

ERBB3 and NGFR mark a distinct skeletal muscle progenitor cell in human development and hPSCs

Michael R. Hicks^{1,2,3}, Julia Hiserodt³, Katrina Paras³, Wakana Fujiwara³, Ascia Eskin^{2,4}, Majib Jan^{1,2,3}, Haibin Xi^{1,2,3}, Courtney S. Young^{1,2,5}, Denis Evseenko⁶, Stanley F. Nelson^{2,4,5}, Melissa J. Spencer^{1,2,5,7}, Ben Van Handel⁶ and April D. Pyle^{1,2,3,5*}

Human pluripotent stem cells (hPSCs) can be directed to differentiate into skeletal muscle progenitor cells (SMPCs). However, the myogenicity of hPSC-SMPCs relative to human fetal or adult satellite cells remains unclear. We observed that hPSC-SMPCs derived by directed differentiation are less functional in vitro and in vivo compared to human satellite cells. Using RNA sequencing, we found that the cell surface receptors ERBB3 and NGFR demarcate myogenic populations, including PAX7 progenitors in human fetal development and hPSC-SMPCs. We demonstrated that hPSC skeletal muscle is immature, but inhibition of transforming growth factor- β signalling during differentiation improved fusion efficiency, ultrastructural organization and the expression of adult myosins. This enrichment and maturation strategy restored dystrophin in hundreds of dystrophin-deficient myofibres after engraftment of CRISPR-Cas9-corrected Duchenne muscular dystrophy human induced pluripotent stem cell-SMPCs. The work provides an in-depth characterization of human myogenesis, and identifies candidates that improve the in vivo myogenic potential of hPSC-SMPCs to levels that are equal to directly isolated human fetal muscle cells.

Directed differentiation of human pluripotent stem cells (hPSCs), including human embryonic stem cells (hESCs) and human induced pluripotent stem cells (hiPSCs), seeks to recapitulate development to form cell lineages that are similar to human in vivo counterparts. Directed differentiation of hPSCs to specific lineages for cell therapies is showing promise in clinical settings and in preclinical animal models for diseases ranging from cardiac myopathy to macular degeneration, diabetes mellitus and Parkinson's disease^{1–3}. However, for many cell lineages, directed differentiation results in progeny that are heterogeneous and functionally immature compared to cells derived during normal human development^{4–6}. Immature cells derived from hPSCs may function inappropriately, diminishing their use for modelling human disease and/or cell replacement therapies⁷.

Although several directed differentiation protocols generate skeletal muscle cells from hPSCs^{8–13}, their developmental stage or functional similarity to fetal or adult human skeletal muscle is unknown. Longstanding protocols that are used to differentiate skeletal muscle from hPSCs require viral-mediated overexpression of transcription factors, such as MYOD^{14,15}, PAX7 (ref. 16) or PAX3 (ref. 17), limiting the generation of truly representative myogenic progenitors. Skeletal myogenesis in vivo relies on tightly controlled spatial and temporal cues to ensure timely embryonic transitions through the presomitic mesoderm, somites and dermomyotome to form the myotome¹⁸. Although recent studies have followed human developmental cues to differentiate hPSCs to somites in vitro^{10,19,20}, none has generated sufficient quantity or quality of skeletal muscle progenitor cells (SMPCs) or satellite cells (SCs) from hPSCs.

SCs are endogenous skeletal muscle stem cells that are responsible for the formation of new muscle and are indispensable for

maintenance and repair²¹, and are thought to arise during secondary fetal myogenesis²². Fetal muscle cells are unique from adult SCs in their ability to retain PAX7 in a niche-independent, non-quiescent state²³. Although a single transplanted SC can give rise to several hundred myofibres²⁴, it remains unclear whether fetal or adult SCs are more regenerative in vivo, or whether directed differentiation of hPSC-SMPCs should seek to attain fetal-like or adult-like SCs. A better understanding of human SC biology and regenerative potential from multiple stages of development could enable derivation of more myogenic hPSC-SMPCs. Cell replacement is a promising therapy for many muscle diseases, including Duchenne muscular dystrophy (DMD), a degenerative muscle disease caused by lack of dystrophin. Autologous transplantation of CRISPR (clustered regularly interspaced short palindromic repeats)–Cas9 (CRISPR-associated protein 9)–corrected hiPSC-SMPCs²⁵ for patients with DMD could repopulate diseased or damaged myofibres with donor SMPCs or SCs to restore muscle function over the patient's lifetime.

To improve our understanding of the developmental and functional status of hPSC-SMPCs, we compared established directed differentiation protocols^{8–11} to muscle cells across multiple stages of human development, including fetal weeks 8–18 and adult SCs. Fetal muscle cells were also profiled to other musculoskeletal tissues by RNA sequencing (RNA-seq), enabling the identification of the muscle-specific receptors ERBB3 and nerve growth factor receptor (NGFR; also known as CD271). Our strategy enabled isolation of PAX7⁺ hPSC-SMPCs with robust engraftment in vivo, including the ability to restore dystrophin in mdx–NOD Scid gamma (NSG) mice, and could potentially improve the delivery of gene-editing therapies for skeletal muscle diseases. Importantly, this work improves our

¹Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, University of California, Los Angeles, CA, USA. ²Center for Duchenne Muscular Dystrophy, University of California, Los Angeles, CA, USA. ³Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, CA, USA. ⁴Department of Human Genetics, University of California, Los Angeles, CA, USA. ⁵Molecular Biology Interdepartmental Program, University of California, Los Angeles, CA, USA. ⁶Department of Orthopaedic Surgery, Keck School of Medicine, Stem Cell Research and Regenerative Medicine, University of Southern California, Los Angeles, CA, USA. ⁷Department of Neurology, University of California, Los Angeles, CA, USA. *e-mail: apyle@mednet.ucla.edu

understanding of human fetal muscle progenitors and provides a developmental identity of hPSC-SMPCs relative to equivalent cells *in vivo*.

Results

hPSC-SMPCs have reduced fusion efficiency compared to their human muscle-derived counterparts. To evaluate the myogenic potential of hPSC-SMPCs relative to fetal and adult SCs, multiple directed differentiation protocols were evaluated^{8–11}, and two were selected for further analysis^{10,11}. Both methods consistently generated muscle cells that expressed the myogenic transcription factors *PAX7*, *MYF5*, *MYOD* and *MYOG*, and spontaneously contracting myotubes (Supplementary Fig. 1). hPSC-SMPCs were dissociated/replated and compared to equivalent numbers of muscle cells isolated from human fetal weeks 9, 14 and 17, and adult skeletal muscle tissues (Fig. 1a). Cells that were isolated from adult muscle formed myotubes most efficiently and contained the most nuclei (24.5 nuclei per myotube, $P < 0.05$). The fusion indices and nuclei per myotube from fetal weeks 14 and 17 were significantly greater than fetal week 9 and hPSC-SMPCs ($P < 0.05$), whereas week 9 fetal muscle cells were not significantly different from hPSC-myotubes ($P = 0.3$).

In myogenic development, cells undergoing proliferation and differentiation co-exist²⁶. To better understand the timing and heterogeneity of myogenic transcription factor expression during myotube differentiation, *PAX7* and *MYOD* were evaluated (Fig. 1b). Indeed, all fetal muscle cultures contained *PAX7*⁺*MYOD*⁺ SMPCs. Myotubes derived from adult SCs expressed a *PAX7*⁺ or *MYOD*⁺ single nucleus, but few *PAX7*⁺*MYOD*⁺ nuclei were present in adult cultures. hPSC-SMPCs from either directed differentiation method contained the largest fraction of *PAX7*⁺*MYOD*⁺ nuclei, and many *MYOD*⁺ nuclei were not contained within *MYHC*⁺ myotubes. Together, these data indicate that hPSC-SMPCs inefficiently differentiate to form myotubes and/or represent a distinct developmental stage.

hPSC-SMPCs have limited engraftment potential *in vivo* compared to human fetal myogenic progenitors. Directed differentiation of hPSCs has yet to generate SMPCs with robust engraftment potential *in vivo*. Previous reports show that cultured human fetal muscle cells also engraft inefficiently²⁷. Notably, directly isolated adult SCs engraft more efficiently than cultured SCs²⁸, but engraftment of directly isolated human fetal muscle cells has never been tested. Thus, we used cultured and uncultured human fetal muscle cells (which were all directly isolated mononuclear cells) to benchmark hPSC-SMPC engraftment potential, and assessed whether the developmental state or culturing of fetal cells correlated with changes in engraftment. In parallel to cultured and uncultured fetal muscle cells, hPSC-SMPCs were transplanted into the tibialis anterior of cardiotoxin-pretreated mdx-NSG mice, which are immunocompromised and lack dystrophin expression (Fig. 1c). At 30 days, several hundred hPSC-SMPCs, marked by human (h)-lamin A/C, were detected in the muscle. However, most hPSC-SMPCs had not fused with the host muscle (as marked by h-lamin A/C⁺spectrin⁺ and h-dystrophin⁺) nor did they reside in the SC position. Instead, most h-lamin A/C⁺ hPSC-SMPCs were found in the perimysium and epimysium of the host tibialis anterior (Fig. 1c). By contrast, both cultured and uncultured fetal cells resulted in an increased number of h-dystrophin⁺ cells compared to hPSC-SMPCs ($P < 0.05$; Supplementary Table 1). Whereas restoration of h-dystrophin⁺ myofibres by cultured fetal muscle was <1% of the host tibialis anterior, uncultured fetal cell engraftment was significantly more efficient, restoring 10–15% of h-dystrophin⁺ myofibres in the host muscle, a level of dystrophin expression that has been reported to result in functional improvement in mice^{29–31}.

To evaluate the *in vivo* myogenic potential of hPSC-SMPCs compared to cultured and directly isolated fetal muscle cells, we counted engrafted myofibres at 11 points (0–10 mm) throughout the tibialis anterior of mdx-NSG mice, and ranked all counts using Kruskal–Wallis analysis of variance tests³² (Fig. 1c). Directly isolated fetal muscle cells had significantly higher engraftment efficiency than cultured fetal cells or hPSC-SMPCs ($P < 0.0001$). These data indicate that generating SMPCs resembling directly isolated fetal muscle could improve the inferior engraftment potential of hPSC-SMPCs.

Enriching for HNK1[−]NCAM⁺ SMPCs increases myogenic cell numbers but does not increase myogenicity *in vivo*. To potentially enrich for cells with increased myogenicity, we applied a previously described sorting strategy that removes HNK1⁺ cells (a neuroectodermal marker) and selects for neural cell adhesion molecule (NCAM; also known as CD56)^{8,33}. NCAM is also expressed by human fetal muscle²⁷ and adult SCs³⁴. We isolated HNK1[−]NCAM⁺ cells and evaluated their myogenic potential using two directed differentiation methods^{10,11}. HNK1[−]NCAM⁺ enrichment increased *PAX7* and *MYF5* expression by ~1.7-fold compared to unsorted hPSC-SMPCs ($P < 0.05$; Fig. 2a). In both protocols, HNK1[−]NCAM⁺ SMPCs could be grown in skeletal muscle basal media 2 (SkBM-2, Lonza) with fibroblast growth factor 2 (FGF2), and when induced to differentiate, the number of *MYHC*⁺ cells increased compared to dissociated/replated SMPCs ($P < 0.05$; Fig. 2b).

hPSC line variability is known to greatly affect the propensity of *in vitro* differentiation^{35,36}, and some reports suggest that DMD hiPSCs inefficiently differentiate to skeletal muscle *in vitro*³³. After NCAM sorting, hiPSC-SMPCs from a wild-type, DMD and a previously generated CRISPR–Cas9-corrected DMD line²⁵, could be differentiated to produce equivalent numbers of myotubes (Fig. 2c). All hiPSC lines differentiated less well than the H9 hESC line. Thus, the absence or presence of dystrophin did not affect myogenic differentiation *in vitro*.

We next evaluated whether sorting of HNK1[−]NCAM⁺ could improve the *in vivo* engraftment potential of hPSC-SMPCs. However, unlike *in vitro*, NCAM⁺ sorting did not improve engraftment compared to unsorted SMPCs (Fig. 2d). Irradiation was unable to improve hPSC-SMPC engraftment potential (Supplementary Table 1). In summary, enriching hPSC-SMPCs for HNK1[−]NCAM⁺ does not increase the engraftment potential, despite modestly increasing the myogenic potential *in vitro*.

