ACVR1^{R206H} increases osteogenic/ECM gene expression and impairs myofiber formation in human skeletal muscle stem cells

Emilie Barruet^{1,2,3*}, Steven M. Garcia², Jake Wu², Blanca M. Morales¹, Stanley Tamaki², Tania Moody¹, Jason H. Pomerantz^{2,3} and Edward C. Hsiao^{1*}

¹Division of Endocrinology and Metabolism, Department of Medicine, and the Institute for Human Genetics, University of California, San Francisco, CA 94143, U.S.A.

²Departments of Surgery and Orofacial Sciences, Division of Plastic and Reconstructive Surgery, Program in Craniofacial Biology, Eli and Edythe Broad Center of Regeneration Medicine, University of California San Francisco, 94143, USA.

³Department of Surgery, Division of Plastic and Reconstructive Surgery, University of California, San Francisco, 94143, USA.

Email addresses:

Emilie.Barruet@ucsf.edu

Edward.Hsiao@ucsf.edu

*Co-Corresponding Authors:

Emilie Barruet, PhD

Edward Hsiao, MD, PhD

Division of Endocrinology and Metabolism, Department of Medicine

Institute for Human Genetics

University of California at San Francisco

513 Parnassus Ave., HSE 901G

San Francisco, CA 94143-0794

Phone: 415-476-9732

Fax: 415-476-1356

26 **Abstract:**

27 Abnormalities in skeletal muscle repair lead to poor function and complications such as scarring
 28 or heterotopic ossification (HO). Here, we use fibrodysplasia ossificans progressiva (FOP), a
 29 disease of progressive HO caused by ACVR1^{R206H} (Activin receptor type-1 receptor) mutation, to
 30 elucidate how ACVR1 affects skeletal muscle repair. Rare and unique primary FOP human muscle
 31 stem cells (Hu-MuSCs) isolated from cadaveric skeletal muscle demonstrated increased ECM
 32 marker expression, and showed skeletal muscle-specific impaired engraftment and regeneration
 33 ability. Human induced pluripotent stem cell (iPSC)-derived muscle stem/progenitor cells
 34 (iMPCs) Single cell transcriptome analyses from FOP also revealed unusually increased ECM and
 35 osteogenic marker expression compared to control iMPCs. These results show that iMPCs can
 36 recapitulate many aspects of Hu-MuSCs for detailed in vitro study, that ACVR1 is a key regulator
 37 of Hu-MuSC function and skeletal muscle repair; and that ACVR1 activation in iMPCs or Hu-
 38 MuSCs contributes to HO by changing the local tissue environment.

39 **Introduction:**

40 Human diseases of skeletal muscle are major medical problems. Aberrant repair after muscle
41 injury results in scarring or heterotopic ossification (HO; bone formation in an inappropriate site)
42 which can be devastating. Disrupted signaling of bone morphogenetic proteins (BMPs), originally
43 identified by their ability to induce bone formation when injected into muscle (Urist, 1965),
44 changes muscle homeostasis (Ono et al., 2011) by controlling the proliferation and differentiation
45 of satellite cells (SCs) (Ono et al., 2011; Stantzou et al., 2017). SCs marked by PAX7 (Paired Box
46 7) (Seale et al., 2000), are a prerequisite for skeletal muscle regeneration (Mauro, 1961) and are
47 thought to be the main human muscle stem cells (Hu-MuSCs) of postnatal skeletal muscle. Upon
48 injury, activated SCs give rise to myoblasts, which form new myofibers or fuse to existing muscle
49 fibers to repair muscle damage (Kuang, Kuroda, Le Grand, & Rudnicki, 2007). A subset of SCs
50 does not differentiate and serves to replenish the SC pool. However, how the BMP pathway
51 regulates muscle repair or SC function in conditions of HO remains unclear and is a major
52 knowledge gap.

53 Recent protocols to create Hu-MuSCs-like cells from human induced pluripotent stem cells
54 (hiPSCs) (Takahashi et al., 2007) could generate skeletal myogenic lineage cells, but with limited
55 muscle regenerative capacity (Borchin, Chen, & Barberi, 2013; Magli et al., 2017; Shelton et al.,
56 2014; van der Wal et al., 2018; Xi et al., 2017). PAX7 cell engraftment capability was often not
57 reported or showed limited success. Transgene-induced *PAX7* or *PAX3* expression in hiPSCs can
58 increase engraftment of *PAX7*⁺ cells and contribution to the SC pool (Al Tanoury et al., 2020; Wu
59 et al., 2018); however, these engineered cells may not reflect physiology due to the genetic
60 manipulation of these master transcription factors.

61 BMP signaling is one pathway that could be manipulated to make Hu-MuSCs-like cells from
62 hiPSCs (iMPCs) (Chal et al., 2016). Although inhibiting the BMP pathway is an important step
63 that promotes iMPC formation, the BMP pathway is also critical for maintaining *PAX7* expression

in primary MuSCs and for preventing commitment to myogenic differentiation (Friedrichs et al., 2011). Abrogation of BMP signaling in SCs slowed myofiber growth (Stantzou et al., 2017), and increased BMP4 levels in Duchenne's muscular dystrophy (DMD) can exacerbate the disease (Shi, de Gorter, Hoogaars, t Hoen, & ten Dijke, 2013).

Fibrodysplasia ossificans progressiva (FOP), a congenital disease of abnormal skeletal muscle regeneration and severe HO, provides a unique opportunity to understand how changes in BMP signaling affect Hu-MuSC and iMPC formation and function. The FOP iMPCs and Hu-MuSCs carry the classical ACVR1^{R206H} (c.617>A) mutation (Shore et al., 2006) that causes hyperactivation of the BMP-SMAD signaling (Billings et al., 2008) and aberrant responses to Activin A (Hatsell et al., 2015). Primary FOP Hu-MuSCs can engraft and regenerate injured muscle of NSG mice, but at lower efficiency than Hu-MuSCs from non-FOP donors, and that the source of FOP Hu-MuSCs (biceps vs. diaphragm) appears to impact their engraftment efficiency. A new non-transgenic strategy for creating human iMPCs was developed and applied to multiple iPSC lines from control subjects and subjects with FOP. These iMPCs shared transcriptional similarities with primary Hu-MuSCs, and abnormal activation of the BMP pathway by ACVR1^{R206H} changed the transcription profiles of the FOP iMPCs compared to controls.

Results:

1- Lower efficiency engraftment of Primary Human FOP Hu-MuSCs

Muscle tissue samples were obtained from FOP cadavers which allowed the study of markers *in situ* and the isolation of primary human cells with endogenous activated ACVR1 signaling. Hematoxylin and eosin and alcian blue staining of FOP primary muscle samples from two deceased FOP subjects (**Figure 1-Source Data 1**) in a region without HO showed no gross defects. FOP muscle tissue near a HO lesion showed increased ECM proteoglycan components (alcian blue staining) in the interstitial space of the muscle fibers near the HO lesion (**Figure 1A**). Primary Hu-MuSCs carrying the ACVR1^{R206H} activating mutation were isolated from unfixed muscle tissue from two FOP autopsies. Biceps muscle, commonly affected by HO, and diaphragm muscle, one of the rare skeletal muscle sites spared from HO in patients with FOP, were analyzed. PAX7 staining confirmed the presence of SCs in human FOP biceps (**Figure 1B**, top and middle). Interestingly, a small subregion showed Collagen Type 1 expression with embedded satellite cells (**Figure 1B**, bottom) suggesting possible early HO formation despite absence of gross HO.

Sufficient Hu-MuSCs (**Figure 1-figure supplement 1A**) were sorted for transplant and gene expression analysis (**Figure 1-Source Data 2**). FOP Hu-MuSCs showed lower *PAX7* expression compared to control Hu-MuSCs. *COL1A1* and *ID3* were increased in affected FOP muscle (biceps) (**Figure 1C**). However, *ID1* was increased in both the FOP diaphragm and biceps compared to control muscles (**Figure 1C**). Five weeks after transplantation, biceps and diaphragm FOP Hu-MuSCs had engrafted and formed human fibers (**Figure 1D, E**). However, the number of human DYSTROPHIN-positive fibers was significantly lower with FOP Hu-MuSCs vs. control Hu-MuSCs (**Figure 1F**). The number of engrafted PAX7⁺ cells was qualitatively lower with FOP vs. control Hu-MuSCs (**Figure 1F** and **Figure 1-figure supplement 1B**). Ten weeks after re-injury with bupivacaine at week 5 (**Figure 1D**), the number of human DYSTROPHIN fibers was significantly decreased when FOP biceps Hu-MuSCs were transplanted compared to control Hu-

Figure 1: Regenerative properties of primary Hu-MuSCs isolated from FOP patients. (A) Hematoxylin and eosin staining of muscle cross sections from FOP subjects, with no heterotopic ossification (left) and with heterotopic ossification, depicted by the white dashed line (right, 100 μ m scale bar). (B) Immunofluorescence staining for PAX7 and COL1 of muscle cross sections from control and FOP subjects. White arrowheads mark SCs, 50 μ m scale bar. (C) Heat map of normalized gene expression of sorted human satellite cells from 2 control subjects and 2 different muscles of 1 FOP subject. (D) Schematic and experimental time course. SCs from biceps and diaphragm muscles of a deceased FOP patient were sorted and transplanted into NSG mice. (E) Immunofluorescence staining 5 weeks after xenotransplantation. White arrowheads mark SCs, 50 μ m scale bar. (F) Quantification of human DYSTROPHIN fibers and human PAX7⁺ cells 5 weeks after transplant. (G) Mice were re-injured 5 weeks after transplantation with bupivacaine. Immunofluorescence staining was at week 10. White arrows mark SCs, 50 μ m scale bar. (H) Quantification of human DYSTROPHIN fibers and human PAX7⁺ cells after re-injury at week 10. n=1, biological replicates, n $\geq$ 3 technical replicates. Error bar represent mean and SD. *, p<0.05, **, p<0.01. Muscle specimen details are listed in **Figure 1-Source Data 1**.

105 MuSCs, but not when FOP diaphragm Hu-MuSCs were transplanted (**Figure 1G, H**). No
 106 differences in the number of human PAX7 cells were identified (**Figure 1G, H**). No radiologic

evidence of HO was found in any mice (**Figure 1-figure supplement 1C**).

Thus, primary FOP Hu-MuSCs can engraft and regenerate injured muscle of NSG mice, but at lower efficiency than unaffected MuSCs, and the source of FOP Hu-MuSCs (biceps or diaphragm) may impact MuSC engraftment efficiency.

