

Identification of biomarkers indicative of functional skeletal stem cells

Z. Qiryaqoz^{1*}

S. Timilsina^{1*}

*D. Czarnowski*¹

*P.H. Krebsbach*²

*L.G. Villa-Diaz*¹

¹Department of Biological Sciences, Oakland University, Rochester, MI 48309

²School of Dentistry, University of California, Los Angeles, Los Angeles, CA 90095

* Authors contributed equally

Correspondence:

Luis Villa-Diaz, Department of Biological Sciences, Oakland University, Rochester, MI

Email: luisvilladiaz@oakland.edu

Number of words in abstract

Number of words in full manuscript

Structured Abstract

Objectives: Skeletal stem cells (SkeSC) are characterized by expression of cell surface biomarkers, and potential towards bone, cartilage and fat differentiation. However, these biomarkers are not cell-type specific and are not indicative of the differentiation status of these cells, and therefore are poorly reliable. Our objective was to identify alternative cell surface biomarkers and transcription factors commonly present between SkeSCs isolated from the bone marrow (BM) and those derived from pluripotent stem cells (PSC)

Materials and Methods: Human PSCs were induced into SkeSCs. FACS and qRT-PCR was used to determine differences in expression of cell surface biomarkers and transcription factors between SkeSCs derived from PSCs and isolated from BM, in differentiating cells, in cells from early and late passage, and in fibroblasts.

Results: Our results demonstrated a significant reduction in proliferation and differentiation capacity of SkeSCs to into adipocytes and osteocytes after 3 passages. Protein and mRNA analysis indicated that commonly used biomarkers remain highly expressed in non functional cells. However, integrin $\alpha 6$ (CD49f) and the transcription factors GATA6, PRDM16, SIM2 and SOX11 were significantly upregulated in SkeSCs in relation to fibroblasts, while decreased significantly their expression in skeletal stem cells in early stages of adipogenic and osteogenic differentiation and in late passages cells, which have lost proliferation and differentiation capabilities.

Conclusions: Our results suggest that CD49f and the transcription factors GATA6, PRMD16, SIM2 and SOX11 identify functional SkeSCs.

KEYWORDS

Skeletal stem cells, transcription factors, cell surface biomarkers, differentiation, stem cells

Introduction

Mesenchymal stem cells, recently defined as skeletal stem cells (SkeSC), are resident cells of the bone marrow (BM) and are capable of self-renewal and give rise to adipocytes, bone, cartilage, and hematopoiesis-supportive stroma (1, 2). The following characteristics have been implemented to define SkeSCs *in vitro*: adherence to plastic, positive expression of cell surface markers CD73, CD90 and CD105, negative for hematopoietic and endothelial markers, and capacity for *in vitro* differentiation into bone, cartilage, and fat (3). Using these minimal criteria, the isolation of potential SkeSCs has been reported from adipose tissue (4), dental pulp (5), endometrial stroma (6), palatine tonsil (7) and umbilical cord blood (8), suggesting their widespread presence in the body. However, it has been argued that compared to SkeSCs derived from BM, other stem cell populations are developmentally different and do not contribute to skeletal development, or postnatal physiology (2). **An alternative source for SkeSCs comes from the differentiation of human pluripotent stem cells (PSC), which include embryonic stem cells (ESC) (9), and induced pluripotent stem cells (iPSCs) (10), which are obtained after genetic reprogramming of somatic cells into a cell that have capacity to give rise to all cells in the body (11), as ESCs .**