RNA-seq reveals that HNK1[−]NCAM⁺ hPSC-SMPCs are heterogeneous and more differentiated than fetal muscle progenitor cells. To better understand the functional differences between fetal myogenic cells, hPSC-SMPCs and their differentiated progeny, we performed RNA-seq on unsorted, directly isolated fetal muscle cells (which were all mononuclear), NCAM⁺ cultured fetal muscle cells and hPSC-SMPCs, and myotubes generated from both fetal muscle and hPSCs (Fig. 3a). Genes related to myogenic progenitor identity, including *PAX7* and *MYF5*, were enriched in cultured fetal muscle cells compared to hPSC-SMPCs (Supplementary Table 2). By contrast, hPSC-SMPCs demonstrated substantial enrichment of transcription factors and structural proteins related to myogenic differentiation, such as *MYOG*. Principle component analysis confirmed that hPSC-SMPCs clustered more closely with myotubes (Fig. 3b). Moreover, Gene Ontology analysis identified multiple categories associated with development that were enriched in NCAM⁺ hPSC-SMPCs, suggesting that this population contains embryonic cell types. NCAM⁺ cultured fetal cells were also enriched for gene categories related to cell migration (Supplementary Table 2). Comparison of directly isolated fetal muscle cells to hPSC-SMPCs generated similar results, and Gene Ontology analysis revealed potential factors, including *HGF* (which encodes hepatocyte growth factor) and *LIF* (which encodes leukaemia inhibitory factor),

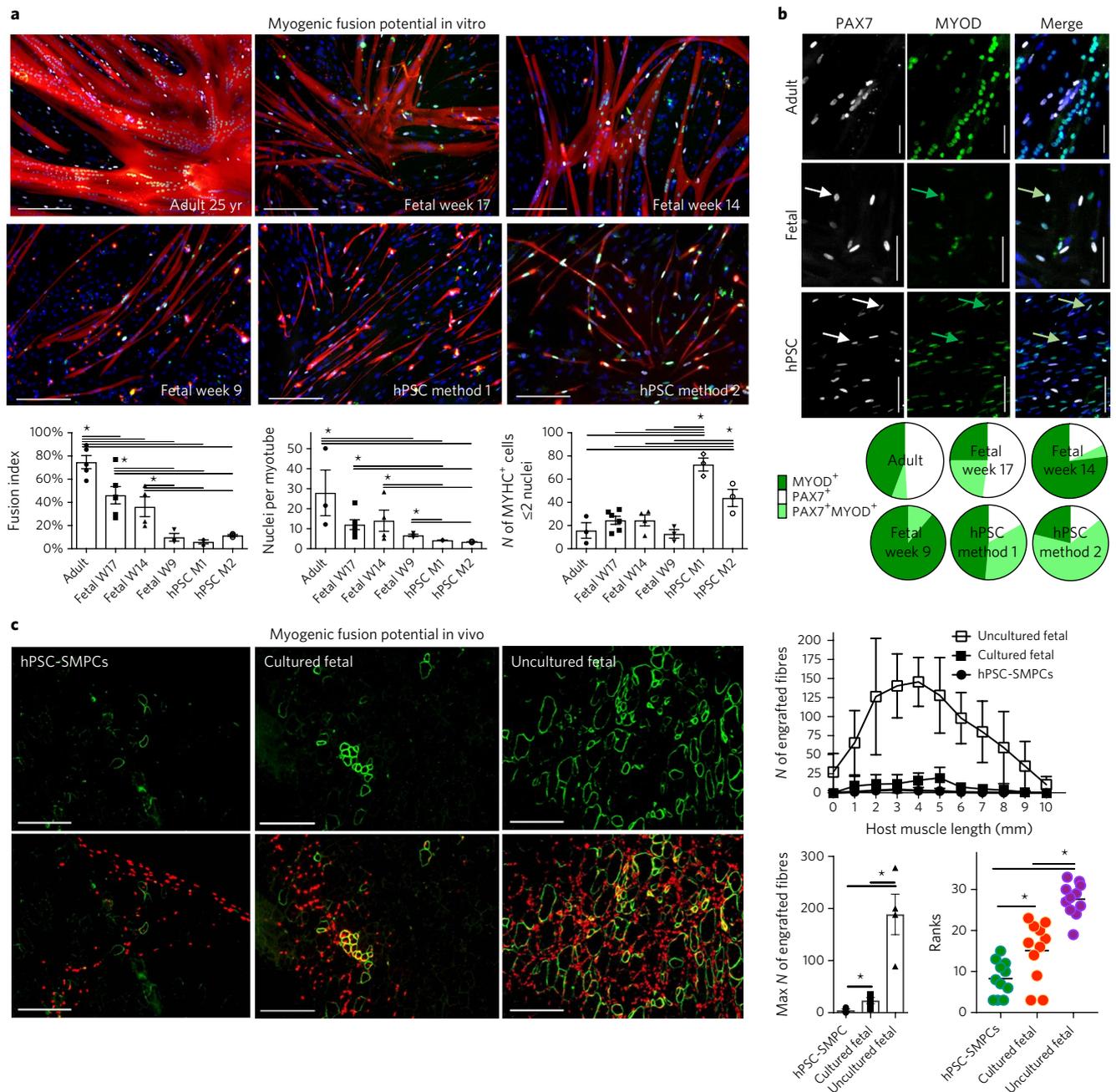


Fig. 1 | hPSC-SMPCs have reduced in vitro and in vivo myogenic potential relative to fetal or adult SCs. a, Human muscle obtained from fetal weeks 9, 14 and 17, an adult 25 years of age or hPSCs differ in their ability to form MYHC⁺ myotubes in vitro (red). PAX7 (white), MYOD (green) and DAPI (blue) are also shown. Scale bars, 200 μ m. Fusion index (the percentage of nuclei within MYHC⁺ cells containing ≥ 3 nuclei per total nuclei) and nuclei per myotube are greatest in adult and fetal muscle, whereas hPSCs primarily had ≤ 2 MYHC nuclei per mm² (mean \pm s.e.m.; $N=3$ adult, $N=5$ fetal tissues and $N=3$ hPSC from independent directed differentiations; one-way ANOVA post-hoc Tukey: $*P < 0.05$). **b**, Co-localization of PAX7 (white) and MYOD (green) differ across developmental stages (shown by arrows). Pie charts show the proportion of PAX7 and MYOD expression in each cell type ($N=3$ adult, $N=5$ fetal tissues and $N=3$ hPSC from independent directed differentiations). Scale bars, 50 μ m. **c**, To quantify engraftment, we assessed the total number of human cells (h-lamin A/C⁺, red), as well as the number of fused human myofibres (h-lamin A/C⁺spectrin⁺, red) and h-dystrophin (green) 30 days post-enugraftment. Graphs quantify the mean \pm s.e.m. number of fused human myofibres from multiple cross-sections along the length of the mdx-NSG tibialis anterior ($N=3$ mice engrafted with uncultured fetal tissue, $N=4$ mice engrafted with cultured fetal tissue and $N=4$ mice engrafted with hPSC-SMPC; Kruskal-Wallis ranks tests: $*P < 0.05$), and the maximum number of engrafted myofibres in a single cross-section (one-way ANOVA post-hoc Tukey: $*P < 0.05$). Scale bars, 200 μ m.

associated with supporting fetal muscle (Supplementary Table 3). Although evaluation of all mononuclear cells from directly isolated fetal muscle compared to cultured cells provided a comprehensive analysis of cells contained within the highest in vivo myogenic

potential (Supplementary Table 4), identification of potential surface markers was not feasible because of contaminating cell types (such as endothelial and haematopoietic cells) present in the unsorted fetal muscle cell samples.

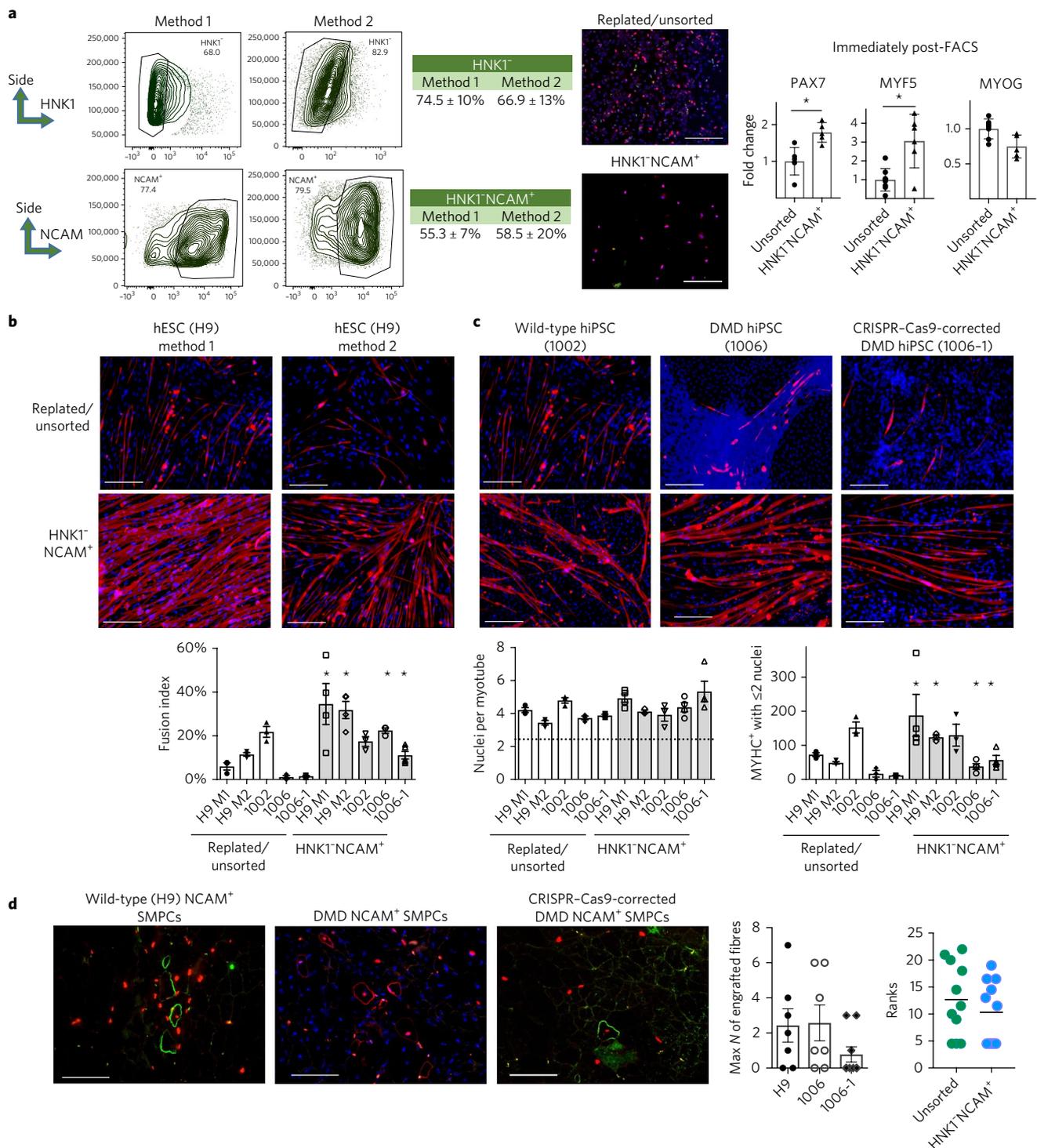


Fig. 2 | HNK1-NCAM⁺ increases myogenic cell numbers but does not increase myogenicity in vivo. **a**, HNK1-NCAM⁺ FACS-sorted hPSC-SMPCs have increased PAX7 (red) and MYF5 expression compared to replated/unsorted day 50 SMPCs by immunofluorescence and qPCR (method 1; *N* = 6 unsorted or *N* = 5 NCAM hPSC-SMPCs from independent directed differentiations, mean ± s.e.m., two-tailed *t*-test: **P* < 0.05). Immunofluorescence also shows staining for DAPI (blue). Scale bars, 100 μm. FACS plots show mean ± s.d. of HNK1⁻ and HNK1-NCAM⁺ as a percentage of live, non-doublet cells. **b**, HNK1-NCAM⁺ cells show increased myotube differentiation when hPSC-SMPCs are derived from two methods (M1, method 1; M2, method 2). **c**, hPSC myotubes differentiate independent of dystrophin expression. Immunofluorescence (**b,c**) shows staining for MYHC (red) and DAPI (blue). Scale bars (**b,c**), 200 μm. Graphs show quantification of myotube fusion from all hPSC lines (*N* = 3 from independent hPSC-myotube experiments, mean ± s.e.m., two-tailed *t*-test of NCAM versus unsorted for each hPSC line (**b,c**): **P* < 0.05). **d**, HNK1-NCAM⁺ wild-type, DMD and CRISPR-Cas9-corrected hPSC-SMPCs all engraft inefficiently in vivo. Immunofluorescence and quantification of h-lamin A/C⁺h-spectrin⁺ (red) and h-dystrophin (green) are shown. Graphs quantify the maximum number of engrafted myofibres in a single cross-section (left; mean ± s.e.m., two-tailed *t*-test: not significant (NS); *P* > 0.05)), and the mean ± s.e.m. of engrafted myofibres (*N* = 7 mice per group) from multiple cross-sections along the length of the muscle (right; Mann-Whitney *U*-test: NS (*P* > 0.05)). Scale bars, 100 μm.