2- Human FOP iPSCs can differentiate into skeletal muscle cells

The rarity of FOP disease and difficulty obtaining human tissue samples from patients with FOP makes it difficult to obtain a reliable source of muscle stem cells. Therefore, we used established and fully characterized control hiPSCs (Wtc11, 1323-2, and BJ2) and ACVR1^{R206H} hiPSCs (F1-1, F2-3, F3-2) lines previously derived from patients with FOP. The control hiPSCs yielded PAX7 and MYOGENIN-expressing cells (**Figure 2A,B** and **Figure 2-figure supplement 1A,B**) as early as day 25 (**Figure 2-figure supplement 1A,B**), expressed DYSTROPHIN (**Figure 2B**), and formed contractile myotubes (**Figure 2C** and **Video1-2**).

Since BMPs control skeletal muscle differentiation from hiPSCs (Chal et al., 2016; Xi et al., 2017), we investigated if genetic activation of the BMP pathway via ACVR1^{R206H} could alter the myogenic differentiation of hiPSCs (F1-1, F2-3, F3-2) derived from patients with FOP (Matsumoto et al., 2013). All three FOP iPSC lines formed contractile myotubes with cells expressing DYSTROPHIN and PAX7 (**Figure 2C,D** and **Figure 2-figure supplement 1A,B**). *PAX7*, *MYOGENIN*, and *DYSTROPHIN* (**Figure 2E** and **Figure 2-figure supplement 1C**) gene

Figure 2: Skeletal muscle differentiation of human control and FOP iPSC lines. (A) Differentiation schematic of hiPSCs into skeletal muscle stem cell-like cells (iMPCs). (B) Immunofluorescence staining of iMPCs at day 50. White arrows show PAX7 expressing cells, 50µm scale bar. (C) iMPCs can form contractile myotubes, 100µm scale bar. (D) FOP iMPCs form DYSTROPHIN and PAX7 expressing cells (day 50). White arrows show PAX7 expressing cells, 50µm scale bar. (E) *PAX7* and *MYOGENIN* gene expression at day 25 and day 50 of differentiation (n=3 biological replicates and n≥3 technical replicates). Error bars represent mean ± SD. No statistically significant differences were detected. (F) Schematic describing the addition of the BMP pathway inhibitor LDN193189 (LDN). (G) Representative immunofluorescence staining of 2 control and 2 FOP iMPCs for MHC (Myosin Heavy Chain, green) and nuclear stain DAPI (blue) at day 50 +/- LDN addition, 200µm scale bar

126 expression showed some heterogeneity among the different hiPSC lines but these differences were
 127 not statistically significant. Adding a SMAD inhibitor of the BMP pathway (LDN193189) into the

differentiation protocol (**Figure 2F**) improved the differentiation of the control hiPSC lines. LDN also improved differentiation of the FOP F3-2 (**Figure 2G**) line, which had shown lower PAX7 and MYOGENIN expression (**Figure 2E**). These findings show that individual hiPSC lines are heterogeneous in differentiation to skeletal muscle lineages, similar to other protocols (Volpato & Webber, 2020), that FOP hiPSCs can form skeletal muscle cells and Hu-MuSC-like cells expressing PAX7 despite upregulation of the BMP pathway by ACVR1^{R206H}, and that chemical blockade of the BMP pathway can improve the formation of Hu-MuSC-like cells from the FOP iPSC line that showed the lowest efficiency.

3- PAX7 expressing cells can be isolated from myogenic differentiation

To test the regenerative properties of PAX7-expressing MuSCs, we used FACS to purify HNK1⁻ CD45⁻CD31⁻ cells co-expressing CD29, CXCR4, and CD56 markers present on human PAX7⁺ cells (Garcia et al., 2018) (**Figure 3A** and **Figure 3-figure supplement 1A,B**). FACS analysis identified intermediate CD56 cells expressing high PAX7 and low MYOGENIN (**Figure 3B** and **Figure 3-figure supplement 1C**), consistent with Hu-MuSC expression profiles. All HNK1⁻ CD45⁻CD31⁻ CXCR4⁺CD29⁺CD56^{dim} cells formed myotubes expressing MHC (**Figure 3-figure supplement 1D**) when cultured in terminal differentiation media demonstrating isolation of functional iMPCs with satellite cell characteristics from the cultures.

Figure 3: Isolation and transplantation of PAX7-expressing cells from iPSC muscle differentiation culture. hiPSCs were differentiated into iMPCs until day 50 and sorted via flow cytometry. **(A)** Gating strategy. **(B)** Myogenic gene expression of CD56^{dim} and CD56^{hi} cells (n=3 biological and technical replicates, *** $p < 0.001$, **** $p < 0.0001$ by two-way ANOVA test). **(C)** Representative human DYSTROPHIN (top, 200 μ m scale bar), and total LAMININ, human LAMIN A/C, human SPECTRIN, and PAX7 (bottom, 100 μ m scale bar) immunohistochemistry of NSG mice anterior tibialis, where sorted iMPCs were transplanted. White arrows show engrafted hiPSC-derived muscle stem cells. **(D)** Quantification of human DYSTROPHIN fibers and human PAX7 cells at week 5 after transplant (n=3 biological replicates).

4- *Isolated iMPCs can regenerate injured mouse muscle and form human fibers*

To assess iMPC regenerative capacity *in vivo*, we injected 1,000-10,000 iMPCs derived from control or FOP hiPSCs (**Figure 3-Source Data 1**) into the tibialis anterior (TA) muscle of whole-body irradiated NSG immunocompromised mice (to hinder endogenous satellite cells) previously injured with bupivacaine (Garcia, Tamaki, Xu, & Pomerantz, 2017). The bupivacaine step induces myofiber injury and promotes engraftment of donor SCs. New fibers expressing human DYSTROPHIN and PAX7 cells were found after 5 weeks, showing that iMPCs could engraft and promote muscle regeneration (**Figure 3C**). However, the number of human fibers and human PAX7⁺ cells remained low (**Figure 3D**) compared to primary Hu-MuSCs. While some iMPC transplants yielded up to 60 new human fibers, some did not yield any human fibers. By comparison, 2,000 primary non-FOP Hu-MuSCs resulted in an average of 155 human fibers based on prior assessments using the same assay (Garcia et al., 2018). No significant differences between control and FOP iMPCs were identified, though some individual FOP samples showed higher engraftment (**Figure 3D**).

These results showed that iMPCs can engraft into a muscle injury site in mice, but engraftment efficiency may be lower than primary Hu-MuSCs or be the result of differences in experimental conditions. Also, ACVR1^{R206H} did not significantly impact muscle fiber regeneration in this assay.

5- *Transcriptional profiling of iMPCs*

The lower engraftment of iMPCs compared to primary Hu-MuSCs suggested that the FACS-purified population was still heterogeneous or that iMPCs do not fully recapitulate adult primary Hu-MuSCs. Single cell RNA sequencing (scRNAseq) from control (1323-2) and FOP (F3-2) iMPCs (the lines were selected based on their lower intra-line variability, **Figure 2**) were analyzed (**Figure 4A; Figure 4-Source Data 1**). Cell populations for both samples were defined by the dimension reduction technique of uniform manifold approximation and projection (UMAP) (Becht

Figure 4: Single cell RNA sequencing of HNK1⁺CD45⁺CD31⁺CXCR4⁺CD29⁺CD56^{dim} cells. (A) Schematic of differentiating hiPSCs, sorting of iMPCs, and scRNAseq. (B) UMAP visualization plots of the control (2,051 cells) and FOP (5,306 cells) samples. Each dot represents one cell colored by cluster identification. The myogenic cell type was assigned according to expression of a combination of marker genes (Figure 4-figure supplement 1). (C) UMAP of cells combined from both control and FOP samples. (D) Feature expression plots showing the localization of cells expressing myogenic markers, mesenchymal, and neural cell marker. (E) Dot plot displaying expression genes associated with myogenesis, neurogenesis and mesenchymal markers. (F) Proportion of cells per cluster for each sample.

et al., 2018) and unsupervised clustering with Seurat v3 package (Stuart et al., 2019) (Figure 4B).

Both control (Figure 4-figure supplement 1A) and FOP (Figure 4-figure supplement 1B)

samples had clusters expressing myogenic genes (*PAX7* and *MYOD*); mesenchymal genes

(*PDGFRA*); and neuronal genes (*SOX2*). Merged analysis to allow direct comparison identified 13

distinct clusters (Figure 4C and Figure 4-figure supplement 1C). Clusters 0-2, 4-9, and 12

consisted of cells expressing neuronal stem/progenitor cell genes such as *EFNB3*, *SOX2*, and *ASCL1* (**Figure 4D,E** and **Figure 4-figure supplement 1D,E**). Mesenchymal genes (*PDGFRA*, *ASPN*, and *COL1A1*) were expressed in cluster 8 (**Figure 4D,E** and **Figure 4-figure supplement 1D,E**). Cells expressing muscle markers (*PAX7*, *MYF5*, *MYOD* and *MYOG* (*MYOGENIN*)) were found in clusters 3, 10, and 11 (**Figure 4D,E** and **Figure 4-figure supplement 1D,E**). The frequency of muscle cells (clusters 3, 10, 11) was higher in FOP (14%, 2.5%, and 1.7%) compared to controls (4.2%, 0.5%, and 0.5%) (**Figure 4F**) suggesting that FOP hiPSCs may be more efficient at making muscle progenitor cells. Thus, FACS-purified iMPC cultures contain muscle stem/progenitor cells, but other cell types such as mesenchymal, neuronal progenitor cells, and myoblasts persist in the culture. Furthermore, the frequency of iMPCs appears to be higher in FOP vs. control cell cultures.

6- FACS- sorted iMPC transcriptome is heterogeneous

Typical muscle progenitor cultures are expected to contain cells undergoing expansion, differentiation, and maturation. Sub-clustering (**Figure 5A,B**) identified 5 new myogenic subpopulations (**Figure 5B**). We ordered the myogenic cells into 3 major branches using the Monocle analysis package (Trapnell et al., 2014) and constructed pseudotime differentiation trajectories (**Figure 5C**). *PAX7* and *MYF5* were upregulated in branch B but downregulated in branches A and C (**Figure 5C,D**). *MYOD* and *MYOG* showed higher expression patterns early in branch A and late in branch C (**Figure 5C,D**). Quiescent iMPCs were at the intersection of the 3 branches (clusters 0, 1) while distal parts of the 3 branches, notably comprised of clusters 3 and 4, contained more mature cells with transcriptional profiles similar to activated MuSCs, progenitor cells, or myoblasts. Cluster 4 expressed higher levels of *MYOD*, *MYOG*, *SOX8*, and *MEF2C* (markers of differentiated MuSCs/myoblasts/myocytes), while clusters 0-3 expressed higher levels

Figure 5: Transcriptional profile of iMPCs. (A) Identified myogenic clusters (in red) were sub-clustered from the other clusters (neurogenic and mesenchymal) and re-analyzed. (B) UMAP of the new myogenic subclusters was generated. (C) Pseudotime trajectory plot generated via Monocle analysis depicting all myogenic clusters. (D) Gene plots displaying the expression of specific myogenic genes as a function of pseudotime. Arrows represents the direction and major branches of the pseudotime. Branches A-C are marked with letters along the trajectories. (E) Feature expression of cells expressing myogenic markers. (F) Genes expression associated with stemness, quiescence, and activation for the described Hu-MuSCs subsets. (G) Feature expression plots of additional top differentially expressed marker for each cluster. (H-I) Violin plots of myogenic (H) and chondro/osteogenic genes (I) that were significantly differentially expressed. Violin plot width depicts the larger probability density of cells expressing each particular gene at the indicated expression level. *, significantly different following differential expression testing using the Wilcoxon rank sum test per cluster. (J) SMAD and P38MAPK pathway markers significantly differentially expressed between control and FOP myogenic cells. *p* values for H-J are in **Figure 5-Source Data 1**.

