

ARTICLE

Open Access

# Establishment of stably expandable induced myogenic stem cells by four transcription factors

Eun-Joo Lee<sup>1</sup>, Minhyung Kim<sup>2</sup>, Yong Deuk Kim<sup>1,3</sup>, Myung-Jin Chung<sup>1,3</sup>, Ahmed Elfadl<sup>1,3</sup>, H. M. Arif Ulah<sup>1,3</sup>, Dongsu Park<sup>4,5,6</sup>, Sunray Lee<sup>7</sup>, Hyun-Sook Park<sup>7</sup>, Tae-Hwan Kim<sup>1</sup>, Daehee Hwang<sup>2,8,9</sup> and Kyu-Shik Jeong<sup>1,3</sup>

## Abstract

Life-long regeneration of healthy muscle by cell transplantation is an ideal therapy for patients with degenerative muscle diseases. Yet, obtaining muscle stem cells from patients is very limited due to their exhaustion in disease condition. Thus, development of a method to obtain healthy myogenic stem cells is required. Here, we showed that the four transcription factors, Six1, Eya1, Esrrb, and Pax3, converts fibroblasts into induced myogenic stem cells (iMSCs). The iMSCs showed effective differentiation into multinucleated myotubes and also higher proliferation capacity than muscle derived stem cells both in vitro and in vivo. The iMSCs do not lose their proliferation capacity though the passaging number is increased. We further isolated CD106-negative and  $\alpha 7$ -integrin-positive iMSCs (sort-iMSCs) showing higher myogenic differentiation capacity than iMSCs. Moreover, genome-wide transcriptomic analysis of iMSCs and sort-iMSCs, followed by network analysis, revealed the genes and signaling pathways associated with enhanced proliferation and differentiation capacity of iMSCs and sort-iMSCs, respectively. The stably expandable iMSCs provide a new source for drug screening and muscle regenerative therapy for muscle wasting disease.

## Introduction

Dysfunction of muscle stem cells causes muscle diseases. Although there is much progress in understanding the genetic defects in degenerative muscle diseases, the diseases remain incurable. Therefore, the transplantation of stem cells to damaged muscle tissue has been considered as an ideal therapeutic strategy.

Autologous stem cell transplantation is by far the most common stem cell transplantation. In degenerative muscle diseases, however, application of muscle derived stem cells (MDSC) from patients has been technically challenging. It is very hard to isolate myoblasts or satellite

cells from muscular dystrophy or sarcopenia patients, as these patients show difficulties in regenerating muscle and have less muscle than normal individuals. Additionally, muscle laceration for sampling is very invasive, especially in case of patients with muscular dystrophy or sarcopenia. Further, MDSCs, as multipotent, have the differentiation capacity that is not confined only to myogenesis, but also to adipogenesis, chondrogenesis and osteogenesis<sup>1</sup>. An animal study also showed that satellite cells under aging become fibrogenic rather than myogenic via the activation of Wnt signaling<sup>2</sup>. Thus, a new source for myogenic stem cells that can be used in stem cell therapy is necessary.

It has been revealed that terminally differentiated somatic cells can be directly converted into totally different cell types by forcing ectopic expression of specific transcription factors (TFs), referred to as direct conversion. Various cell types, including neurons<sup>3,4</sup>, hepatocytes<sup>5,6</sup>, cardiomyocytes<sup>7-9</sup>, and blood progenitor cells<sup>10</sup>

Correspondence: Kyu-Shik Jeong (jeongks@knu.ac.kr)

<sup>1</sup>Department of Veterinary Pathology, College of Veterinary Medicine, Kyungpook National University, Daegu 41566, Republic of Korea

<sup>2</sup>School of Interdisciplinary Bioscience and Bioengineering, Pohang University of Science and Technology, Pohang 37673, Republic of Korea

Full list of author information is available at the end of the article

These authors contributed equally: Eun-Joo Lee, Minhyung Kim

Edited by Y. Yang

© The Author(s) 2018



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

were induced from completely different somatic cells using tissue-specific TFs.

The concept that ectopic expression of cell type-specific genes alters cell properties was first reported by Davis et al.<sup>11</sup> in 1987. They revealed that transfected MyoD cDNA converts C3H10T1/2 embryonic fibroblasts into myoblasts. MyoD, a well-known muscle specific TF, converts primary cells including dermal fibroblasts, chondroblasts, smooth muscle, and retinal pigmented epithelial cells into myoblasts and myotubes<sup>12</sup>. Also, Liu et al. suggested that ectopic expression of MyoD and Cx43 make fibroblasts differentiate into muscle fibers<sup>13</sup>. MyoD can induce myogenic differentiation from non-myogenic cells. However, MyoD inhibits proliferation of these transdifferentiated cells and these cells do not have self-renewing and expansion capacities. The proliferation capacity is important for the use of the generated cells because sufficient cells are required for cell transplantation and drug screening. No studies have established stably expandable myogenic stem cells by direct reprogramming, although Naoki et al. induced myogenic progenitor cells with the combination of six transcription factors<sup>14</sup>.

*Sine Oculis* (*Six1*) plays a critical role in muscle development. Loss of *Six1* leads to lethality of fetus. The fetuses of *Six1* mutant mice exhibit impaired primary myogenesis and muscle defects in the diaphragm, forelimb, and hindlimb<sup>15</sup>. Overexpression of *Six1* increases the expression of Pax3<sup>16</sup>. *Six1* promotes the proliferation of Pax7 (+) cells by upregulating *Smad1/5/8*<sup>17</sup>. *Eyes absent1* (*Eya1*) functions as a cofactor of *Six1* to activate *Six1* target genes including Pax3, MyoD and myogenin<sup>18,19</sup>. *Eya1/Eya2* mutant mice show delayed myogenesis during development<sup>18</sup>. *Estrogen related receptor beta* (*Esrrb*) is a TF that upregulates the self-renewal of trophoblasts and embryonic stem cells<sup>20,21</sup>. *Paired box 3* (*Pax3*) plays a crucial role in the formation of ventro-lateral dermo-myotome, which develops into the hypaxial body and limb muscle during the development<sup>22–24</sup>.

In this study, by using a novel combination of these four TFs involved in myogenesis or self-renewal, we have established stably expandable induced myogenic stem cells from fibroblasts and revealed its *in vitro* and *in vivo* proliferation and myogenic differentiation capacity.

## Materials and methods

### Lentivirus preparation

The *Six1* open reading frame was subcloned into the pLJM lentiviral vector (Addgene no. 19319), which have a puromycin resistance gene. *Eya1* and *Pax3* were subcloned into the FUW-tet-O plasmid (Addgene no. 20321). A plasmid containing *Esrrb* (Addgene no. 40798) was purchased from Addgene. A total of  $5 \times 10^6$  of 293 FT cells were plated onto a 100 mm dish. When the 293

FT cells were 90% confluent, they were transfected with 3  $\mu$ g of lentiviral vector delivered by 36  $\mu$ L Lipofectamine 2000 (Invitrogen), 5 mL of opti-MEM, and 9  $\mu$ g of ViraPower lentiviral packaging mixture (Invitrogen). Viral supernatant was harvested 48 h after transfection, centrifuged at 3000 rpm for 15 min at 4°C, and filtered through a 0.45  $\mu$ m filter (Millipore). The titration of viruses was greater than  $5 \times 10^5$  IFU/mL according to measurement with Lenti-X GoStix (Clontech).

### Establishment of iMSCs

To isolate mouse embryonic fibroblasts (MEFs), mouse embryos (C57BL/6J) were isolated from the uterus of pregnant female mice at E13.5 days. Head and red organs were eliminated and remaining whole body parts were cut and minced. The minced tissues were incubated for 10 min at 37°C with 0.05% of trypsin-EDTA for enzymatic digestion and then the enzyme was neutralized. The pellet was suspended in media and filtered through 0.45  $\mu$ m filters. MEFs were maintained in MEF media (10% fetal bovine serum (FBS), 1% penicillin/streptomycin in high glucose DMEM). The MEFs were maintained until passage 3 and seeded at  $5 \times 10^4$ . The MEFs were incubated with *Six1* lentivirus for 48 h. Two days after transduction, the culture media was replaced with 1.5 mL of myogenic growth media (10% FBS, 10% horse serum, 5 ng/mL murine basic fibroblast growth factor (FGF) and 1% penicillin/streptomycin in high glucose DMEM) and the cells were treated with 5  $\mu$ g/mL puromycin for the next 25 days. The media was changed every 2 days. After finishing puromycin selection, the *Six1* transduced cells were seeded at  $5 \times 10^4$  and transduced with the lentivirus of *Esrrb*, *Eya1*, and *Pax3*. After 48 h of transduction with lentivirus of *Esrrb*, *Eya1*, and *Pax3*, the media containing lentivirus was replaced with myogenic growth media. The cells were maintained for 7 days with doxycycline (2  $\mu$ g/mL) to turn on the tet-O system. When the cells showed robust proliferation, single-cell sorting was performed using the FACS Aria and single-cell lines were established. Each single-cell line was propagated and after cell growth and RNA purification, PCR was performed for Pax7 and Myf5.

### Real-time reverse transcriptase-polymerase chain reaction (real-time RT-PCR) analysis

RNA was isolated using Trizol following the manufacturer's protocol. cDNA was synthesized using 0.5  $\mu$ g of the RNA with reverse transcriptase (Invitrogen). For real-time RT-PCR, 2  $\times$  SYBR Green PCR mastermix was used. The primers for Pax7 (QT00147728), Myf5 (QT00199507), MyoD (QT00101983), myogenin (QT00112378), MHC (QT0106850), and GAPDH (QT01658692) were purchased from Qiagen. Rotorgene Q (Qiagen) was used to perform real-time RT-PCR and analyze the results.

### Immunofluorescence (IF)

Cells were grown on 2% gelatin-coated multi-well plates and fixed for 15 min at room temperature in 4% paraformaldehyde. After aspirating fixative and washing with phosphate-buffered saline (PBS), samples were incubated for 10 min at  $-20^{\circ}\text{C}$  in 100% methanol for permeabilization and then rinsed three times with PBS for 5 min each. Samples were then blocked for 1 h at room temperature with 5% BSA. After aspirating the blocking buffer, the samples were incubated with primary antibodies against Pax7 (1:50, Developmental Studies Hybridoma Bank), Myf5 (1:200, Santa Cruz), MyoD (1:100, Santa Cruz), myogenin (1:25, Developmental Studies Hybridoma Bank), and MHC (1:100, Santa Cruz) for overnight at  $4^{\circ}\text{C}$ . The samples were rinsed three times with PBS for 5 min each. The samples were incubated in alexar fluor 594- or TRITC-conjugated secondary antibodies (1:400) for 1 h at room temperature. For counter staining, the samples were incubated in 4', 6-diamidino-2-phenylindole (DAPI) for 10 min and mounted with coverslip.

### Immunoblot analysis

Cells were washed with cold PBS and harvested with RIPA buffer containing complete protease inhibitor cocktail (Roche). Protein concentration was measured according to the Bradford method<sup>25</sup>. Next, 50  $\mu\text{g}$  of total protein was incubated at  $100^{\circ}\text{C}$  for 10 min with electrophoresis sample buffer. The samples were loaded on each well and separated by 6–20% SDS-polyacrylamide gel. Proteins were transferred to a polyvinylidene difluoride (PVDF) membrane. The blot was blocked with 5% skim milk for 1 h at room temperature and incubated overnight at  $4^{\circ}\text{C}$  with primary antibodies against Six1 (1:1000, Cell Signaling), Eya1 (1:500, Abcam), Pax3 (1:500, R&D Systems), Esrrb (1:1000, R&D Systems), GRB2 (1:1000, Santa Cruz), and GAPDH (1:2000, Cell Signaling). After incubation with secondary antibodies (1:2000, Cell Signaling) conjugated to horseradish peroxidase, signals were detected by Super Signal West Dura Extended Duration Substrate (Pierce). The signals were visualized using an Image Analyzer (UVP) or exposed to medical X-ray film.