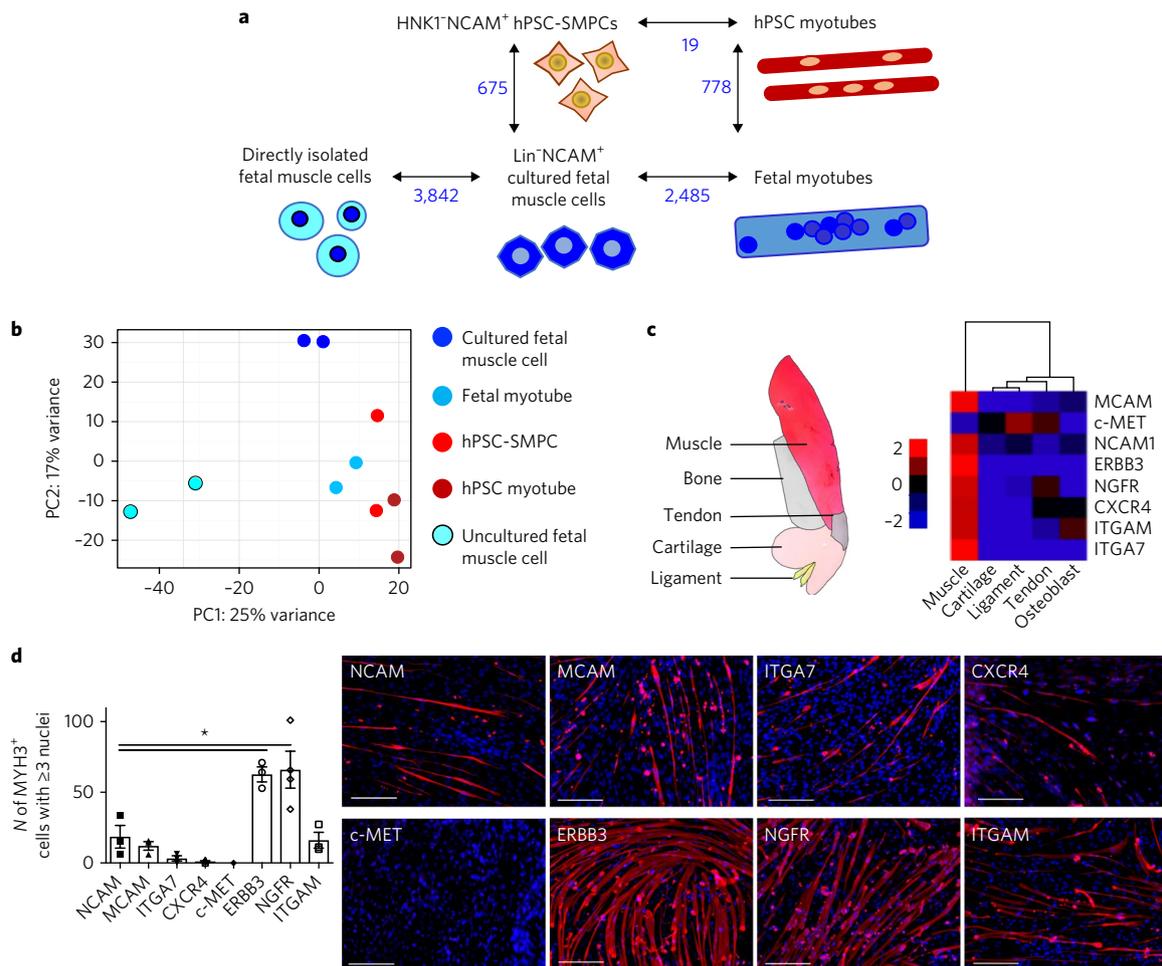


Fig. 3 | RNA-seq identifies unique gene signatures in fetal and hPSC-derived muscle. a, Differential gene expression using CuffDiff of all profiled cell types are shown ($N=2$, $q < 0.05$; blue). **b**, Principal component (PC) analysis of the five cell types. Gene lists and key biological processes that are upregulated in hPSC-SMPCs, cultured fetal or directly isolated fetal muscle cells are shown in Supplementary Tables 2–4 ($N=2$, $q < 0.05$). **c**, Illustration of profiled fetal musculoskeletal tissue and WGCNA identification of candidate cell surface receptors enriched in fetal muscle cells versus other musculoskeletal tissues are shown. **d**, Screen for fetal muscle receptors identifies ERBB3 and NGFR that enrich for hPSC-SMPCs. CRISPR-Cas9-corrected DMD hiPSC-SMPCs (1006-1) were sorted on eight candidate subpopulations and were fused in vitro and stained for MYHC (red) and DAPI (blue). Graph shows quantification of myotube differentiation ($N=3$ from independent hPSC-myotube experiments, mean \pm s.e.m., one-way ANOVA post-hoc Dunnett: $*P < 0.05$). Scale bars, 200 μ m.

Muscle tissue-specific surface markers enable enrichment of myogenic cells from hPSC-SMPCs. To identify putative surface markers enriched on myogenic cells versus other skeletogenic lineages, we used the results of a Weighted Gene Co-Expression Network Analysis (WGCNA) performed across five fetal musculoskeletal tissues: muscle, bone, cartilage, ligament and tendon (Fig. 3c; accession code: [GSE106292](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE106292)). In WGCNA, as modules of genes are defined based on their expression levels in each cell type, we chose candidate cell surface markers that were associated with myogenesis and/or enriched in the muscle module for further evaluation as possible markers of hPSC-SMPCs (Fig. 3c). We next screened candidate receptors for the ability to increase myogenic potential using the 1006-1 CRISPR-Cas9-corrected DMD hiPSC line. Of the populations tested, cells positive for NGFR and ERBB3 were the most myogenic and were able to form myotubes more efficiently than NCAM⁺ cells in vitro ($P < 0.05$; Fig. 3d). Thus, we chose NGFR and ERBB3 as two candidate surface markers to improve isolation of hPSC-SMPCs.

ERBB3 and NGFR surface marker levels demarcate a switch between primary and secondary fetal myogenesis and enrich for PAX7⁺ fetal muscle progenitor cells. RNA-seq identified two

markers specific to week 17 fetal muscle that were capable of enriching hPSC-SMPCs. Based on our data, hPSC-SMPCs most closely resemble weeks 8–9 (late primary) muscle progenitors in their functional capacity. We hypothesized that ERBB3⁺ and NGFR⁺ cells would be present throughout human developmental myogenesis. As early as fetal week 8, a myogenic population of ERBB3⁺NGFR⁺ cells were identified that persisted throughout the time points examined (Fig. 4a and Supplementary Fig. 2). At weeks 8–9, the ERBB3⁺NGFR⁺ population enriched for several myogenic transcription factors ($P < 0.05$, $N=3$). This is important because other known fetal markers, NCAM, MCAM (melanoma cell adhesion molecule, also known as CD146) and CD82, were either non-specific to muscle or not expressed at these time points, respectively. NCAM⁺MCAM⁺ and CD82 began to co-express with ERBB3⁺NGFR⁺ cells at fetal week 11.5 (Fig. 4 and Supplementary Fig. 2).

By fetal week 17, two distinct ERBB3 subpopulations were identified containing myogenic activity (Fig. 4b). ERBB3⁺NGFR⁺ cells were enriched for *PAX7* and *MYF5* by 8–10-fold compared to ERBB3⁻NGFR⁻ cells ($P < 0.01$, $N=5$), whereas ERBB3⁺NGFR⁻ cells were enriched for *MYOD* and *MYOG* by 40-fold ($P < 0.05$); NGFR^{low}ERBB3⁻ cells weakly expressed myogenic transcription

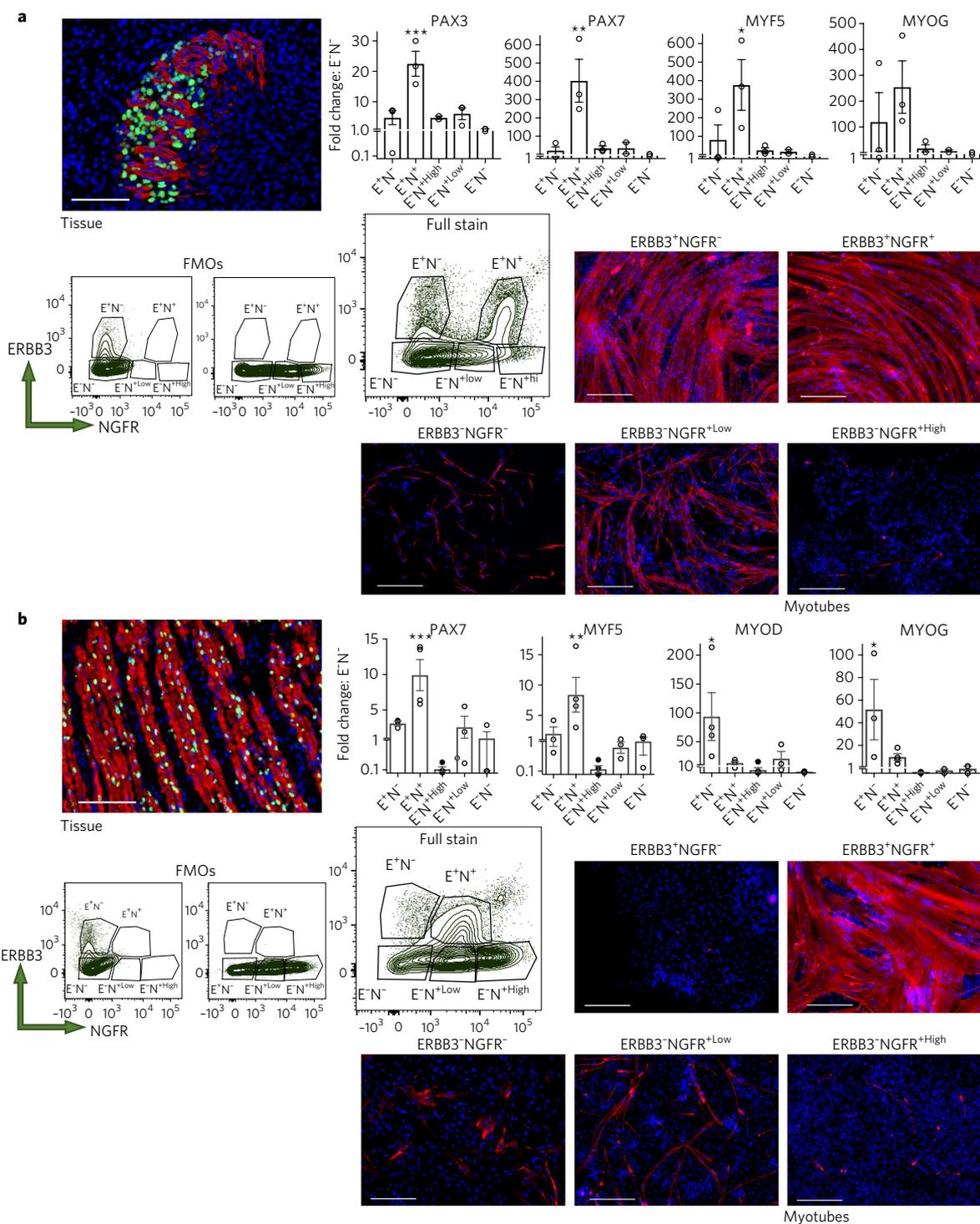


Fig. 4 | Increased myogenic ability resides in the ERBB3⁺NGFR⁺ fraction of human fetal muscle during primary and secondary myogenesis. a, b, Five subpopulations of human myogenic cells were identified based on ERBB3 and NGFR expression during primary myogenesis (fetal weeks 8–10) (**a**) and secondary myogenesis (fetal weeks 16–18) (**b**). Immunofluorescence of fetal week 9 skeletal muscle tissue (**a**) demonstrated the establishment of the limb myofibres, and week 17 tissue (**b**) demonstrated maturation to multinucleated myofibres; immunofluorescence shows staining for MYHC (red), PAX7 (green) and DAPI (blue). Scale bars of immunofluorescence images of skeletal muscle tissue, 100 μ m. FACS plots of isotype controls and full staining (live, non-doublet CD45⁻CD31⁻CD235a⁻ cells) show that ERBB3⁺NGFR⁺ populations lose NGFR^{High} positivity during the transition from primary to secondary myogenesis. Statistical analyses of all fetal populations identified by FACS are shown in Supplementary Fig. 2. Graphs show fold change of myogenic gene expression of immediately sorted subpopulations compared to ERBB3⁻NGFR⁻ cells ($N=3$ (weeks 8–10 tissue) and $N=5$ (weeks 16–18 tissue), mean \pm s.e.m., one-way ANOVA post-hoc Dunnett (increase): * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). Sorted populations were expanded in SkBM-2 for 24 h and were induced to differentiate to myotubes in N2 media for 5 days, highlighting differences in the in vitro myogenic potential between progenitors and subpopulations at different stages of human myogenesis. Myotubes were stained with MYHC (red) and DAPI (blue) ($N=3$ from independent hPSC-myotube experiments). Scale bars in immunofluorescence images of myotubes, 200 μ m.

factors, but did contain a population with some myogenic potential. Primary and secondary fetal muscle progenitors could also be distinguished by their positivity for ERBB3 and NGFR, as expression of NGFR was much higher during weeks 8–9 (Supplementary Fig. 2). Thus, the combination of these markers distinguished fetal muscle progenitors at different stages of development and differentiation. Upon sorting and differentiation to form myotubes, the ERBB3⁺NGFR⁺ fraction reached close to 100% fusion efficiency within 4 days, whereas NGFR^{Low}ERBB3⁻ generated few myotubes (Fig. 4b). ERBB3⁺NGFR⁻ cells did not proliferate or form myotubes. We concluded that fetal ERBB3⁺NGFR⁺ cells are specifically enriched for PAX7-expressing myogenic progenitors during fetal myogenesis.

