HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2017 July 26.

Published in final edited form as:

Cell Rep. 2017 June 27; 19(13): 2867–2877. doi:10.1016/j.celrep.2017.06.005.

PAX7 Targets, CD54, Integrin α9β1 and SDC2, Allow Isolation of Human ES/iPS Cell-Derived Myogenic Progenitors

Alessandro Magli[1,*], **Tania Incitti**[1,*], **James Kiley**[1], **Scott A Swanson**[2], **Radbod Darabi**[3], **Fabrizio Rinaldi**[1], **Sridhar Selvaraj**[1], **Ami Yamamoto**[1], **Jakub Tolar**[4,5], **Ce Yuan**[5], **Ron Stewart**[2], **James A Thomson**[2], and **Rita CR Perlingeiro**[1,§]

[1]Lillehei Heart Institute, Department of Medicine, University of Minnesota, Minneapolis, MN, USA

[2]Morgridge Institute for Research, Madison, WI, USA

[3]Center for Stem Cell and Regenerative Medicine, Dept. of Medicine, University of Texas, Houston-TX, USA

[4]Department of Pediatrics, University of Minnesota, Minneapolis, MN, USA

[5]Stem Cell Institute, University of Minnesota, Minneapolis, MN, USA

Abstract

Pluripotent stem (PS) cell-derived cell types hold promise for treating degenerative diseases. However, PS cell differentiation is intrinsically heterogeneous therefore clinical translation requires development of practical methods for isolating progenitors from unwanted and potentially teratogenic cells. Muscle-regenerating progenitors can be derived through transient PAX7 expression. To better understand the biology, and to discover potential markers for these cells, here we investigate PAX7 genomic targets and transcriptional changes in human cells undergoing PAX7-mediated myogenic commitment. We identify CD54, integrin α9β1 and Syndecan2 (SDC2) as surface markers on PAX7-induced myogenic progenitors. We show that these markers allow for the isolation of myogenic progenitors using both fluorescent- and cGMP-compatible magnetic-based sorting technologies, and that CD54+α9β1+SDC2+ cells contribute to long-term muscle regeneration *in vivo*. These findings represent a critical step towards enabling translation of PS cell-based therapies for muscle diseases.

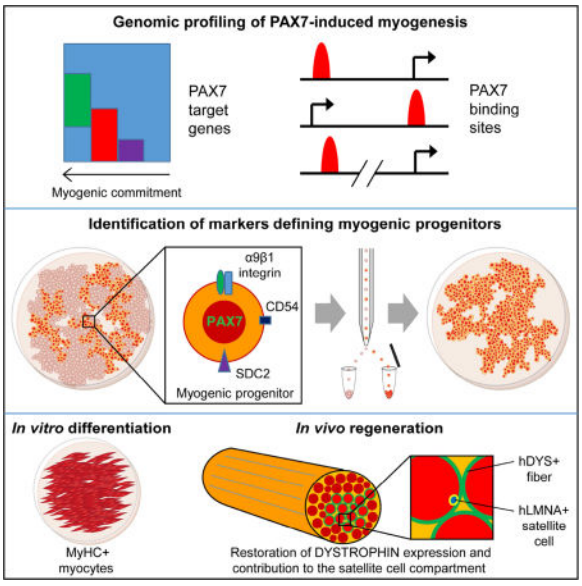
Graphical abstract

[§]**Lead Contact** Rita C.R. Perlingeiro PhD, Lillehei Heart Institute, University of Minnesota, 4-128 CCRB, 2231 6th St. SE, Minneapolis, MN 55455, USA, perli032@umn.edu, Phone: 612 625 4984, Fax: 612 301 8298.

[*]These authors contributed equally to this manuscript

CONTRIBUTIONS

A.M. and T.I. designed and performed the research, analyzed the data and wrote the manuscript; J.K. performed the research and analyzed the data; S.A.S and C.Y. performed the bioinformatic analyses; R.S. and J.A.T. supervised the bioinformatic analyses; R.D., F.R., S.S. and A.Y. performed research; J.T. supervised iPSC facility; R.C.R.P. contributed with experimental design, interpretation of the data and wrote the manuscript.



Keywords

Stem cell therapy; skeletal myogenesis; muscular dystrophy; RNA-sequencing; muscle regeneration; ChIP-sequencing; PAX7; CD54; integrin α9β1; SDC2

INTRODUCTION

Pluripotent stem cells are an attractive option for regenerative medicine as they can repeatedly produce large amounts of differentiated tissue, thus representing an unlimited source of cells for therapeutic application. Nevertheless, a major requirement remains the development of protocols, compatible with Current Good Manufacturing Practice (cGMP) standards, for the isolation of specific cell populations with proven *in vivo* regenerative potential whilst preventing the presence of contaminating teratogenic undifferentiated pluripotent stem cells.

Cell-based therapies are particularly attractive for muscular dystrophies (Tedesco and Cossu, 2012). These are a heterogeneous group of genetic diseases characterized by relentless and catastrophic progression of muscle wasting leading to premature death in high risk disorders, such as Duchenne Muscular Dystrophy (DMD) (Emery, 2002). Following the expansion of the PS cell field, several pioneering studies (Darabi et al., 2012; Darabi et al., 2008; Darabi et al., 2011a; Darabi et al., 2011b; Filareto et al., 2012; Filareto et al., 2013), have demonstrated functionally beneficial engraftment and many others (Borchin et al., 2013; Chal et al., 2015; Choi et al., 2016; Shelton et al., 2014; Xi et al., 2017; Xu et al., 2013), have demonstrated skeletal myogenic differentiation *in vitro* of both embryonic (ES) and induced-pluripotent (iPS) stem cells. We have shown that upon controlled expression of PAX7 followed by GFP-mediated sorting, human ES/iPS derived myogenic progenitors proliferate robustly *in vitro* and contribute to muscle regeneration *in vivo*, as transplanted cells generate functional myofibers while also seeding the muscle stem compartment (Darabi et al., 2012).

To successfully translate this approach toward the clinic, we envisioned the need to identify specific surface markers for the isolation of this progenitor cell population, which would avoid potentially immunogenic transgenes (Ansari et al., 2016) (e.g. GFP or mCherry) as well as establishing quality control for the cGMP process. For this, we performed whole transcriptome sequencing analysis, in which we systematically evaluated different time points along the PAX7-dependent myogenic commitment from human pluripotent stem cells. Our results revealed a subset of genes differentially expressed at various stages of this differentiation process, including a discrete number of surface markers. After Fluorescence Activated Cell Sorting (FACS)-mediated screening, we identified integrin α9β1, CD54 and SDC2, as potential surface markers to be used for the prospective isolation of human PS cell-derived myogenic progenitors. We demonstrate that these surface molecules reproducibly allow the isolation of myogenic progenitors from multiple human ES/iPS cell lines and that α9β1+CD54+SDC2+ (triple+) cells represent a homogenous population of PAX7+ cells endowed with *in vivo* muscle regeneration potential. These findings provide a clinically relevant method for the purification of PS cell-derived muscle progenitors for therapeutic applications.

RESULTS

Identification of PAX7-induced transcriptional changes during human myogenesis

Controlled expression of PAX7 in differentiating human ES/iPS cells results in the specification of a highly proliferative population of myogenic progenitors with considerable regenerative potential *in vivo*. However in order for this approach to be suitable for clinical translation, strategies for the purification of PAX7-induced (referred as $PAX7^+$ hereafter) myogenic progenitors that do not rely on fluorescent reporters (e.g. GFP), but rather specific native surface markers that are compatible with clinical-grade purification methods, are required. To address this important issue, we performed a systematic analysis to identify the gene expression changes induced by PAX7 during the commitment of the skeletal myogenic lineage from human H9 ES cells *in vitro*. As schematized in Fig. 1A, we focused on the following time points: PAX7+ mesodermal cells (day 14), PAX7+ proliferating myogenic progenitors (approximately day 23), and differentiated myocytes (differentiation stage – around day 30; 7 days in the absence of PAX7 induction). Since PAX7 expression is doxycycline inducible, we also collected uninduced control samples at the same time points (termed mesodermal cells and proliferating cells; Fig. 1A, left). RNA-sequencing (RNA-seq) identified about 2000 genes whose expression levels displayed a fold change >4 and TPM (Transcripts Per Million reads) >64 in at least one of the replicates (Fig. 1B). Unsupervised analysis of these differentially expressed genes identified 4 distinct clusters representing cells undergoing PAX7-dependent myogenic commitment (PAX7+ mesodermal cells, PAX7+ proliferating myogenic progenitors, and differentiation stage) as well as uninduced cells (Fig. 1B). Importantly, upon principal component analysis (PCA), these clusters were easily separated according to PAX7 expression (dox-treatment – on PC2 axis), and progression toward the differentiated skeletal myogenic state (on PC1 axis) (Fig. 1C). As shown by the heat map in Fig. 1B and Spearman correlations map in Fig. S1A, gene expression changes follow a pattern suggesting the gradual activation of the skeletal myogenic program induced by PAX7, which is also discernible at the single gene level (Fig.