Adding to the debate concerning the above definition of SkeSCs, are the observations that the defining biomarkers are also expressed in fibroblasts (12), and remain highly expressed in SkeSCs that no longer have capacity for differentiation (Fig 1 and Fig 3). This suggests that these biomarkers are not cell-type specific, nor are they truly indicative of the undifferentiated state of SkeSCs. Other biomarkers are being proposed to identify SkeSCs, including CD146, CD271 and CD49f. CD146, also known

as melanoma-associated cell adhesion molecule (MCAM), is expressed in self-renewing osteoprogenitor subendothelial cells of the BM stroma (13) and in perivascular regions (14). Upon transplantation to heterotopic sites, CD146⁺ cells isolated from the BM are able to create a hematopoietic microenvironment (13). CD271, also known as nerve growth factor receptor, selectively identifies SkeSCs from the BM (15). CD49f, also known as integrin $\alpha 6$, is highly expressed in fetal SkeSCs isolated from BM (16), while its expression is lower in adult cells (17). During *in vitro* expansion of SkeSCs, CD49f expression gradually decreases after each passage, and its knockdown reduces the differentiation potential (17).

A refined phenotypic analysis of SkeSCs should include expression of transcription factors, **since they are part of the regulatory network that controls the transcription of thousands of genes that give identity to cells**. Towards this end, some studies have reported the transcriptome of SkeSCs isolated from the BM and other sources (18, 19). However, due to the heterogeneity of the sample origins, a consensus of transcription factors present in SkeSCs has not been reached. Here, we aimed to define the phenotypic characteristic of SkeSCs combining the expression of cell surface markers and transcription factors. To better define this phenotype we compared SkeSCs isolated from the BM as a gold standard to SkeSCs derived from human ESCs and iPSCs, (9, 10). Fibroblasts were included in our analysis to identify non-specific SkeSC's biomarkers. Cells induced into adipocytes and osteoblasts were included to identify biomarkers specific for undifferentiated SkSCs.

Materials and Methods

Culture and differentiation of human pluripotent stem cells

All experiments were repeated at least in triplicates and with multiple human ESC and human iPSC lines. NIH approved human ESCs lines H1, H7 and H9 (WiCell Research Institute) and CHB8 and CHB10 (Children's Hospital Corporation, Boston, MA) and three iPSCs derived in our laboratory (hGF2-iPSCs, hGF4-iPSCs, hFF iPSCs) (20) were cultured on Matrigel® hESC-qualified Matrix (Corning®) with StemFlex medium (Gibco® Life Technologies), in incubators with high humidity and 5% CO₂ at 37°C. Undifferentiated colonies were passaged using TrypLE Express (Gibco® Life Technologies).

Differentiation of PSCs into SkeSCs was induced as described previously (9). In subsequent culture, SkeSCs were seeded at a density of 7×10^3 cells/cm². To induced osteogenic differentiation, 3×10^3 SkeSCs were seeded per well of a six-well plate and incubated in StemMACS™ OsteoDiff media (Miltenyi Biotec) for 10 days. Cells were fixed with 10% formalin for 20 minutes at RT and stained for Alkaline phosphatase (SIGMAFAST BCIP/NBT, SIGMA-ALDRICH) to verify osteoblast differentiation. For adipogenic differentiation, 5×10^3 skeletal stem cells were seeded per well of a six-well plate and incubated in Stem MACS™ AdipoDiff Media (Miltenyi Biotec) for 21 days. To verify adipocyte differentiation, cells were staining with Oil red O.

Human bone marrow skeletal stem cell culture

Human BM was collected from patients undergoing iliac bone graft procedures with University of Michigan IRB approval and informed patient consent. SkeSCs were isolated and cultured from the BM extracts as previously described (21). Cells at passage 4 were used as controls for the experiment described here.

Cell doubling time calculations

The doubling time was calculated with the following formula: $((\text{Duration of cell culture (hours)}) * (\log(2))) / (\log(\text{Final concentration}) - \log(\text{Initial Concentration}))$, using Roth V. 2006 Doubling Time Computing (<http://www.doubling-time.com/compute.php>).

Flow cytometry analysis

Cells were harvested and prepared for flow cytometry analysis as described previously (10). The following PE-conjugated antibodies were used: CD29, CD49f, CD73, CD104, CD271, and CD45. At least 10,000 events were acquired for each sample using a FACSCalibur instrument (Becton Dickinson) and cell flow cytometry data were analyzed using CELLQUEST software (Becton Dickinson). The value (percentage) of positive cells for each antibody was calculated by subtracting the isotype control value from the detected value of each antibody.