201 of *PAX7* and *MYF5* (markers of more quiescent MuSCs) (**Figure 5E,F**). Quiescence (*SPRY1*,

HEY1, *HES1*) and cell cycle (*TOP2A* or *KI67*) markers were increased in clusters 0-2 and cluster 3 respectively (**Figure 5F,G** and **Figure 5-figure supplement 1A,B**). Cluster 3 had a higher proportion of cells in G2M and S phase (**Figure 5-figure supplement 1B**). The cell cycle distribution was similar in control **Figure 5: Transcriptional profile of iMPCs. (A)** Identified myogenic clusters (in red) were sub-clustered from the other clusters (neurogenic and mesenchymal) and re-analyzed. **(B)** UMAP of the new myogenic subclusters was generated. **(C)** Pseudotime trajectory plot generated via Monocle analysis depicting all myogenic clusters. **(D)** Gene plots displaying the expression of specific myogenic genes as a function of pseudotime. Arrows represents the direction and major branches of the pseudotime. Branches A-C are marked with letters along the trajectories. **(E)** Feature expression of cells expressing myogenic markers. **(F)** Genes expression associated with stemness, quiescence, and activation for the described Hu-MuSCs subsets. **(G)** Feature expression plots of additional top differentially expressed marker for each cluster. **(H-I)** Violin plots of myogenic **(H)** and chondro/osteogenic genes **(I)** that were significantly differentially expressed. Violin plot width depicts the larger probability density of cells expressing each particular gene at the indicated expression level. *, significantly different following differential expression testing using the Wilcoxon rank sum test per cluster. **(J)** SMAD and P38MAPK pathway markers significantly differentially expressed between control and FOP myogenic cells. *p* values for **H-J** are in **Figure 5-Source Data 1**.

and FOP (**Figure 5-figure supplement 1C**). Cells expressing *APOE* and *KRT17* defined cluster 0. *CDH15*, a niche regulator of SCs quiescence (Goel, Rieder, Arnold, Radice, & Krauss, 2017) was enriched in cluster 1. Cluster 2 consisted of cells expressing high levels of *FOS* (**Figure 5F,G** and **Figure 5-figure supplement 1A**).

The proportion of cells in clusters 3 and 4 were similar in both control (10% and 12.7%) and FOP (14.3% and 9%) samples. The proportion of cells in clusters 0 and 3 was higher in FOP (31.4%

and 23.4%) compared to control (20% and 5.5%), while the proportion of cells in cluster 2 was increased in control (51.8% vs 21.9%) (**Figure 5-figure supplement 1D**).

Thus, hiPSC differentiation cultures contain subpopulations of iMPCs showing the expected spectrum of quiescence, activation, and differentiation with FOP cultures having a higher proportion of cells in the stem cell/progenitor and proliferating phases and fewer mature myoblasts.

7- *FOP iMPCs cells express increased chondro/osteogenic and ECM markers*

Differential expression analysis on the transcriptional profiles of the sub-clustered myogenic cells (**Figure 5A,B** and **Figure 5-figure supplement 1E**) was used to see if ACVR1^{R206H} altered transcriptional signatures. While *PAX7* was significantly increased in control vs. FOP cells in cluster 3 only (**Figure 5H**), *MYF5* was significantly increased in FOP cells from cluster 1 compared to control cells (**Figure 5-Source Data 1**). Since Hu-MuSCs show heterogeneous levels of *PAX7* and *MYF5* expression (Kuang et al., 2007), this suggests the ACVR1^{R206H} mutation may favor one sub-population over another. Interestingly, *SPRY1*, a known regulator of quiescence (Shea et al., 2010) which decreases with age (Bigot et al., 2015), was significantly downregulated in FOP cells of clusters 0-4, while *DLK1*, which act as a muscle regeneration inhibitor (Andersen et al., 2013) was significantly increased in FOP cells (**Figure 5H** and **Figure 5-Source Data 1**).

Since the ACVR1^{R206H} mutation increases BMP pathway activity and expression of chondrogenic and osteogenic markers in multiple lineages (Barruet et al., 2016; Culbert et al., 2014; Matsumoto et al., 2013), we examined the expression levels of extracellular matrix, fibrogenic, chondrogenic, and osteogenic genes. *ITM2A*, *COL1A1*, *DCN*, *CD15*, *SPARC*, and *TIMP1* were significantly increased in FOP cells (**Figure 5I** and **Figure 1-Source Data 1**). ECM proteoglycans known to be involved in inflammation, including *BGN* (Nastase, Young, & Schaefer, 2012) and *LUM* (Nikitovic, Papoutsidakis, Karamanos, & Tzanakakis, 2014), *TAGLN*

[regulates osteogenic differentiation (Elsafadi et al., 2016)], and *IGBP5* [increased in aged satellite cells (Soriano-Arroquia, McCormick, Molloy, McArdle, & Goljanek-Whysall, 2016)], were also increased in FOP cells (**Figure 5-figure supplement 1F** and **Figure 1-Source Data 1**).

Finally, gene expression of target genes of the BMP pathway (*ID1*, *ID3*, *BMPs*, and *SMADs*) and the p38MAPK pathway (**Figure 5J** and **Figure 5-figure supplement 1G**) was assessed to see if activated ACVR1 altered these pathways. *ID1* was significantly higher (clusters 0, 1, 3, and 4) in control cells while *ID3* was significantly higher in FOP cells (clusters 1, 2). In addition, the BMP/TGF β pathway downstream target gene *SMAD5* was significantly higher in cluster 2 of control cells. The p38 pathway components *JUN* (clusters 1, 3), *JUNB* (clusters 0-2), and *FOS* (clusters 1, 3) were significantly increased in FOP cells. Within the known ACVR1 co-receptors, *BMP2* (clusters 3, 4) and *ACVR2B* (clusters 0, 2, 3) expression were significantly higher in control cells while *ACVR2A* expression was significantly higher in FOP cells (cluster 1, 4). *ACVR2B* was significantly higher in control cells (cluster 0, 1, 4) (**Figure 5J**). Similar to the primary FOP Hu-MuSCs (**Figure 1**), these results suggest that FOP iMPCs have a chondrogenic/osteogenic signature, increased *ID3* expression, and also showed higher p38 pathway activity and higher levels of the *ACVR2A* co-receptor at different stages of myogenic differentiation.

8- iMPC transcriptome shows similarities to primary Hu-MuSCs

Comparing the iMPC scRNAseq to primary Hu-MuSCs data of sorted satellite cells from a human vastus lateralis muscle (Barruet et al., 2020) was used to identify if their lower engraftment efficiency was due to transcriptional differences. The merged data UMAP (**Figure 6A**) showed that clusters 0, 1, 4, 6, and 7 contained myogenic cells (**Figure 6B,C**). *PAX7*⁺ cells were identified in clusters 0, 1, 4, and 7 (**Figure 6C**). Myocyte contaminants present in the primary sorted cells constituted cluster 6. Although iMPCs expressed *PAX7* and *MYF5*, the gene expression levels were

higher in primary Hu-MuSCs. In contrast, *MYOD* was higher in iMPCs (**Figure 6C** and **Figure 6-figure supplement 1**).

Detailed analysis of the myogenic cell subset (cluster 0, 1, 4, 6 and 7, **Figure 6D**, cells labeled in red, and **Figure 6E**) was done using pseudotime trajectory analysis to elucidate the states of the iMPCs with respect to primary Hu-MuSCs. Cells from clusters 0, 1, and 4 were distributed along branches B and C. Myocytes ordered at the distal end of branch A. iMPCs ordered away from primary Hu-MuSCs (*PAX7*⁺) and primary myocytes (*MYL1*⁺) (**Figure 6F-H**). Branch expression analysis modeling (BEAM) allowed us to investigate significant gene that are branch-dependent in their expression (**Figure 6-Source Data 1**). Branch B, consisting of iMPCs and subset of primary Hu-MuSCs, expressed significantly higher levels of genes associated with mesenchymal, fibrogenic, chondrogenic, osteogenic lineages and extracellular matrix (**Figure 6H,I**). Thus, iMPCs retained strong bi-potency compared to primary Hu-MuSCs, suggesting that iMPCs may not be as committed to the muscle lineage as primary adult muscle stem cells (Xi et al., 2020).

Figure 6: iMPC transcriptional signature compared to human primary muscle stem cells. (A-B) UMAP of cells combined from human primary muscle stem cells (vastus muscle) and the control and FOP samples. (A) UMAP showing the distribution of cells per sample, and (B) with clusters labeled. (C) Feature expression plots of cells expressing myogenic markers. (D) Myogenic cells (in red) are comprised in cluster 0, 1, 4, 6, and 7. (E) UMAP of the myogenic clusters only used for the pseudotime analysis. (F-G) Pseudotime trajectory plot generated via Monocle analysis depicting all myogenic clusters (F) and samples (G). Branches are annotated A-C in (F). (H) Level of expression of ECM/osteogenic genes along the cell trajectories. (I) Heatmap representing genes that are significantly branch dependent using the BEAM analysis (Figure 6-Source Data 1) and also genes that have similar lineage-dependent expression patterns. Cell fate (branches) are shown in the lower right panel.

Discussion:

Developing optimal strategies for skeletal muscle regeneration and repair requires a detailed understanding of how these processes are regulated. Especially for diseases like FOP where muscle injury can trigger disease progression or severe complications, generating large numbers of human iMPCs using transgene-free protocols, holds promise for understanding pathological mechanisms and finding new therapeutic targets. This study revealed several novel findings, including deficiencies in the muscle repair capacity of FOP Hu-MuSCs, abnormal production of extracellular matrix components by iMSCs that likely contribute to FOP disease progression, and important differences between primary Hu-MuSCs and iMPCs that may impact other studies. In addition, our results suggest that there are muscle specific differences in Hu-MuSC repair capacity, potentially explaining why some skeletal muscles appear to be protected from developing HO.