### Myogenic differentiation in vitro

To evaluate myogenesis in vitro, cells were seeded onto a 2% gelatin-coated dish and incubated until the cells occupied 90% of the culture dish area. Next, the dishes were washed with PBS and the media was changed to myogenic differentiation media. The myogenic differentiation media contained 2% horse serum and 1% penicillin/streptomycin in low-glucose DMEM. Cells were incubated for 3 days in myogenic differentiation media. The media was changed every day. Differentiated cells were analyzed by real-time RT-PCR and IF.

### Proliferation test in vitro

To evaluate proliferation capacity,  $3 \times 10^4$  cells were seeded onto a 2% gelatin-coated dish. The trypsinized cells were stained with trypan blue. The number of cells was counted every 24 h using a hemocytometer and trypan blue. This experiment was repeated three times.

### Animals

Mice were maintained in a room at  $22 \pm 3^{\circ}\text{C}$ , relative humidity  $50 \pm 10\%$ , and 12 h light-dark cycle and were given food and water ad libitum. Animal experiments were performed in accordance with the NIH guidelines for the care and use of laboratory animals and approved by institutional animal care and use committee of Kyungpook National University (Approval no. 2014–0167).

### Myogenic differentiation in vivo

For in vivo myogenic differentiation analysis, 4-week-old male mdx mice were used in vivo. We injected notexin into the middle of tibialis anterior (TA) muscle at a dose of 50  $\mu\text{g}$  to injure the muscle one day before cell transplantation. The iMSCs were prepared and 30,000 cells were suspended in 20  $\mu\text{L}$  of PBS. The cells were transplanted using an insulin syringe. The tibialis anterior muscle was harvested from mice 4 weeks after cell transplantation. To detect dystrophin, staining was performed as previously described<sup>26</sup>. Briefly, frozen sections were blocked overnight at  $4^{\circ}\text{C}$  with 10% horse serum in PBS, and then probed overnight at  $4^{\circ}\text{C}$  with rabbit anti-dystrophin antibody (Abcam, 1:100). TRITC-conjugated anti-rabbit IgG was used as the secondary antibody (Invitrogen, 1:400). Images were obtained by confocal fluorescence microscopy.

### Intravital imaging of iMSCs

pWPXL-based lentiviral vector expressing tdTomato (here after Tomato) was kindly provided by Xiang Zhang from Baylor College Medical School. To generate iRFP670-expressing lentivirus, the EGFP gene in pWPXL lentiviral vector (Addgene) was removed by BamHI/EcoRI digestion and full-length iRFP670 from the piRFP670 plasmid (Addgene) obtained by BamHI/MfeI digestion was cloned into BamHI/EcoRI sites of pWPXL. pWPXL-based lentivirus was produced using HEK 293T cells. To visualize transplanted iMSCs in muscle regeneration, in vitro cultured MDSCs and iMSCs were transduced with iRFP670-expressing lentivirus or tdTomato-expressing lentivirus respectively. Tomato<sup>+</sup> iMSCs and iRFP670<sup>+</sup> MDSCs were sorted by FACS and expanded for transplantation analysis. Muscle injury and stem cell transplantation was performed as described previously<sup>27,28</sup> with minor modifications. Briefly, right TA muscle was injured with 50  $\mu\text{l}$  of barium chloride (1.2%) with 10–20 multiple needle punctures for stem cell transplantation. Twenty-four hours after injury,

in vitro cultured Tomato<sup>+</sup> iMSCs and iRFP670<sup>+</sup> MDSCs and iMSCs ( $1 \times 10^5/50 \mu\text{L}$ ) were transplanted into injured muscle. To sequentially image the muscle after cell engraftment in vivo, mice were anesthetized and prepared for customized multiphoton/confocal hybrid microscope system (Leica TCS SP8 MP platform) specifically designed for live animal imaging as described previously<sup>29</sup>. After the mouse was mounted on a motorized stage, the exposed and injured TA muscle was scanned for second harmonic generation (SHG by femto-second titanium:sapphire laser pulses tuned to 880 nm: Chameleon Vision Laser, Coherent) from muscle fibers. Tomato- (554 nm excitation, 581 nm detection) or iRFP670-expressing cells (630 nm excitation, 670 nm detection) in combination with myofibers (440 nm s harmonic signals) were simultaneously imaged by intravital microscopy. To identify the imaging location of the TA muscle and to repeat the imaging sequentially, the xyz coordinates of imaging position from the lateral tibia head (a landmark position) were recorded.

#### Tumor formation assay

Six-week-old male nude mice (balb/c nu/nu mice) were used in this assay. First,  $5 \times 10^6$  iMSCs and iPSCs were suspended in PBS. The cells were injected subcutaneously into the dorsal part of the nude mice. After 3 weeks, mice were sacrificed and the tumor tissues from the injected site were isolated for histopathological analysis. Tumor tissues were fixed in 10% neutral-buffered formalin, processed routinely, and embedded in paraffin wax. The sections were cut to 4- $\mu\text{m}$  thickness and then deparaffinized in toluene and rehydrated in a graded alcohol series. The sections were stained with hematoxylin and eosin (H&E).

#### mRNA microarray experiments

Total RNA was isolated from MEFs, MDSCs, iMSCs, and sort-iMSC cells using the RNeasy mini kit (Qiagen). RNAs were obtained from two independent biological replicates. RNA integrity was assessed using the Agilent 2100 Bioanalyzer, and RNA integrity numbers for all samples were above 8.0. RNA was reverse-transcribed, amplified, and hybridized to the Agilent SurePrint G3 mouse GE 8  $\times$  60 K microarray, including 62,976 probes corresponding to 23,853 genes, according to Agilent's protocols. Probe intensities were obtained using the Agilent G2565BA microarray scanner and then normalized using the quantile normalization method<sup>30</sup>. The microarray data were deposited to the gene expression omnibus (GEO) database (Accession ID: GSE94506).

#### Statistical analysis of gene expression data

We identified differentially expressed genes (DEGs) from the three comparisons (MDSC vs. MEF, iMSC vs. MEF, and sort-iMSC vs. MEF) using a previously reported

statistical test<sup>31</sup>. Briefly, a *T*-value was computed for each gene. An empirical distribution of the null hypothesis (i.e., a gene is not differentially expressed) was estimated by calculating *T*-values for the genes after randomly permuting the samples and by applying the Gaussian density estimation method to the *T*-values obtained from the random permutations. For each gene, the adjusted *P*-value was computed by performing the two-tailed *T*-test for its *T*-value using the empirical null distribution. DEGs were selected as the genes with (1) adjusted *P*-values  $\leq 0.05$  and 2) absolute  $\log_2$ -fold-changes  $\geq 1$ . Finally, the enrichment analysis of gene ontology biological processes (GOBPs) was performed for a list of genes using DAVID software<sup>32</sup> and the GOBPs represented by the genes were identified as those with the enriched *P*-values  $< 0.05$  (EASE test in DAVID).

#### Network analysis

To build a network model for muscle differentiation defined by iMSCs and sort-iMSC cells, we first selected the cell cycle- and differentiation-related GOBPs represented by the DEGs. We then obtained the DEGs involved in these processes and identified KEGG pathways represented by such DEGs: Calcium, Jak-STAT, MAPK, Notch, TGF- $\beta$ , VEGF, and Wnt signaling pathways, as well as the pathway for the regulation of the actin cytoskeleton. Next, we built a network model describing the interactions among the DEGs involved in these pathways using interaction data in the KEGG pathways and also protein-protein interactions (PPIs) collected from the following PPI databases: the Biological General Repository for Interaction Datasets (BioGRID)<sup>33</sup>, the Database of Interacting Proteins (DIP)<sup>34</sup>, High confidence protein-protein interactions (HitPredict)<sup>35</sup>, the IntAct molecular interaction database (IntAct)<sup>36</sup>, the Molecular INTeraction database (MINT)<sup>37</sup>, and functional protein association networks (STRING)<sup>38</sup>. The nodes in the network model were arranged into the above pathways based on the information in the KEGG pathway database.

#### Statistical analysis

All values are presented as the mean  $\pm$  S.E.M. Statistical analyses were determined using one-way analysis of variance (ANOVA) followed by Tukey's multiple comparison tests. The value of statistical significance was set at  $*P < 0.05$ ,  $**P < 0.01$ , or  $***P < 0.001$ .

## Results

#### Combination of four TFs is critical to establish iMSCs

To induce the direct conversion of fibroblasts to myogenic stem cells, we selected ten TFs, MyoD, Lin28, Pax3, Eya1, Bmi1, Esrrb, Lbx1, Ezh2, Dppa2 and Six1, that are known to be important for myogenic development and stem cell function and tested the function of the TFs in

the direct reprogramming of myogenic stem cells from MEFs. Among them, we found the combination of four master TFs (that will be noted 4F), Six1, Eya1, Esrrb and Pax3, reprogrammed MEFs into myogenic stem cells.

To identify key TFs that reprogram MEFs into myogenic stem cells among the four TFs, we transduced MEFs with only three of the four TFs: 4F-Six1 represented cells transduced with four transcriptional factors minus Six1 (Eya1, Esrrb, Pax3), 4F-Esrrb represented cells transduced with four transcriptional factors minus Esrrb (Six1, Eya1, Pax3), 4F-Eya1 represented cells transduced with four transcriptional factors minus Eya1 (Six1, Esrrb, Pax3) and 4F-Pax3 represented cells transduced with four transcriptional factors minus Pax3 (Six1, Eya1, Esrrb). We evaluated the expression level of myogenic factors of the transduced cells by real-time PCR and IF. In real-time PCR analysis, 4F-Six1 and 4F-Esrrb showed extremely low expression of Pax7, Myf5 and MyoD, which were similar to those of MEFs grown in myogenic growth media. 4F-Eya1 also expressed low levels of myogenic genes. However, Pax7, Myf5 and MyoD were highly expressed in both 4F-Pax3 and 4F. Additionally, 4F-Pax3 showed higher expression than 4F, except for Myf5 (Fig. 1a–c).

These real-time RT-PCR results were coincided with those of IF. According to the IF of Pax7 and MyoD, no or only a few Pax7- and MyoD-positive cells were observed in 4F-Six1 (Pax7 0%, Myf5 0%, MyoD 0%), 4F-Esrrb (Pax7 0%, Myf5 0%, MyoD 0%), and 4F-Eya1 (Pax7 8%, Myf5 10%, MyoD 16%). However, both 4F-Pax3 and 4F had a lot of positive cells for Pax7, Myf5, and MyoD. However, both 4F-Pax3 and 4F showed large numbers of cells positive for Pax7, Myf5, and MyoD. 4F-Pax3 showed 42% of Pax7-positive cells and 57% of MyoD-positive cells. 4F showed 31% of Pax7 positivity and 35% MyoD positivity. 4F-Pax3 showed higher expression of Pax7 and MyoD compared to 4F. However, the expression of Myf5 was higher in 4F (62%) than in 4F-Pax3 (38%) (Fig. 1d).

Interestingly, 4F-Pax3 showed minimal differentiation into muscle fibers when incubated in myogenic differentiation media (DM) for 15 days (Fig. 1e), although early myogenic differentiation markers (Pax7, Myf5, and MyoD) were highly expressed in 4F-Pax3 (Fig. 1a–d). Most 4F-Pax3 cells died in 1 week during incubation in DM, while the viability of 4F was high under the same myogenic DM condition (Fig. 1e). Only a few surviving 4F-Pax3 cells elongated and formed myotubes in myogenic DM. We plate the same number of cells on 2 well chamber slides and wait until the confluency reached 90%. When the cell confluency reached 90%, we incubate each cells in myogenic differentiation media. Thus, the initial cell number of 4F and 4F-Pax3 was same. However almost 60% of the initial cells were disappeared during the medium exchange.