1D and Fig. S1B). In agreement with our previous studies using mouse ES cells (Magli et al., 2012; Magli et al., 2014), these results show that genes associated with paraxial mesoderm development, such as DLL1, MEOX1 and LFNG, are also up-regulated by PAX7 expression, indicating the ability of this transcription factor to induce a population of paraxial mesodermal progenitors (Fig. 1D). Moreover, taking advantage of recently published RNAseq data from developing human embryos (Xi et al., 2017), we demonstrate that both PAX7+ mesodermal cells and PAX7+ proliferating myogenic progenitors are more closely related to cells isolated from somites (SM and SM_DEV) than presomitic mesoderm (PSM - Fig S1C), thus supporting the capability of our system to recapitulate cellular processes occurring during human development.

After validating selected candidates by RT-qPCR (Fig. S1D), we chose sets of genes from each cluster that are up-regulated with respect to the uninduced samples (boxed areas, Fig. 1b) and applied Gene Ontology (GO)-based functional annotation to each set. This revealed that each cluster is characterized by up-regulated genes with distinct biological function based on GO classification (Fig. 1E and Fig. S1E). As expected, while both PAX7+ mesodermal cells and PAX7+ proliferating myogenic progenitors are enriched for genes involved in cell motility/adhesion and communication (Fig. 1E), differentiated myocytes (differentiation stage) are characterized by the expression of genes involved in terminal differentiation (Fig. S1E).

To gain further insight on the PAX7-mediated activation of the skeletal myogenic program, we also performed Chromatin-immunoprecipitation followed by sequencing (ChIP-seq) for PAX7 in PAX7+ proliferating myogenic progenitors. As expected, we identified putative PAX7 binding sites in several differentially regulated genes (approximately 50%), including known PAX3/7 targets MET, MYF5 and MYOD (Bajard et al., 2006; Epstein et al., 1996; Soleimani et al., 2012) (Fig. 1F and Fig. S1F). Following annotation, we observed that PAX7 peaks are located within introns and in intergenic regions (Fig. 1G), while the gene type distribution highlighted a preference for protein-coding genes and long non-coding RNAs (Fig. S1G). Analysis of the top 3000 peaks showed enrichment of both the PAX Paired- and homeo-domain dependent motifs (Fig. 1H), demonstrating the specificity of our approach in identifying *bona fide* PAX7 binding sites. These data are in agreement with previous work in murine primary myoblasts (Soleimani et al., 2012) and, collectively, represent a detailed analysis of PAX7 function during human skeletal myogenesis, allowing us to interrogate specific questions such as the molecular identity of the myogenic progenitor population endowed with muscle regenerative potential.

Human pluripotent-derived myogenic progenitors express CD54, integrin α9β1 and SDC2

Next, we focused on the identification of candidate surface markers specific for cells undergoing PAX7-dependent myogenic commitment. GO analysis and manual inspection of the RNAseq data (available in Table S1) identified a group of 50 genes associated with cell adhesion, which were found up-regulated following PAX7 induction (Table S2) in PAX7+ mesodermal cells and PAX7+ proliferating myogenic progenitors. RNA molecules are subjected to post-transcriptional regulation and protein expression might not reflect such variations. Therefore, using available commercial antibodies, we screened several candidate

surface markers to identify changes in protein expression levels between PAX7+ mesodermal cells (identified based on GFP expression, which is encoded by the bicistronic transcript PAX7-ires-GFP) and control uninduced cells (PAX7-negative). As shown in Fig. 2A and Fig. S2A, FACS analysis confirmed distinct up-regulation of 3 surface markers following PAX7 induction (compare GFP+ and GFP-cells). These surface markers are the Intercellular Adhesion Molecule 1 (ICAM1 or CD54), Syndecan 2 (SDC2 or CD362), and Alpha9 Integrin (ITGA9) subunit. Since Integrin B1 is the main and almost exclusive beta integrin subunit expressed in these cells (Fig. S2B), we studied the expression of ITGA9 using the antibody for the integrin α9β1 dimer (referred as α9β1 hereafter).

Because RNA-seq data showed time dependent changes in gene expression during the PAX7-mediated myogenic commitment, we next evaluated the expression of these markers over time. Following PAX7 induction, we detected a population of CD54+ (bright) cells in PAX7+ mesodermal fraction (day 14; Fig. 2B, left). This population was also found positive for α9β1 integrin, and it was clearly distinct from control populations (no dox) in the PAX7+ mesodermal and PAX7+ proliferating myogenic progenitor stages (day 23; Fig. 2B, right). Quantification of the frequency of these populations confirmed the up-regulation of these surface markers upon PAX7 induction (Fig. 2C). Similar results were obtained when cells were analyzed based on PAX7 expression using the GFP reporter (GFP+ *vs.* GFP−) (Fig. S2C–D). Not surprisingly, upon inspection of our ChIP-seq data, we were able to identify PAX7 binding at the 5' region of CD54 (ICAM1) and in the intronic region of SDC2 (Fig. S1F), strongly supporting the notion that PAX7 directly regulates these genes. Because the PAX7-expressing population positive for these markers is clearly separated from the negative fraction at the proliferating myogenic progenitor stage (Fig. 2B, right panel), we focused our purification efforts at this specific time point. Of note, CD54+ α9β1+ proliferating myogenic progenitors are ~95% SDC2+ (Triple+; Fig. 2D–E), and retrospective analysis of GFP expression shows that Triple+ cells are nearly 100% GFP+ (Fig. 2D–E), thus suggesting that these markers can replace the GFP transgene for the purification of human PS cell-derived PAX7+ myogenic progenitors.

To confirm this hypothesis, we FACS-purified different cell populations from dox-treated and untreated cultures using different combinations of these three antibodies, and assessed their ability to generate Myosin Heavy Chain positive (MYHC+) skeletal myocytes *in vitro*. Remarkably, MYHC+ terminal differentiation was robust and indistinguishable between the dox-treated CD54+α9β1+SDC2+ (Triple+) cell fraction and the control GFP+ (PAX7+) cell population (Fig. 2F and Fig S2F). In contrast, dox-treated CD54-α9β1- or CD54-α9β1-SDC2-(Triple-) cells and CD54-α9β1-, CD54+α9β1+, SDC2+ or Triple+ cell fractions from uninduced cultures showed virtually no myogenic potential, as demonstrated by the extremely low number of MYHC+ cells (Fig. 2F and Fig S2E–F). These results highlight the transcriptional importance of PAX7 in the specification of the skeletal myogenic lineage since cells without PAX7 expression have no myogenic potential. Importantly, both GFP+ and Triple+ purified cells displayed equivalent proliferative potential (Fig. S2G). Also of extreme relevance, these surface markers were detected upon PAX7 induction in multiple different differentiating pluripotent stem cell lines (Fig. S3A–B), and were consistently found to purify myogenic progenitors with robust skeletal myogenic differentiation potential

(Fig. S3C). These data clearly demonstrate that the isolation of human PS cell-derived PAX7+ myogenic progenitors can be achieved using CD54, α9β1 and SDC2.

CD54, α9β1 and SDC2 identify PAX7-induced myogenic progenitors independent of culture conditions

A distinct population of CD54+α9β1+SDC2+ cells was also identified upon PAX7 induction in pluripotent stem cells differentiated using a chemically-defined serum-free monolayer protocol, (Fig. 3A–B), which was confirmed to be endowed with high myogenic potential (Fig. 3C and Fig S3D), and indistinguishable from cells sorted based on GFP expression. These data indicate that PAX7 is capable of inducing a population of CD54+α9β1+SDC2+ myogenic progenitors independently of the culture conditions used.

Since the use of lentiviral vectors is required for the consistent *in vitro* generation of PAX7+ proliferating myogenic progenitors, we have tested the pCCL lentiviral backbone vector system, which is a third generation self-inactivating vector (Dull et al., 1998) demonstrated to be safer than standard lentiviral vectors (Aiuti et al., 2013; Biffi et al., 2013; Romero et al., 2013a) in recent gene therapy trials. Validation studies for doxycycline-dependent PAX7 expression in the absence of a GFP reporter in pluripotent stem cells transduced with this construct (pCCL-PAX7) show indistinguishable expression of CD54, α9β1 and SDC2 compared to inducible PAX7-ires-GFP cells and importantly, allow FACS-mediated isolation of myogenic progenitors, which robustly differentiate into MYHC+ cells (Fig S3E–F).