The iMPC and primary Hu-MuSC single cell transcriptomes revealed clear subsets of cells at all stages of muscle differentiation (quiescent, activated, and differentiated). Satellite cell subtype markers [e.g. *COLs*, *DLK1*, *ID3*, and *HES1*, (Barruet et al., 2020)] were also highly expressed in iMPCs suggesting that *in vitro* differentiation may favor specific subtypes of satellite cells. iMPCs also expressed high levels of mesenchymal/ECM markers, similar to that in human fetal muscle progenitor cells (Xi et al., 2020). These results suggest that iMPCs may retain a more progenitor-like phenotype as compared to Hu-MuSCs. In addition, we showed that PAX7⁺ iMPCs can engraft in muscle injury models in mice, despite heterogeneity in the cultures and lower expression of typical satellite cells markers (e.g. *PAX7*, *MYF5*).

This iMPC system revealed that activation of the BMP pathway by the FOP ACVR1^{R206H} mutation induced changes that could contribute to abnormal muscle healing. As expected, the classical SMAD pathways were active in FOP cells; however, it was unexpected to find that the p38 pathway via JUN/FOS was more active in the FOP iMPCs, considering that FOP cultures had a higher proportion of cells in the stem cell/progenitor and proliferating phases. While increased

SMAD activity (Billings et al., 2008) or misinterpretation of the Activin A ligand (Hatsell et al., 2015) occur with ACVR1^{R206H} mutation, abnormal p38 signaling may be critical in some cells like macrophages (Barruet et al., 2018). Since p38 is a major regulator of Hu-MuSC function (Segales, Perdiguero, & Munoz-Canoves, 2016) and has been associated with inflammation and ECM accumulation in aged muscle (Cosgrove et al., 2014), further studies are needed to determine how this p38 signaling contributes to satellite cell subtype specification (Barruet et al., 2020) and affect healing in FOP.

Importantly, our iMPC and Hu-MuSC transplant studies showed no HO in the recipient mice, consistent with prior FOP genetic studies (Dey et al., 2016; Lees-Shepard et al., 2018). However, FOP iMPCs showed increased chondrogenic/osteogenic and ECM gene expression, a feature seen in other bone-related cell types in FOP (Barruet et al., 2016; Culbert et al., 2014; Lees-Shepard et al., 2018). Thus, FOP Hu-MuSCs may contribute to HO formation indirectly by modulation of the osteogenic environment such as contributing to changes in muscle stiffness, as previously reported in mice (Stanley, Heo, Mauck, Mourkioti, & Shore, 2019). In addition, while our results suggest no major differences or possibly a slight increase in the ability to form skeletal muscle progenitors, activated ACVR1 by the R206H mutation decreased (but did not abrogate) in vitro formation of mature myoblasts and in vivo muscle repair after transplant. This impaired formation of mature skeletal muscle likely contributes to the abnormal skeletal muscle healing in FOP and may predispose patients with FOP to HO formation.

Comparing primary FOP Hu-MuSCs from muscles that develop HO (biceps) and non-affected muscle (diaphragm) suggests that the source of the Hu-MuSCs impacts engraftment efficiency. The re-injury model showed that engraftment of Hu-MuSCs from FOP biceps, but not from diaphragm, remained significantly decreased. Thus, it is intriguing to consider that the clinical sparing of the diaphragm from HO in patients with FOP may result from a less impaired or unimpaired muscle repair process in diaphragm satellite cells. Further delineation of muscle-

specific Hu-MuSC properties will be revealing to understand this observation. Our finding that primary FOP Hu-MuSCs have lower engraftment ability provides a potential explanation for the poor skeletal muscle repair observed in patients with FOP (Shore, 2012). However, the FOP and control iMPCs showed no major differences in engraftment, possibly due to decreased assay sensitivity from the lower engraftment efficiency of iMPCs in general, because iMPCs may be more immature than primary Hu-MuSCs, or that iMPCs represent a subtype of cells that may be more reflective of non-ossifying skeletal muscle like diaphragm.

Figure 7. Testing the role of activated ACVR1 signaling in muscle repair. FOP iMPCs and FOP primary human muscle stem cells are used for modeling muscle stem cell engraftment and regeneration properties.

This study has several limitations. Our transplant experiments used immunocompromised mice where mature B and T cells are absent and macrophages are defective (Shultz et al., 2005). This may dampen the engraftment/regenerative phenotypes, particularly as immune cells are important in muscle regeneration (Furrer & Handschin, 2017) and there is growing awareness of the role of the immune system in the pathogenesis of FOP (Barruet et al., 2018; Convente et al., 2018). Also, while our FACS strategy can isolate Hu-MuSCs and iMPCs without the need of transcription factors or markers, we noted that the engraftment efficiency still varied among the iMPC lines

despite FACS purification. This heterogeneity is an ongoing problem among all published iMPC protocols to date. In our case, this was counter-balanced by our consistent findings across multiple lines and between iMPCs and the availability of primary Hu-MuSCs. Although our primary cell studies were limited by the rarity of FOP (estimated at 1 in 1.4 million people) and even rarer suitable cadaveric samples, the combination of iPSC-derived lineages with the rare primary samples provided multiple avenues for supporting our conclusions. Finally, the ACVR1^{R206H} mutation increases BMP signaling but also introduces neofunction to Activin A. Our studies are not able to distinguish between these two contributing pathways. Our finding of increased p38 activity in the FOP iMPCs was also seen in FOP subject monocyte-derived macrophages (Barruet et al., 2018), suggesting that this alternate signaling pathway by ACVR1 should also be investigated further. Future studies will help elucidate the different factors that contribute to the hiPSC-line specific effects, including potential roles for disease modifier genes or BMP pathway modulators.

This study shows that human iPSC-derived muscle stem cells can be a valuable tool to model musculoskeletal diseases of skeletal muscle injury and repair. Correlating the findings in iMPCs with primary Hu-MuSCs revealed an indirect role for skeletal muscle progenitors in HO formation, as well as subtypes of Hu-MuSCs that may contribute to skeletal muscle specific regenerative capacity. These studies highlight the importance of skeletal muscle regeneration in disease pathogenesis and establish a foundation for understanding how skeletal muscle repair and osteogenesis are linked.

389 Materials and Methods:

Key Resources Table				
Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
NSG mice	NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ	https://www.jax.org/strain/005557	005557	8-12 week-old
Sequenced-based reagent	Human RT-PCR Primers	Applied Biosystems Taqman Assays	B-ACTIN Hs01060665_g1	
Sequenced-based reagent	Human RT-PCR Primers	Applied Biosystems Taqman Assays	DYSTROPHIN Hs007758098_m1	
Sequenced-based reagent	Human RT-PCR Primers	Applied Biosystems Taqman Assays	PAX7 Hs00242962_m1	
Sequenced-based reagent	Human RT-PCR Primers	Applied Biosystems Taqman Assays	CD56 HS00941830_m1	
Sequenced-based reagent	Human RT-PCR Primers	Applied Biosystems Taqman Assays	MYOGENIN Hs01072232_m1	
Sequenced-based reagent	Human RT-PCR Primers	Applied Biosystems Taqman Assays	COL1A1 Hs01076780_g1	
Sequenced-based reagent	Human RT-PCR Primers	Applied Biosystems Taqman Assays	ID1 Hs03676575_s1	
Sequenced-based reagent	Human RT-PCR Primers	Applied Biosystems Taqman Assays	ID3 Hs00954037_g1	
antibody	Mouse monoclonal anti-Human DYSTROPHIN	DSHB	MANDYS104(7F7)	IF(1:10)
antibody	Mouse monoclonal anti-Human/Mouse DYSTROPHIN	Thermofisher	PA5-16734	IF(1:500)
antibody	Mouse monoclonal anti-Human/Mouse PAX7	DSHB	PAX7	IF(1:10)
antibody	Mouse monoclonal anti-Human/Mouse MHC	DSHB	MF20	IF(1:100)

antibody	Rabbit polyclonal anti-LAMININ	Sigma-Aldrich	L9393	IF(1:250)
antibody	Mouse monoclonal anti-Human SPECTRIN	Leica Microsystems	NCL-SPEC1	IF(1:100)
antibody	Mouse monoclonal anti-Human LAMIN A/C	Vector Laboratories	VP-L550	IF(1:100)
antibody	Mouse monoclonal anti-Human CD31 (Beads)	Miltenyi Biotec	130-091-935	FACS
antibody	Mouse monoclonal anti-Human CD45 (Beads)	Miltenyi Biotec	130-045-801	FACS
antibody	Mouse monoclonal anti-Human CD31 AF450 (WM-59)	Ebioscience	48-0319-42	FACS
antibody	Mouse monoclonal anti-Human CD34 eFluor450 (4H11)	Ebioscience	48-0349-42	FACS
antibody	Mouse monoclonal anti-Human CD45 AF450 (30-F11)	Ebioscience	48-0451-82	FACS
antibody	Mouse monoclonal anti-Human CD29 FITC (TS2/16)	Ebioscience	11-0299-41	FACS
antibody	Recombinant human anti-CD56 APC-vio-770 (REA196)	Miltenyi Biotec	130-114-548	FACS
antibody	Mouse monoclonal anti-Human CXCR4 PE (12G5)	Ebioscience	12-9999-41	FACS
antibody	Mouse monoclonal anti-Human HNK1 PE (TB01)	Ebioscience	12-057742	FACS
antibody	Mouse monoclonal anti-Human CD45 PE (30-F11)	Ebioscience	12-04551-82	FACS
antibody	Mouse monoclonal anti-Human CD31 PE (390)	Ebioscience	12-0311-82	FACS

antibody	Mouse monoclonal anti-Human CD56 APC (CMSSB)	Ebioscience	47-0567-42	FACS
antibody	Mouse monoclonal anti-Human CXCR4 PE-Cy7 (12G5)	Ebioscience	25-9999-42	FACS
antibody	FcR block	Miltenyi Biotec	130-059-901	
software, algorithm	GraphPad Prism	GraphPad Prism (https://graphpad.com)		
software, algorithm	Seurat (3.1.5)	https://satijalab.org/seurat/		
software, algorithm	Monocle (2.12.0)	http://cole-trapnell-lab.github.io/monocle-release/		
software, algorithm	cellranger	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/feature-bc		
software, algorithm	FlowJo	https://www.flowjo.com		

390

391 ***Cell Culture and Differentiation***

392 Pluripotent hiPSC lines derived from control and FOP fibroblasts (Matsumoto et al., 2013; Spencer
393 et al., 2014) were cultured in mTeSR1 medium (StemCell Technologies) on irradiated SNL feeder
394 cells (McMahon & Bradley, 1990) as described previously. hiPSCs were passaged at least once on
395 Matrigel (Corning)-coated plates (150-300µg/ml) to remove the SNLs before use in differentiation
396 assays. ROCK inhibitor Y-27632 (10µM, StemCell Technologies) was added to mTeSR1 when
397 cells were split and removed the following day.