ERBB3 and NGFR enrich for PAX7 and myogenic capacity in hPSC-SMPCs. Based on ERBB3 and NGFR expression on fetal myogenic progenitors, we evaluated the ability of these markers to enrich hPSC-SMPCs. In all directed differentiation protocols tested, an ERBB3⁺NGFR⁺ SMPC population was detected as early as day 27 (Supplementary Fig. 3). By 50 days of directed differentiation, both protocols produced two distinct ERBB3 subpopulations, distinguished by NGFR expression, that resembled secondary myogenesis (Fig. 5). Sorting these subpopulations revealed that ERBB3⁺NGFR⁺ hPSC-SMPCs enriched for PAX7 and MYF5 by 20-fold compared to ERBB3⁻NGFR⁻ cells ($P < 0.001$, $N = 4$), whereas ERBB3⁺NGFR⁻ cells were enriched for MYOD and MYOG ($P < 0.001$). As expected, ERBB3⁺NGFR⁺ SMPCs could be induced to form homogeneous myotubes. These data demonstrate a striking similarity between the surface marker phenotypes of fetal myogenic progenitors and hPSC-SMPCs undergoing directed differentiation.

To test the versatility of these markers, four hPSC lines, including DMD hiPSCs and isogenic CRISPR–Cas9-corrected DMD hiPSCs, were evaluated. Across all lines, myogenic potential was highly enriched in the ERBB3⁺NGFR⁺ subpopulation, as defined by the expression of myogenic transcription factors and the differentiation potential in vitro (Fig. 5). However, double positivity based on NGFR levels varied across cell lines. Of note, upon passaging, NGFR levels increase and are associated with decreased myogenic potential in fetal and hPSC-SMPCs. Thus, this previously unrecognized combination of surface markers informs on the SMPC differentiation state, in which the ERBB3⁺ fraction was consistently the most myogenic.

Transforming growth factor- β inhibition improves hPSC-SMPC differentiation and myotube fusion. Despite enabling myogenic enrichment, differentiation of ERBB3⁺NGFR⁺ populations isolated from hPSCs yielded myotubes that were often thinner and contained fewer nuclei than later-stage fetal or adult counterparts. To understand the basis for these differences, we further analysed the RNA-seq data comparing hPSC-myotubes and fetal myotubes to their progenitors. These data revealed that activators of transforming growth factor- β (TGF- β) signalling decreased during the course of fetal myotube differentiation (Fig. 6a). By contrast, hPSC-SMPCs had higher expression of TGF- β signalling genes (for example, *TGFB1*, *ACVR1B* and *MSTN*), and hPSC myotubes failed to downregulate TGF- β signalling (*MSTN*, *INHBA* and *TGFB2*) relative to fetal myotubes.

To test the function of TGF- β signalling in hPSC-derived myotube differentiation, we manipulated TGF- β during hPSC-SMPC to myotube formation. Compared to N2 media alone, the addition of TGF- β 1 completely inhibited myotube formation ($P < 0.05$). By contrast, TGF- β inhibition (TGF- β i) using SB-431542 or A83-01 significantly increased hPSC-myotube fusion and produced morphology that was similar to late-stage fetal myotubes (Fig. 6 and Supplementary Fig. 4). Single PAX7⁺ cells remained in hPSC-myotube cultures in the presence of either TGF- β 1 or TGF- β i (Fig. 6b), suggesting TGF- β i does not deplete PAX7⁺ SMPCs.

TGF- β regulation of hPSC myotube differentiation is independent of dystrophin expression. In two independent DMD and isogenic CRISPR–Cas9-corrected DMD hiPSC lines (1006 and 1003), the fusion indices of hPSC-SMPCs increased by 2–5-fold with TGF- β i compared to no treatment ($P < 0.05$; Supplementary Fig. 4). These data demonstrate that inhibition of improper TGF- β levels during hPSC-SMPC differentiation can improve myogenesis in vitro.

TGF- β i promotes hPSC-myotube maturation in vitro. In mouse, TGF- β signalling is a potent negative regulator of fetal muscle maturity^{22,37}; however, TGF- β has never been evaluated in human development. To test whether ectopic TGF- β signalling inhibited hPSC-SMPCs from forming more mature myotubes, we measured embryonic (MYH3), fetal (MYH8), and adult (MYH1) myosins at both the RNA and the protein levels³⁸. We found that without TGF- β i, hPSC-myotubes predominately expressed embryonic and fetal myosins. Although ERBB3⁺NGFR⁺ enrichment increased the expression of embryonic and fetal myosins relative to NCAM sorting ($P < 0.01$), enrichment did not increase adult myosin expression (Supplementary Fig. 5).

Upon TGF- β i, the expression of all myosins increased in hPSC-myotubes, and the expression of *MYH8* and *MYH1* in ERBB3⁺NGFR⁺ hPSC-myotubes was similar to or higher than that of fetal myotubes, respectively (Supplementary Fig. 5). Western blot analysis confirmed that TGF- β i also increased MYH8 and MYH1 protein levels in hPSC-myotubes relative to adult myotubes ($P < 0.05$; Fig. 6c and Supplementary Fig. 7). Interestingly, TGF- β i was more effective at inducing myosin expression in hPSC-SMPCs than in fetal or adult myogenic progenitors, which is in agreement with mouse embryonic and fetal myogenic progenitors²².

As another measure of maturation, we evaluated ultrastructural changes in hPSC-myotube morphology relative to fetal myotubes by transmission electron microscopy. In the absence of TGF- β i, NGFR⁺ hPSC-myotubes displayed disorganized sarcomere patterning (Fig. 6d), whereas TGF- β i promoted the formation of better-organized sarcomeres and Z-disk patterning. By contrast, untreated fetal myotubes showed limited organization. These data demonstrate that TGF- β i facilitates hPSC-myotube maturity and additively increases the myogenicity of enriched hPSC-SMPC subpopulations.

In vivo engraftment of ERBB3⁺ hiPSC-SMPCs restores dystrophin to levels approaching uncultured fetal muscle. We hypothesized that the lack of engraftment by NCAM⁺ hPSC-SMPCs was due to the immature nature and heterogeneity of transplanted cells. Thus, we tested whether ERBB3 or NGFR enrichment could improve engraftment in mdx-NSG mice. Single-marker enrichment strategies using CRISPR–Cas9-corrected DMD hiPSC lines²⁵ were tested, as these cell lines may have more therapeutic relevance. Compared to NCAM⁺ sorting, both NGFR⁺ and ERBB3⁺ SMPCs significantly increased the number of engrafted myofibres as shown by h-lamin A/C⁺spectrin⁺dystrophin⁺ positivity ($P < 0.05$; Supplementary Table 1). These data demonstrate that ERBB3⁺ SMPCs are strongly enriched for cells that are capable of generating muscle in vivo.

As shown in Fig. 6, myotubes generated from enriched SMPCs are immature due to precocious TGF- β signalling, and TGF- β i additively increased SMPC myogenicity. Co-delivery of TGF- β i during injection and for 2 weeks following transplantation increased the numbers of h-lamin A/C⁺spectrin⁺dystrophin⁺ myofibres in all SMPCs subpopulations tested (Supplementary Table 1). Mice injected with ERBB3⁺ hiPSC-SMPCs and treated with TGF- β i, improved the maximum number of h-dystrophin⁺ fibres to 137 ± 22 per cross-section (Fig. 7a). Across all points measured, this represented a 50-fold greater engraftment than NCAM⁺ cells that were also treated with TGF- β i ($P < 0.001$). A percentage of h-lamin A/C⁺

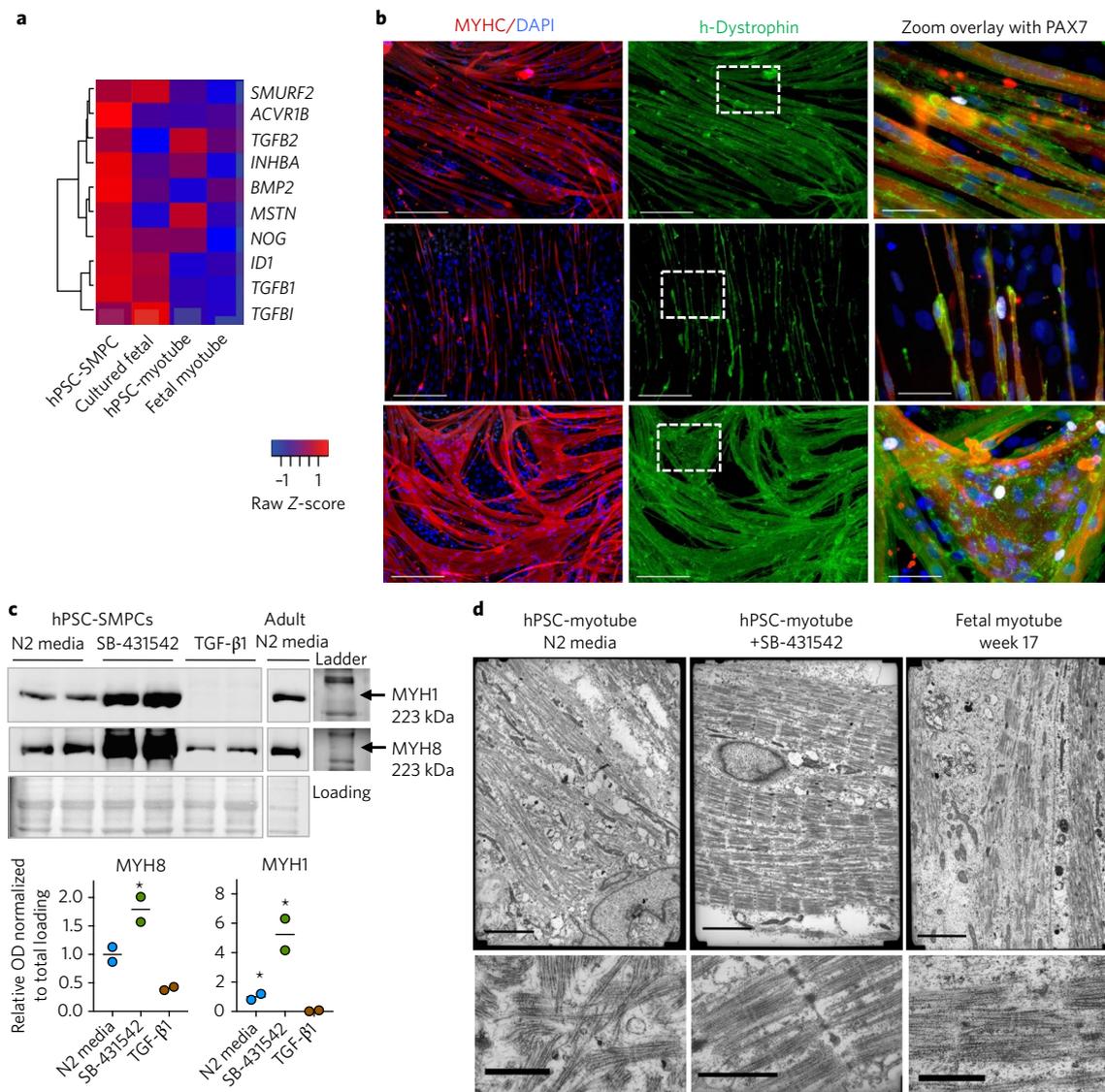


Fig. 6 | Inhibition of TGF- β signalling induces hPSC skeletal muscle maturation. **a**, Heatmap of TGF- β signalling genes reveals high levels of TGF- β in hPSC myotubes as determined by RNA-seq FPKM (fragments per kilobase of transcript per million mapped reads) values. **b**, ERBB3⁺ SMPCs from 1006-1 CRISPR-Cas9-corrected DMD hiPSCs were differentiated in N2 media alone (top row), with recombinant TGF- β 1 (middle row) or in the presence of the TGF- β inhibitor SB-431542 (bottom row; $N=6$). hPSC-myotubes were stained with MYHC, h-dystrophin, PAX7 and DAPI. Scale bars, 200 μ m. Zoom images show no change in PAX7 expression with either treatment, but an increase in myotube formation was seen with TGF- β 1. **c**, Western blot analysis showing that SB-431542-treated hPSC-SMPCs have an increased expression of MYH1 and MYH8 compared to N2 media alone or after treatment with recombinant TGF- β 1 ($N=2$ from independent myotube experiments; Student's t -test: $*P < 0.05$). Ponceau S-stained gel is shown for loading control. Ladder and unprocessed blots are included in Supplementary Fig. 7. OD, optical density. **d**, Transmission electron microscopy of fetal week 17, or NGFR⁺ hPSC-myotubes with or without TGF- β 1 demonstrates increased sarcomere organization and Z-disk patterning after SB-431542 treatment (from two independent myotube experiments). Scale bars, 3 μ m (top) and 1.67 μ m (bottom).

hPSC-SMPCs, NCAM, is expressed on many cell types during development, including neural cells, muscle progenitors, myotubes and other mesodermal progenitors⁴¹, and is not sufficient for progenitor enrichment in directed differentiation cultures. Enrichment of muscle progenitors using the surface markers identified in this study and targeting dysregulated signal pathways identified in RNA-seq data, enabled the maturation of hPSC-SMPCs with significantly greater myogenic potential in vitro and in vivo.