CD54 enables GMP-compatible purification of human PS cell-derived myogenic progenitors

As shown above, myogenic progenitors are homogeneously defined by the expression of CD54, α9β1 and SDC2. Nevertheless, because methods compatible with clinical application involve the use of single marker purification by magnetic beads, we predicted that sequential purification of 3 surface markers would be inefficient for this purpose. Although CD54 is expressed in uninduced cells, this receptor is clearly upregulated following PAX7 induction (Fig. 2A), such that cells prepared for FACS can be visually identified based on doxycycline treatment (Fig S3G). Considering that $CD54^{bright}$ cells are also positive for both α9β1 and SDC2 (Fig. 2D), we investigated whether CD54 could be used as a single marker for the isolation of PS cell-derived myogenic progenitors. For this, we isolated CD54-labeled human PS cell-derived myogenic progenitors using either FACS or magnetic cell sorting (MACS). Retrospective analysis of GFP expression revealed that both CD54 FACS- and MACS-purified cells are highly enriched for GFP (Fig. 3D–E) and, as expected, both methods yielded a population of myogenic progenitors endowed with robust myogenic differentiation potential, comparable to the GFP+ (PAX7+) control cell population (Fig. 3F). Altogether these data demonstrate that a single step purification strategy can provide cultures with optimal enrichment of human PS cell-derived PAX7+ proliferating myogenic progenitors and support the translational potential of these identified markers for future therapeutic applications.

One of the main concerns associated with potential cell therapies involving PS cell-derived cultures is the presence of contaminating undifferentiated pluripotent cells which may give rise to teratomas. We investigated whether the cell fraction negative for the aforementioned identified markers displayed pluripotency features, which would associate the negative fraction with teratoma risk. FACS-purified CD54-α9β1-SDC2- (Triple-) cells maintained in culture conditions specific to PS cells were not able to form human pluripotent stem cell-like colonies (Fig. S4A), and consistently, both CD54-α9β1-SDC2- and CD54+α9β1+SDC2+ cells lacked expression of the pluripotency markers OCT3/4, SOX2 and NANOG (Fig. S4B). To further demonstrate the safety of human PS cell-derived myogenic progenitors, we injected CD54+ MACS-sorted (representing the potential cGMP product) and CD54- FACS-sorted (to specifically evaluate their teratoma-forming ability) cell preparations into immunodeficient (NSG) mice to determine their teratoma-forming potential. To date, 12 months post-injection, no signs of potential teratoma-forming activity have been observed in mice injected with CD54+ or CD54- cells. This is in contrast to undifferentiated human PS cells, which gave rise to evident teratomas 2 months after cell injection (Fig. S4C).

Skeletal muscle regeneration of CD54+α9β1+SDC2+ myogenic progenitors

A fundamental requirement for a prospective cell population to be used for therapy is the demonstration of *in vivo* regenerative potential using animal models. We previously demonstrated that human ES/iPS-derived myogenic progenitors efficiently participate in muscle regeneration by producing fibers positive for both human DYSTROPHIN (hDYS) and human LAMIN A/C. To assess whether human myogenic progenitors isolated using these identified surface markers have the same regenerative potential as their GFP+ control counterparts, we injected CD54+α9β1+SDC2+ sorted myogenic progenitors into cardiotoxin-injured Tibialis Anterior (TA) muscles of immunodeficient (*NSG*) mice. Whereas no detection of human DYS and LAMIN A/C was observed in PBS-injected controls (Fig. 4A, left), TA muscles that had been transplanted with myogenic progenitors purified based on the expression of CD54, α9β1 and SDC2 displayed large numbers of hDYS+/hLAMIN A/C+ fibers (Fig. 4a, right panel) both 8 weeks and 10 months post-intramuscular cell delivery. Similar results were obtained upon transplantation of these cells in the DMD mouse model *NSG-mdx*^{4cv} (Fig. S4D). Importantly, the regenerative potential of CD54+α9β1+SDC2+ sorted myogenic progenitors was found equivalent to that observed with the GFP+ (PAX7+) control cell population (Fig. 4A–B).

One of the most attractive features of stem cell-based therapies is the contribution to the tissue-resident stem cell pool responsible for the long-term homeostasis. We show CD54+α9β1+SDC2+ sorted myogenic progenitors provide long-term regenerative potential, as demonstrated by the detection of hDYS+/hLAMIN A/C+ fibers as late as 10 months post-transplantation (Fig 4A, **lower panels**). Consistently, we observed the presence of human donor-derived (hLAMIN A/C+) cells occupying the satellite cell position and co-expressing the muscle stem cell markers M-cadherin (Cooper et al., 1999; Donalies et al., 1991; Irintchev et al., 1994) (Fig. 4C) or Pax7 (Seale et al., 2000) (Fig. S4E). All together, these data support the application of this purification strategy for clinically-compatible production of human PS cell-derived myogenic progenitors endowed with long term regenerative potential.

DISCUSSION

The genetic nature of myopathies such as Duchenne Muscular Dystrophy requires therapeuti**c** approaches able to both repair the existing myofibers and repopulate the stem cell pool. After initial attempts involving myoblast transplantation (Huard et al., 1991; Partridge et al., 1989) (Mendell et al., 1995; Partridge et al., 1998; Tremblay et al., 1993; Vilquin, 2005), which produced no clinical benefit, the recognition that stemness is lost upon *ex vivo* culture has led many to focus on satellite cells, particularly their self-renewal and efforts to enable their *ex vitro* expansion (Arpke et al., 2013; Bernet et al., 2014; Cosgrove et al., 2014; Dumont et al., 2015; Gilbert et al., 2010; Quarta et al., 2016; Sousa-Victor et al., 2014). In addition, several other adult stem cell populations have been reported to be endowed with myogenic potential (Mitchell et al., 2010; Sampaolesi et al., 2003; Torrente et al., 2004). Although this is encouraging, definitive demonstration of their clinical efficacy remains to be proven.

Since the isolation of human ES cells in 1998 (Thomson et al., 1998), regenerative medicine based on PS cells has attracted remarkable interest from both academia and industry, which promoted the rapid expansion of this field and, among various achievements, resulted in the development of iPS cell technology (Takahashi and Yamanaka, 2006). This has the potential to revolutionize the field of regenerative medicine as it potentially enables for the use of autologous genome-edited or allogeneic HLA-matched iPS cells. Importantly, the future success of regenerative medicine relies on the safe and reproducible production of tissue-specific progenitors from PS cells (Carpenter and Rao, 2015). Methods for the isolation of the target cell population represent one of the crucial requirements for clinical application of PS cell-derived stem cell products, and their development requires a detailed characterization of the PS cell differentiation process. Based on our earlier findings, in this study we took advantage of the conditional PAX7 expression system to investigate the process leading to the specification of the skeletal myogenic lineage by performing RNA- and ChIP-seq experiments. Of note, this work represents the first detailed investigation of PAX7 function in human myogenesis and identifies several potential regulators of the muscle lineage. Specifically, these analyses revealed a group of differentially expressed genes accounting for the unbiased separation of the samples into 4 independent clusters and whose localization on the PCA graph is reminiscent of the time-dependent commitment toward the skeletal myogenic lineage (Fig. 1B–C). Furthermore, many up-regulated genes contains putative PAX7 binding sites (Fig. 1F and Fig. S1F) and upon functional classification of these data, we observed that genes involved in cell adhesion and motility represent the most significant category up-regulated by PAX7 (Fig. 1E). Using this knowledge, we first performed a screen to identify markers enriched on the surface of PS cell-derived myogenic progenitors (Fig. 2A–B and Fig. S2A), and then we demonstrated that CD54, α9β1 and SDC2 mark and allow for the isolation of these progenitors from cultures of differentiating pluripotent stem cells (Fig. 2C–E and Fig. 3A–C). Importantly, myogenic progenitors isolated based on the expression of these identified markers are indistinguishable from the ones sorted using intracellular fluorescent reporters, and accordingly are endowed with robust *in vitro* proliferative potential and *in vivo* regenerative capacity, including contribution to the satellite cell compartment (Fig. 4 and Fig. S4D–E). Issues related to

safety represent significant hurdles when translating initial studies at the bench into viable clinical protocols. Defining the global PAX7 DNA-binding and transcriptional profiles allowed us to discover and use surface markers to significantly purify the cell product. In addition, these markers eliminate the need for a transduction reporter (GFP- Fig. S3E–F) and guarantee the selection of myogenic cells in which we know PAX7 is functionally active.