The 4F cells elongated and fused together to form myotubes in myogenic DM (Fig. 1e). After incubation in myogenic DM, 4F showed numerous MHC-positive cells, while 4F-Pax3 showed a few MHC-positive cells (Fig. 1f). Without Pax3, the transduced cells could not differentiate into muscle fibers, although Pax7, Myf5, and MyoD expressions were high, suggesting that the four factors are necessary for giving myogenic differentiation capacity.

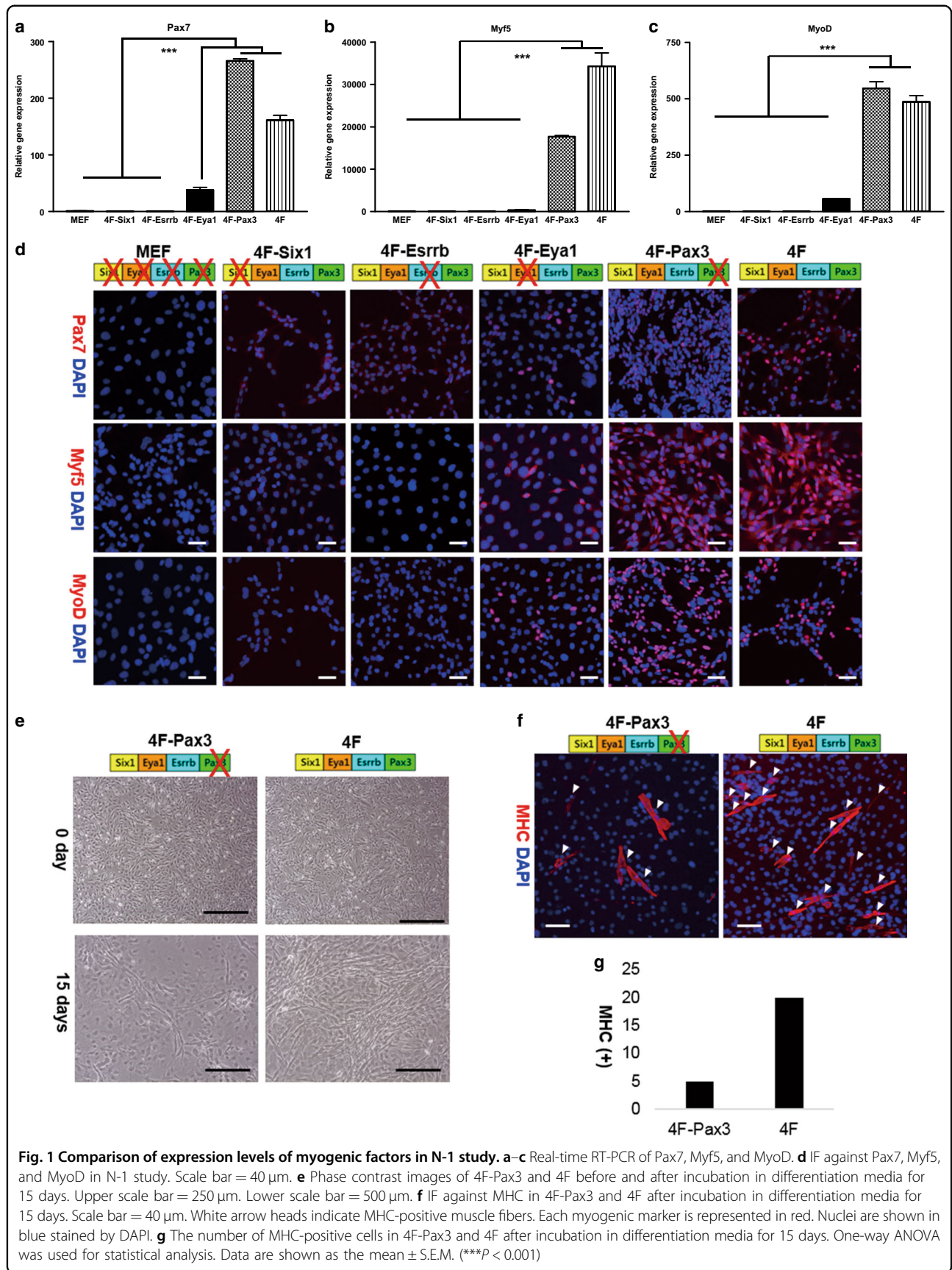
### Establishment of iMSCs with myogenic potentials

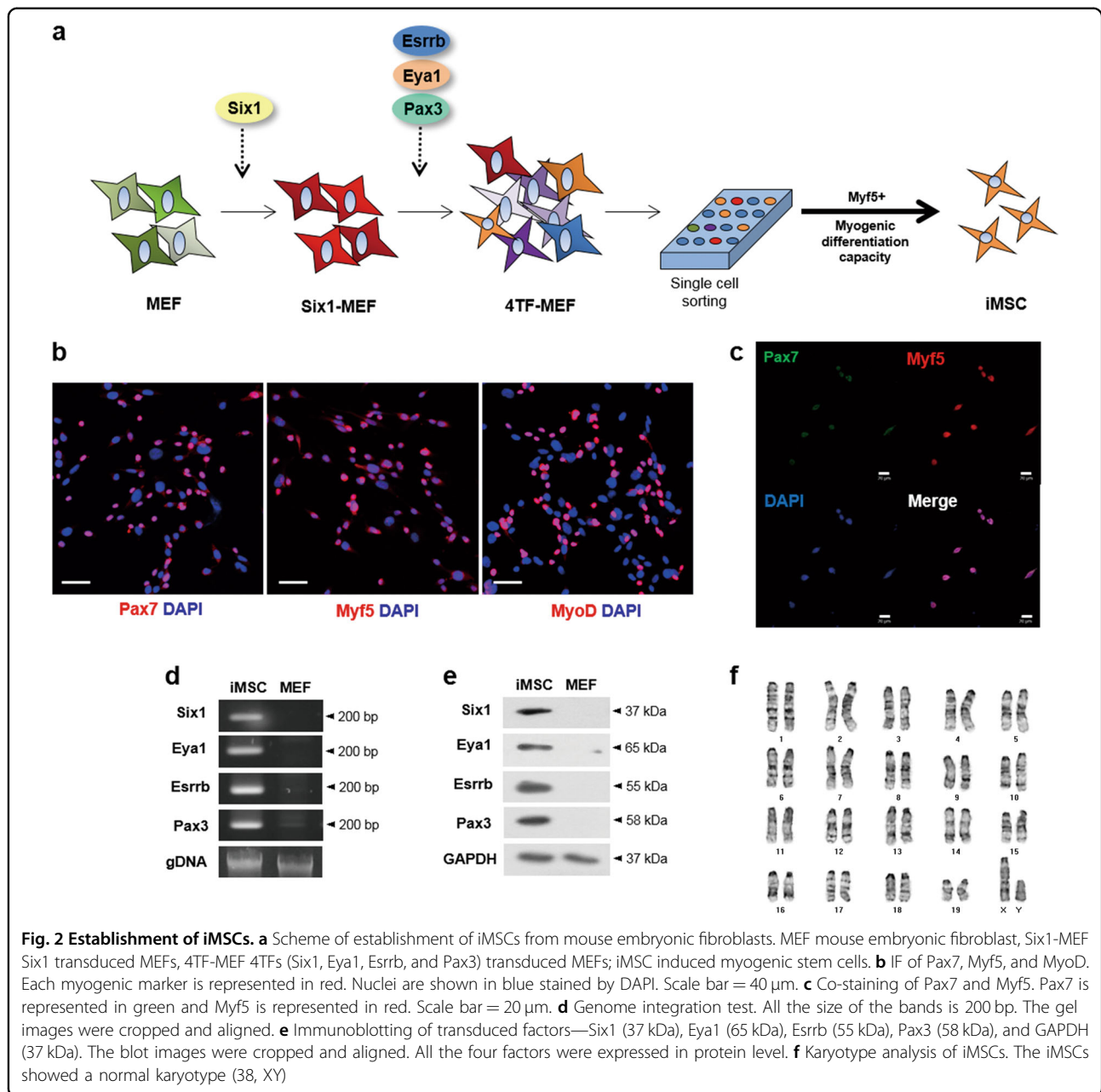
After transduction of the four TFs, we isolated single-cell colonies, since it is hard to distinguish myogenic stem cells from MEFs by morphology. Among the 99 single-cell lines, we sorted out 22 single-cell lines showing high expression of both Pax7 and Myf5. To clarify the myogenic differentiation capacity in vitro, we incubated each single-cell line in myogenic differentiation media. The nine single-cell lines among the 22 of Myf5 + lines represented high myogenic differentiation capacity showing multinucleated muscle fibers. We refer the cells as induced myogenic stem cells (iMSCs). The established iMSCs were positive for Pax7 (55.83%), Myf5 (44.53%), and MyoD (56.28%) in IF (Fig. 2b). And all the Pax7-positive cells were positive for Myf5 (Fig. 2c). The four TFs were integrated stably into the genome of iMSCs and expressed at the protein level (Fig. 2d, e). The karyotype of iMSCs was normal (Fig. 2f).

### iMSCs are more proliferative than MDSCs and MEFs

Prolonged proliferation of stem cells in vitro is a key feature for the therapeutic purpose of stem cells. We, therefore, measured the proliferation capacity of iMSCs. When we compared the cell growth curve of iMSC at early (P12) and late (P50) passages, iMSCs at each two passages showed no difference in their proliferation capacity (Fig. 3a). The doubling time also had no significant difference between the two passages ( $p = 0.322$ ), suggesting that the iMSCs are stably expandable without losing their proliferative capacity though the passage number is increased (Fig. 3b).

When we compare the proliferation capacity of iMSCs to that of MDSCs and MEFs, the iMSCs showed higher proliferation capacity than the others. To measure the proliferation capacity, we seeded  $3 \times 10^4$  of MEFs, MDSCs, and iMSCs in each well and counted the number of cells every 24 h. After 72 h of incubation, the number of iMSCs was  $98.25 \times 10^4 \pm 8.66$ , while the numbers of MEFs and MDSCs were  $12.92 \times 10^4 \pm 5.83$  and  $49.17 \times 10^4 \pm 2.79$ . Thus, these data indicate that iMSCs have higher proliferation capacity than MDSCs and MEFs (Fig. 3c). Additionally, the doubling time of iMSCs (16.76 h) was shorter than MEFs (27.08 h) and MDSCs (21.16 h) (Fig. 3d).