As we found that hPSC-SMPCs are embryonic, we evaluated progenitors from primary and secondary human myogenesis to identify more relevant enrichment markers. We demonstrated that ERBB3⁺NGFR⁺ expression enriches for PAX7⁺ and MYF5⁺ cells during the first and second trimesters of human development,

including at time points when other surface markers cannot distinguish fetal muscle progenitors⁴². Shifts in ERBB3⁺NGFR⁺ expression levels reliably demarcated transitions between early and late waves of human fetal myogenesis, and correlated with the establishment of primary limb myofibres or maturation of secondary fetal myofibres. ERBB3⁺NGFR⁺ expression was also able to demarcate progenitors from more-differentiated MYOD⁺ cells; thus, expression levels of this marker combination may serve as an important tool for studying human SMPC biology. Upon searching a recent fetal data set, we found that ERBB3 is also expressed by human fetal MCAM⁺ cells⁴². Likewise, we found that these markers enrich for myogenic SMPCs from heterogeneous hPSC cultures across multiple directed differentiation protocols^{10,11,13,19}.

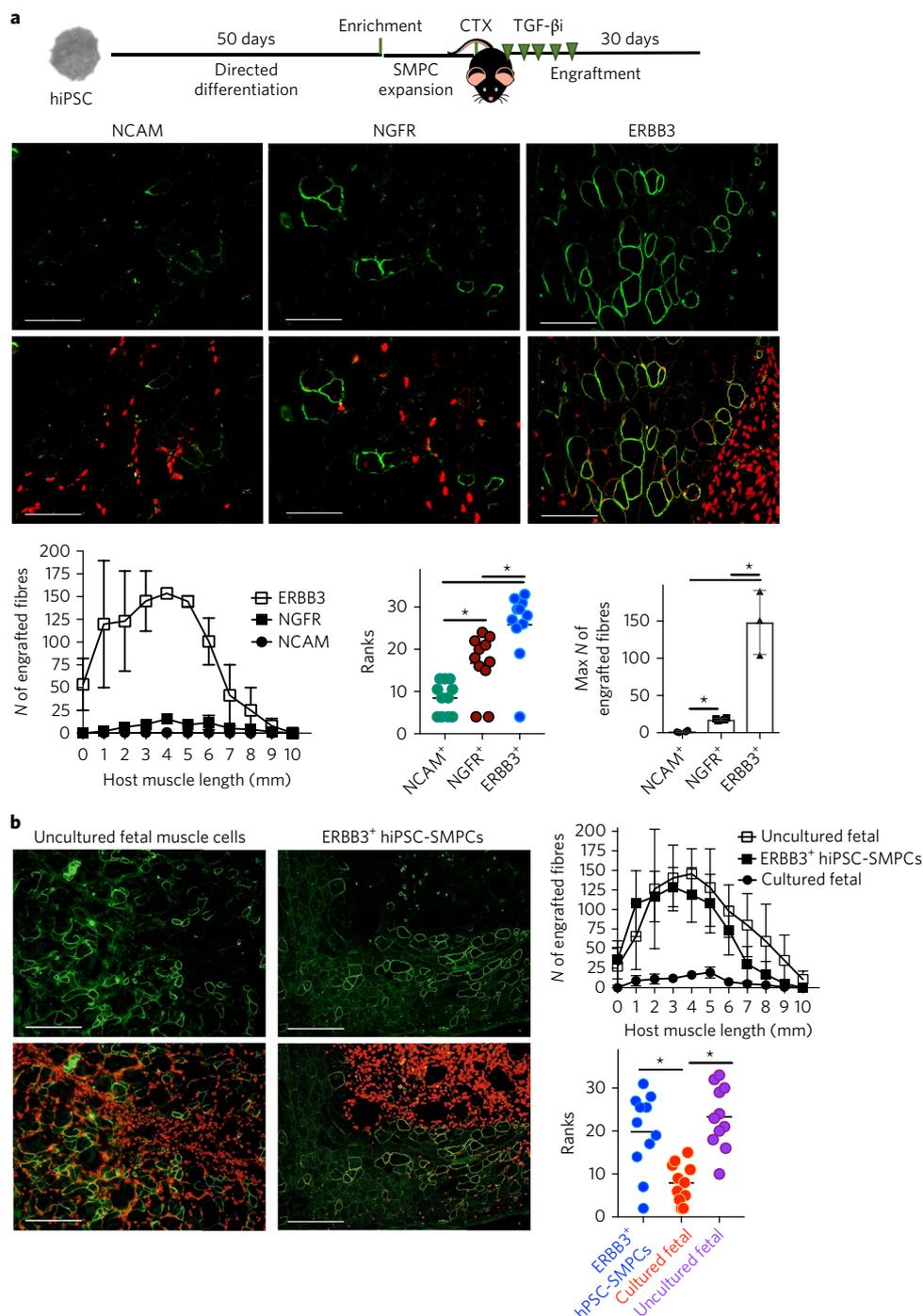


Fig. 7 | In vivo engraftment of CRISPR-Cas9-corrected DMD hiPSC-SMPCs restores dystrophin to levels approaching uncultured fetal muscle.

a, CRISPR-Cas9-corrected DMD hiPSC-SMPCs were enriched for surface markers, and upon engraftment, were co-delivered with the TGF-β inhibitor SB-431542. Immunofluorescence of h-lamin A/C (red) denotes human cells, and h-spectrin (red) and h-dystrophin (green) denote areas fused with mdx-NSG muscle fibres 30 days post-enugraftment. Graphs show the average ± s.e.m. number of engrafted human plus myofibres from multiple cross-sections along the length of the muscle (N = 3 NCAM, N = 4 NGFR and N = 4 ERBB3 independent enriched SMPCs from directed differentiation engrafted into mice per group; Kruskal-Wallis ranks tests: *P < 0.05), and the maximum number of engrafted myofibres in a single cross-section (one-way ANOVA post-hoc Tukey: *P < 0.05). Scale bars, 100 μm. CTX, cardiotoxin. **b**, CRISPR-Cas9-corrected ERBB3⁺ SMPCs plus TGF-βi engraft equivalent to uncultured (directly isolated) fetal muscle cells. Graphs show the mean ± s.e.m. number of engrafted myofibres from multiple cross-sections along the length of the muscle (N = 3 fetal SMPCs and N = 4 ERBB3⁺ independent enriched SMPCs from directed differentiation engrafted mice per group; Kruskal-Wallis ranks tests: *P < 0.05). Scale bars, 200 μm.

Although ERBB3 has been identified in the initial phases of adult mouse SC activation^{43,44}, the mechanism by which ERBB3 regulates muscle progenitors during human fetal myogenesis prior to adult

SC generation or quiescence, and in hPSC-SMPCs, is unknown. In mouse development, deficiencies in neuregulin-1 signalling, the primary ligand of ERBB3, result in loss of self-renewal of

fetal muscle progenitors⁴⁵ and suggest that ERBB3 signalling may regulate the balance of fetal myogenic cell fate by preventing precocious SMPC differentiation. We hypothesize that neuregulin-1 and ERBB3 signalling may also have important roles in specification and support of PAX7 cells, including in hPSC-SMPC cultures. In mouse studies, muscle progenitor cells expressing ERBB3 or NGFR are responsive to paracrine signals secreted by neural crest or neuronal cells^{45,46}. During hPSC directed differentiation, SMPCs may be supported by a local microenvironment that includes three-dimensional cellular structures and neuronal cells⁴⁷. As several biological processes associated with neuronal cells and transmission of neural impulses were identified by RNA-seq in hPSC-SMPC cultures, we expect that neuronal cells are important for hPSC myogenesis throughout directed differentiation.

A critical bottleneck to all hPSC directed differentiation protocols is the need to mature progenitors towards functionally relevant cell types. In mouse development, activation of TGF- β signalling prevents embryonic muscle progenitor cells from undergoing secondary myogenesis^{37,48}. In a similar manner, we found that several members of the TGF- β superfamily were dysregulated in hPSC-myotubes compared to human fetal myotubes. Our data provide multiple lines of evidence that TGF- β 1 is a major driver of hPSC maturation towards secondary or tertiary myogenesis. Additional strategies or factors that promote SMPC transition to SCs will be required, as TGF- β 1 is only important for enhancing the maturation of differentiated myotubes in vitro or myofibres in vivo.

DMD is a devastating muscle disease caused by out-of-frame mutations in the gene encoding dystrophin. We have shown that CRISPR-Cas9-mediated genetic deletion of DMD hiPSCs can restore the dystrophin reading frame²⁵. However, targeting muscle stem cells with CRISPR-Cas9 in vivo is inefficient⁴⁹. Correction of patient-derived hiPSCs ex vivo and subsequent differentiation to a human SC equivalent offers one potential route for the delivery of SMPCs resulting in dystrophin restoration after transplantation. We have shown that directly isolated human fetal muscle cells can restore several hundred dystrophin⁺ myofibres to levels speculated to be required for functional clinical gain^{29–31}. We also demonstrated proof of concept that CRISPR-Cas9-corrected DMD hiPSC-SMPCs can be directed to differentiate and restore dystrophin in mdx-NSG mice to levels approaching those of directly isolated fetal cells. It will be important to understand differences in PAX7⁺ cell states during development, as fetal cells maintain PAX7 expression independent of the adult SC niche²³. Identifying ways to expand and support PAX7⁺ cells from fetal, adult or hPSC-SMPCs without activating MYOD, leading to loss of self-renewal, will be required to develop a cell replacement therapy for diseases including DMD⁵⁰.

In summary, directed differentiation protocols that more closely model human development may better recapitulate cell-specific lineages and, therefore, lead to better in vitro models or improved clinical use. We used human fetal myogenesis to identify previously unreported candidates for enriching and maturing hPSCs myogenic activity in vitro and in vivo. Furthering our understanding of human developmental myogenesis will provide inroads to regenerative approaches for muscle diseases including in combination with gene-edited patient-derived hiPSCs.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at <https://doi.org/10.1038/s41556-017-0010-2>.