Although transgene-free monolayer cultures for the differentiation of hPS cells toward the skeletal myogenic lineage have been reported, findings produced so far have not been sufficient to justifying their application to regenerative medicine. Overall, these protocols allow for the derivation of myocytes, however their translational applicability is uncertain since engraftment capability, in most cases, was not reported (Borchin et al., 2013; Chal et al., 2015; Loh et al., 2016; Shelton et al., 2014; Xi et al., 2017) or was very limited (Xu et al., 2013). Notably, in absence of cell purification, these protocols generate heterogeneous populations including neural cells (Chal et al., 2016; Kim J, 2017). Moreover, based on MYOGENIN expression (Chal et al., 2016; Choi et al., 2016; Shelton et al., 2014), these PS cell-derived myogenic cells more resemble myoblasts, a population already tested in several clinical trials without beneficial improvements to date (reviewed by (Tedesco and Cossu, 2012). In contrast, a differentiation strategy incorporating inducible PAX7 expression produces homogeneous cultures of proliferating PAX7+ myogenic progenitors, with minimum spontaneous terminal differentiation, as previously demonstrated by negligible expression levels of MYOGENIN and MHC in the expansion phase (Darabi et al, 2012). Consistently, these PAX7+ myogenic progenitors are capable of *in vivo* skeletal muscle regeneration and contribution to the muscle stem cell pool, essential aspects for safe and effective cell therapy for muscle diseases.

Based on these findings, we postulate these surface markers will be instrumental for the large scale cGMP-compatible production of PS cell-derived myogenic progenitors. This purification strategy combined with PAX7 expression obtained using third generation self-inactivating lentiviral vectors (Fig. S3E–F), currently in gene therapy clinical trials for hematological disorders with encouraging results in terms of safety (Aiuti et al., 2013; Biffi et al., 2013; Romero et al., 2013b), or expression cassettes integrated in a safe harbor locus (DeKelver et al., 2010; Hockemeyer et al., 2009) (e.g. AAVS1), has the potential to enable clinical translation of this myogenic progenitor cell population in patients with muscular dystrophy.

MATERIAL AND METHODS

Cell cultures and differentiation

Inducible PAX7 Human H9 ES cells and PLZ *wt* iPS cells were generated previously (Darabi et al., 2012) and they were maintained on matrigel-coated flasks using mTeSR (Stem Cell Technology). Inducible PAX7 DMD1705 and DMD0907 and LGMD2a-1509 human iPS cells were generated by lentiviral transduction of the pSAM2-PAX7-ires-GFP and pFUGW-rtTA constructs (Darabi et al., 2012). The third generation lentiviral construct pCCL-c-MCS (Zufferey et al., 1998) was kindly provided by Donald Kohn. The tetracycline-responsive-element (TRE) followed by PAX7 or PAX7-ires-GFP sequences and Ubiquitin-C (UBC) promoter followed by the rtTA sequences was subcloned into the pCCL

vector to generate respectively pCCL-PAX7, pCCL-PAX7-ires-GFP and pCCL-rtTA. Details on the cloning strategy are available upon request. These constructs were used to transduce H9 human ES cells as described above. Cells were passaged at 90% of plate density using Accumax (3 min incubation) followed by gentle resuspension. After centrifuge for 5 min at 300g, cell pellets were resuspended in mTeSR supplemented with 10μM of the ROCK inhibitor Y-27632 and replated in matrigel-coated flasks or used for differentiation. For skeletal myogenic differentiation (protocol described in (Darabi and Perlingeiro, 2016)), cells were plated in 6cm non-adherent Petri dishes and incubated at 37°C, 5% CO_2 on a shaker at 60 RPM. After 2 days, mTeSR was replaced with EB Myogenic (EBM) Media supplemented with 10μM Y-27632 and 5μM GSK3 inhibitor (CHIR990217). 2 days later, media was replaced to remove GSK3 inhibitor and the formed EBs were cultured in suspension on the shaker up to day 7. After plating EBs on gelatin-coated flasks using EBM + 10ng/ml human basic FGF (bFGF), EBs were let to adhere on the flasks for 3 days and then supplemented with 1μg/ml doxycycline (dox) for PAX7 induction. At day 14, cells were harvested using Trypsin+EDTA solution and FACS sorted based on GFP expression. Alternatively, unsorted single cells were re-seeded on gelatin-coated flasks using EBM supplemented with 5ng/ml bFGF, 1μg/ml dox and expanded as monolayer until day 23 for FACS sorting using specific antibodies. GFP+ or Marker+ cells were then plated at 1.5*10^6 cells per T25 on gelatin-coated flasks using EBM supplemented with 5ng/ml bFGF, 1μg/ml dox and 5μM ROCK inhibitor (ROCK inhibitor was removed the day after) and cultured as monolayer. At ~90% cell density, cells were passaged using Trypsin+EDTA and replated on new gelatin-coated flasks. Terminal differentiation was induced by switching 100% confluent cultures to KOSR differentiation media (KnockOut™ DMEM + 20% KnockOut™ Serum Replacement + 1% Penicillin/Streptomycin + 2mM GlutaMAX™ - all products from Gibco). Serum-free differentiation using inducible-PAX7 cell lines is based on the protocol published by Chal and colleagues (Chal et al., 2015). At day 9 of the protocol, cultures were supplemented with 1μg/ml doxycycline (dox) for PAX7 induction. Cells were passaged 2 times and then FACS sorted using specific markers or GFP. Sorted cells were then plated at 1.5*10^6 cells per T25 on matrigel-coated flasks using DMEM-F12 + KOSR + Insulin-Transferrin-Selenium (ITS) supplemented with 10ng/ml HGF, 2ng/ml IGF, 1μg/ml dox and 5μM ROCK inhibitor (ROCK inhibitor was removed the day after). Terminal differentiation was induced by withdrawing doxycycline from the medium in 100% confluent cultures. Data presented in the manuscript are originated from biological replicates (at least 3 independent experiments for at least 3 independent cell lines).

FACS analysis, sorting and MACS

A detailed description is provided in Supplemental Experimental Procedures.

RNA isolation, library preparation and sequencing

A detailed description is provided in Supplemental Experimental Procedures.

Transcriptome analysis

Genes were filtered to select only those with median-normalized TPM (Transcripts Per Million reads) greater than 64 in at least one sample, and a fold change greater than 4 between the median-normalized Expected Counts (EC) of any two samples. Heatmaps of

relative expression display for each gene G and sample S the value log2 ((1+median-normalized EC(G,S))/(1+mean over all Si median-normalized EC(G, Si)). Functional annotation of genes (From Supplemental file 1) was performed using the online tool DAVID (Huang da et al., 2009). Graphs represent the output of PANTHER biological process category. Additional information is available in Supplemental Experimental Procedures.

Chromatin-immunoprecipitation, library generation, sequencing and data analysis

A detailed description is provided in Supplemental Experimental Procedures.

Transplantation studies

Animal experiments were carried out according to protocols approved by the University of Minnesota Institutional Animal Care and Use Committee. Both tibialis anterior (TA) muscles of 6–8 weeks old male NOD-scid IL2Rg^{null} (NSG - Jackson lab) or NOD-scid IL2Rg^{null} mdx^{4cv} (Arpke et al., 2013) (*NSG-mdx^{4cv}*) mice were pre-injured with 15 µl of cardiotoxin 10 µM (Latoxan). 24 hrs after cardiotoxin, cells were collected at approximately 60–80% cell density using Enzyme-free cell dissociation buffer (Gibco), then washed with PBS and resuspended in PBS supplemented with 10ng/ml bFGF at the concentration of 5*10^5 cells/10µl.

Cells were injected in one TA per mouse while the contralateral TA received the same volume of PBS as internal control. 8 weeks and 10 months after transplantation, mice were euthanized and TAs collected for immunostaining analysis as previously described (Darabi et al., 2008). Briefly, muscles were frozen in isopentane cooled in liquid nitrogen, and serial 10 µm-thick cryosections were collected. Teratoma studies were performed by injecting 1.5×10^6 H9 ES cells (positive control) or 5×10^6 MACS-purified CD54+ (representing the potential cGMP product) or FACS-purified CD54- (to specifically evaluate their teratoma-forming ability) cells in the quadriceps of NSG mice. Before injection, cells were resuspended in 1:1 solution DMEM-F12 and matrigel (final volume including cells: 65µl).

Immunostaining and quantification

A detailed description is provided in Supplemental Experimental Procedures.

Statistical analysis

Differences between samples were assessed by using the unpaired two-tailed Student's t-test. p values < 0.05 were considered significant.

Data availability

Raw and processed RNA-seq and ChIP-seq data have been deposited in the NCBI Gene Expression Omnibus (GEO) database under accession code GSE98976.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This project was supported by NIH grants R01 AR055299 (R.C.R.P.), U01 HL100407R01 (R.C.R.P. and J.A.T.), the Greg Marzolf Jr Foundation (R.C.R.P.), ADVault, Inc and MyDirectives.com (R.C.R.P.) and by Regenerative Medicine Minnesota (A.M.). The cytogenetic analyses were performed in the Cytogenomics Shared Resource at the University of Minnesota with support from the comprehensive Masonic Cancer Center NIH Grant #P30 CA077598-09.

The authors thank Dr. Jaemin Kim for his assistance with the serum-free differentiation protocol and Dr. Nelio Oliveira for his assistance with the teratoma experiments.