398 hiPSC lines were differentiated into skeletal muscle cells using modifications based on prior
399 protocols (Chal et al., 2016; Shelton et al., 2014). Our hiPSC lines differentiated better with a
400 lower cell number seeding and a longer time of recovery between the seeding and the start of the

differentiation (2 days, data not shown). Cells were seeded at 7.5×10^5 cells per well of a 12-well plate on Matrigel two days before the differentiation medium (E6 medium, supplemented with either 10 μ M CHIR99021 (Tocris) for 2 days or with 3 μ M CHIR99021 and 0.5 μ M LDN193189 for six days). Cells were then grown in un-supplemented E6 media until day 12, then changed to StemPro-34 media supplemented with 10ng/ml bFGF until day 20. The medium was then replaced by E6 medium until day 35, when DMEM/F12 supplemented with N2 (Gibco) and Insulin-Transferrin-Selenium (ITS-A, 100X Gibco) was added. Media was changed daily until harvest at day 50. After sorting, cells were plated in satellite cell media [DMEM/F12 (Gibco), 20% FBS (Hyclone), 1X ITS (Gibco), 1X Penicillin/ Streptomycin (Gibco)] for further functional assays (**Figure 2A,F**). Once cells reached confluence, cells were cultured in differentiation media (DMEM, 2% horse serum, 1X Penicillin/ Streptomycin; Gibco).

Flow cytometry

HNK1⁻CD45⁻CD31⁻CXCR4⁺CD29⁺CD56^{dim} cells were sorted from skeletal muscle differentiation of control and FOP hiPSCs (described in detail in Supplemental Experimental Procedures). Human primary satellite cells were isolated and sorted as described (Garcia et al., 2018; Garcia et al., 2017). Human muscle was freshly harvested and stored in DMEM with 30% FBS at 4°C overnight or for 2 extra days (delay due to shipping). Muscle samples were digested, erythrocytes were lysed, and hematopoietic and endothelial cells were depleted with magnetic column depletion using CD31, CD34, and CD45 (eBioscience). Cells were further gated as described in **Figure 1-figure supplement 1A** and sorted for CXCR4⁺/CD29⁺/CD56⁺ and collected for subsequent experimentation.

Animal care and Transplantation studies

All mouse studies were performed using protocols approved by the UCSF Institutional Animal Care and Use Committee. Mice were either bred and housed in a pathogen-free facility at UCSF or purchased from The Jackson Laboratory. 8-12 week-old NSG (NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ) mice were randomized to all experimental groups by sex and littermates. The TA of each mice was irradiated with 18 Gy before transplantation. Isolated primary human satellite cells (Hu-MuSCs) or iMPCs were injected with 50 μ l 0.5% bupivacaine directly into the TA muscle of one leg as described (Garcia et al., 2018; Garcia et al., 2017) and summarized in the Supplemental Experimental Procedures. The TA for each mouse was harvested at week 5 or week 10 after transplantation and frozen in O.C.T. compound in 2-methylbutane chilled in liquid nitrogen. Serial 6 μ m transverse frozen sections were analyzed or stored at -80°C.

Cell Immunostaining and NSG Tibialis Anterior Analysis

iMPCs were fixed with 4%PFA/PBS for 10min at room temperature, permeabilized with 0.1%Triton-100X (Sigma-Aldrich), and blocked with 5% BSA (Sigma-Aldrich). Cells were stained overnight with primary antibodies for PAX7, MYOGENIN, DYSTROPHIN, and MHC. Cells were then incubated for 1 hr at room temperature in the dark with secondary antibodies Alexa488-conjugated goat anti-mouse IgG and Alexa546-conjugated goat anti-mouse IgG (Invitrogen). Nuclei were stained with DAPI (Sigma-Aldrich). Images were taken on a Nikon Eclipse E800 or Leica DMI 4000B.

Immunohistochemistry and Immunofluorescence of Human Muscle Samples

Human muscle samples were fixed in neutral buffered formalin for 24 h and then placed in 70% ethanol for at least 24 h. The sample with heterotopic bone was decalcified in 10% EDTA (pH 7.2-7.4) before paraffin embedding and sectioning. Sections were stained with hematoxylin and eosin

(J. David Gladstone Institutes Histology Core) or for alcian blue (pH 1.0) for cartilage and nuclear red stain for nuclei.

Freshly harvested human muscle was stored in DMEM with 30% FBS at 4°C, or snap in frozen in O.C.T. compound in 2-methylbutane chilled in liquid nitrogen. Serial 6µm transverse frozen sections were analyzed or stored at -80°C and processed similarly to the mouse TA samples above. Sections were stained with PAX7 (DSHB) and mouse monoclonal anti-Collagen Type I (Millipore-Sigma). Details about specimens are in **Figure 1-Data Source 1**.

RT-PCR and Quantitative Analysis

Tissues were collected in TRI Reagent (Sigma-Aldrich) to isolate total RNA using the ArcturusTM PicoPureTM RNA kit (Applied Biosystems) as previously described for small samples (Schepers, Hsiao, Garg, Scott, & Passegue, 2012). 0.2 to 0.5 µg of RNA were transcribed into cDNA with VeriScript cDNA synthesis kit (Affymetrix). cDNA was then pre-amplified with GE PreAmp Master Mix (Fluidigm Inc). Real-time quantitative PCR was performed in triplicated with either VeriQuest Probe qPCR Master Mix (Affymetrix) or Taqman Universal PCR Master Mix (Life Technologies) on either a Viia7 thermocycler (Life Technologies) or on a BioMark 48.48 dynamic array nanofluidic chip (Fluidigm, Inc) according to manufacturers' instructions. Beta actin was used for normalization as endogenous control.

Single cell RNA Sequencing and Analysis

scRNAseq was performed using the Chromium Single Cell 3' Reagent Version 2 Kit from 10X Genomics. 45,000 (FOP) and 30,000 (control) HNK1⁻CD45⁻CD31⁻CXCR4⁺CD29⁺CD56^{dim} cells isolated from iMPC differentiations were analyzed, using the sequencing procedures and analytical strategies described in Supplemental Experimental Procedures.

Statistical analysis

The data were analyzed utilizing GraphPad Prism v.7 software (GraphPad) using one-way (transplant) and two-way ANOVA with post hoc Tukey's or Sidak's multiple comparison test (gene expression). The Sidak test was used when comparing means between control and FOP, and the Tukey test was used when means of both control and FOP were compared together with other groups for the gene expression data. For the transplantation studies, at least 3 mice were used per group. At least three biological replicates were performed for each experiment unless indicated otherwise. All error bars are depicted as standard deviation, p-values are (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001).

Human Specimen Procurement

Human samples were collected through the UCSF Biospecimens and Skeletal Tissues for Rare and Orphan Disease Genetics (BSTROnG) Biobank, using protocols approved by the UCSF Institutional Review Board. All participants provided written consent. For the transplantation studies (**Figure 6D-H** and **Figure 6-figure supplement 1**), biopsy of the control subject was obtained from a 44 yo female healthy individual undergoing surgery at UCSF, the muscle from the FOP patient was obtained at autopsy from a 55 yo. Written informed consent was obtain from all subjects or their families.

Acknowledgments:

The authors thank Kelly Wentworth and Samuel Kou for their assistance collecting the autopsy samples and Francesco Tedesco for his support on the AFM grant. This work was supported by a NIH/NIAMS R01AR066735 to ECH, a French Muscular Association (AFM-Telethon) Trampoline grant to ECH and EB, the Radiant Hope Foundation to ECH, and the UCSF Cohort

Development Grant to ECH; the California Institute for Regenerative Medicine Fellowship Program to UCSF (TG2-01153) to EB, and the UCSF Program for Breakthrough Biomedical Research (PBBR) to EB; and the NIH R01AR072638-03 to JHP. Finally, the authors would like to thank the patients and their families for their generous specimen donations.

Author contributions:

Conceptualization, ECH, JHP, EB; Methodology, EB; Software Analyses, EB, TM; Validation, EB, SMG; Formal Analysis, EB; Investigation, EB, SMG, JW, BMM, ST; Resources, ECH, JHP; Data Curation, EB, TM; Writing – Original Draft, EB, ECH; Writing – Review and Editing, EB, ECH, JHP; Visualization, EB; Supervision, ECH, JHP; Project Administration, ECH, EB; Funding, ECH, EB, JHP; clinical samples, ECH, JHP.

Competing interests:

ECH receives clinical trial research funding from Clementia Pharmaceuticals, an Ipsen company, and Neurocrine Biosciences, Inc., through his institution. ECH received prior funding from Regeneron Pharmaceuticals, through his institution. ECH serves in an unpaid capacity on the international FOP Association Medical Registry Advisory Board, on the International Clinical Council on FOP, and on the Fibrous Dysplasia Foundation Medical Advisory Board. These activities pose no conflicts for the presented research.

Data and materials availability:

Single cell gene expression data have been deposited (GSE151918). The dataset used for the primary Hu-MuSCs can be found here,

- 521 <https://datadryad.org/stash/landing/show?id=doi%3A10.7272%2FQ65X273X>. Detailed scripts
- 522 can be found here, https://github.com/EmilieB12/FOP_muscle/tree/main.

REFERENCES

- Al Tanoury, Z., Rao, J., Tassy, O., Gobert, B., Gapon, S., Garnier, J. M., . . . Pourquie, O. (2020). Differentiation of the human PAX7-positive myogenic precursors/satellite cell lineage in vitro. *Development*, 147(12). doi:10.1242/dev.187344
- Andersen, D. C., Laborda, J., Baladron, V., Kassem, M., Sheikh, S. P., & Jensen, C. H. (2013). Dual role of delta-like 1 homolog (DLK1) in skeletal muscle development and adult muscle regeneration. *Development*, 140(18), 3743-3753. doi:10.1242/dev.095810
- Barruet, E., Garcia, S. M., Striedinger, K., Wu, J., Lee, S., Byrnes, L., . . . Pomerantz, J. H. (2020). Functionally heterogeneous human satellite cells identified by single cell RNA sequencing. *Elife*, 9. doi:10.7554/eLife.51576
- Barruet, E., Morales, B. M., Cain, C. J., Ton, A. N., Wentworth, K. L., Chan, T. V., . . . Hsiao, E. C. (2018). NF-kappaB/MAPK activation underlies ACVR1-mediated inflammation in human heterotopic ossification. *JCI Insight*, 3(22). doi:10.1172/jci.insight.122958
- Barruet, E., Morales, B. M., Lwin, W., White, M. P., Theodoris, C. V., Kim, H., . . . Hsiao, E. C. (2016). The ACVR1 R206H mutation found in fibrodysplasia ossificans progressiva increases human induced pluripotent stem cell-derived endothelial cell formation and collagen production through BMP-mediated SMAD1/5/8 signaling. *Stem Cell Res Ther*, 7(1), 115. doi:10.1186/s13287-016-0372-6
- Becht, E., McInnes, L., Healy, J., Dutertre, C. A., Kwok, I. W. H., Ng, L. G., . . . Newell, E. W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*. doi:10.1038/nbt.4314
- Bigot, A., Duddy, W. J., Ouandaogo, Z. G., Negroni, E., Mariot, V., Ghimbovski, S., . . . Duguez, S. (2015). Age-Associated Methylation Suppresses SPRY1, Leading to a Failure of Re-

quiescence and Loss of the Reserve Stem Cell Pool in Elderly Muscle. *Cell Rep*, 13(6), 1172-1182. doi:10.1016/j.celrep.2015.09.067