Received: 14 July 2017; Accepted: 16 November 2017;
Published online: 18 December 2017

References

- Chong, J. J. H. et al. Human embryonic-stem-cell-derived cardiomyocytes regenerate non-human primate hearts. *Nature* **510**, 273–277 (2014).
- Schwartz, S. D. et al. Human embryonic stem cell-derived retinal pigment epithelium in patients with age-related macular degeneration and Stargardt's macular dystrophy: follow-up of two open-label phase 1/2 studies. *Lancet* **385**, 509–516 (2014).
- Steinbeck, J. A. et al. Optogenetics enables functional analysis of human embryonic stem cell-derived grafts in a Parkinson's disease model. *Nat. Biotechnol.* **33**, 204–209 (2015).
- Hrvatin, S. et al. Differentiated human stem cells resemble fetal, not adult, β cells. *Proc. Natl Acad. Sci. USA* **111**, 3038–3043 (2014).
- Witty, A. D. et al. Generation of the epicardial lineage from human pluripotent stem cells. *Nat. Biotechnol.* **32**, 1026–1035 (2014).
- Yang, X., Pabon, L. & Murry, C. E. Engineering adolescence: maturation of human pluripotent stem cell-derived cardiomyocytes. *Circ. Res.* **114**, 511–523 (2014).
- Jonsson, M. K. et al. Application of human stem cell-derived cardiomyocytes in safety pharmacology requires caution beyond hERG. *J. Mol. Cell. Cardiol.* **52**, 998–1008 (2012).
- Borchin, B., Chen, J. & Barberi, T. Derivation and FACS-mediated purification of PAX3⁺/PAX7⁺ skeletal muscle precursors from human pluripotent stem cells. *Stem Cell Rep.* **21**, 620–631 (2013).
- Xu, C. et al. A zebrafish embryo culture system defines factors that promote vertebrate myogenesis across species. *Cell* **155**, 909–921 (2013).
- Chal, J. et al. Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. *Nat. Biotechnol.* **33**, 962–969 (2015).
- Shelton, M. et al. Derivation and expansion of PAX7⁺-positive muscle progenitors from human and mouse embryonic stem cells. *Stem Cell Rep.* **3**, 516–529 (2014).
- Chal, J. et al. Generation of human muscle fibers and satellite-like cells from human pluripotent stem cells in vitro. *Nat. Protoc.* **11**, 1833–1850 (2016).
- Swartz, E. W. et al. A novel protocol for directed differentiation of C9orf72-associated human induced pluripotent stem cells into contractile skeletal myotubes. *Stem Cells Transl. Med.* **5**, 1461–1472 (2016).
- Abujarour, R. et al. Myogenic differentiation of muscular dystrophy-specific induced pluripotent stem cells for use in drug discovery. *Stem Cells Transl. Med.* **3**, 149–160 (2014).
- Kimura, E. et al. Cell-lineage regulated myogenesis for dystrophin replacement: a novel therapeutic approach for treatment of muscular dystrophy. *Hum. Mol. Genet.* **17**, 2507–2517 (2008).
- Darabi, R. et al. Human ES- and iPSC-derived myogenic progenitors restore dystrophin and improve contractility upon transplantation in dystrophic mice. *Cell Stem Cell* **10**, 610–619 (2012).
- Darabi, R. et al. Functional skeletal muscle regeneration from differentiating embryonic stem cells. *Nat. Med.* **14**, 134–143 (2008).
- Sambasivan, R. & Tajbakhsh, S. Skeletal muscle stem cell birth and properties. *Semin. Cell Dev. Biol.* **18**, 870–882 (2007).
- Xi, H. et al. In vivo human somitogenesis guides somite development from hiPSCs. *Cell Rep.* **18**, 1573–1585 (2017).
- Loh, K. M. et al. Mapping the pairwise choices leading from pluripotency to human bone, heart, and other mesoderm cell types. *Cell* **166**, 451–467 (2016).
- Yin, H., Price, F. & Rudnicki, M. A. Satellite cells and the muscle stem cell niche. *Physiol. Rev.* **93**, 23–67 (2013).
- Biressi, S. et al. Myf5 expression during fetal myogenesis defines the developmental progenitors of adult satellite cells. *Dev. Biol.* **379**, 195–207 (2013).
- Tierney, M. T. et al. Autonomous extracellular matrix remodeling controls a progressive adaptation in muscle stem cell regenerative capacity during development. *Cell Rep.* **14**, 1940–1952 (2016).
- Sacco, A., Doyonnas, R., Kraft, P., Vitorovic, S. & Blau, H. M. Self-renewal and expansion of single transplanted muscle stem cells. *Nature* **456**, 502–506 (2008).
- Young, C. S. et al. A single CRISPR-Cas9 deletion strategy that targets the majority of DMD patients restores dystrophin function in hiPSC-derived muscle cells. *Cell Stem Cell* **18**, 533–540 (2016).
- Tierney, M. T. & Sacco, A. Satellite cell heterogeneity in skeletal muscle homeostasis. *Trends Cell Biol.* **26**, 434–444 (2016).
- Castiglioni, A. et al. Isolation of progenitors that exhibit myogenic/osteogenic bipotency in vitro by fluorescence-activated cell sorting from human fetal muscle. *Stem Cell Rep.* **2**, 92–106 (2014).
- Montarras, D. et al. Direct isolation of satellite cells for skeletal muscle regeneration. *Science* **309**, 2064–2067 (2005).
- Cerletti, M. et al. Highly efficient, functional engraftment of skeletal muscle stem cells in dystrophic muscles. *Cell* **134**, 37–47 (2008).
- Godfrey, C. et al. How much dystrophin is enough: the physiological consequences of different levels of dystrophin in the mdx mouse. *Hum. Mol. Genet.* **24**, 4225–4237 (2015).

31. Sharp, P. S., Bye-a-Jee, H. & Wells, D. J. Physiological characterization of muscle strength with variable levels of dystrophin restoration in mdx mice following local antisense therapy. *Mol. Ther.* **19**, 165–171 (2011).
32. Chan, Y. & Walmsley, R. P. Learning and understanding the Kruskal–Wallis one-way analysis-of-variance-by-ranks test for differences among three or more independent groups. *Phys. Ther.* **77**, 1755–1762 (1997).
33. Choi, I. Y. et al. Concordant but varied phenotypes among Duchenne muscular dystrophy patient-specific myoblasts derived using a human iPSC-based model. *Cell Rep.* **15**, 2301–2312 (2016).
34. Xu, X. et al. Human satellite cell transplantation and regeneration from diverse skeletal muscles. *Stem Cell Rep.* **5**, 419–434 (2015).
35. Cahan, P. & Daley, G. Q. Origins and implications of pluripotent stem cell variability and heterogeneity. *Nat. Rev. Mol. Cell Biol.* **14**, 357–368 (2013).
36. Osafune, K. et al. Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat. Biotechnol.* **26**, 313–315 (2008).
37. Cusella-De Angelis, M. G. et al. Differential response of embryonic and fetal myoblasts to TGF beta: a possible regulatory mechanism of skeletal muscle histogenesis. *Development* **120**, 925–933 (1994).
38. Schiaffino, S., Rossi, A. C., Smerdu, V., Leinwand, L. A. & Reggiani, C. Developmental myosins: expression patterns and functional significance. *Skelet. Muscle* **5**, 22 (2015).
39. Pagliuca, F. W. et al. Generation of functional human pancreatic β cells in vitro. *Cell* **159**, 428–439 (2014).
40. Maroof, A. M. et al. Directed differentiation and functional maturation of cortical interneurons from human embryonic stem cells. *Cell Stem Cell* **12**, 559–572 (2013).
41. Evseenko, D. et al. Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. *Proc. Natl Acad. Sci. USA* **107**, 13742–13747 (2010).
42. Alexander, M. S. et al. CD82 is a marker for prospective isolation of human muscle satellite cells and is linked to muscular dystrophies. *Cell Stem Cell* **19**, 800–807 (2016).
43. Figeac, N., Serralbo, O., Marcelle, C. & Zammit, P. S. ErbB3 binding protein-1 (Ebp1) controls proliferation and myogenic differentiation of muscle stem cells. *Dev. Biol.* **386**, 135–151 (2014).
44. Golding, J. P., Calderbank, E., Partridge, T. A. & Beauchamp, J. R. Skeletal muscle stem cells express anti-apoptotic ErbB receptors during activation from quiescence. *Exp. Cell Res.* **313**, 341–356 (2007).
45. Van Ho, A. T. et al. Neural crest cell lineage restricts skeletal muscle progenitor cell differentiation through neuregulin1–ErbB3 signaling. *Dev. Cell* **21**, 273–287 (2011).
46. Deponti, D. et al. The low-affinity receptor for neurotrophins p75^{NTR} plays a key role for satellite cell function in muscle repair acting via RhoA. *Mol. Biol. Cell* **20**, 3620–3627 (2009).
47. Esteves de Lima, J., Bonnin, M. A., Birchmeier, C. & Duprez, D. Muscle contraction is required to maintain the pool of muscle progenitors via YAP and NOTCH during fetal myogenesis. *eLife* **5**, e15593 (2016).
48. Manceau, M. et al. Myostatin promotes the terminal differentiation of embryonic muscle progenitors. *Genes Dev.* **22**, 668–681 (2008).
49. Arnett, A. L. et al. Adeno-associated viral (AAV) vectors do not efficiently target muscle satellite cells. *Mol. Ther. Methods Clin. Dev.* **1**, 14038 (2014).
50. Quarta, M. et al. An artificial niche preserves the quiescence of muscle stem cells and enhances their therapeutic efficacy. *Nat. Biotechnol.* **34**, 752–759 (2016).

Acknowledgements

We thank S. Younesi, J. Marshall, M. Emami, E. Korsakova, K. Saleh, V. Rezek, J. Wen and C. Kumagai-Cresse for helping with stem cell culture and mouse experiments. Furthermore, we thank E. Mokhonova for performing the western blots. We also thank J. Morgan's laboratory for training on cell engraftments. The following cores were used: CDMD Muscle Phenotyping and Imaging Core, High Throughput and Cell Repository Core, and Bioinformatics and Genomics Core; the Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research at UCLA (BSCRC) Flow Cytometry Core; the UCLA Technology Center for Genomics and Bioinformatics, JCCC Electron Microscopy Core, the CFAR Flow Cytometry Core (NIH P30CA016042, 5P30AI028697); and the UCLA Humanized Mouse Core (CFAR, NIAID AI028697). M.R.H. is the recipient of a BSCRC and Schaffer Fellowship, a CDMD–Cure Duchenne Fellowship and a CDMD–NIH Paul Wellstone Center Training Fellowship (U54 AR052646). D.E. was funded by an NIH grant K01AR061415, Department of Defense (DoD) grant W81XWH-13-1-0465 and California Institute for Regenerative Medicine (CIRM) grant RB5-07230. Funding was provided by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) R01AR064327 to A.D.P., and NIAMS 5P30AR05723 to M.J.S., the CDMD at UCLA, the NIH/NCATS, the UCLA CTSI (UL1TR000124), the BSCRC Research Award (A.D.P.), the Rose Hills Foundation Research Award to A.D.P., and a CIRM Inception and CIRM Quest (DISC1-08823 and DISC2-08824 to A.D.P.).

Author contributions

M.R.H., B.V.H. and A.D.P. conceived the study. M.R.H., B.V.H. and M.J.S. designed the experiments. A.E., B.V.H., S.F.N. and M.R.H. analysed the data. M.R.H., B.V.H., J.H., K.P., W.F., M.J., H.X. and C.S.Y. conducted the experiments. M.R.H. and A.D.P. wrote the manuscript. M.R.H., A.D.P., B.V.H., J.H., K.P., C.S.Y., M.J.S. and S.F.N. edited the manuscript. Funding acquisition, M.R.H., A.D.P. and D.E.; Resources, D.E.; A.D.P. supervised the study.

Competing interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41556-017-0010-2>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to A.D.P.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

hPSC lines and cell culture techniques. All hPSC work was approved by the Embryonic Stem Cell Research Oversight (ESCRO) Committee. H9 hESCs were obtained from WiCell under material transfer agreement (MTA). Fibroblasts were taken from patient skin biopsies at the Center for Duchenne Muscular Dystrophy (CDMD) and were reprogrammed to derive 1002 (wild type), 1006 (DMD) and 1003 (DMD) hiPSC lines, as previously described²⁵. To establish a corrected DMD hiPSC (1006-1 or 1003-49), CRISPR–Cas9 gene-editing was used to remove exons 45–55 and restore dystrophin expression, as previously described²⁵. Four independent hPSC lines and two isogenic CRISPR–Cas9-corrected hiPSC lines were used to replicate all findings.

hPSCs were grown and maintained on hESC-qualified matrigel-coated plates in mTESR medium (Stem Cell Technologies) containing 0.4% penicillin/streptomycin (P/S; Hyclone), as previously described^{19,25}. All cell lines were tested for mycoplasma and were negative for mycoplasma contamination.

Directed differentiation protocols to generate SMPCs. Method 1 (ref. ¹¹) and method 2 (ref. ¹⁰) were selected to evaluate the developmental and functional status of hPSC-SMPCs. Method 1 is used for all figures throughout the paper. Method 2 was included in Figs. 1, 2 and 5 and Supplementary Fig. 3. Other directed differentiation protocols^{13,19} also confirmed the effects of TGF- β inhibition on myotube differentiation (SB-431542; 10 μ m) and ERBB3+NGFR+ enrichment of hPSC-SMPCs.

Myogenic gene expression of *PAX3*, *PAX7*, *MYF5*, *MYOD*, *MYOG* and *MYH3* were measured at multiple time points by quantitative PCR (qPCR) (Supplementary Fig. 1). Day 50 resulted in the greatest myogenic gene expression for both methods. By day 50 of method 1, muscle cell contraction was observed in all four independent cell lines tested. Directed differentiation using method 2 generated cells expressing skeletal muscle markers by day 27 (Fig. 5); however, to keep comparisons consistent, day 50 was used for analyses.