References

Aiuti A, Biasco L, Scaramuzza S, Ferrua F, Cicalese MP, Baricordi C, Dionisio F, Calabria A, Giannelli S, Castiello MC, et al. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. Science. 2013; 341:1233151. [PubMed: 23845947]

Ansari AM, Ahmed AK, Matsangos AE, Lay F, Born LJ, Marti G, Harmon JW, Sun Z. Cellular GFP Toxicity and Immunogenicity: Potential Confounders in in Vivo Cell Tracking Experiments. Stem Cell Rev. 2016; 12:553–559. [PubMed: 27435468]

Arpke RW, Darabi R, Mader TL, Zhang Y, Toyama A, Lonetree CL, Nash N, Lowe DA, Perlingeiro RC, Kyba M. A new immuno-, dystrophin-deficient model, the NSG-mdx(4Cv) mouse, provides evidence for functional improvement following allogeneic satellite cell transplantation. Stem Cells. 2013; 31:1611–1620. [PubMed: 23606600]

Bajard L, Relaix F, Lagha M, Rocancourt D, Daubas P, Buckingham ME. A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. Genes & development. 2006; 20:2450–2464. [PubMed: 16951257]

Bernet JD, Doles JD, Hall JK, Kelly Tanaka K, Carter TA, Olwin BB. p38 MAPK signaling underlies a cell-autonomous loss of stem cell self-renewal in skeletal muscle of aged mice. Nature medicine. 2014; 20:265–271.

Biffi A, Montini E, Lorioli L, Cesani M, Fumagalli F, Plati T, Baldoli C, Martino S, Calabria A, Canale S, et al. Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. Science. 2013; 341:1233158. [PubMed: 23845948]

Borchin B, Chen J, Barberi T. Derivation and FACS-mediated purification of PAX3+/PAX7+ skeletal muscle precursors from human pluripotent stem cells. Stem Cell Reports. 2013; 1:620–631. [PubMed: 24371814]

Carpenter MK, Rao MS. Concise review: making and using clinically compliant pluripotent stem cell lines. Stem Cells Transl Med. 2015; 4:381–388. [PubMed: 25722426]

Chal J, Al Tanoury Z, Hestin M, Gobert B, Aivio S, Hick A, Cherrier T, Nesmith AP, Parker KK, Pourquie O. Generation of human muscle fibers and satellite-like cells from human pluripotent stem cells in vitro. Nat Protoc. 2016; 11:1833–1850. [PubMed: 27583644]

Chal J, Oginuma M, Al Tanoury Z, Gobert B, Sumara O, Hick A, Bousson F, Zidouni Y, Mursch C, Moncuquet P, et al. Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. Nat Biotechnol. 2015

Choi IY, Lim H, Estrellas K, Mula J, Cohen TV, Zhang Y, Donnelly CJ, Richard JP, Kim YJ, Kim H, et al. Concordant but Varied Phenotypes among Duchenne Muscular Dystrophy Patient-Specific Myoblasts Derived using a Human iPSC-Based Model. Cell Rep. 2016; 15:2301–2312. [PubMed: 27239027]

Cooper RN, Tajbakhsh S, Mouly V, Cossu G, Buckingham M, Butler-Browne GS. In vivo satellite cell activation via Myf5 and MyoD in regenerating mouse skeletal muscle. J Cell Sci. 1999; 112(Pt 17):2895–2901. [PubMed: 10444384]

Cosgrove BD, Gilbert PM, Porpiglia E, Mourkioti F, Lee SP, Corbel SY, Llewellyn ME, Delp SL, Blau HM. Rejuvenation of the muscle stem cell population restores strength to injured aged muscles. Nature medicine. 2014; 20:255–264.

Darabi R, Arpke RW, Irion S, Dimos JT, Grskovic M, Kyba M, Perlingeiro RC. Human ES- and iPS-derived myogenic progenitors restore DYSTROPHIN and improve contractility upon transplantation in dystrophic mice. Cell Stem Cell. 2012; 10:610–619. [PubMed: 22560081]

Darabi R, Gehlbach K, Bachoo RM, Kamath S, Osawa M, Kamm KE, Kyba M, Perlingeiro RC. Functional skeletal muscle regeneration from differentiating embryonic stem cells. Nature medicine. 2008; 14:134–143.

Darabi R, Pan W, Bosnakovski D, Baik J, Kyba M, Perlingeiro RC. Functional Myogenic Engraftment from Mouse iPS Cells. Stem Cell Rev. 2011a; 7:948–957. [PubMed: 21461712]

Darabi R, Perlingeiro RC. Derivation of Skeletal Myogenic Precursors from Human Pluripotent Stem Cells Using Conditional Expression of PAX7. Methods Mol Biol. 2016; 1357:423–439. [PubMed: 25403466]

Darabi R, Santos FN, Filareto A, Pan W, Koene R, Rudnicki MA, Kyba M, Perlingeiro RC. Assessment of the myogenic stem cell compartment following transplantation of Pax3/Pax7-induced embryonic stem cell-derived progenitors. Stem Cells. 2011b; 29:777–790. [PubMed: 21374762]

DeKelver RC, Choi VM, Moehle EA, Paschon DE, Hockemeyer D, Meijsing SH, Sancak Y, Cui X, Steine EJ, Miller JC, et al. Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. Genome Res. 2010; 20:1133–1142. [PubMed: 20508142]

Donalies M, Cramer M, Ringwald M, Starzinski-Powitz A. Expression of M-cadherin, a member of the cadherin multigene family, correlates with differentiation of skeletal muscle cells. Proc Natl Acad Sci U S A. 1991; 88:8024–8028. [PubMed: 1840697]

Dull T, Zufferey R, Kelly M, Mandel RJ, Nguyen M, Trono D, Naldini L. A third-generation lentivirus vector with a conditional packaging system. J Virol. 1998; 72:8463–8471. [PubMed: 9765382]

Dumont NA, Wang YX, Rudnicki MA. Intrinsic and extrinsic mechanisms regulating satellite cell function. Development. 2015; 142:1572–1581. [PubMed: 25922523]

Emery AE. The muscular dystrophies. Lancet. 2002; 359:687–695. [PubMed: 11879882]

Epstein JA, Shapiro DN, Cheng J, Lam PY, Maas RL. Pax3 modulates expression of the c-Met receptor during limb muscle development. Proc Natl Acad Sci U S A. 1996; 93:4213–4218. [PubMed: 8633043]

Filareto A, Darabi R, Perlingeiro RC. Engraftment of ES-Derived Myogenic Progenitors in a Severe Mouse Model of Muscular Dystrophy. J Stem Cell Res Ther. 2012:10.

Filareto A, Parker S, Darabi R, Borges L, Iacovino M, Schaaf T, Mayerhofer T, Chamberlain JS, Ervasti JM, McIvor RS, et al. An ex vivo gene therapy approach to treat muscular dystrophy using inducible pluripotent stem cells. Nat Commun. 2013; 4:1549. [PubMed: 23462992]

Gilbert PM, Havenstrite KL, Magnusson KE, Sacco A, Leonardi NA, Kraft P, Nguyen NK, Thrun S, Lutolf MP, Blau HM. Substrate elasticity regulates skeletal muscle stem cell self-renewal in culture. Science. 2010; 329:1078–1081. [PubMed: 20647425]

Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKelver RC, Katibah GE, Amora R, Boydston EA, Zeitler B, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol. 2009; 27:851–857. [PubMed: 19680244]

Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009; 4:44–57. [PubMed: 19131956]

Huard J, Bouchard JP, Roy R, Labrecque C, Dansereau G, Lemieux B, Tremblay JP. Myoblast transplantation produced dystrophin-positive muscle fibres in a 16-year-old patient with Duchenne muscular dystrophy. Clin Sci (Lond). 1991; 81:287–288. [PubMed: 1653675]

Irintchev A, Zeschnigk M, Starzinski-Powitz A, Wernig A. Expression pattern of M-cadherin in normal, denervated, and regenerating mouse muscles. Developmental dynamics : an official publication of the American Association of Anatomists. 1994; 199:326–337. [PubMed: 8075434]

Kim JMA, Chan SSK, Oliveira VKP, Wu J, Darabi R, Kyba M, Perlingeiro RCR. Expansion and Purification are Critical for the Therapeutic Application of Pluripotent Stem Cell-Derived Myogenic Progenitors. Stem Cell Reports. 2017 In press.