**Fig. 2 Establishment of iMSCs.** **a** Scheme of establishment of iMSCs from mouse embryonic fibroblasts. MEF mouse embryonic fibroblast, Six1-MEF Six1 transduced MEFs, 4TF-MEF 4TFs (Six1, Eya1, Esrrb, and Pax3) transduced MEFs; iMSC induced myogenic stem cells. **b** IF of Pax7, Myf5, and MyoD. Each myogenic marker is represented in red. Nuclei are shown in blue stained by DAPI. Scale bar = 40  $\mu$ m. **c** Co-staining of Pax7 and Myf5. Pax7 is represented in green and Myf5 is represented in red. Scale bar = 20  $\mu$ m. **d** Genome integration test. All the size of the bands is 200 bp. The gel images were cropped and aligned. **e** Immunoblotting of transduced factors—Six1 (37 kDa), Eya1 (65 kDa), Esrrb (55 kDa), Pax3 (58 kDa), and GAPDH (37 kDa). The blot images were cropped and aligned. All the four factors were expressed in protein level. **f** Karyotype analysis of iMSCs. The iMSCs showed a normal karyotype (38, XY)

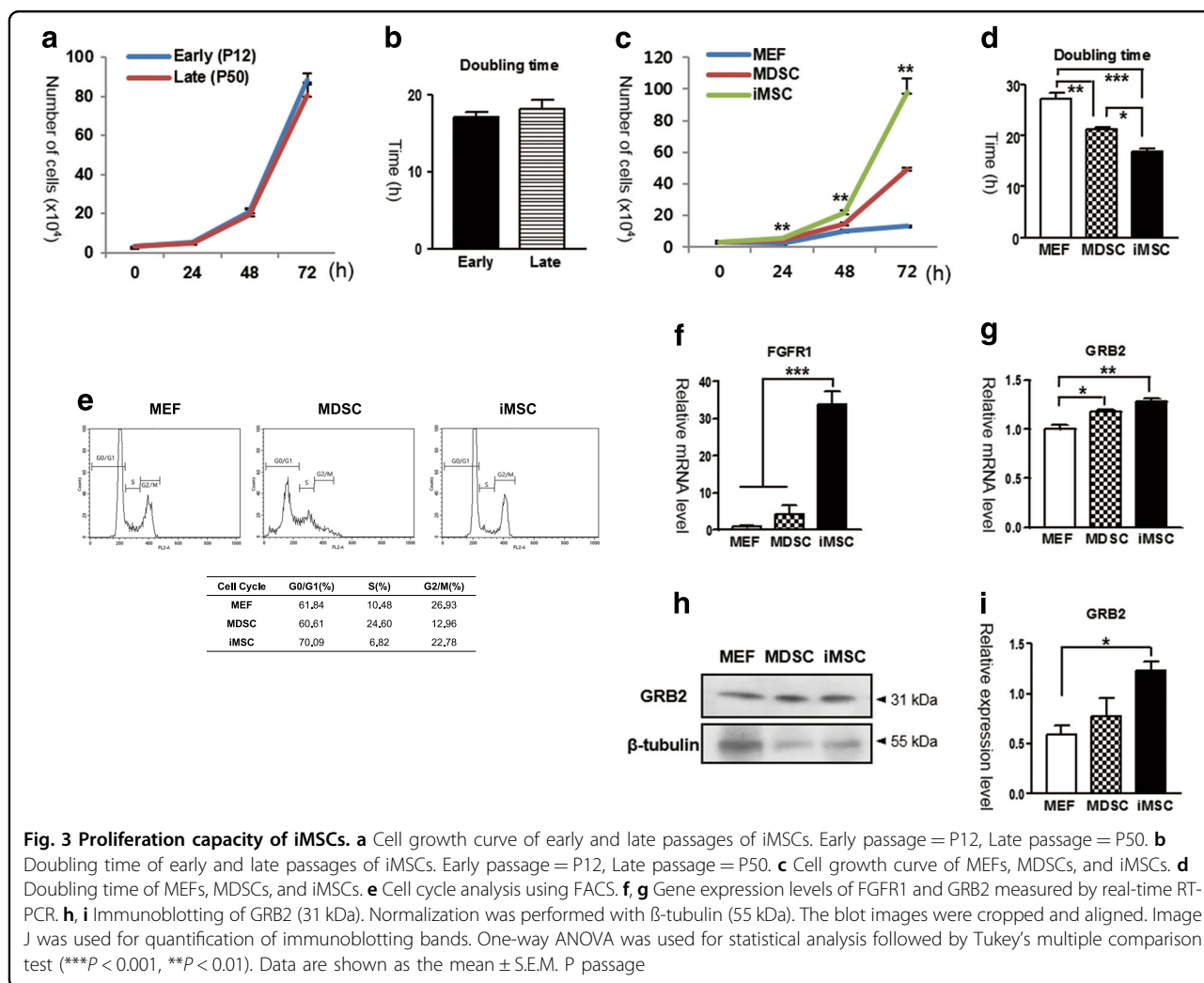
When we evaluated the cell cycle of iMSCs, 70.09% of iMSCs were at G0/G1 phase, which was higher than the proportions of MDSCs (60.61%) and MEFs (61.84%) at G0/G1 phase, whereas the proportions of iMSCs at the S phase and G2/M phase were 6.82% and 22.78%, respectively, which were relatively low (Fig. 3e).

To investigate a high proliferation capacity of iMSCs, we evaluated the expression of fibroblast growth factor receptor1 (FGFR1) and growth factor receptor bound protein 2 (GRB2) in iMSCs. According to real-time RT-PCR, iMSCs showed significantly higher expression of FGFR1 than MEFs and MDSCs ( $***P < 0.001$ ) (Fig. 3f).

Also, iMSCs showed higher mRNA and protein levels of GRB2 than MEFs and MDSCs (Fig. 3g–i).

#### More myogenic iMSCs can be further enriched by CD106-negative and $\alpha$ 7-integrin-positive fraction

To enhance myogenic differentiation capacity, we sorted the iMSCs with the cell surface markers, CD106 and  $\alpha$ 7-integrin, based on the FACS analysis results (Fig. S1). A subset of iMSCs was found to be CD106 (-) and  $\alpha$ 7-integrin (+) ( $14 \pm 2.14$  %) (Fig. 4a). We referred to these as sort-iMSCs. When we maintained the iMSCs and sort-iMSCs in growth media, the sort-iMSCs were more

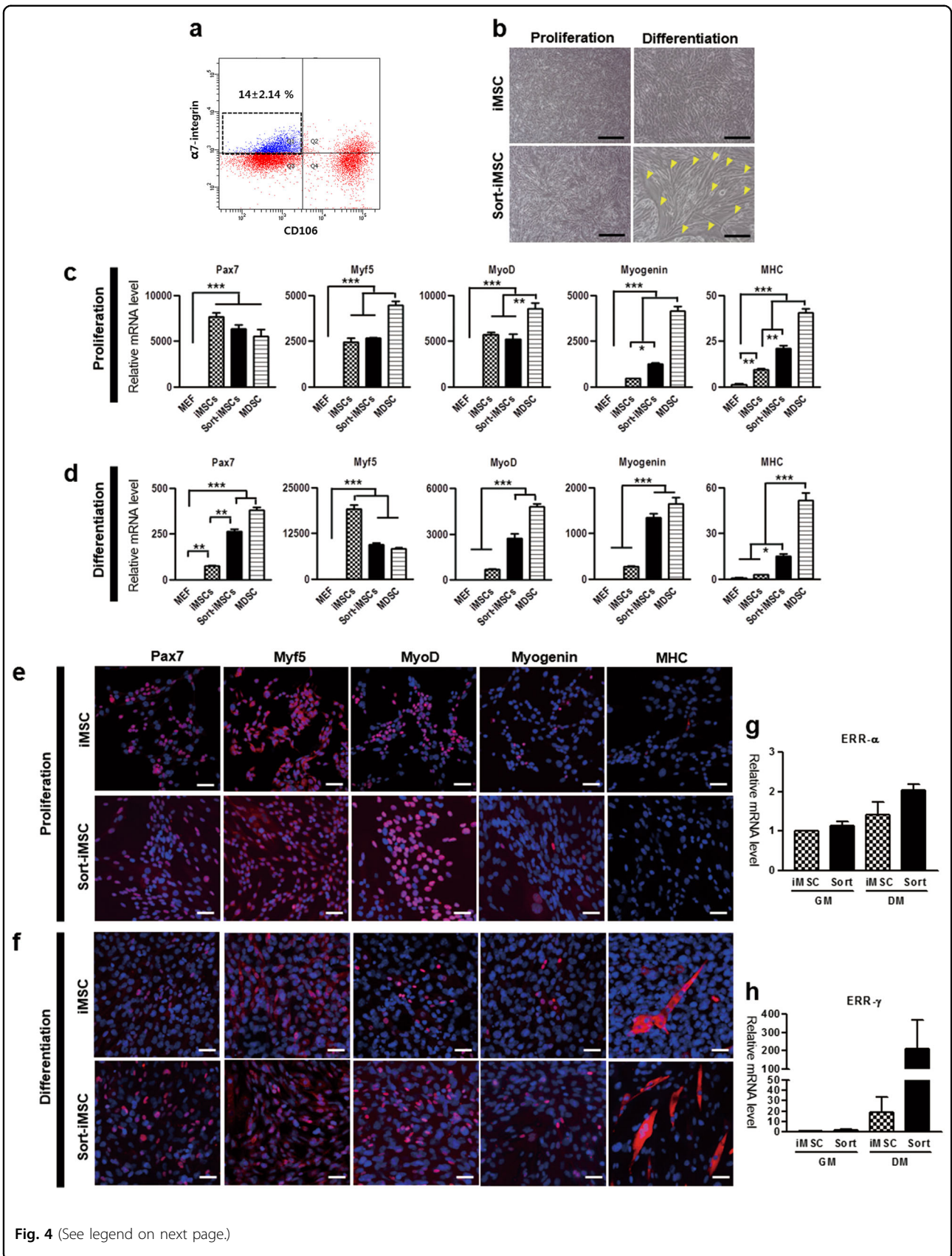


refractive and smaller than iMSCs. In 1 day after incubation of sort-iMSCs in myogenic DM, the cells began to elongate and fuse together, causing the nucleus to come into line. In 3 days after incubation in myogenic DM, sort-iMSCs formed multinucleated branch-shaped myotubes. The iMSCs also showed such differentiation in myogenic DM, but only a limited differentiation to few multinucleated myotubes, which are a late myogenic differentiation marker (Fig. 4b).

To clarify the myogenic differentiation capacity of iMSCs and sort-iMSCs, we compared the expression of Pax7, Myf5, MyoD, myogenin, and MHC under both proliferation and differentiation conditions. Both iMSCs and sort-iMSCs showed high expression levels of Pax7 comparable to MDSCs in proliferation media. Both iMSCs and sort-iMSCs showed higher expression of Myf5, MyoD, myogenin, and MHC in proliferation media than MEFs, but lower expression than MDSCs (Fig. 4c–e). On the other hand, sort-iMSCs showed higher expression of myogenin and MHC than iMSCs, although iMSCs and

sort-iMSCs showed similar expression of Pax7, Myf5, and MyoD (Fig. 4c). We maintained the cells in myogenic DM for 3 days. In 3 days after incubation in DM, sort-iMSCs showed higher expression of all the myogenic regulatory genes except Myf5 compared to iMSCs. Additionally, sort-iMSCs showed comparable expression of Myf5 and myogenin to MDSCs in DM (Fig. 4d–f). Additionally, sort-iMSCs showed comparable expression of Myf5 and myogenin to MDSCs in DM (Fig. 4d–f). Taken together, iMSCs showed upregulation of early myogenic regulatory factors including Pax7, Myf5, and MyoD in proliferation media. In myogenic differentiation media, sort-iMSCs showed upregulation of late myogenic regulatory factors including myogenin and MHC. Thus, sort-iMSCs exhibited higher myogenic differentiation capacity than iMSCs. To clarify the reason for higher myogenic differentiation capacity of sort-iMSCs, we checked the expression of ERR-α and ERR-γ, known as regulators of myogenic differentiation through regulation of mitochondrial biogenesis in skeletal muscle<sup>39</sup>. ERR-γ and ERR-α in sort-iMSCs





(see figure on previous page)

**Fig. 4 Comparison of iMSCs and sort-iMSCs in both GM and DM.** **a** FACS sorting of iMSCs using CD106(-) and  $\alpha$ 7-integrin(+). **b** Phase contrast images of iMSCs and sort-iMSCs. iMSCs and sort-iMSCs were maintained in proliferation or myogenic differentiation media. Yellow arrows indicate multinucleated fibers, indicating late-stage muscle differentiation in vitro. Scale bar = 500  $\mu$ m. **c** Gene expressions of MEF, MDSCs, iMSCs, and sort-iMSCs under GM. **d** Gene expressions of MEF, MDSC, iMSCs, and sort-iMSCs under DM. **e** IF for Pax7, Myf5, MyoD, myogenin, and MHC of iMSCs and sort-iMSCs in GM. **f** IF for Pax7, Myf5, MyoD, myogenin, and MHC of iMSCs and sort-iMSCs in DM. Each myogenic marker is represented in red. Nuclei are shown in blue stained by DAPI. Scale bar = 40  $\mu$ m. **g, h** Expression levels of ERR- $\alpha$  and ERR- $\gamma$  of iMSCs and sort-iMSCs in GM and DM measured by real-time RT-PCR. Data are shown as the mean  $\pm$  S.E.M (\*\* $P$  < 0.001, \*\* $P$  < 0.01, \* $P$  < 0.05). GM growth media, DM differentiation media

were upregulated in DM condition, compared to iMSCs (Fig. 4g, h). The increased expression of ERR- $\gamma$  and ERR- $\alpha$  in sort-iMSCs indicates that sort-iMSCs are more potent in oxidation, leading to upregulation of myogenic regulatory factors, including myogenin, and myogenic differentiation.