Dissociation of skeletal muscle cells from fetal and adult tissues. Human adult skeletal muscle between 25 and 67 years of age were obtained from donor autopsy of quadriceps provided by the National Disease Research Interchange (NDRI), and human fetal skeletal muscle weeks 8–18 were obtained from Novogenix, Inc. or the UCLA Center for AIDS Research (CFAR) Gene and Cellular Therapy Core using institutional review board (IRB)-approved de-identified and consented human fetal tissues. Use of human tissues was IRB exempt by the UCLA Office of the Human Research Protection Program (IRB no. 15-000959). Skeletal muscles were dissected from quadriceps, but whole limbs were taken from human fetal tissue weeks 8–9, as the musculature at this stage is not well established. Muscle was cleaned in DMEM/F12 media that contained Ca²⁺ and Mg²⁺, 0.4% P/S, 0.2% Primocin (Invivogen) and minced in digestion buffer containing 500 U per ml collagenase II (Worthington). Muscle tissue were incubated in digestion buffer at 37 °C, 5% CO₂ and titrated every 10 min until dissociated cells could pass through p1000 pipettes. Digested tissues were diluted in PBS and filtered through 100 μ m meshes to remove debris. Tissue were centrifuged at 600 g for 5 min, and re-suspended in sterile PBS containing 2% fetal bovine serum (FBS) for cell counting.

Analyses of hPSC-SMPCs from directed differentiation or muscle progenitors from tissues in vitro. hPSC-SMPCs from day 50 cultures were dissociated using TrypLE (Life Technologies) and gently titrated to break apart large three-dimensional structures. hPSC-SMPCs were then filtered through 100 μ m meshes to acquire single cells. hPSC-SMPCs from methods 1 and 2 were pelleted and seeded at 100,000 cells per well in 24-well plates in SkBM-2.

SCs from human adult quadriceps represented only 1–2% of all cells. Thus, adult SCs were sorted for CD45⁺CD31⁺NCAM⁺, as previously described¹⁴. For FACS gating, adult cells were first sorted on 4,6-diamidino-2-phenylindole (DAPI)⁺ to remove dead cells. Forward scatter and side scatter were used to remove debris and doublets. SCs were sorted using BD FACSAriaII cell sorters in the Jonsson Comprehensive Cancer Center (JCCC) at UCLA. SCs were collected in SkBM-2 and plated at 5,000–25,000 cells per 24-well plates. SCs were grown to 70% confluence in SkBM-2.

hPSC-SMPCs, fetal and adult muscle cells were induced to differentiate in N2 media for 7 days. To evaluate muscle progenitor or SC ability to differentiate and fuse to form myotubes, cultures were immunostained with MYHC, MYOD and PAX7, or isotype controls. Primary end point measurements included: (1) the number of myotubes per square millimetre defined as MYHC⁺ cells containing ≥ 3 nuclei, (2) the number of myocytes per square millimetre defined as MYHC⁺ cells containing ≤ 2 nuclei, (3) the number of nuclei contained within each myotube or myocyte, and (4) the fusion index defined as nuclei in myotubes as a percentage of the total nuclei population. At least three random $\times 10$ images (0.621 mm²) were taken per well, and images were used to count end point measurements using ImageJ. Image quantification data were then evaluated using one-way analysis of variance (ANOVA) with post-hoc Tukey to calculate significance ($P < 0.05$).

Assessment of population heterogeneity and timing of cell sorting to improve hPSC-SMPC directed differentiation and expansion. To identify lineage heterogeneity, candidate markers were evaluated between days 35 and 50 of

directed differentiation (method 1) and between days 24 and 50 (method 2). FACS analysis was used to assess mesenchymal, neuronal, neural crest, cardiac muscle, haematopoietic and endothelial cell lineages via the expression of CD73 and platelet-derived growth factor receptor- α (PDGFR α), N-cadherin, HNK1, vascular endothelial growth factor receptor 2 (VEGFR2), CD45 and CD31, respectively. Of these markers, HNK1⁺ neural crest cells were most abundant (Fig. 2a), and HNK1⁺ cells were subsequently removed from all experiments, including Figs. 2–7 and Supplementary Figs. 3–6.

To identify the most appropriate time point to enrich hPSC-SMPCs, a pilot study that contained three independent wells of FACS-sorted HNK1⁺NCAM⁺ SMPCs from days 35, 40 and 50 (method 1), and from days 24, 30 and 50 (method 2) was conducted. For method 1, sorted cells at day 50 produced the greatest number of MYHC⁺ cells, and for method 2, days 30 and 50 produced the greatest number of MYHC⁺ cells.

To evaluate PAX7 expression, sorted hPSC-SMPCs were seeded on 12-mm round bottom coverslips (GG-12-PLL, Neuvitro) coated with poly-L-lysine and matrigel. Sorted hPSC-SMPCs were allowed to settle for 3 h and then fixed and immunostained for PAX7 or mouse IgG1 isotype. Alternatively, cells were collected in SkBM-2 and immediately lysed for qPCR, and PAX7 gene expression was measured (Fig. 2a).

Use of growth factors to improve hPSC-SMPC directed differentiation or after sorting. To improve directed differentiation, candidate growth factors FGF2 (20 ng/ml), HGF (200 ng/ml), insulin-like growth factor 1 (IGF1; 100 ng/ml), and combinations were tested⁵¹. At day 50, myogenic genes *PAX7*, *MYF5*, *MYOD*, *MYOG* and *MYH3* were measured by qPCR and expressed as fold change versus day 35 of directed differentiation. The addition of IGF1, but not HGF, increased the expression of myogenic markers. Subsequently, IGF1 was added to N2 media in all experiments (Figs. 2–7 and Supplementary Figs. 3–6). We found that SkBM-2 media (containing epidermal growth factor (EGF), insulin, dexamethasone and 10% FBS) was better able to support PAX7⁺ cells after passaging. The addition of FGF2 in SkBM-2 further helped to decrease *MYOD* and *MYOG* expression (data not shown). Subsequently, FGF2 was added to SkBM-2 media in experiments (Figs. 2–7 and Supplementary Figs. 3–6).

Engraftment procedures. Mdx/C57BL10 mice were crossed to NSG mice. Each pup was bred and genotyped for homozygosity (mdx-NSG) using Transnetyx. Mdx-NSG mice were housed in the Humanized Mouse Core at UCLA, an immunocompromised core facility. All animal work was conducted under protocols approved by the UCLA Animal Research Committee (ARC) in the Office of Animal Research Oversight (OARO). Mdx-NSG mice 6–8 weeks of age were subjected to a muscle injury by intramuscular injection of 50 μ l of 10 μ m cardiotoxin (*Naja Mossambica-Mossambica*, Sigma), 24 h before cell injections.

To prepare cells for transplantation, hPSC-SMPCs were dissociated using TrypLE and titrated to break apart three-dimensional structures. SMPCs were filtered through 100 μ m meshes, centrifuged at 600 g, and resuspended in Hank's Balanced Salt Solution (HBSS) at 1×10^6 cells per 5 μ l. FACS-sorted hPSC-SMPCs were expanded in SkBM-2 plus FGF2 for less than three passages, or <7 days to prevent precocious MYOD activation. Cultured fetal muscle were dissociated from tissue and cultured in SkBM-2 plus FGF2 for three passages before engraftment. Directly isolated fetal muscle cells were dissociated as described, and immediately transplanted. Pelleted cells were kept at 4 °C and transported to the Mouse Core. Mice were anaesthetized using 2% isoflurane, and 5–10 μ l of cells in solution were injected into the tibialis anterior using Hamilton micro-syringes.

For irradiation experiments, a custom chamber was built using 3-cm thick cerrobend shielding block, which allowed mouse hind limbs to be selectively irradiated. Mice were subjected to 18-gray gamma irradiation over a 12.2-min period using a ¹³⁷Cs irradiator. Mice were anaesthetized with ketamine during the procedure and treated with 500 cc saline daily to control for weight loss. Cardiotoxin was performed 24 h post-irradiation, and cells were injected 48 h post-irradiation. We did not observe an increase in HNK1⁺NCAM⁺ hPSC-SMPC engraftment following irradiation ($N = 9$ mice; Supplementary Table 1).

For engraftment experiments using TGF- β i, 10 μ m SB-431542 was added to cell suspension in HBSS and co-delivered during cell injections. 30 μ l of 10 μ m SB-431542 was injected intramuscularly every third day for 2 weeks post-enugraftment. As a control, HBSS was injected at corresponding time points. All mice survived engraftment procedures and TGF- β i treatment.

After 30 days, mdx-NSG mice were euthanized, and the tibialis anterior muscles were dissected and immediately embedded in optimal cutting temperature (OCT) compound and flash-frozen in isopentane cooled by liquid nitrogen. Embedded muscles were stored at –80 °C until sectioned in 9–10 μ m slices using a Leica microcryotome (LSXII). To determine the position of each cross-section for cell engraftment quantification, 1,200 μ m of tibialis anterior muscle was removed and 20 sections were collected on positively charged microscope slides in 50 μ m intervals, and 300 μ m of tissue was collected for RNA or western blots. Serial quadrants were collected until the entire tibialis anterior was sectioned. Microscope slides were kept in boxes covered with aluminum foil covered at –20 °C.

Immunofluorescence staining for engrafted muscle tissues. Immunofluorescence staining was performed as previously described⁵². For a list of antibodies and concentrations used see Supplementary Table 6. The human dystrophin antibody MANDYS106 was kindly supplied by G. Morris⁵³.

Sample collection for RNA-seq and sequencing analyses. Week 17 fetal muscle cells were directly isolated from limb muscles, and RNA was immediately collected. Alternatively, cultured week 17 fetal skeletal muscle cells after two passages were labelled with CD45⁻CD31⁻NCAM⁺, or hPSC-SMPCs after 50 days of directed differentiation were labelled with HNK1⁻NCAM⁺ (H9 line, method 1), and enriched by FACS. RNA was immediately isolated using RNeasy Microkits (Qiagen). NCAM⁺-sorted fetal muscle cells or hPSC-SMPCs were also differentiated to myotubes in N2 media for 7 days and lysates were collected for RNA. All samples were collected in duplicate, $N=2$. RNA was taken to the UCLA JCCC for sequencing on the Illumina HiSeq 3000 using KAPA amplification.

Samples were aligned in TopHat (v2.0.8b)⁵⁴ to the human genome (hg19, University of California, Santa Cruz (UCSC)) and Ensembl Homo Sapiens GTF file (GRCh37). The average alignment rate per sample was 88.82%, and average reads aligned per sample was 15,365,455. Cufflinks (v2.1.1)⁵⁵ was used to assess gene expression per sample, and Cuffdiff from Cufflinks was used to evaluate differential gene expression. Genes with a false discovery rate (FDR) Q -value threshold of 0.05 were considered significant. These genes were submitted to the Database for Annotation, Visualization and Integrated Discovery (DAVID)^{56,57} to find biological meaning among the samples. Select Gene Ontology terms are shown (Fig. 3), and all Gene Ontology $P < 0.05$ are included in Supplementary Tables 2–4. Heatmaps were created in R version 3.2.3, using packages 'gplots' and 'pheatmap'. Data were loaded into Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) with the accession number GSE87365. Gene Set Enrichment Analysis (GSEA)⁵⁸ was used to determine whether an a priori defined set of genes was significant between groups. Gene set permutation was used.

Utilization of WGCNA from fetal musculoskeletal tissues to identify candidate cell surface markers. The raw sequencing data for the muscle-specific gene set that were used to generate the WGCNA in Fig. 3c across musculoskeletal tissues are available in GEO using the accession number GSE106292. WGCNA data were deposited in GSE106292. To identify muscle-specific cell surface receptors, we used a WGCNA data set of fetal week 17 musculoskeletal tissue including: chondrocytes (Lin⁻CD146⁺BMPRI1B⁺), osteoblasts (Lin⁻ALP⁺), myoblasts (Lin⁻CD56⁺CD146) and ligament and tendon (Lin⁻) isolated and purified by flow cytometry (Fig. 3c). WGCNA produced a list of >400 genes associated with muscle-specific receptors or membrane proteins. To identify candidate enrichment markers, genes were cross-referenced to known myogenic markers in PubMed (National Center for Biotechnology Information (NCBI)). From these data, NCAM, MCAM, integrin $\alpha 7$ (ITGA7), ERBB3, NGFR, CXC-chemokine receptor type 4 (CXCR4) and integrin αM (ITGAM; also known as CD11B) were selected for further analysis. Although not enriched in the fetal muscle module, c-MET and CD82 (ref. 42) were included in the analyses.