Loh KM, Chen A, Koh PW, Deng TZ, Sinha R, Tsai JM, Barkal AA, Shen KY, Jain R, Morganti RM, et al. Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types. Cell. 2016; 166:451–467. [PubMed: 27419872]

Magli A, Schnettler E, Rinaldi F, Bremer P, Perlingeiro RC. Functional Dissection of Pax3 in Paraxial Mesoderm Development and Myogenesis. Stem Cells. 2012

Magli A, Schnettler E, Swanson SA, Borges L, Hoffman K, Stewart R, Thomson JA, Keirstead SA, Perlingeiro RC. Pax3 and Tbx5 specify whether PDGFRalpha+ cells assume skeletal or cardiac muscle fate in differentiating embryonic stem cells. Stem Cells. 2014; 32:2072–2083. [PubMed: 24677751]

Mendell JR, Kissel JT, Amato AA, King W, Signore L, Prior TW, Sahenk Z, Benson S, McAndrew PE, Rice R, et al. Myoblast transfer in the treatment of Duchenne's muscular dystrophy. N Engl J Med. 1995; 333:832–838. [PubMed: 7651473]

Mitchell KJ, Pannerec A, Cadot B, Parlakian A, Besson V, Gomes ER, Marazzi G, Sassoon DA. Identification and characterization of a non-satellite cell muscle resident progenitor during postnatal development. Nat Cell Biol. 2010; 12:257–266. [PubMed: 20118923]

Partridge T, Lu QL, Morris G, Hoffman E. Is myoblast transplantation effective? Nat Med. 1998; 4:1208–1209. [PubMed: 9809519]

Partridge TA, Morgan JE, Coulton GR, Hoffman EP, Kunkel LM. Conversion of mdx myofibres from dystrophin-negative to -positive by injection of normal myoblasts. Nature. 1989; 337:176–179. [PubMed: 2643055]

Quarta M, Brett JO, DiMarco R, De Morree A, Boutet SC, Chacon R, Gibbons MC, Garcia VA, Su J, Shrager JB, et al. An artificial niche preserves the quiescence of muscle stem cells and enhances their therapeutic efficacy. Nat Biotechnol. 2016; 34:752–759. [PubMed: 27240197]

Romero Z, Urbinati F, Geiger S, Cooper AR, Wherley J, Kaufman ML, Hollis RP, de Assin RR, Senadheera S, Sahagian A, et al. beta-globin gene transfer to human bone marrow for sickle cell disease. J Clin Invest. 2013a; 123:3317–3330.

Romero Z, Urbinati F, Geiger S, Cooper AR, Wherley J, Kaufman ML, Hollis RP, de Assin RR, Senadheera S, Sahagian A, et al. beta-globin gene transfer to human bone marrow for sickle cell disease. J Clin Invest. 2013b

Sampaolesi M, Torrente Y, Innocenzi A, Tonlorenzi R, D'Antona G, Pellegrino MA, Barresi R, Bresolin N, De Angelis MG, Campbell KP, et al. Cell therapy of alpha-sarcoglycan null dystrophic mice through intra-arterial delivery of mesoangioblasts. Science. 2003; 301:487–492. [PubMed: 12855815]

Seale P, Sabourin LA, Girgis-Gabardo A, Mansouri A, Gruss P, Rudnicki MA. Pax7 is required for the specification of myogenic satellite cells. Cell. 2000; 102:777–786. [PubMed: 11030621]

Shelton M, Metz J, Liu J, Carpenedo RL, Demers SP, Stanford WL, Skerjanc IS. Derivation and expansion of PAX7-positive muscle progenitors from human and mouse embryonic stem cells. Stem Cell Reports. 2014; 3:516–529. [PubMed: 25241748]

Soleimani VD, Punch VG, Kawabe Y, Jones AE, Palidwor GA, Porter CJ, Cross JW, Carvajal JJ, Kockx CE, van IWF, et al. Transcriptional dominance of Pax7 in adult myogenesis is due to high-affinity recognition of homeodomain motifs. Dev Cell. 2012; 22:1208–1220. [PubMed: 22609161]

Sousa-Victor P, Gutarra S, Garcia-Prat L, Rodriguez-Ubreva J, Ortet L, Ruiz-Bonilla V, Jardi M, Ballestar E, Gonzalez S, Serrano AL, et al. Geriatric muscle stem cells switch reversible quiescence into senescence. Nature. 2014; 506:316–321. [PubMed: 24522534]

Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006; 126:663–676. [PubMed: 16904174]

Tedesco FS, Cossu G. Stem cell therapies for muscle disorders. Curr Opin Neurol. 2012; 25:597–603. [PubMed: 22918488]

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. Embryonic stem cell lines derived from human blastocysts. Science. 1998; 282:1145–1147. [PubMed: 9804556]

Torrente Y, Belicchi M, Sampaolesi M, Pisati F, Meregalli M, D'Antona G, Tonlorenzi R, Porretti L, Gavina M, Mamchaoui K, et al. Human circulating AC133(+) stem cells restore dystrophin

expression and ameliorate function in dystrophic skeletal muscle. J Clin Invest. 2004; 114:182–195. [PubMed: 15254585]

Tremblay JP, Malouin F, Roy R, Huard J, Bouchard JP, Satoh A, Richards CL. Results of a triple blind clinical study of myoblast transplantations without immunosuppressive treatment in young boys with Duchenne muscular dystrophy. Cell Transplant. 1993; 2:99–112. [PubMed: 8143083]

Vilquin JT. Myoblast transplantation: clinical trials and perspectives. Acta Myol. 2005; 24:119–127. [PubMed: 16550929]

Xi H, Fujiwara W, Gonzalez K, Jan M, Liebscher S, Van Handel B, Schenke-Layland K, Pyle AD. In Vivo Human Somitogenesis Guides Somite Development from hPSCs. Cell Rep. 2017; 18:1573–1585. [PubMed: 28178531]

Xu C, Tabebordbar M, Iovino S, Ciarlo C, Liu J, Castiglioni A, Price E, Liu M, Barton ER, Kahn CR, et al. A zebrafish embryo culture system defines factors that promote vertebrate myogenesis across species. Cell. 2013; 155:909–921. [PubMed: 24209627]

Zufferey R, Dull T, Mandel RJ, Bukovsky A, Quiroz D, Naldini L, Trono D. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. J Virol. 1998; 72:9873–9880. [PubMed: 9811723]

HIGHLIGHTS

- Genome-wide analyses reveal gene expression changes during human skeletal myogenesis
- CD54, integrin α9β1 and SDC2 identify PAX7-induced myogenic progenitors
- CD54+α9β1+SDC2+ myogenic progenitors have long-term *in vivo* regenerative potential
- CD54 alone allows for GMP-compatible purification of myogenic progenitors

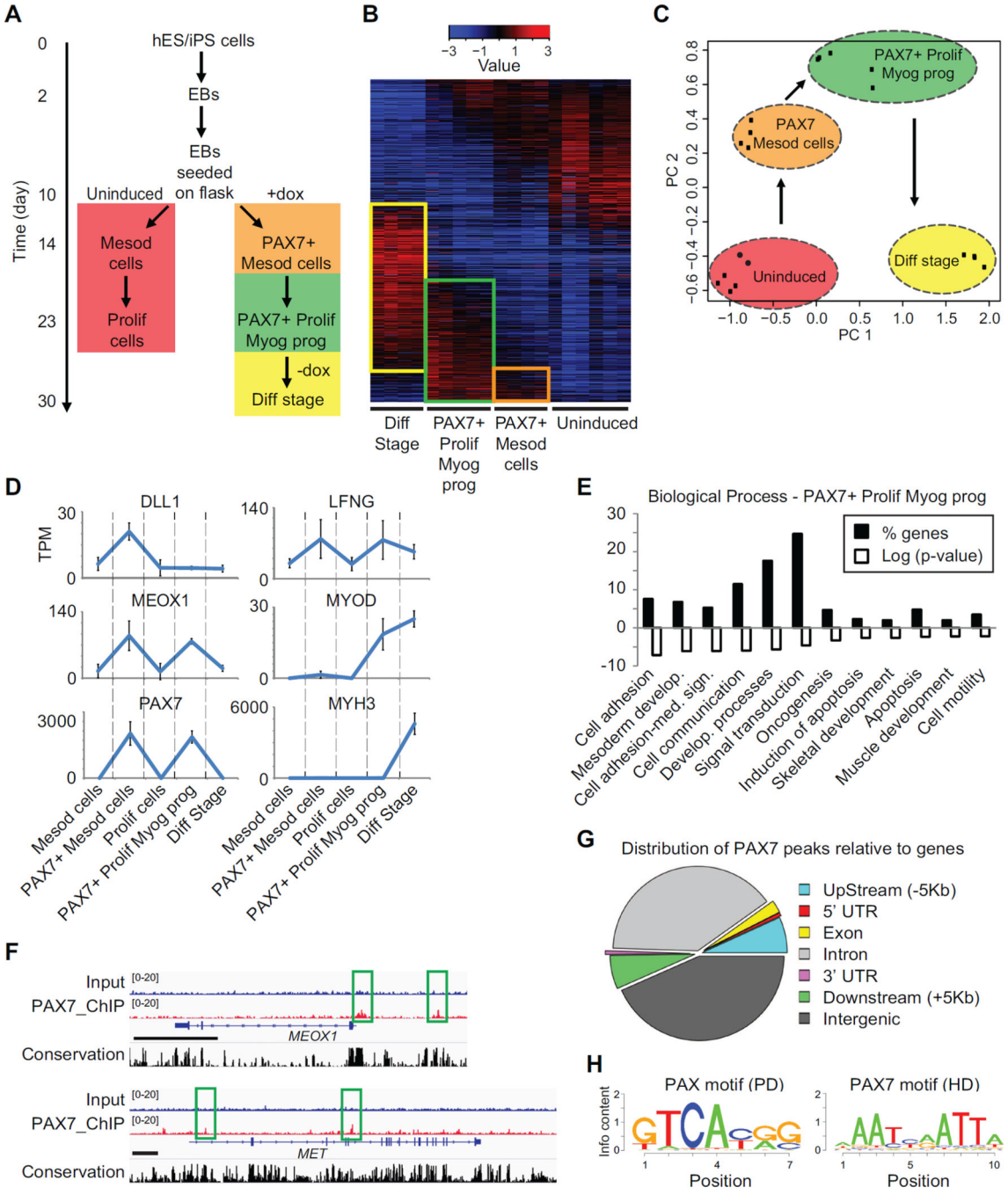


Figure 1.