#### Sort-iMSCs restore more dystrophin than MDSCs and MEF in mdx mice

So far, we checked the characteristics of iMSCs and sort-iMSCs in vitro. We evaluated myogenic differentiation capacity of sort-iMSCs in vivo, since sort-iMSCs showed more potent myogenic differentiation capacity than iMSCs in vitro. To evaluate myogenic differentiation of sort-iMSCs in vivo, we injected the cells into mdx mice that do not express dystrophin. In 4 weeks after cell injection, the sort-iMSCs injected muscles were regenerated and appeared normal in gross findings without any neovascularization and granulation tissue (Fig. 5a). The sort-iMSCs transplanted-TA muscle contained dystrophin-positive muscle fibers, whereas either MEFs or MDSCs transplanted-TA muscle showed few or only a few dystrophin-positive fibers. These data indicate that the injected sort-iMSCs were engrafted successfully, survived in the injured muscle, and differentiated into muscle fibers. We confirmed that the myogenic differentiation capacity of sort-iMSCs was higher than MEFs and MDSCs in vivo (Fig. 5b, c). In H&E staining, we detected calcification area (arrow heads in Fig. 5c) in all the notexin injected-mdx mice except sort-iMSCs group. Reduced calcification area in sort-iMSCs group represents that more recovery of damaged muscle cells is achieved in sort-iMSCs group than other group (Fig. 5c). In Azan staining, unlike deep-blue stained MDSC injected group, the sort-iMSCs injected group showed weak stainability representing less collagen fibers produced in sort-iMSCs injected group. None of the injected mice showed apparent immune reactions and infiltration of inflammatory cells (Fig. 5c).

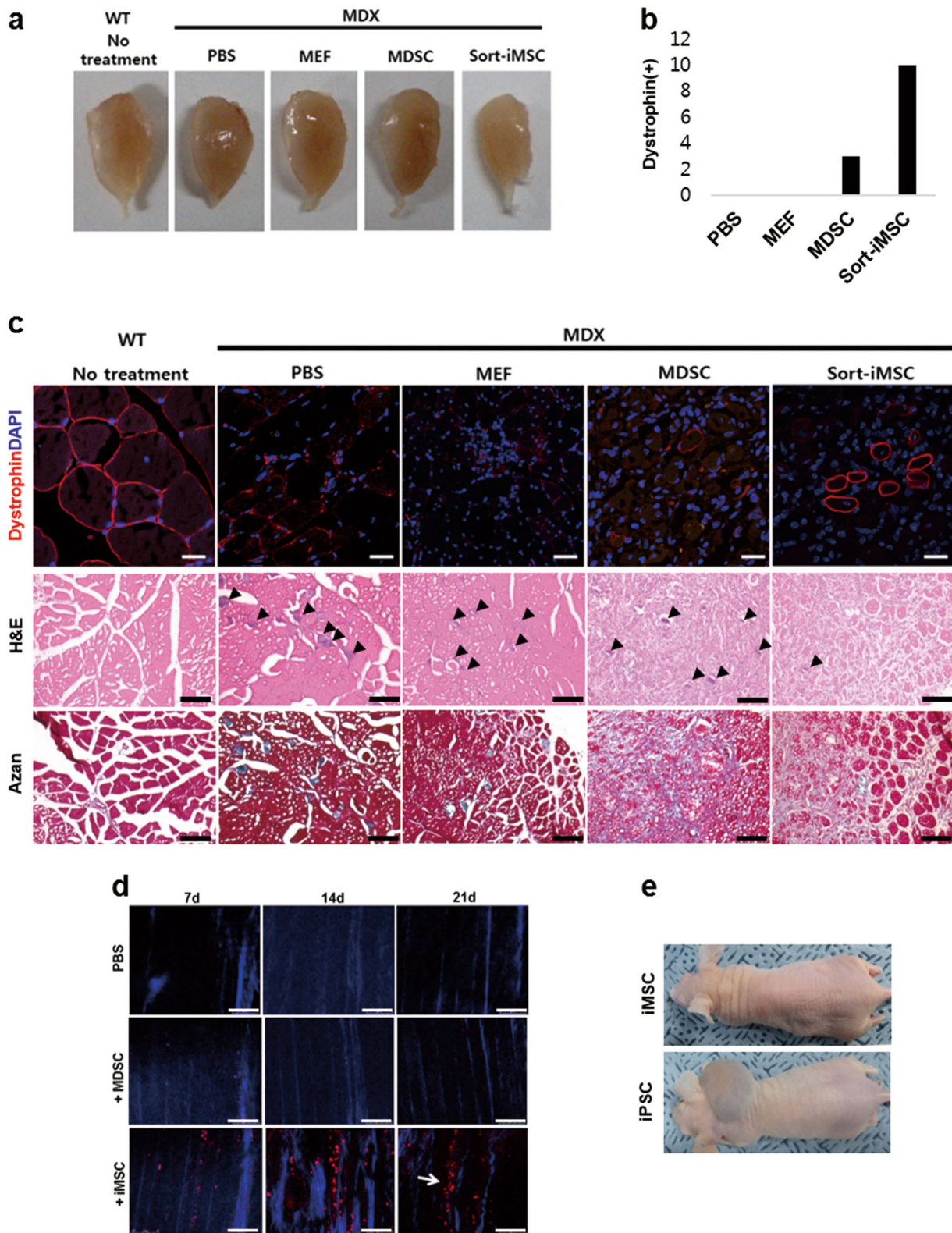
#### iMSCs are more proliferative than MDSCs in mdx mice

To evaluate the early engraftment and proliferation of iMSCs in vivo, we transplanted Tomato + iMSCs ( $5 \times 10^4$  cells/mouse) or iRFP + MDSCs ( $5 \times 10^4$  cells/mouse as a

control) into injured muscles and track them in vivo real-time by using an *intravital* imaging system sequentially. At 7 days after transplantation, the Tomato + iMSCs engrafted successfully and aligned with existing fibers were observed. Sequential *intravital* imaging at 14 and 21 days after transplantation revealed that the number of transplanted tomato + iMSCs was increased and they incorporated into regenerating fibers at 21 days, indicating iMSCs are transplantable and retain their myogenic ability even after transplantation in vivo in injured muscle. The engraftment efficiency and in vivo proliferation capacity were much higher in iMSCs compared to in MDSCs (Fig. 5d). Because of the high proliferation capacity of iMSCs in vivo, we evaluated whether the iMSCs have tumorigenicity. No tumor mass was observed in iMSCs injected mice, whereas iPSC-injected nude mice showed teratoma in 3 weeks (Fig. 5e). We confirmed that the iMSCs did not induce tumor formation when injected in vivo.

#### Genome-wide mRNA expression analysis of iMSCs and sort-iMSCs

To understand the molecular nature underlying differentiation and proliferation capacity of iMSCs and sort-iMSCs, we performed genome-wide gene expression profiling of MDSCs, MEFs, iMSCs, and sort-iMSCs using the Agilent SurePrint G3 mouse GE  $8 \times 60$  K microarray. Using the gene expression profiles, we identified a total of 5,111 differentially expressed genes (DEGs) from the following three comparisons: (1) MDSC vs. MEF (3,543 DEGs with 1560 upregulated and 1983 downregulated), (2) iMSCs vs. MEF (2,687 DEGs with 1,064 upregulated and 1,623 downregulated), and (3) sort-iMSCs vs. MEF (3,486 DEGs with 1,382 upregulated and 2,104 downregulated) (Fig. 6a and Supplementary Table S1). To systematically explore these DEGs, we categorized them into 23 clusters (C1-23; Supplementary Table S2) based on their up- and downregulation in the tree comparisons. Of them, we focused on 10 clusters (C1 and C4-12) showing differential expression in the two comparisons of iMSCs vs. MEFs and sort-iMSCs vs. MEFs. These 10 clusters were further grouped into six groups (G1-6) based on their up- and downregulation patterns in the two comparisons (Fig. 6b and Supplementary Table S1).



**Fig. 5** (See legend on next page.)

(see figure on previous page)

**Fig. 5 Myogenic differentiation and proliferation capacity in mdx mice.** **a** In vivo differentiation capacity of sort-iMSCs. Gross observation of cell-transplanted T.A muscle of mdx mice. **b** The number of dystrophin-positive cells in IF staining against dystrophin. **c** IF staining against dystrophin (upper panel). Representative images of transverse section of T.A muscle. Dystrophins are represented in red. Nuclei are shown in blue stained by DAPI. Scale bar = 40  $\mu$ m. Hematoxyline and eosin staining (middle panel). Arrow heads indicate calcification area. Scale bar = 100  $\mu$ m. Azan staining (lower panel). Scale bar = 100  $\mu$ m. **d** In vivo proliferation capacity of iMSCs. Sequential images of TA muscle repair transplanted with  $1 \times 10^5$  iMSCs or MDSCs at 7, 14, and 21 days after transplantation. Collagen fibrils of the muscle fiber are shown in blue. The transplanted cells are represented in red. Arrow indicates regenerating fiber cells. Scale bar = 75  $\mu$ m. **e** Tumor formation test of iMSCs in nude mice. WT wild type, iPSC induced pluripotent stem cell

G1 and G5 comprised clusters up- and downregulated in iMSCs, respectively, but not in sort-iMSCs, compared to MEFs, whereas G3 and G6 comprised clusters up- and downregulated in sort-iMSCs, respectively, but not in iMSCs. G2 and G5 comprised clusters up- and downregulated commonly in both iMSCs and sort-iMSCs, respectively.

To understand the cellular processes associated with G1-6, we performed the enrichment analysis of GOBPs for the genes in G1-6 using DAVID software (Fig. 6c and Supplementary Tabls S3). G1 upregulated in iMSCs, compared to MEFs, significantly ( $P < 0.05$ ) represented cell cycle-related processes (cell cycle and mitotic cell cycle), whereas G4 downregulated in iMSCs represented apoptosis-related processes (programmed cell death and apoptosis). Interestingly, of these processes, only the apoptosis-related processes were represented by the downregulated genes in MDSCs, compared to MEFs, but the cell cycle-related were represented by neither the up- nor downregulated genes in MDSCs. Collectively, these data suggest that iMSCs show higher proliferation capacity than MEFs and even than MDSCs, consistent to our findings in Fig. 3c, d. Moreover, G3 and G5 up- and downregulated in sort-iMSCs, respectively, compared to MEFs, represented differentiation-related processes (tissue morphogenesis, epithelial tube morphogenesis, blood vessel development, bone development, and cell morphogenesis). All of these processes were also represented by the DEGs in MDSCs, compared to MEFs. These data suggest that sort-iMSCs show differentiation capacity compatible to MDSCs, which is consistent with our findings shown in Fig. 4d. Moreover, two differentiation-related processes (tissue and epithelial tube morphogenesis) were represented by G3, but by neither G1 nor G2, consistent to enhanced differentiation capacity of sort-iMSCs compared to iMSCs, consistent to our finding in vitro shown in Fig. 4.