Billings, P. C., Fiori, J. L., Bentwood, J. L., O'Connell, M. P., Jiao, X., Nussbaum, B., . . . Kaplan, F. S. (2008). Dysregulated BMP signaling and enhanced osteogenic differentiation of connective tissue progenitor cells from patients with fibrodysplasia ossificans progressiva (FOP). *J Bone Miner Res*, 23(3), 305-313. doi:10.1359/jbmr.071030

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(6), 620-631. doi:10.1016/j.stemcr.2013.10.007

Chal, J., Al Tanoury, Z., Hestin, M., Gobert, B., Aivio, S., Hick, A., . . . Pourquie, O. (2016). Generation of human muscle fibers and satellite-like cells from human pluripotent stem cells in vitro. *Nat Protoc*, 11(10), 1833-1850. doi:10.1038/nprot.2016.110

Choi, I. Y., Lim, H., Estrellas, K., Mula, J., Cohen, T. V., Zhang, Y., . . . Lee, G. (2016). Concordant but Varied Phenotypes among Duchenne Muscular Dystrophy Patient-Specific Myoblasts Derived using a Human iPSC-Based Model. *Cell Rep*, 15(10), 2301-2312. doi:10.1016/j.celrep.2016.05.016

Convente, M. R., Chakkalakal, S. A., Yang, E., Caron, R. J., Zhang, D., Kambayashi, T., . . . Shore, E. M. (2018). Depletion of Mast Cells and Macrophages Impairs Heterotopic Ossification in an *Acvr1*(R206H) Mouse Model of Fibrodysplasia Ossificans Progressiva. *J Bone Miner Res*, 33(2), 269-282. doi:10.1002/jbmr.3304

Cosgrove, B. D., Gilbert, P. M., Porpiglia, E., Mourkioti, F., Lee, S. P., Corbel, S. Y., . . . Blau, H. M. (2014). Rejuvenation of the muscle stem cell population restores strength to injured aged muscles. *Nat Med*, 20(3), 255-264. doi:10.1038/nm.3464

Culbert, A. L., Chakkalakal, S. A., Theosmy, E. G., Brennan, T. A., Kaplan, F. S., & Shore, E. M. (2014). *Alk2* regulates early chondrogenic fate in fibrodysplasia ossificans progressiva

572 heterotopic endochondral ossification. *Stem Cells*, 32(5), 1289-1300.
573 doi:10.1002/stem.1633

574 Dey, D., Bagarova, J., Hatsell, S. J., Armstrong, K. A., Huang, L., Ermann, J., . . . Yu, P. B. (2016).
575 Two tissue-resident progenitor lineages drive distinct phenotypes of heterotopic
576 ossification. *Sci Transl Med*, 8(366), 366ra163. doi:10.1126/scitranslmed.aaf1090

577 Elsafadi, M., Manikandan, M., Dawud, R. A., Alajez, N. M., Hamam, R., Alfayez, M., . . .
578 Mahmood, A. (2016). Transgelin is a TGFbeta-inducible gene that regulates osteoblastic
579 and adipogenic differentiation of human skeletal stem cells through actin cytoskeleton
580 organization. *Cell Death Dis*, 7(8), e2321. doi:10.1038/cddis.2016.196

581 Friedrichs, M., Wirsdoerfer, F., Flohe, S. B., Schneider, S., Wuelling, M., & Vortkamp, A. (2011).
582 BMP signaling balances proliferation and differentiation of muscle satellite cell
583 descendants. *BMC Cell Biol*, 12, 26. doi:10.1186/1471-2121-12-26

584 Furrer, R., & Handschin, C. (2017). Optimized Engagement of Macrophages and Satellite Cells in
585 the Repair and Regeneration of Exercised Muscle. In B. Spiegelman (Ed.), *Hormones,*
586 *Metabolism and the Benefits of Exercise* (pp. 57-66). Chamcham.

587 Garcia, S. M., Tamaki, S., Lee, S., Wong, A., Jose, A., Dreux, J., . . . Pomerantz, J. H. (2018).
588 High-Yield Purification, Preservation, and Serial Transplantation of Human Satellite Cells.
589 *Stem Cell Reports*, 10(3), 1160-1174. doi:10.1016/j.stemcr.2018.01.022

590 Garcia, S. M., Tamaki, S., Xu, X., & Pomerantz, J. H. (2017). Human Satellite Cell Isolation and
591 Xenotransplantation. *Methods Mol Biol*, 1668, 105-123. doi:10.1007/978-1-4939-7283-
592 8_8

593 Goel, A. J., Rieder, M. K., Arnold, H. H., Radice, G. L., & Krauss, R. S. (2017). Niche Cadherins
594 Control the Quiescence-to-Activation Transition in Muscle Stem Cells. *Cell Rep*, 21(8),
595 2236-2250. doi:10.1016/j.celrep.2017.10.102

596 Hatsell, S. J., Idone, V., Wolken, D. M., Huang, L., Kim, H. J., Wang, L., . . . Economides, A. N.
597 (2015). ACVR1R206H receptor mutation causes fibrodysplasia ossificans progressiva by
598 imparting responsiveness to activin A. *Sci Transl Med*, 7(303), 303ra137.
599 doi:10.1126/scitranslmed.aac4358

600 Hicks, M. R., Hiserodt, J., Paras, K., Fujiwara, W., Eskin, A., Jan, M., . . . Pyle, A. D. (2018).
601 ERBB3 and NGFR mark a distinct skeletal muscle progenitor cell in human development
602 and hPSCs. *Nat Cell Biol*, 20(1), 46-57. doi:10.1038/s41556-017-0010-2

603 Kuang, S., Kuroda, K., Le Grand, F., & Rudnicki, M. A. (2007). Asymmetric self-renewal and
604 commitment of satellite stem cells in muscle. *Cell*, 129(5), 999-1010.
605 doi:10.1016/j.cell.2007.03.044

606 Lees-Shepard, J. B., Yamamoto, M., Biswas, A. A., Stoessel, S. J., Nicholas, S. E., Cogswell, C.
607 A., . . . Goldhamer, D. J. (2018). Activin-dependent signaling in fibro/adipogenic
608 progenitors causes fibrodysplasia ossificans progressiva. *Nat Commun*, 9(1), 471.
609 doi:10.1038/s41467-018-02872-2

610 Magli, A., Incitti, T., Kiley, J., Swanson, S. A., Darabi, R., Rinaldi, F., . . . Perlingeiro, R. C. R.
611 (2017). PAX7 Targets, CD54, Integrin alpha9beta1, and SDC2, Allow Isolation of Human
612 ESC/iPSC-Derived Myogenic Progenitors. *Cell Rep*, 19(13), 2867-2877.
613 doi:10.1016/j.celrep.2017.06.005

614 Matsumoto, Y., Hayashi, Y., Schlieve, C. R., Ikeya, M., Kim, H., Nguyen, T. D., . . . Hsiao, E. C.
615 (2013). Induced pluripotent stem cells from patients with human fibrodysplasia ossificans
616 progressiva show increased mineralization and cartilage formation. *Orphanet J Rare Dis*,
617 8, 190. doi:10.1186/1750-1172-8-190

618 Mauro, A. (1961). Satellite cell of skeletal muscle fibers. *J Biophys Biochem Cytol*, 9, 493-495.
619 doi:10.1083/jcb.9.2.493

- McMahon, A. P., & Bradley, A. (1990). The Wnt-1 (int-1) proto-oncogene is required for development of a large region of the mouse brain. *Cell*, 62(6), 1073-1085.
- Nastase, M. V., Young, M. F., & Schaefer, L. (2012). Biglycan: a multivalent proteoglycan providing structure and signals. *J Histochem Cytochem*, 60(12), 963-975. doi:10.1369/0022155412456380
- Nikitovic, D., Papoutsidakis, A., Karamanos, N. K., & Tzanakakis, G. N. (2014). Lumican affects tumor cell functions, tumor-ECM interactions, angiogenesis and inflammatory response. *Matrix Biol*, 35, 206-214. doi:10.1016/j.matbio.2013.09.003
- Ono, Y., Calhabeu, F., Morgan, J. E., Katagiri, T., Amthor, H., & Zammit, P. S. (2011). BMP signalling permits population expansion by preventing premature myogenic differentiation in muscle satellite cells. *Cell Death Differ*, 18(2), 222-234. doi:10.1038/cdd.2010.95
- Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., . . . Human Cell Atlas Meeting, P. (2017). The Human Cell Atlas. *Elife*, 6. doi:10.7554/eLife.27041
- Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., & Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol*, 33(5), 495-502. doi:10.1038/nbt.3192
- Schepers, K., Hsiao, E. C., Garg, T., Scott, M. J., & Passegue, E. (2012). Activated Gs signaling in osteoblastic cells alters the hematopoietic stem cell niche in mice. *Blood*, 120(17), 3425-3435. doi:10.1182/blood-2011-11-395418
- Seale, P., Sabourin, L. A., Girgis-Gabardo, A., Mansouri, A., Gruss, P., & Rudnicki, M. A. (2000). Pax7 is required for the specification of myogenic satellite cells. *Cell*, 102(6), 777-786.
- Segales, J., Perdiguero, E., & Munoz-Canoves, P. (2016). Regulation of Muscle Stem Cell Functions: A Focus on the p38 MAPK Signaling Pathway. *Front Cell Dev Biol*, 4, 91. doi:10.3389/fcell.2016.00091