Enrichment of ERBB3⁺NGFR⁺ hPSC-SMPCs and human fetal muscle. Human fetal muscle at weeks 9–18 were dissociated as described, or hPSC-SMPCs were taken from directed differentiation. Cells were filtered through 100 μm meshes, washed in PBS and re-suspended in FACS buffer (2% FBS in PBS) at a concentration of 5×10^6 cells per ml. Human Fc block was used to prevent non-specific binding, but additional brilliant violet (BV) dye blocking was not used. In samples containing BV antibodies, far-red live/dead cell viability dye (ThermoFisher L10120) was used, instead of DAPI, to prevent fluorescent spillover. In all FACS experiments, fetal and hPSC-SMPCs were stained for ERBB3, NGFR, MCAM⁵⁹, CD82 and NCAM, and lineage negative markers CD235a, CD45 and CD31 were used to exclude non-muscle cells. Samples were labelled with antibodies for 45 min at 4°C, filtered through 40 μm tubes, and analysed by the Becton Dickinson Fortessa, or sorted using an ARIA-II flow cytometers in the UCLA Broad Stem Cell Research Center Flow Cytometry Core. Fluorescence minus one (FMO) controls were used as gating controls in all experiments. Samples were analysed in FlowJo-OSX version 10.

For hPSC-SMPC experiments, 200,000 ERBB3⁺NGFR⁺ cells were typically isolated per 6-well plates of directed differentiation culture, but as few as 10,000 cells could be successfully cultured. Cells were collected in SkBM-2 plus FGF2, seeded on matrigel-coated culture plates, and were allowed to recover from FACS for 3–5 days. Cells were then passaged for in vitro experiments, engrafted for in vivo experiments, or cryopreserved using 10% DMSO in 90% FBS. Four directed differentiation protocols were tested for their ability to enrich HNK1⁻ERBB3⁺NGFR⁺ SMPCs^{10,11,13,19}. Although variability in myogenic potential exists, in all protocols, ERBB3⁺NGFR⁺ cells were the most myogenic (Fig. 5). When using a single marker, ERBB3 was best able to enrich for muscle progenitors with in vitro and in vivo myogenic potential (Fig. 7).

We assessed for myogenic activity in vitro for up to 10 passages or in vivo for <4 passages. Longer culturing periods decreased myogenic activity in vitro and in vivo. Upon culturing, we found that double-positive ERBB3⁺NGFR⁺ cell populations increased NGFR positivity (ERBB3⁺NGFR^{High}) by FACS analysis.

When sorted on ERBB3⁺NGFR^{High}, both cultured/passaged week 17 fetal cells and passaged hPSC-SMPCs were enriched for myogenic markers, but expressed significantly higher levels of *MYOG* compared to uncultured fetal muscles or non-passaged hPSC-SMPCs. Thus, ERBB3⁺NGFR^{High}-expressing cultured cells activate after continuous passaging, suggesting that these muscle progenitors cannot be expanded or supported as myogenic cells for long periods.

TGF- β regulates skeletal muscle differentiation and maturation across three independent hPSC lines. To test the effects of TGF- β on SMPC differentiation, small-molecule inhibitors of TGF- β signalling (SB-431542 (ref. 60) and A83-01 (ref. 61)) were evaluated. SB-431542 was more potent than A83-01 (1–20 μm , data not shown), and SB-431542 (10 μm) is shown in Figs. 6 and 7 and Supplementary Figs. 5 and 6. Recombinant TGF- β 1 (10 ng per ml, Peprotech) was also evaluated. NCAM⁺ or ERBB3⁺ hPSC-SMPCs were differentiated for 5–7 days in N2 media with or without TGF- β 1 (Supplementary Figs. 5 and 6). To evaluate maturation status, wild-type hiPSC-SMPCs (1002), fetal weeks 14 and 17 muscle progenitors, and adult SCs were cultured in parallel. Three subpopulations of hPSC-SMPCs, negatively sorted for HNK1⁻, were tested: NCAM⁺, ERBB3⁺ and double-positive ERBB3⁺NGFR⁺. All muscle cells were treated with TGF- β 1 (SB-431542, 10 μm) or TGF- β (10 ng per ml) in N2 media containing IGF1 for 7 days (Supplementary Fig. 6). Cells were immunostained or evaluated by qPCR.

Transmission electron microscopy. NGFR⁺ hPSC subpopulations and fetal muscle cells were seeded on electron microscopy (EM)-grade plastic coverslips coated in matrigel 1:50, and differentiated with or without TGF- β 1 for 7 days. Cells were fixed in 4% paraformaldehyde and 1% glutaraldehyde overnight. Cells were taken to the UCLA Electron Microscopy Core Facility for processing using a Reichert Ultracut ultramicrotome. Cells were imaged using a JEOL 100CX transmission electron microscope.

RT-PCR validation and primers. Primers were designed using NCBI primer blast or based on published work. Primers (10 μm) were validated using cells known to express the genes described in Supplementary Table 5. Complementary DNA (cDNA) concentrations were 5-fold serially diluted starting at 5 ng μl^{-1} . Primers with 0.9–1.1 efficiency were used for experiments. See Supplementary Table 5 for primer sequences.

Statistics and reproducibility. Experiments were repeated at least three times with similar results unless a different number of repeats is stated in the legend. Statistical testing was performed using the unpaired two-tailed Student's t -test or ANOVA, as stated in the figure legends. The method used, P -values and N numbers are indicated in the figure legends. No statistical method was used to predetermine sample size. All images for myotube counts were randomly selected for quantification. Both male and female mdx-NSG mice were selected at random for engraftment experiments.

We used four independent hPSC lines to validate the reproducibility of findings of this study. hPSC lines shown in figures are as follows: Fig. 1: H9; Fig. 2: H9, CDMD 1002, CDMD 1006 and CDMD 1006-1; Fig. 3: H9 and CDMD 1006-1; Fig. 5: H9, CDMD 1002, CDMD 1003 and CDMD 1003-49; Fig. 6: 1006-1; Fig. 7: 1006-1; Supplementary Fig. 1: H9; Supplementary Fig. 3: H9; Supplementary Fig. 5: CDMD 1006 and CDMD 1006-1; and Supplementary Fig. 6: CDMD 1002.

To quantify cell engraftments, all h-lamin A/C⁺spectrin⁺ and h-dystrophin⁺ myofibres from >10 cross-sections were counted. Plots were made by distributing counts along the recipient muscle length from distal (knee) to proximal (foot), where $y = N$ of engrafted myofibres and $x =$ location in millimetres. Plots were then statistically compared using non-parametric Mann–Whitney U -tests (two samples) or Kruskal–Wallis tests (three or more samples), which allows for the analysis of non-parametric (distribution independent) data in which engrafted cells are unevenly distributed at points distant from the injection site, and enables the comparison of three or more samples⁶².

Life Sciences Reporting Summary. Further information on experimental design is available in the Life Sciences Reporting Summary.

Data availability. RNA-seq data that support the findings of this study have been deposited in the GEO under accession codes GSE87365 (for data referring to Figs. 3a,b and 6) and GSE106292 (for data referring to Fig. 3c). Flow cytometry data have been deposited in FlowRepository.org under the accession code FR-FCM-ZYD9. Source data have been provided as Supplementary Table 7. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

References

- Rommel, C. et al. Mediation of IGF-1-induced skeletal myotube hypertrophy by PI(3)K/Akt/mTOR and PI(3)K/Akt/GSK3 pathways. *Nat. Cell Biol.* **3**, 1009–1013 (2001).
- Young, C. S. et al. Creation of a novel humanized dystrophic mouse model of Duchenne muscular dystrophy and application of a CRISPR/Cas9 gene editing therapy. *J. Neuromuscul. Dis.* **4**, 139–145 (2017).

53. Nguyen, H. T. & Morris, G. E. Use of epitope libraries to identify exon-specific monoclonal antibodies for characterization of altered dystrophins in muscular dystrophy. *Am. J. Hum. Genet.* **52**, 1057–1066 (1993).
54. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
55. Trapnell, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
56. da Huang, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
57. da Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
58. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
59. Lapan, A. D., Rozkalne, A. & Gussoni, E. Human fetal skeletal muscle contains a myogenic side population that expresses the melanoma cell-adhesion molecule. *Hum. Mol. Genet.* **21**, 3668–3680 (2012).
60. Inman, G. J. et al. SB-431542 is a potent and specific inhibitor of transforming growth factor- β superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. *Mol. Pharmacol.* **62**, 65–74 (2002).
61. Tojo, M. et al. The ALK-5 inhibitor A-83-01 inhibits Smad signaling and epithelial-to-mesenchymal transition by transforming growth factor- β . *Cancer Sci.* **96**, 791–800 (2005).

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

▶ Experimental design

1. Sample size

Describe how sample size was determined.

Sample sizes were chosen based on pilot studies of the experiment being performed. Please see supplemental methods pages 4, 5, and 7. For comparisons between hPSC-SMPCs and fetal or adult muscle, drastic differences were immediately apparent and thus for studies a small sample size (N=3-5) was sufficient to determine differences. For in vivo experiments, using HNK1-NCAM+ sorted hPSC-SMPCs little to no engraftment was observed. Across cell lines, we engrafted 18 mice with HNK1-NCAM+ sorted hPSC-SMPCs, see table S1. Once new subpopulations were discovered, we observed large differences in the abilities of NGFR and ERBB3+ hPSC-SMPCs to engraft. Thus 3-6 animals were enough to reach statistical significance.

2. Data exclusions

Describe any data exclusions.

No data were excluded from the analyses. In instances where cells did not engraft in the mice, these are still reported in the data table S1.

3. Replication

Describe whether the experimental findings were reliably reproduced.

While, line-to-line variability is observed in experimental results. We have reliably reproduced the significant findings of this study across multiple cell lines and directed differentiation protocols as highlighted in the supplemental information Figure S5.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Treatment groups were processed identically throughout all experiments so that each group had equal probability of responding to treatment and to limit bias of results. For transplantation studies 50% males and 50% female mdx-NSG mice were used. Mice were selected at random for engraftment experiments. All animals were housed in the same area and no preference was given to which mouse received a given treatment/transplantation.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Depending on the experiment, analyses were performed in a blinded fashion. Microscope slides labels were covered with tape and recorded as i.e. A, B, C, etc. Scoring of image quantification data was recorded so that multiple users could reevaluate data and derive the same results. For animal studies, sectioning and staining were blinded with animal numbers. After image quantification of h-Lamin AC and h-Dystrophin, key was used to identify groups. Investigators were blinded to analysis on image quantifications.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- The test results (e.g. P values) given as exact values whenever possible and with confidence intervals noted
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

We used GraphPad PRISM 6.0 to perform statistical comparisons. FlowJo was used for FACS analyses. ImageJ was used to assist with image quantification. For RNA-SEQ samples were aligned in TopHat (v2.0.8b) to the human genome (hg19, UCSC) and Ensembl Homo Sapiens GTF file (GRCh37). Cuffdiff from Cufflinks was used to evaluate differential gene expression. For additional information on all software used please see extended experimental procedures.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). [Nature Methods guidance for providing algorithms and software for publication](#) provides further information on this topic.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

There are no restrictions on the availability of the material.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

All antibodies used in this study are commercially available and are described in Supplementary Table 6.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

Cell lines used included hESC line: H9 (Wicell), and hiPSC lines: CDMD 1002, CDMD 1003, CDMD 1006 and CRISPR/Cas9 Corrected hiPSC lines: CDMD 1003-49 and CDMD 1006-1 (generated in the Pyle lab). Human skeletal muscle myoblasts (HSMM) were purchased from Lonza.

b. Describe the method of cell line authentication used.

All hiPSC lines were validated by karyotype analyses and teratoma formation.

c. Report whether the cell lines were tested for mycoplasma contamination.

Yes, all cell lines were tested and were mycoplasma free.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

No commonly misidentified cell lines were used.

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

Mdx/C57BL10 mice were crossed to Nod scid gamma (NSG) mice. Each pup was bred and genotyped for homozygosity (mdx-NSG) using Transnetyx. Mdx-NSG mice were housed in the Humanized Mouse Core at UCLA, an immunocompromised core facility. Both males and females were used equally. All animal work was conducted under protocols approved by the UCLA Animal Research Committee in the Office of Animal Research Oversight. Mdx-NSG mice ages 6-8 weeks were used for cell injections.

Policy information about [studies involving human research participants](#)

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

Human adult skeletal muscle between 25-67 years of age were obtained from donor autopsy of quadriceps provided by the National Disease Research Interchange (NDRI), and human fetal skeletal muscle were obtained from Novogenix, Inc. or the UCLA CFAR Gene and Cellular Therapy Core using IRB approved de-identified and consented human fetal tissues in accordance with institutional guidelines. Use of human tissues was IRB exempt by the UCLA Office of the Human Research Protection Program (IRB #15-000959).