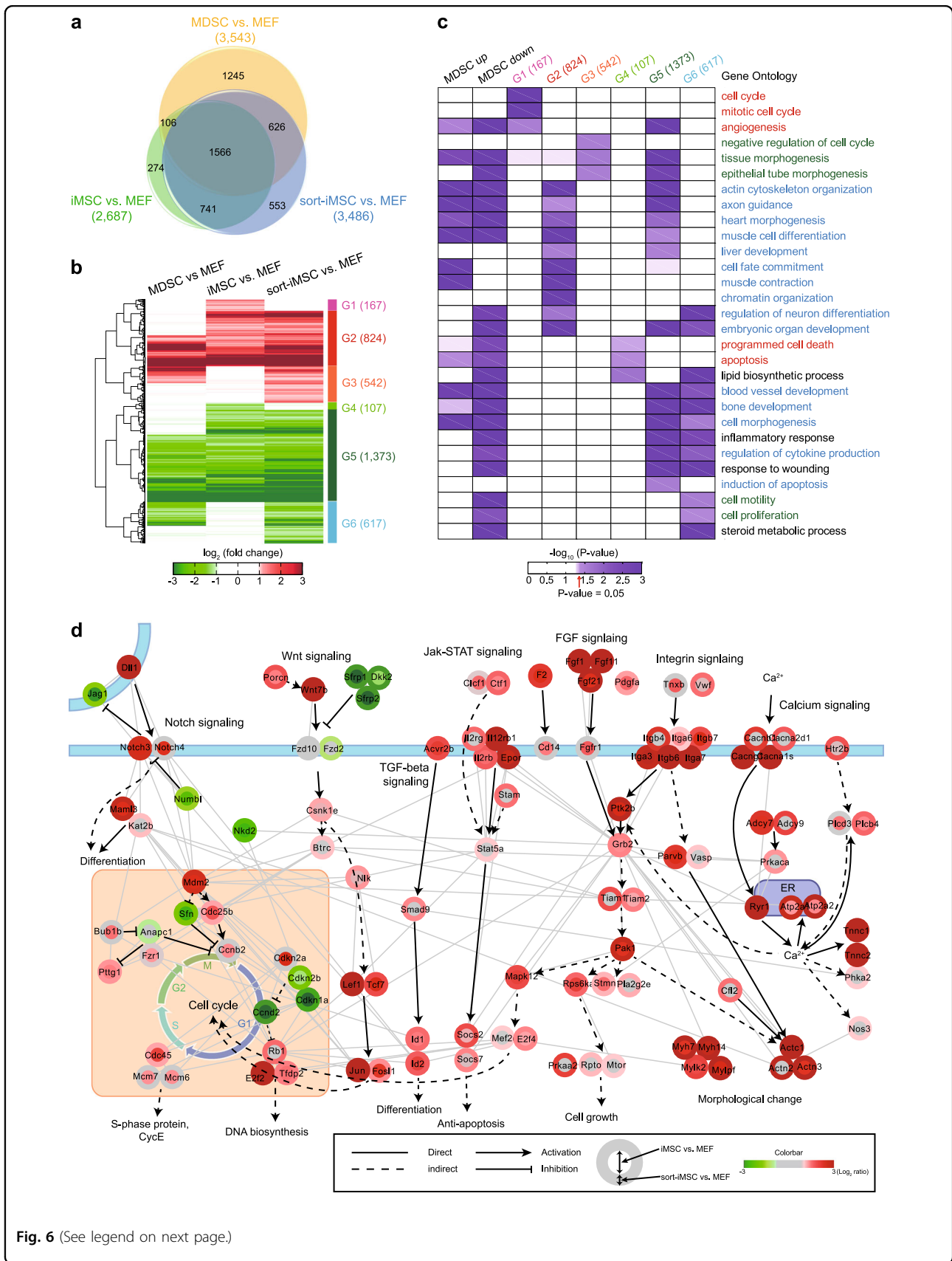
#### Network model describing signaling pathways associated with enhanced proliferation and differentiation capacity of iMSCs and sort-iMSCs

To examine the intracellular signaling pathways associated with proliferation and differentiation capacity of iMSCs and sort-iMSCs, we first selected the DEGs

involved in the aforementioned cell cycle- and differentiation-related cellular processes represented by G1-G6 and then identified KEGG signaling pathways represented by these DEGs. Next, we reconstructed a network model describing the interactions among the DEGs involved in these signaling pathways (Fig. 6d). First, Notch and TGF- $\beta$  signaling pathways were activated in iMSCs and sort-iMSCs, as indicated by upregulation of Notch3/4, Dll1, and Maml3 (Notch signaling) and Acvr2b, Smad9, and Id1/2 (TGF- $\beta$  signaling) in iMSCs or sort-iMSCs, compared to MEFs. Second, Wnt and Jak-stat signaling pathways were activated in iMSCs and sort-iMSCs as indicated by upregulation of Wnt7b, Porcn, CsnK1e, Btrc, Lef1, and Tcf7 (Wnt signaling) and Il2rg/b, Il12rb1, Stam, Stat5a, and Socs2/7 (Jak-stat signaling) in iMSCs or sort-iMSCs, compared to MEFs. These pathways are known to be involved in cell proliferation and/or differentiation, suggesting their contribution to high proliferation capacity of iMSCs (Fig. 3c) and high differentiation capacity of sort-iMSCs (Fig. 5b). Third, FGF and integrin signaling pathways were activated in iMSCs and sort-iMSCs as indicated by upregulation of Fgf1/11/21 and Fgfr1 (Fgf signaling) and Itga3/6/7 and Itgb4/6/7 (Integrin signaling), as well as upregulation of downstream signaling molecules in the actin reorganization pathway (Ptk2b, Grb2, Tiam1/2, Pak1, Parvb, and Vasp) in iMSCs or sort-iMSCs, compared to MEFs. Finally, calcium signaling was activated in iMSCs and sort-iMSCs as indicated by upregulation of calcium transporters (Cacna1s/a2d1/b1/g1, Ryr1, and Atp2a1/2). The FGF, integrin, and calcium signaling pathways are known to activate myosins (Myh7/14 and Mylk2/pf), actins (Actc1/n2/n3) and troponins (Tnnnc1/2), suggesting their contribution to morphological changes required for proliferation of iMSCs and differentiation of sort-iMSCs. Collectively, the network model suggests that these signaling pathways can play key roles in conferring higher proliferation and differentiation capacity to iMSCs and sort-iMSCs, respectively.

#### Discussion

In this study, we established induced myogenic stem cells (iMSCs) by ectopic expression of Six1, Eya1, Esrrb, and Pax3. The iMSCs have myogenic differentiation



**Fig. 6** (See legend on next page.)

(see figure on previous page)

**Fig. 6 Comparison of genome-wide mRNA expression of MDSCs, MEFs, iMSCs, and sort-iMSCs.** Differentially expressed genes in iMSCs and sort-iMSCs and their associated cellular processes and signaling pathways. **a** Relationship between the DEGs from the three comparisons, MDSCs vs. MEFs (3543 DEGs), iMSCs vs. MEFs (2687 DEGs), and sort-iMSCs vs. MEF (3486 DEGs). **b** Differential expression patterns of the DEGs in the six groups (G1-6). Numbers in parenthesis denote the sizes of the DEGs in the corresponding groups. Colors in the heat map represent up- (red) and downregulation (green) of the genes (rows in the heat map). Color bar, gradient of  $\log_2$ -fold-changes of DEGs in the three comparisons. The dendrogram was generated by performing a hierarchical clustering of the  $\log_2$ -fold-changes of DEGs in the three comparisons (Euclidean distance as a dissimilarity measure and complete linkage). **c** GOBPs represented by the DEGs in G1-6, as well as upregulated (MDSCs up) and downregulated (MDSCs down) genes in MDSCs, compared to MEFs. Color gradient represents the significance,  $-\log_{10}(P\text{-value})$ , of the corresponding GOBPs being enriched by the DEGs in G1-6, MDSC up and MDSC down, where  $P$ -value is the enrichment  $P$ -value computed by DAVID. Font colors of GOBP labels represent groups of GOBPs: red, cell cycle-related processes represented by DEGs in iMSCs, but not by DEGs in sort-iMSCs; green, differentiation-related processes represented by DEGs in sort-iMSCs, but not by DEGs in iMSCs; and blue, cell cycle and differentiation-related processes represented by DEGs in both iMSCs and sort-iMSCs. **d** Signaling network model associated with proliferation and differentiation of iMSCs and sort-iMSC cells. Node and node border colors represent  $\log_2$ -fold changes of their corresponding genes in iMSCs vs. MEFs and sort-iMSCs vs. MEFs, respectively. Color gradient represents  $\log_2$ -fold-changes of DEGs in the comparisons indicated in the legend (box at bottom right). The edges represent direct activation (arrows), repression (inhibition symbols), indirect activation through intermediate molecules not shown (dotted line) and PPIs (gray solid line), respectively. The types of interactions were obtained from the KEGG pathway database. Plasma membrane is denoted as the thick blue lines

capacity both in vitro and in vivo. Also, the iMSCs represented higher proliferative capacity even than MDSCs. The iMSCs indeed do not lose proliferative capacity until the passage was up to 50 (Fig. 3a). The ectopic expression of four TFs upregulates endogenous myogenic regulatory genes.

To clarify the role of the four TFs in the establishment of iMSCs, we transduced MEFs with three transcriptional factors omitting one of the four factors. Six1 and Esrrb play a pivotal role in converting MEFs to myogenic stem cells. Without six1 or esrrb, the increase of pax7, myf5, and myoD was minimal and establishment of iMSCs was hard to achieve (Fig. 1a–d). It is studied that Pax3 and Pax7 are key factors that confer early myogenic capacity leading to myotome during mouse muscle development<sup>23</sup>. However, overexpression of Pax3, without Six1 or Esrrb, is not sufficient to upregulate Pax7 and convert MEFs to myogenic lineage cells in 4F-Six1 and 4F-Esrrb.

The expression of myogenic factors was slightly increased in 4F-Eya1. Eya1 has dual role as a protein tyrosine phosphatase and transcriptional cofactor<sup>40</sup>. Eya1 acts as phosphatase and turns on Six1 from repression to activation<sup>19,40</sup>. Without Eya1, Six1 cannot function as much as interacting with Eya1. The interaction of Six1 and Eya1 directly activates Pax3 in limb muscle development, suggesting that the expression of Six1 and Eya1 increase the expression of Pax3<sup>18,19</sup>.

Interestingly, 4F-Pax3 cells represented higher expression in Pax7 than 4F cells, while similar expression was observed for MyoD and lower for Myf5 (Fig. 1a–d). The increase of Pax7 in 4F-Pax3 may be caused by the interaction of Six1 and Eya1. However, without Pax3, the cells could not differentiate into myotubes, even when myogenic expressions were increased (Fig. 1). Pax7, a paralogue of Pax3, can substitute for Pax3 because it shares most myogenesis functions with Pax3, including the

regulation of MyoD<sup>41</sup>. Although the increased expression of Pax7 in 4F-Pax3 can replace ectopic Pax3 in terms of regulating myogenic factors including Myf5 and MyoD under the proliferation condition, it is not enough to trigger myogenic differentiation.