- 643 Shea, K. L., Xiang, W., LaPorta, V. S., Licht, J. D., Keller, C., Basson, M. A., & Brack, A. S.
644 (2010). Sprouty1 regulates reversible quiescence of a self-renewing adult muscle stem cell
645 pool during regeneration. *Cell Stem Cell*, 6(2), 117-129. doi:10.1016/j.stem.2009.12.015
- 646 Shelton, M., Metz, J., Liu, J., Carpenedo, R. L., Demers, S. P., Stanford, W. L., & Skerjanc, I. S.
647 (2014). Derivation and expansion of PAX7-positive muscle progenitors from human and
648 mouse embryonic stem cells. *Stem Cell Reports*, 3(3), 516-529.
649 doi:10.1016/j.stemcr.2014.07.001
- 650 Shi, S., de Gorter, D. J., Hoogaars, W. M., t Hoen, P. A., & ten Dijke, P. (2013). Overactive bone
651 morphogenetic protein signaling in heterotopic ossification and Duchenne muscular
652 dystrophy. *Cell Mol Life Sci*, 70(3), 407-423. doi:10.1007/s00018-012-1054-x
- 653 Shore, E. M. (2012). Fibrodysplasia ossificans progressiva: a human genetic disorder of
654 extraskelatal bone formation, or--how does one tissue become another? *Wiley Interdiscip*
655 *Rev Dev Biol*, 1(1), 153-165. doi:10.1002/wdev.9
- 656 Shore, E. M., Xu, M., Feldman, G. J., Fenstermacher, D. A., Cho, T. J., Choi, I. H., . . . Kaplan, F.
657 S. (2006). A recurrent mutation in the BMP type I receptor ACVR1 causes inherited and
658 sporadic fibrodysplasia ossificans progressiva. *Nat Genet*, 38(5), 525-527.
659 doi:10.1038/ng1783
- 660 Shultz, L. D., Lyons, B. L., Burzenski, L. M., Gott, B., Chen, X., Chaleff, S., . . . Handgretinger,
661 R. (2005). Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R
662 gamma null mice engrafted with mobilized human hemopoietic stem cells. *J Immunol*,
663 174(10), 6477-6489. doi:10.4049/jimmunol.174.10.6477
- 664 Soriano-Arroquia, A., McCormick, R., Molloy, A. P., McArdle, A., & Goljanek-Whysall, K.
665 (2016). Age-related changes in miR-143-3p:Igfbp5 interactions affect muscle regeneration.
666 *Aging Cell*, 15(2), 361-369. doi:10.1111/accel.12442

Spencer, C. I., Baba, S., Nakamura, K., Hua, E. A., Sears, M. A., Fu, C. C., . . . Conklin, B. R. (2014). Calcium transients closely reflect prolonged action potentials in iPSC models of inherited cardiac arrhythmia. *Stem Cell Reports*, 3(2), 269-281. doi:10.1016/j.stemcr.2014.06.003

Stanley, A., Heo, S. J., Mauck, R. L., Mourkioti, F., & Shore, E. M. (2019). Elevated BMP and Mechanical Signaling Through YAP1/RhoA Poises FOP Mesenchymal Progenitors for Osteogenesis. *J Bone Miner Res*, 34(10), 1894-1909. doi:10.1002/jbmr.3760

Stantzou, A., Schirwis, E., Swist, S., Alonso-Martin, S., Polydorou, I., Zarrouki, F., . . . Amthor, H. (2017). BMP signaling regulates satellite cell-dependent postnatal muscle growth. *Development*, 144(15), 2737-2747. doi:10.1242/dev.144089

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., 3rd, . . . Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell*, 177(7), 1888-1902 e1821. doi:10.1016/j.cell.2019.05.031

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131(5), 861-872. doi:10.1016/j.cell.2007.11.019

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

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., . . . Rinn, J. L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*, 32(4), 381-386. doi:10.1038/nbt.2859

Urist, M. R. (1965). Bone: formation by autoinduction. *Science*, 150(3698), 893-899. doi:10.1126/science.150.3698.893

van der Wal, E., Herrero-Hernandez, P., Wan, R., Broeders, M., In 't Groen, S. L. M., van Gestel, T. J. M., . . . Pijnappel, W. (2018). Large-Scale Expansion of Human iPSC-Derived

- Skeletal Muscle Cells for Disease Modeling and Cell-Based Therapeutic Strategies. *Stem Cell Reports*, 10(6), 1975-1990. doi:10.1016/j.stemcr.2018.04.002
- Volpato, V., & Webber, C. (2020). Addressing variability in iPSC-derived models of human disease: guidelines to promote reproducibility. *Dis Model Mech*, 13(1). doi:10.1242/dmm.042317
- Wu, J., Matthias, N., Lo, J., Ortiz-Vitali, J. L., Shieh, A. W., Wang, S. H., & Darabi, R. (2018). A Myogenic Double-Reporter Human Pluripotent Stem Cell Line Allows Prospective Isolation of Skeletal Muscle Progenitors. *Cell Rep*, 25(7), 1966-1981 e1964. doi:10.1016/j.celrep.2018.10.067
- Xi, H., Fujiwara, W., Gonzalez, K., Jan, M., Liebscher, S., Van Handel, B., . . . Pyle, A. D. (2017). In Vivo Human Somitogenesis Guides Somite Development from hPSCs. *Cell Rep*, 18(6), 1573-1585. doi:10.1016/j.celrep.2017.01.040
- Xi, H., Langerman, J., Sabri, S., Chien, P., Young, C. S., Younesi, S., . . . Pyle, A. D. (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*. doi:10.1016/j.stem.2020.04.017
- Xu, X., Wilschut, K. J., Kouklis, G., Tian, H., Hesse, R., Garland, C., . . . Pomerantz, J. H. (2015). Human Satellite Cell Transplantation and Regeneration from Diverse Skeletal Muscles. *Stem Cell Reports*, 5(3), 419-434. doi:10.1016/j.stemcr.2015.07.016
- Zheng, G. X., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., . . . Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nat Commun*, 8, 14049. doi:10.1038/ncomms14049

717 **Abbreviations:**

718	PAX7: Paired Box 7
719	hiPSCs: Human induced pluripotent stem cells
720	BMP: bone morphogenetic protein
721	ACVR1/2A: activin receptor type 1/2A
722	MYOD1: Myogenic Differentiation 1
723	HNK1: Human Natural Killer-1
724	CXCR4: C-X-C chemokine receptor type 4
725	MYF5: Myogenic Factor 5
726	COL1A1: Collagen Type I Alpha 1 Chain
727	PDGFRA: Platelet Derived Growth Factor Receptor Alpha
728	SOX2: SRY-Box 2
729	EFNB3: Ephrin B3
730	MAP2: Microtubule Associated Protein 2
731	ASCL1: Achaete-Scute Family BHLH Transcription Factor 1
732	ASPN: Asporin
733	APOE: Apolipoprotein E
734	OGN: Osteoglycin
735	TOP2A: DNA Topoisomerase II Alpha
736	FABP7: Fatty Acid Binding Protein 7
737	SPRY1: Sprouty 1
738	MYL1: Myosin Light Chain 1
739	DLK1: Delta Like Non-Canonical Notch Ligand 1
740	ITM2A: Integral Membrane Protein 2A
741	DCN: Decorin

742 SPARC: Secreted Protein Acidic and Cysteine Rich

743 TIMP1: Tissue Inhibitor of Metalloproteinase 1

744 BGN: Biglycan

745 LUM: Lumican

746 TAGLN: Transgelin

747 IGBP5: Insulin Like Growth Factor Binding 5

748 ID1/3: Inhibitor of DNA Binding 1/3

749 TGFb: Transforming Growth Factor b

750 BMPR2: Bone Morphogenetic Protein Receptor 2

751 CAV1: Caveolin 1

752 ECM: Extracellular Matrix

753

754

755 **Supplementary Materials**

756

757 **Figure 1-figure supplement 1. Transplanted satellite cells from biceps and diaphragm of FOP**
758 **patient do not form bone.**

759 **Figure 2-figure supplement 1. Early expression of muscle markers during the myogenic**
760 **differentiation.**

761 **Figure 3-figure supplement 1. Differentiated cells types in the myogenic differentiation varies**
762 **between cell lines, sorted cells express PAX7 and can differentiate into myotubes.**

763 **Figure 4-figure supplement 1. Markers expressed in sorted control and FOP HNK1⁻CD45⁻**
764 **CD31⁻CXCR4⁺CD29⁺CD56^{dim} cells.**

765 **Figure 5-figure supplement 1. Analysis of the myogenic sub-cluster.**

Figure 6-figure supplement 1. Expression of myogenic markers across primary and hiPS-derived myogenic cells.

Figure 1-Source Data 1. Muscle specimen information.

Figure 1-Source Data 2. Hu-MuSCs transplantation details.

Figure 3-Source Data 1: hiPSC-derived HNK1⁻CD45⁻CD31⁻ CXCR4⁺CD29⁺CD56^{dim} cell transplants

Figure 4-Source Data 1: Quality control information for each sample.

Figure 5-Source Data 1. Differential expression analysis of control and FOP myogenic cells.

Figure 6-Source Data 1. BEAM analysis.

Video 1. Myotube contraction of differentiated Control (1323-2) hiPSCs.

Video 2. Myotube contraction of differentiated FOP (F1-1) hiPSCs.

Data File S1- hiPSCs samples integration.

Data File S2- Myogenic sub-clustering.

Data File S3- Pseudotime analyses.

Data File S4- Primary Hu-MuSCs/hiPSCs integration.

Figure 1-figure supplement 1. Transplanted satellite cells from biceps and diaphragm of FOP patient do not form bone. (A) Flow cytometry analysis sorting strategy to isolate human satellite cells. **(B)** Representative immunostaining of human satellite cells. White arrows indicate PAX7+ satellite cells (scale bar, 50µm). **(C)** X-ray of transplanted mice at 5 and 10 weeks after initial transplantation (scale bars, 4mm (week 5), 7mm (week 10)) showing no heterotopic ossification. Details about muscle specimens used in this figure are in **Figure 2-Source Data1**.

Figure 2-figure supplement 1. Early expression of muscle markers during the myogenic differentiation. (A) MYOGENIN (scale bar, 200µm) and (B) PAX7 immunofluorescence staining showing that MYOGENIN and PAX7 are expressed as early as day 25 during the myogenic differentiation in control and BMP impaired hiPSCs (scale bar, 200µm). (C) *DYSTROPHIN* gene expression at day 25 and day 50 of differentiation (n=3 biological replicates and n≥3 technical replicates). Error bar represent mean and SD.

Figure 3-figure supplement 1. Differentiated cells types in the myogenic differentiation varies between cell lines, sorted cells express PAX7 and can differentiate into myotubes. Related to Figure 3. (A) Representative flow cytometry profiles of negative (CD45 and CD31) and positive (CD29 and CD56) markers used to purify muscle stem cells. **(B)** Quantification of the percentage of cells positive for CD45, CD31, CD29, and CD56 at day 20. (n=3 biological replicates and n=3 technical replicates). Error bar represent mean and SD. **(C)** FACS analysis of the myogenic differentiation at day 50. Muscle stem cells stained for PAX7, CD29, and CD56. **(D)** Representative immunofluorescence staining for MHC showing sorted muscle stem cells after being cultured in differentiation media for 7 days (scale bar, 200µm).

Figure 4-figure supplement 1. Markers expressed in sorted control and FOP HNK1⁺CD45⁺CD31⁺ CXCR4⁺CD29⁺CD56^{dim} cells. (A-B) Feature expression plots showing the localization of cells expressing myogenic markers (*PAX7*, *MYF5*) mesenchymal marker (*PDGFRA*) and neuronal progenitor marker (*SOX2*) in the control (A) and FOP (B) sorted cells. (C) UMAP visualization plots of merged samples with cells colored by samples. (D) Feature expression plots showing the localization of cells expressing myogenic (*MYOG*), mesenchymal (*ASPN*) and neuronal markers (*MAP2*, *ASCL1*) in the merged samples. (E) Heat map of the top 5 differentially expressed genes in each cluster of the merged sample analysis.