Identification of gene expression changes during human skeletal myogenesis. (**A**) Schematic representation of the skeletal myogenic differentiation protocol of human pluripotent stem cells using inducible PAX7 expression. Samples were collected at day 14 (Uninduced Mesod cells and PAX7+ Mesod cells), day 23 (Uninduced Prolif cells and dox-treated GFP+ cells, also referred as PAX7+ Proliferating Myogenic prog) and day 30 Day 14 sorted (Diff stage - 7 days post dox withdrawal). (**B**) Heat map representing differentially expressed transcripts (expressed as TPM – Transcripts Per Million reads from RSEM package) between control

and PAX7-induced differentiating human pluripotent stem cells. (**C**) Principal component analysis of differentially expressed genes identified by RNA-seq. PC1 accounts for 78.5% of variance, PC2 for 18.3%. (**D**) Expression profiles of selected genes from RNA-seq. TPM: Transcripts Per Million reads. (**E**) Gene Ontology classification of up-regulated genes based on Biological Process performed using DAVID web resource. Histogram reports % genes and log (p-value) (shown on the negative y axis). (**F**) Snapshot of tracks from the Integrative Genome Viewer (IGV) browser showing PAX7 peaks (green boxes) at the MEOX1 and MET loci. Conservation (Phastcons scale 0–1). Bar: 10Kb. (**G**) Distribution of PAX7 peaks relative to genes following annotation using PAVIS web resource. (**H**) Enriched motifs identified at the PAX7 ChIP-seq peaks. PAX (Paired-domain dependent) zscore: −32.228 and PAX7 (Homeodomain dependent) zscore: −21.562.

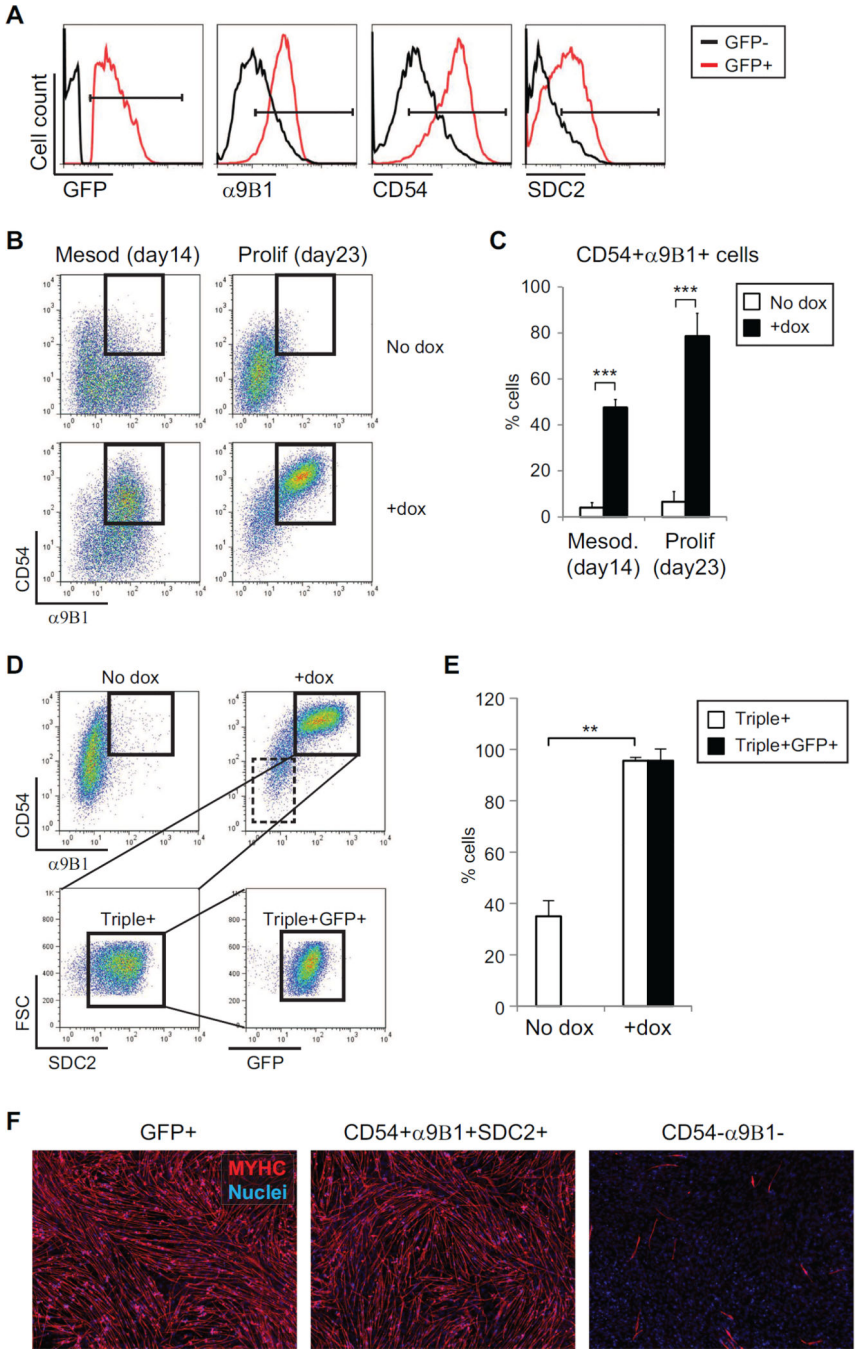


Figure 2.

Human pluripotent-derived myogenic progenitors express CD54, α9β1 and SDC2. (**A**) Histograms from FACS analysis of differentiating pluripotent stem cells using the indicated antibodies. PAX7-expressing cells (GFP+; red line) are compared to control cells (GFP-; black line). (**B**) Dot plot FACS analysis of CD54 and α9β1 at day 14 of differentiation and proliferating myogenic progenitors following PAX7 induction. (**C**) Quantification of FACS data from panel B. (**D**) FACS analysis of proliferating myogenic progenitors and retrospective quantification of GFP expression in CD54+α9β1+SDC2+ fraction. Dashed

box in the upper right FACS plot indicates the fraction of CD54-α9β1- tested for myogenic potential in panel E. Triple+ box in lower left FACS plot indicates the fraction of CD54+α9β1+SDC2+ tested for myogenic potential in panel E. (**E**) Quantification of FACS data (n=5) from panel D. (**F**) Representative immunostaining (n=5) for the skeletal myogenic differentiation marker MYHC (red) following terminal differentiation of CD54+α9β1+SDC2+, GFP+ and CD54-α9β1-(from panel D +dox, dashed box) sorted cells. Nuclei: blue. Bar 100μm. ** $p<0.01$, *** $p<0.001$.