The increased expression of Pax7 may explain the insufficiency of muscle differentiation of 4F-Pax3 in the differentiation environment (Fig. 1e, f). For muscle differentiation, Pax7 expression must be downregulated<sup>42</sup>. Activated satellite cells undergo sequential processes of proliferation, downregulation of Pax7, and then differentiation into muscle fibers. Pax7 activates proliferation of satellite cells but delays the expression of myogenin leading to delays of differentiation<sup>42</sup>. Particularly, during Myf5 expression, which is the determinant of myogenesis, 4F cells shows higher expression level than 4F-Pax3, as Myf5 is the direct target of Pax3<sup>24</sup>. To differentiate into muscle fibers, the cells needs proper cell number and confluency to connect adjacent cells. The myogenic differentiation media contains minimum nutrients and the only the cells with myogenic differentiation potential can be survived in that such a harsh niche. The cells that do not survive in the myogenic differentiation medium, like 4F-Pax3, are cells that lack the ability to differentiate into muscle functionally.

Thus, all the four factors are necessary to induce myogenic stem cells, which have myogenic differentiation capacity in DM not only express myogenic factors. Without one of the four TFs, MEFs could not convert into myogenic stem cells which express myogenic regulatory factors and differentiate into myogenic fibers under differentiation condition. Various combinations of candidate TFs can be used to induce direct conversion into myogenic stem cells. Naoki et al.<sup>14</sup> suggested that the combination of six transcription factors induce direct reprogramming of skeletal muscle progenitors from

embryonic fibroblasts. However, the combination of TFs for direct conversion should be accurate because it is very delicate. Induction with or without one specific TF can trigger reprogramming of totally different type of cells. For example, without Pax3, in this study, the cells represented proliferative capacity with weak myogenic differentiation. Thus, accurate combination of TFs is important to establish myogenic stem cells with both proliferation and myogenic differentiation capacities.

The established iMSCs with the four factors represented high expression levels of early myogenic regulatory factors including Pax7, Myf5, and MyoD (Fig. 2b). Interestingly, iMSCs represent even 1.38 times higher Pax7 expression than MDSCs and 7,634 times higher than MEF ( $P < 0.001$ ). Compared with the fact that <30% of MDSC is Pax7-positive, about 55.83% of iMSCs are positive for Pax7, which is quite high expression<sup>43</sup>. Pax7 promotes self-renewal of satellite cells and maintains the satellite cell pool<sup>44</sup>. It is revealed in Pax7 mutant mice that Pax7 have anti-apoptotic function<sup>41</sup>. Also, repressed expressions of late differentiation markers including myogenin and MHC enhanced the proliferation of iMSCs under proliferation condition. Thus, the increased expressions of early myogenic factors are related to the high proliferation capacity of iMSCs. Additionally, iMSCs are expandable up to passage 90 (data not shown). This is a major difference between iMSCs and MDSCs. The low expressions of myogenin and MHC maintain the elevated expression levels of Pax7 and Myf5 and proliferation capacity in GM condition.

In the FACS analysis, the iMSCs showed distinct expression patterns with MDSCs except for  $\alpha 7$ -integrin. The iMSCs were strongly positive for  $\alpha 7$ -integrin ( $97.23 \pm 0.43$  %) which is compatible to satellite cells.  $\alpha 7$ -integrin is one of the most well-known cell surface markers of satellite cells and myoblasts<sup>45</sup>.  $\alpha 7$ -integrin can be used to distinguish myoblasts from fibroblasts<sup>46</sup>. The iMSCs were intermediately positive for CD34 ( $14.20 \pm 2.92$ %) and Syndecan 4 ( $7.57 \pm 3.56$ %) (Fig. S1). Satellite cells are positive for  $\alpha 7$ -integrin, CD34, CXCR motif receptor R-4 (CXCR) and CD106 and negative for CD45, CD31, CD11b, and Sca-1<sup>47–49</sup>. The iMSCs can be distinguished from satellite cells based on their reactivity against CXCR, CD34, CD106, and Sca-1. Although iMSCs showed myogenic differentiation capacity, the cells differed from satellite cells in terms of cell surface markers.

Interestingly, the FACS analysis of iMSCs revealed two distinct peaks for CD106, Sca-1, and CD73, indicating that the cells were composed of two populations (Fig. S1). All the three markers are mesenchymal stem cell markers<sup>50,51</sup>, suggesting that the cells contained both mesenchymal stem cell stage- and non-mesenchymal stem cell stage-populations. We revealed that sorting with  $\alpha 7$ -integrin and the absence of CD106 increased both the

myogenic differentiation capacity and myogenic regulatory factor gene expressions (Fig. 4). It has been revealed that FACS sorting with specific cell surface markers enhances the myogenic differentiation capacity<sup>52</sup>.

We compared the myogenic capacity of iMSCs and sort-iMSCs under both proliferation and differentiation condition. The iMSCs exhibited a spindle-shaped cytoplasm and round nucleus, which were similar to mesenchymal stem cells in proliferation media. The sort-iMSCs showed robust myogenic differentiation capacity when incubated in myogenic differentiation media for 3 days (Fig. 4b). In the myogenic differentiation media, the sort-iMSCs represented higher expressions of myogenic regulatory factors than iMSCs, excepting for Myf5 (Fig. 4d). Particularly, myogenin expression in sort-iMSCs increased significantly, comparable to MDSCs. Down-regulation of Myf5 in sort-iMSCs may lead to upregulation of late myogenic differentiation factors including MyoD, myogenin, and MHC in DM. Thus, the sort-iMSCs may be more reactive to the myogenic differentiation niche than iMSCs, considering that they share similar myogenic expression pattern in proliferation media (Pax7, Myf5, and MyoD in Fig. 4c), but completely different expression in differentiation media (Fig. 4d). It suggests that the sort-iMSCs have greater myogenic differentiation capacity than iMSCs. Thus, we obtain enough number of iMSCs in myogenic growth media and sort the cells using the cell surface markers when we apply in vivo test or check myogenic differentiation capacity.

The high myogenic differentiation ability of sort-iMSCs may be related to its increased expression of ERR- $\gamma$  and ERR- $\alpha$  in DM (Fig. 4g, h). In the previous study, muscle specific ERR- $\gamma$   $-/-$  represented immature myotube formation<sup>39</sup>. It is revealed that increase of mitochondrial biogenesis is needed to induce myogenic differentiation<sup>53</sup> and ERR- $\gamma$  increased mitochondrial activity and oxidative capacity in skeletal muscle<sup>54</sup>. And ERR- $\alpha$  regulates myogenic differentiation cooperating with PGC-1 $\alpha$ , which is one of master regulator of mitochondrial biogenesis in skeletal muscle<sup>55–57</sup>.

We evaluated the myogenic differentiation properties of sort-iMSCs in vivo. The sort-iMSCs differentiated into myotubes and expressed dystrophin when transplanted into mdx mice (Fig. 5a–c). It suggests that the sort-iMSCs can be successfully engrafted, survive, and differentiate into myotubes in vivo. Although satellite cells injected-mdx mice showed a few dystrophin-positive fibers, the number was much smaller than sort-iMSCs injected one. Sort-iMSCs injection restored more dystrophin (+) fibers than MEFs or MDSCs in mdx mice indicating that the sort-iMSCs are more potent than MEFs and even MDSCs when transplanted in vivo (Fig. 5c). In most studies, researchers destroy the endogenous satellite cells of the recipient mdx mice through the 18G irradiation, and then

transplant satellite cells or myoblasts into the muscle of recipient<sup>58,59</sup>. This can avoid competition between endogenous satellite cells and exogenous muscle stem cells, and maximize the engraftment efficiency of transplanted cells. However, in the case of depleting satellite cells by irradiation, it cannot be said that the experiments in which the intramuscular environment is reproduced since all the endogenous satellite cells have been destroyed. Therefore, in this study, we evaluated the ability of sort-iMSCs to survive competition with existing endogenous satellite cells without irradiation. Thus, the positive control, MDSCs, represent lower levels of dystrophin than in previous studies which deplete the endogenous satellite cells in mdx<sup>47,60,61</sup>. Consider that we do not irradiate the mdx mice, sort-iMSCs can be survived in the environment which have been already occupied by endogenous satellite cells.

We established induced myogenic stem cells through ectopic expression of *Six1*, *Eya1*, *Esrrb* and *Pax3*. In contrast, Ito et al. established the induced skeletal muscle progenitor cells (iSkMs) through ectopic expression of a transcriptional factors including *Pax3*<sup>14</sup>. There is a clear difference between induced muscle progenitor cells and induced muscle stem cells. The cell fates of muscle stem cell are proliferation and/or muscle differentiation, whereas the cell fate of muscle progenitor is muscle differentiation. In this study, we confirmed robust proliferation or iMSCs different from previous iSkMs, and it was confirmed that the proliferative capacity was maintained even after the passage, and that it was a muscle stem cell in which differentiated into muscle cells.

The proliferative capacity of iMSCs is superior to MDSCs. The proliferation capacity of MDSCs is low and the proliferation rate is getting slower as the passage number is increased, since they are primary cells. In the proliferation assay, iMSCs represent higher proliferation capacity than MDSCs both *in vitro* and *in vivo* (Figs. 3 and 5c). To reveal the involved mechanisms, we checked FGF signaling which is well-known for its function in cell proliferation<sup>62</sup>. Interestingly, the expression of *FGFR1* was much higher than in both MEFs and MDSCs (\*\* $P < 0.001$ ) (Fig. 3f). In myoblast, proliferation is decreased by the repression of *FGFR1* through *KLF10*<sup>63</sup>. *FGFR1* promotes the proliferation of stem cells through cyclin-dependent kinase inhibitors<sup>64,65</sup>. It is consistent to our cell cycle data that iMSCs showed high percentage of G0/G1 (Fig. 3e). In this study, the expression of *GRB2* was increased in iMSCs (Fig. 3g–i). Further studies are needed to reveal the relationship between proliferation, *GRB2*, and FGF signaling.

Because of the use of monocistronic vectors, ectopic expression of many genes can create various cell populations. To exclude this possibility, single-cell sorting was performed in this experiment. Each cell line was derived

from each cell. Although the iMSC is originated from a single-cell, the iMSC has a heterogeneous characteristic as it continues to pass, suggesting that the iMSC undergoes asymmetric division like satellite cell. This was confirmed by cell morphology and FACS analysis (Fig. S1).

In this study, we established stably expandable induced myogenic stem cells with four defined factors. Especially, the iMSCs have much higher proliferation capacity than MDSCs. The established cells have both myogenic differentiation capacity and robust proliferation properties. The iMSCs can differentiate and form myotubes *in vivo* without tumor formation. The stably expandable iMSCs established in this study provide new source for drug screening and muscle regenerative therapy.

#### Acknowledgements

This research was supported by the Bio-industry technology Development Program, Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries (Grant No. 312062-5), and the National Research Foundation of Korea (NRF-2017R1E1A1A01072781) and the Institute for Basic Science (IBS-R013-G1-2016) funded by Korean ministry of science, ICT, and future planning.

#### Author details

<sup>1</sup>Department of Veterinary Pathology, College of Veterinary Medicine, Kyungpook National University, Daegu 41566, Republic of Korea. <sup>2</sup>School of Interdisciplinary Bioscience and Bioengineering, Pohang University of Science and Technology, Pohang 37673, Republic of Korea. <sup>3</sup>Stem Cell Therapeutic Research Institute, Kyungpook National University, Daegu 41566, Republic of Korea. <sup>4</sup>Department of Molecular Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA. <sup>5</sup>Center for Skeletal Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA. <sup>6</sup>Department of Pathology and Immunology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA. <sup>7</sup>Cell Engineering for Origin Research Center 45-13, Ujeongguk-ro, Jongno-gu, Seoul 03150, Republic of Korea. <sup>8</sup>Center for Plant Aging Research, Institute for Basic Science, Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu 42988, Republic of Korea. <sup>9</sup>Department of New Biology, DGIST, Daegu 42988, Republic of Korea

#### Conflict of interest

The authors declare that they have no conflict of interest.

#### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Supplementary Information** accompanies this paper at (<https://doi.org/10.1038/s41419-018-1114-8>).

Received: 15 May 2018 Revised: 4 September 2018 Accepted: 24 September 2018

Published online: 25 October 2018

#### References

- Asakura, A., Komaki, M. & Rudnicki, M. A. Muscle satellite cells are multipotential stem cells that exhibit myogenic, osteogenic, and adipogenic differentiation. *Differentiation* **68**, 245–253 (2001).
- Brack, A. S. et al. Increased Wnt signaling during aging alters muscle stem cell fate and increases fibrosis. *Science* **317**, 807–810 (2007).
- Vierbuchen, T. et al. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035–1041 (2010).
- Marro, S. et al. Direct lineage conversion of terminally differentiated hepatocytes to functional neurons. *Cell Stem Cell* **9**, 374–382 (2011).



5. Sekiya, S. & Suzuki, A. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. *Nature* **475**, 390–393 (2011).
6. Huang, P. et al. Direct reprogramming of human fibroblasts to functional and expandable hepatocytes. *Cell Stem Cell* **14**, 370–384 (2014).
7. Efe, J. A. et al. Conversion of mouse fibroblasts into cardiomyocytes using a direct reprogramming strategy. *Nat. Cell Biol.* **13**, 215–261 (2011).
8. Ieda, M. et al. Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* **142**, 375–386 (2010).
9. Song, K. et al. Heart repair by reprogramming non-myocytes with cardiac transcription factors. *Nature* **485**, 599–604 (2012).
10. Szabo, E. et al. Direct conversion of human fibroblasts to multilineage blood progenitors. *Nature* **468**, 521–526 (2010).
11. Davis, R. L., Weintraub, H. & Lassar, A. B. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987–1000 (1987).
12. Choi, J. MyoD converts primary dermal fibroblasts, chondroblasts, smooth muscle, and retinal pigmented epithelial cells into striated mononucleated myoblasts and multinucleated myotubes. *Proc. Natl Acad. Sci. USA* **87**, 7988–7992 (1990).
13. Zhongmin, L., Huimin, F., Yang, L. & Song, G. Z. Experimental studies on the differentiation of fibroblasts into myoblasts induced by MyoD genes in vitro. *Int. J. Biomed. Sci.* **4**, 14–19 (2008).
14. Naoki, I., Isao, K., Noriaki, S., Hiroto, T. & Takeda, S. Direct reprogramming of fibroblasts into skeletal muscle progenitor cells by transcription factors enriched in undifferentiated subpopulation of satellite cell. *Sci. Rep.* **7**, 8097 (2017).
15. Laclef, C. Altered myogenesis in Six1-deficient mice. *Development* **130**, 2239–2252 (2003).
16. Grifone, R. et al. Six1 and Six4 homeoproteins are required for Pax3 and Mrf expression during myogenesis in the mouse embryo. *Development* **132**, 2235–2249 (2005).
17. Nord, H., Nygard Skalman, L. & Von Hofsten, J. Six1 regulates proliferation of Pax7-positive muscle progenitors in zebrafish. *J. Cell Sci.* **126**, 1868–1880 (2013).
18. Grifone, R. et al. Eya1 and Eya2 proteins are required for hypaxial somitic myogenesis in the mouse embryo. *Dev. Biol.* **302**, 602–616 (2007).
19. Li, X. et al. Eya protein phosphatase activity regulates Six1-Dach-Eya transcriptional effects in mammalian organogenesis. *Nature* **426**, 247–254 (2003).
20. Latos, P. A. et al. Fgf and Esrrb integrate epigenetic and transcriptional networks that regulate self-renewal of trophoblast stem cells. *Nat. Commun.* **6**, 7776 (2015).
21. Zhang, X., Zhang, J., Wang, T., Esteban, M. A. & Pei, D. Esrrb activates Oct4 transcription and sustains self-renewal and pluripotency in embryonic stem cells. *J. Biol. Chem.* **283**, 35825–35833 (2008).
22. Tremblay, P. et al. A crucial role for Pax3 in the development of the hypaxial musculature and the long-range migration of muscle precursors. *Dev. Biol.* **203**, 49–61 (1998).
23. Relaix, F., Rocancourt, D., Mansouri, A. & Buckingham, M. A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* **435**, 948–953 (2005).
24. Bajard, L. et al. E. A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. *Genes Dev.* **20**, 2450–2464 (2006).
25. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).
26. Liadaki, K., Luth, E. & Kunkel, L. Co-detection of GFP and dystrophin in skeletal muscle tissue sections. *Biotechniques* **42**, 699–700 (2007).
27. Shea, K. L. et al. Sprouty1 regulates reversible quiescence of a self-renewing adult muscle stem cell pool during regeneration. *Cell Stem Cell* **6**, 117–129 (2010).
28. Chakkalakal, J. V., Jones, K. M., Basson, M. A. & Brack, A. S. The aged niche disrupts muscle stem cell quiescence. *Nature* **490**, 355–360 (2012).
29. Park, D. et al. Endogenous bone marrow MSCs are dynamic, fate-restricted participants in bone maintenance and regeneration. *Cell Stem Cell* **10**, 259–272 (2012).
30. Bolstad, B. M., Irazary, R. A., Astrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185–193 (2003).
31. Lee, H. J. et al. Direct transfer of alpha-synuclein from neuron to astroglia causes inflammatory responses in synucleinopathies. *J. Biol. Chem.* **285**, 9262–9272 (2010).
32. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2008).
33. Stark, C. et al. BioGRID: a general repository for interaction datasets. *Nucl. Acids Res.* **34**, D535–D539 (2006).
34. Salwinski, L. et al. The database of interacting proteins: 2004 update. *Nucl. Acids Res.* **32**, D449–D451 (2004).
35. Patil, A., Nakai, K. & Nakamura, H. HitPredict: a database of quality assessed protein-protein interactions in nine species. *Nucl. Acids Res.* **39**, D744–D749 (2011).
36. Orchard, S. et al. The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucl. Acids Res.* **42**, D358–D363 (2014).
37. Licata, L. et al. MINT, the molecular interaction database: 2012 update. *Nucl. Acids Res.* **40**, D8575–61 (2012).
38. Szklarczyk, D. et al. STRINGv10: protein-protein interaction networks, integrated over the tree of life. *Nucl. Acids Res.* **43**, D447–D452 (2015).
39. Murray, J., Auwerx, J. & Huss, J. M. Impaired myogenesis in estrogen-related receptor  $\gamma$  (ERR $\gamma$ )-deficient skeletal myocytes due to oxidative stress. *FASEB J.* **27**, 135–150 (2013).
40. Jemc, J. & Rebay, I. The eyes absent family of phosphotyrosine phosphatases: properties and roles in developmental regulation of transcription. *Annu. Rev. Biochem.* **76**, 513–538 (2007).
41. Relaix, F. et al. Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. *J. Cell Biol.* **172**, 91–102 (2006).
42. Zammit, P. S. et al. Pax7 and myogenic progression in skeletal muscle satellite cells. *J. Cell Sci.* **119**, 1824–1832 (2006).
43. Hideki, O. et al. Differential myocardial infarct repair with muscle stem cells compared to myoblasts. *Mol. Ther.* **12**, 1130–1141 (2005).
44. Kuang, S., Kuroda, K., Le Grand, F. & Rudnicki, M. A. Asymmetric self-renewal and commitment of satellite stem cells in muscle. *Cell* **129**, 999–1010 (2007).
45. George-Weinstein, M., Foster, R. F., Gerhart, J. V. & Kaufman, S. J. In vitro and in vivo expression of alpha 7 integrin and desmin define the primary and secondary myogenic lineages. *Dev. Biol.* **15**, 209–229 (1993).
46. Blanco-Bose, W. E., Yao, C. C., Kramer, R. H. & Blau, H. M. Purification of mouse primary myoblasts based on alpha 7 integrin expression. *Exp. Cell Res.* **265**, 212–220 (2001).
47. Pasut, A., Olynyk, P. & Rudnicki, M. A. Isolation of muscle stem cells by fluorescence activated cell sorting cytometry. *Methods Mol. Biol.* **798**, 53–64 (2012).
48. 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).
49. Maesner, C. C., Almada, A. E. & Wagers, A. J. Established cell surface markers efficiently isolate highly overlapping populations of skeletal muscle satellite cells by fluorescence-activated cell sorting. *Skelet. Muscle* **6**, 35 (2016).
50. Houlihan, D. D. et al. Isolation of mouse mesenchymal stem cells on the basis of expression of Sca-1 and PDGFR-alpha. *Nat. Protoc.* **7**, 2103–2111 (2012).
51. Dominici, M. et al. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy* **8**, 315–317 (2006).
52. Darabi, R. et al. Functional skeletal muscle regeneration from differentiating embryonic stem cells. *Nat. Med.* **14**, 134–143 (2008).
53. Duguez, S., F Asson, L., Denis, C. & Freyssen, D. Mitochondrial biogenesis during skeletal muscle regeneration. *AME J. Physiol.-Endocrinol. Meta.* **282**, E802–E809 (2002).
54. Rangwala, S. M. et al. Estrogen-related receptor  $\gamma$  is a key regulator of muscle mitochondrial activity and oxidative capacity. *J. Biol. Chem.* **285**, 22619–22629 (2010).
55. Murray, J. & Huss, J. M. Estrogen-related receptor  $\alpha$  regulates skeletal myocyte differentiation via modulation of the ERK MAP kinase pathway. *Am. J. Physiol. Cell Physiol.* **301**, C630–C645 (2011).
56. Wu, Z. et al. Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell* **98**, 115–124 (1999).
57. Scarpulla, R. C. Metabolic control of mitochondrial biogenesis through the PGC-1 family regulatory network. *Biochim. Biophys. Acta* **1813**, 1269–1278 (2011).
58. Heslop, L., Morgan, J. E. & Partridge, T. A. Evidence for a myogenic stem cell that is exhausted in dystrophic muscle. *J. Cell Sci.* **113**, 2299–2308 (2000).

59. Boldrin, L., Neal, A., Zammit, P. S., Muntoni, F. & Morgan, J. E. Donor satellite cell engraftment is significantly augmented when the host niche is preserved and endogenous satellite cells are incapacitated. *Stem Cells* **30**, 1971–1984 (2012).
60. Gussoni, E. et al. Dystrophin expression in the mdx mouse restored by stem cell transplantation. *Nature* **401**, 390–394 (1999).
61. Montarras, D. et al. Direct isolation of satellite cells for skeletal muscle regeneration. *Science* **309**, 2064–2067 (2005).
62. Xu, R. -H. et al. Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. *Nature Methods* **2**, 185–190 (2005).
63. Parakati, R. & Dimario, J. X. Repression of myoblast proliferation and fibroblast growth factor receptor 1 promoter activity by KLF10 protein. *J. Biol. Chem.* **288**, 13876–13884 (2013).
64. Dombrowski, C. et al. FGFR1 signaling stimulates proliferation of human mesenchymal stem cells by inhibiting the cyclin-dependent kinase inhibitors p21(Waf1) and p27(Kip1). *Stem Cells* **31**, 2724–2736 (2013).
65. Ohkubo, Y., Uchida, A. O., Shin, D., Partanen, J. & Vaccarino, F. M. Fibroblast growth factor receptor 1 is required for the proliferation of hippocampal progenitor cells and for hippocampal growth in mouse. *J. Neurosci.* **24**, 6057–6069 (2004).