((HX protocol, Xi et al., 2017; JC protocol, Chal et al., 2015; MS protocol, Shelton et al., 2014). Data was acquired from Pyle´s LAB UCLA, (https://aprilpylelab.com/datasets/in-vivo-and-in-vitro-multiple-protocols-smpcs-and-scs/) and published in (Xi et al., 2020).

Off -Target Site	Off-Target Sequence	PAM	Strand	MM	Forward detection Primer	Reverse detection Primer	Gene	Position
Chromosome 15: 42386225 - 42386247	TTTGATCA [TGGGGTATGACT]	CGG	-	0	TTCTCCTTCCCTGGGTTGAC	CTCGGCTGGATTTTGCAACC	CAPN3	Exonic
Chromosome 5: 125465023 -125465045	CTGAGTCA [TGGGGTATGACT]	AGG	-	4	CCCCAAGGAGTGCCAGAGTAA	TGGAGAAGCACTACCTTATGGC	RP11-756H20.1	Intergenic
Chromosome 14: 66857796 - 66857818	GTTGCTGC [TGGGGTATGACT]	GGG	-	4	CCTCTTCCTAAGCGAAACTGA	GTGGTTCCCCTCTGTGTAGGG	GPHN	Intronic
Chromosome 2: 167762690 - 167762712	TTAGAGTT [TGGGGTATGACT]	GGG	+	4	AGCAGGCCACTACTACCTCT	CAGGTCTATTCACCCTCTGGG	CTAGE14P	Intergenic
Chromosome 5: 33894317 - 33894339	GTGGATCA [AAGGGTATGACT]	GGG	-	4	GTATTCTGCCAGCCCCTTCCAA	TCTACGCTGAGGTAACTCCAAG	ADAMTS12	Intergenic
Chromosome 6: 87520187 - 87520209	GAAGATCA [TGGGATATGACT]	GGG	-	4	ACTAAGCCACTGAGAGCAGAT	TCCCTCCCTTGTAGAGCAGT	RARS2	Exonic
Chromosome 14: 57973592 - 57973614	TTGCATTA [TGGGATATGACT]	TGG	+	4	CTGTGGTTCTGTTTTCCTGCTG	GCAGTTTGGGCTTCAGGGAG	SLC35F4	Intronic
Chromosome X: 75062360 - 75062382	CTTCATCA [TAGAGTATGACT]	GGG	+	4	TGGAAAAAGGGGGATAGGCAT	TCTTCCCACCTGCTATGAAGT	ABCB7	Exonic
Chromosome 11: 133745438 - 133745460	TGTTCTCA [TGGGGAATGACT]	AGG	-	4	GGCAGGAGCAGGCAAGAATA	GTAGCAACCCAGGAGTACTGT	RP11-448P19.1	Intergenic
Chromosome 10: 105967626 - 105967648	TTTGATAT [AGAGGTATGACT]	TGG	+	4	AGTGGCCATGGGAGATTGTC	AGCACATAACGCATGCTCAG	N / A	Intergenic
Chromosome 4: 41570773 - 41570795	ATTGGCCA [TGGGGTGTGACT]	TGG	+	4	ATCCTGGTTTGGACCTGTGC	ACCCTAACGAACAGTCCTCCA	LIMCH1	Intronic

Table S1. Predicted off-target effects of the 5´-TTTGATCATGGGGTATGACT -3´ sgRNA sequence using ccTOP CRISPR/Cas9 target online predictor program

Gene	Forward Primer Sequence	Reverse Primer Sequence
OCT4	AGTGAGAGGCAACCTGGAGA	ACACTCGGACCACATCCTTC
SOX2	GCCCTGCAGTACAACTCCAT	GACTTGACCACCGAACCCAT
NANOG	CTGAGATGCCTCACACGGAG	GTGGGTTGTTTGCCTTTGGG
PAX6	GTGAGAGTGGACAGACATCCG	CTGTTCTGCATGCTGGCTCT
BRACHYURY	TTCATAGCGGTGACTGCTTATCA	CACCCCCATTGGGAGTACC
MESOGENIN	CAGGATGAGGACCTTGGCAG	GGATCTTGGTGAGAGGCTGG
TBX6	CCCTACTCGGCTGCATTTCT	GAGCCCACATCCAGATAGCC
HES7	GGAACCCGAAGCTGGAGAAA	CGGAAACCGGACAAGTAGCA
UNCX	GGAGAAGGCGTTCAACGAGA	GGAACCAGACCTGAACTCGG
TBX18	CCACCCCGTGTGTACATTCA	TGGCCTTGGTCATCCAGTTC
MEOX2	CTCTGCAAACCAACTGGCAC	AAGAGTTGGAGCACAGGACG
PAX3	AGACTGGCTCCATACGTCCT	CATGCCCGGGTTCTCTCTTT
PAX7	AACCACATCCGCCACAAGAT	CTCCTGGTAGCGGCAAAGAA
EN1	CAGGAACTCAGCCTCAACGA	ACTCGCTCTCGTCTTTGTCC
SIM1	GAGTGGTGTTCCCAGAAGGG	ATCCAGGGTCTGGAGCAGAT
LBX1	GCGGAGAAGTTACTCGCTGT	CCTTAAACGTCTTGCTGGCG
TFAP2A	CCAAGTCCAACAGCAATGCC	CGACCCGGAACTGAACAGAA
SOX10	CCATCCAGGCCCACTACAAG	GCTCTGGCCTGAGGGGT
RPS16 (Housekeeping)	GCTATCCGTCAGTCCATCTCCAA	CCTTCTTGGAAGCCTCATCCAC

Table S2. qPCR primer pairs applied to detect relative expression on key markers during skeletal muscle organoid development.