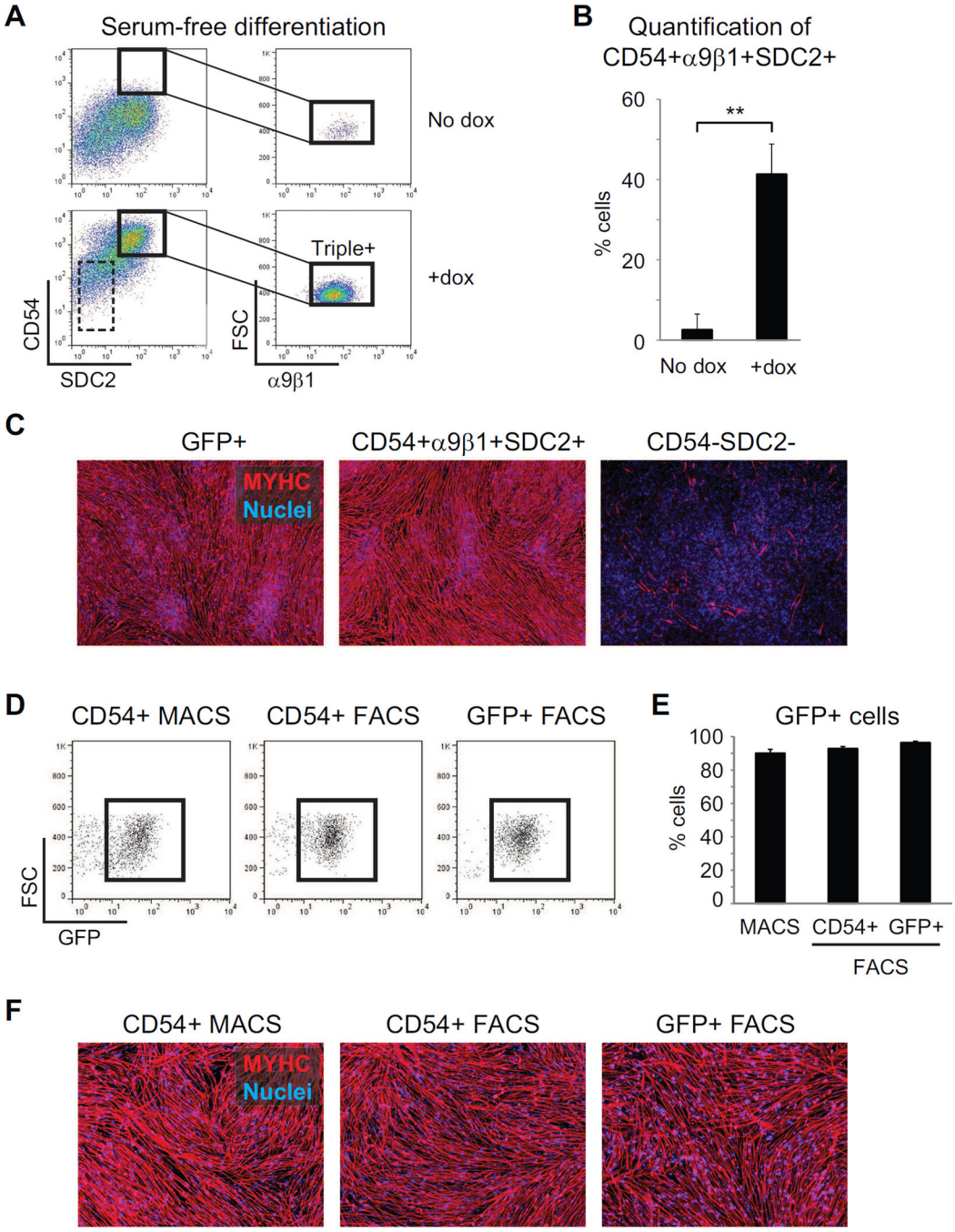


Figure 3.

GMP-compatible purification of human pluripotent-derived myogenic progenitors. (**A**) CD54, α9β1, SDC2 allow purification of myogenic progenitors from pluripotent cells differentiating in serum-free conditions. Representative FACS analysis (from n=4 experiments) of proliferating myogenic progenitors using CD54, α9β1 and SDC2. Dashed box in the lower left FACS plot indicates the fraction of CD54-SDC2- tested for myogenic potential in panel C. (**B**) Quantification of FACS data (n=4) from panel **A**. (**C**) Representative immunostaining (n=4) for the skeletal myogenic differentiation marker

MYHC (red) following isolation of CD54+SDC2+ (from panel **A** +dox, solid box), CD54-SDC2- (from panel **A** +dox, dashed box) and GFP+ cells from serum-free cultures. Nuclei: blue. Bar 100μm. (**D**) Retrospective quantification of GFP expression (n=3) of proliferating myogenic progenitors isolated using MACS or FACS. (**E**) Quantification of FACS data from panel **D**. (**F**) Representative MYHC immunostaining (n=3, red) of CD54- or GFP-sorted myogenic progenitors isolated using MACS or FACS. Nuclei: blue. Bar 100μm. ** $p<0.01$.

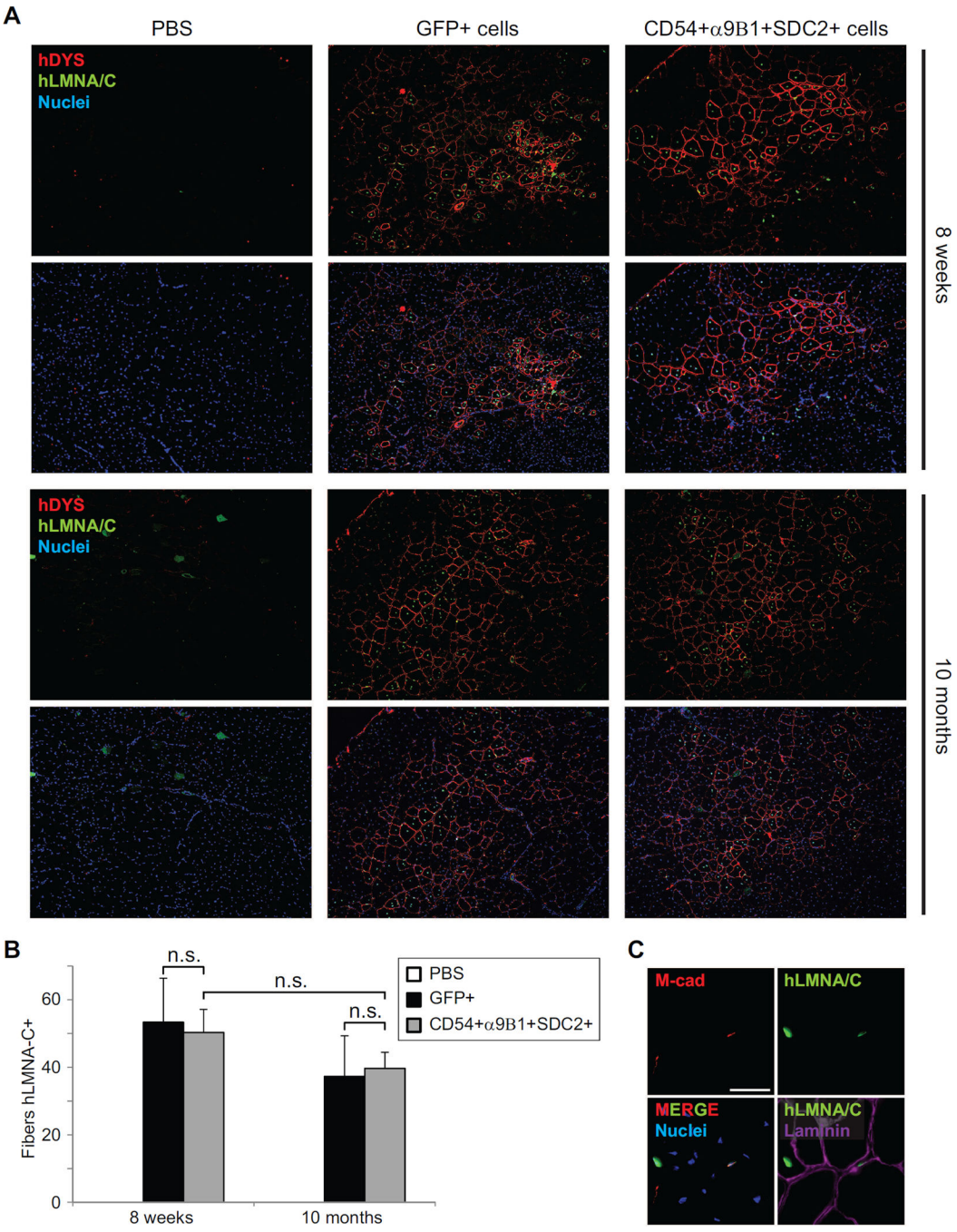


Figure 4.

Skeletal muscle regeneration following transplantation of myogenic progenitors isolated using surface markers. (**A**) Representative immunostaining (n=4 for each group) of cryosections from PBS-injected or GFP+ sorted and CD54+α9β1+SDC2+ cell-transplanted muscles using human-specific antibodies for DYSTROPHIN (red) and LAMIN A/C (green). Nuclei (blue). Bar 100μm. Upper panels show engraftment 8 weeks following intramuscular transplantation whereas lower panels show long-term engraftment at 10 months post-transplantation. (**B**) Quantification of muscle fiber engraftment from panel **A**. Bars represent

Mean ± Standard Error. (**C**) Representative immunostaining of muscle cryosections that had been transplanted with CD54+α9β1+SDC2+ cells, using antibodies for M-cadherin (red), human-specific LAMIN A/C (green) and Laminin α-2 (Violet). Nuclei (blue). Arrowhead indicates a donor derived satellite cell. Bar 20μm. n.s.: not significant.