Figure 5-figure supplement 1. Analysis of the myogenic sub-cluster. (A) Heat map of the top 5 differentially expressed genes in each cluster. (B) Cycle genes were scored for each cluster. Bar plot depicting the proportion of cells in G1/G0, G2/M and S phase for each cluster. (C) Proportion of cells in G1/G0, G2/M and S phase for each sample. (D) Proportion bar graph of cells per cluster for the control and FOP samples. (E) UMAP showing the distribution of the two merged samples. (F) Additional genes significantly differentially expressed by more than 1-fold. (G) Dot plot displaying the expression of genes associated with BMP and P38MAPK pathways for each cluster.

Figure 6-figure supplement 1. Expression of myogenic markers across primary and hiPS-derived myogenic cells. Dot plot displaying the average expression and the percent of cells expressing myogenic genes across clusters for the 3 merged samples.

839 **Figure 1-Source Data 1. Muscle specimen information.**

Figures	Muscle	Age	Sex	Subjects
Figure 1A	Psoas	52	F	FOP
Figure 1A	Intercostal	46	F	FOP
Figure 1B (top)	Vastus Lateralis	59	M	Control
Figure 1B (middle)	Biceps Brachii	21	F	FOP
Figure 1B (bottom)	Biceps Brachii	32	F	FOP
Figure 1C-H and S1	Biceps Brachii	44	F	Control
Figure 1C-H and S1	Biceps Brachii	32	F	FOP
Figure 1C-H and S1	Diaphragm	32	F	FOP

840

841 **Figure 1-Source Data 2. Hu-MuSCs transplantation details.**

Week 5			
Satellite Cells isolated from	# of cells transplanted/mouse	Human Dystrophin fibers	Human PAX7+ cells
Ctrl Biceps	6000	38	4
Ctrl Biceps	6000	30	7
Ctrl Biceps	6000	20	1
FOP Diaphragm	5000	13	1
FOP Diaphragm	5000	8	1
FOP Diaphragm	5000	6	0
FOP Biceps	2800	5	3
FOP Biceps	2800	5	2
FOP Biceps	2800	0	2
Week10			
Ctrl Biceps	6000	133	24
Ctrl Biceps	6000	112	13
Ctrl Biceps	6000	82	19
FOP Diaphragm	5000	98	23
FOP Diaphragm	5000	120	22
FOP Diaphragm	5000	57	10
FOP Diaphragm	5000	42	12
FOP Biceps	2800	0	0
FOP Biceps	2800	58	15
FOP Biceps	2800	12	5

842

843

844

845

Figure 3-Source Data 1: hiPSC-derived HNK1⁺CD45⁻CD31⁻CXCR4⁺CD29⁺CD56^{dim} cell transplants.

Cell line	# of cells transplanted	Human Dystrophin fibers	Human PAX7+ cells
WTC11	8000	6	0
WTC11	8261	0	0
WTC11	12341	0	0
1323-2	8011	0	0
1323-2	7873	0	0
BJ2	15000	2	0
BJ2	887	0	0
BJ2	9767	1	0
F1-1	5508	0	0
F1-1	6084	0	0
F1-1	15221	0	0
F2-3	17638	15	2
F3-2	30000	51	0
F3-2	1300	0	0
F3-2	40000	10	0

Figure 4-Source Data 1: Quality control information for each sample.

	Control (1323-2)	FOP (F3-2)
Estimated Number of Cells	2,054	5,306
Mean Reads per Cell	76,178	32,870
Median Genes per Cell	2,689	2,501
Number of Reads	156,470,954	174,413,305
Valid Barcodes	97.8%	98.0%
Reads Mapped Confidently to Transcriptome	63.1%	66.3%
Reads Mapped Confidently to Exonic Regions	67.0%	70.3%
Reads Mapped Confidently to Intronic Regions	17.4%	15.3%
Reads Mapped Confidently to Intergenic Regions	3.6%	3.4%
Reads Mapped to Antisense to Gene	4.7%	4.9%
Sequencing Saturation	79.0%	54.8%
Fraction Reads in Cells	21,100	22,238

Figure 5-Source Data 1. Differential expression analysis of control and FOP myogenic cells
(see excel file).

Figure 6-Source Data 1. BEAM analysis (see excel file).

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Flow cytometry

Cells were treated with Accutase for 20min at 37C, washed with FACS buffer, and stained for HNK1-PE, CD45-PE, CD31-PE, CD56-APC, CXCR4-PE-Cy7, and CD29-FITC (eBioscience). HNK1⁻CD45⁻CD31⁻ cells [to select against neuronal cells (HNK1, Human Natural Killer-1) (Choi et al., 2016; Hicks et al., 2018), hematopoietic cells (CD45), and endothelial cells (CD31)], co-expressing CD29, CXCR4, and intermediate CD56, markers present on human PAX7⁺ cells (Garcia et al., 2018; Garcia et al., 2017; Xu et al., 2015) were sorted with a FACSAriaIII (BD Biosciences) and Sytox Blue (Life Technologies) was used as a viability marker. Alternatively, cells were permeabilized and fixed (Fix/Perm Buffer Set, BioLegend). Fixed cells were first incubated with primary antibodies (PAX7 (DSHB), CD29 (BD Biosciences) and CD56 (BD Biosciences) following by secondary antibodies (Alexa350-conjugated goat anti-mouse IgG, Alexa488-conjugated goat anti-rat IgG and Alexa546-conjugated donkey anti-goat IgG, Life Technologies).

Cell Immunostaining and NSG Tibialis Anterior Analysis

Collected tibialis anterior frozen cross sections were fixed in 4% PFA for 10 min at room temperature, washed with PBST (PBS with 0.1% Tween-20 (Sigma-Aldrich) and blocked in PBS with 10% goat serum for 1hr at room temperature. Slides were then incubated 4hrs at room temperature with the following: mouse monoclonal anti-human DYSTROPHIN (DSHB), mouse monoclonal IgG1 anti-PAX7 (DSHB), rabbit polyclonal anti-Laminin (Sigma-Aldrich), mouse monoclonal IgG2b anti-human SPECTRIN (Leica Microsystems), and mouse monoclonal IgG2b anti-human LAMIN A/C (Vector Laboratories). After PBST wash, slides were incubated with the following secondary antibodies: Alexa Fluor 555 goat anti-mouse IgG, Alexa Fluor 594 goat anti-mouse IgG1, Alexa Fluor 488 goat anti-mouse IgG2b, and Alexa Fluor 647 goat anti-rabbit (Life Technologies). Finally, sections were mounted with VECTASHIELD mounting media with DAPI (Vector Laboratories). All samples were examined using a Leica upright or DMi8 Leica microscope. Sections with the most human fibers were used for human DYSTROPHIN and PAX7 quantification for each condition.

Single cell RNA Sequencing and Analysis

45,000 (FOP) and 30,000 (control) HNK1⁻CD45⁻CD31⁻ CXCR4⁺CD29⁺CD56^{dim} cells isolated from the iMPC differentiations were loaded onto one well of a 10X chip to produce Gel Bead-in-Emulsions (GEMs). GEMs underwent reverse transcription to barcode RNA before cleanup and cDNA amplification. Libraries were prepared with the Chromium Single Cell 3' Reagent Version 2 Kit. Each sample was sequenced on 1 lane of the NovaSeq 6000 S4. Sequencing reads were processed with Cell Ranger version 2.0.0. using the human reference transcriptome GRCh38. The estimated number of cells, mean reads per cell, median genes per cells, median UMI (Unique Molecular Identifier) counts per cells as well as other quality control information are summarized in **Figure 4-Source Data 1**. Gene-barcoded matrices were analyzed with the R package Seurat v3.1.5 (Satija, Farrell, Gennert, Schier, & Regev, 2015; Stuart et al., 2019; Team, 2014; Zheng et

al., 2017). Gene core matrices from single cell RNA sequencing of primary human satellite cells isolated from a vastus muscle (Barruet et al., 2020) was used when comparing the transcriptional profile of hiPS-derived HNK1⁻CD45⁻CD31⁻CXCR4⁺CD29⁺CD56^{dim} cells. For the comparison with primary Hu-MuSCs, hiPSC-derived cell sequencing reads were re-aligned using the human reference transcriptome hg19. Cells with fewer than 500 genes, greater than 5000 genes and genes expressed in fewer than 5 cells were not included in the downstream analyses. Cells with more than 10% mitochondrial counts were filtered out. Samples were normalized with NormalizeData using default settings. The FindVariableFeatures function was used to determine subset of feature that exhibit high cell-to-cell variation in each dataset based on a variance stabilizing transformation (“vst”). We used the default setting returning 2,000 feature per dataset. These were used for downstream analysis. In the case of the merged data analysis samples were combined utilizing the FindIntegrationAnchors function with the ‘dimensionality’ set at 30. Then, we ran these “anchors” to the IntegratData function for batch correction for all cells enabling them to be jointly analyzed. The resulting outputs were scaled mitochondrial contamination regressed out with the ScaleData function. In addition, while we didn’t regress out heterogeneity associated with cell cycle stage since it is an important factor in determining the state of quiescence of our sorted human muscle stem cells, we regressed out differences between G2/M and S cell cycle stage. PCA was performed with RunPCA, and significant PCs determined based on the Scree plot utilizing the function PCElbowPlot. The resolution parameter in FindClusters was adjusted to 0.5. Clusters were visualized by UMAP with Seurat’s RunUMAP function. We performed differential gene-expression utilizing Seurat v3’s FindMarkers function with default settings which utilizes the Wilcoxon rank-sum test to calculate adjusted p values for multiple comparisons. We used the CellCycleScoring function to assign score based on the expression of G2/M and S phase markers (Regev et al., 2017). Myogenic cells were further analyzed by sub-clustering using the subset function. We then use the FindNeighbors (dims=15) and FindClusters (resolution = 0.4) functions

928 on the myogenic cell subset of the merged hiPSCs samples only to identify sub-clusters
 929 corresponding to different myogenic states. To order the cells in pseudotime based on their
 930 transcriptional similarity we used Monocle 2.12. Variable genes from Seurat analysis were used
 931 as input and clusters were projected onto the minimum spanning tree after ordering. Gene
 932 expression patterns were plotted with `plot_genes_branched_heatmap`,
 933 `plot_genes_branched_pseudotime`, and `plot_multiple_branches_pseudotime`. The BEAM (branch
 934 expression analysis modeling) function was used to score gene significance in a branch-dependent
 935 manner. Cells were re-ordered using the `orderCells` function to set branch A (myocytes) in **Figure**
 936 **5F** as the “root-state”. This allowed us to determine genes that were significantly branch dependent
 937 in branch B (mainly Hu-MuSCs) vs branch C (hiPS-derived cells) using the BEAM analysis.