RNA isolation, preparation, quantitative real-time PCR and reverse transcription PCR

Total RNA extraction, purification, and reversed transcription into cDNA was done as described previously (10). Two microliters of diluted reverse transcribed cDNA were amplified in a polymerase chain reaction (PCR) assay (30 μ l) using TaqMan Gene Expression Master Mix (Applied Biosystems), target primers, and probes (unlabeled PCR primers and a carboxyfluorescein dye-labeled TaqMan minor groove binder probe) (Applied Biosystems). Gene expression was determined by quantitative real-time PCR on an ABI Prism 7700 Sequence Detection System (Applied Biosystems). The relative RNA expression levels of target genes were analyzed by the comparative $\Delta\Delta$ CT method using the housekeeping gene *TBP*, as an internal control. Changes bigger than 2-fold in relative mRNA expression were considered significant. Subsequently, expression levels

of investigated genes were normalized to expression levels of control samples and reported as fold changes. For reverse-transcription PCR, total RNA was reverse transcribed using SuperScript™ One-Step RT-PCR with platinum® Taq (Invitrogen).. Primer sequences for *KI67* are forward: TTGTGCCTTCACTTCCACAT, and reverse: CTGGTAATGCACACTCCACCT; while for *TBP* forward: CTCCCACCCAAAGTCTGATGA and reverse: GCCATAAACCAAGCAGGACG. The cDNA synthesis and pre-denaturation were carried out at 95°C for 2 min. The PCR amplification was performed for 35 cycles at 95°C for 30 sec, 55°C for 30 sec, and 72°C for 30 sec. The final extension cycle was run at 72°C for 10 min. Finally, 14µl of PCR reaction product were loaded onto a 2.0% agarose gel. Band densitometry analysis was done using ImageLab 6.0.

Statistical analysis

Experiments were performed in triplicate and data are expressed as mean value $\pm$ S.E.M., and analyzed by an unpaired *t* test. Levels of statistical significance were set at $p < 0.05$.

Results

It was observed that SkeSCs derived from human PSCs (PSC-SkeSCs) maintained a spindle morphology during the eight passages, however, the capacity for adipocyte and osteoblast differentiation become significantly reduced from passage 3 to passage 6, and was almost non-existent by passage 8 (Fig 1). Calculations for cell doubling time indicated a significant decrease in proliferation rates between passage 3 compared to passages 6 and 8, which was validated by a 30% decrease in *KI67* RNA level (Fig. 2). This indicated that differentiation and proliferation was reduced in SkeSCs during *in vitro* propagation.

Flow cytometry analysis of PSC-SkeSCs at passages 3, 6 and 8 indicated that CD105 and CD29 biomarkers remained highly expressed in SkeSCs during all the passages, while CD146 and CD271 were consistently lowly expressed (Fig. 3). CD49f and CD73 were highly expressed at passage 3, decreased significantly ($p < 0.05$) at passage 6 and returned to high levels at passage 8. CD45 expression was low at passage 3, and increased with subsequent subculture. Integrin $\alpha 6$ (CD49f) heterodimerizes with integrin $\beta 1$ (CD29) and integrin $\beta 4$ (CD104), therefore, we investigated the expression of the latter biomarker and determined that it was expressed consistently at low percentages in SkeSCs. The expression of all these biomarkers was comparable to levels identified in foreskin and gingival fibroblasts with the exception of CD49f, which were low ($< 10\%$) in fibroblasts (data not shown). Difference in expression of integrin $\alpha 6$ (CD49f) between fibroblasts and SkeSCs was validated at RNA level, observing a 88.9 ± 37.9 (mean $\pm$ SEM) fold increase in SkeSCs in relation to fibroblasts. These data suggest that CD49f is a more indicative biomarker of SkeSCs compared to other tested biomarkers, as it is not

expressed in fibroblasts and its expression becomes reduced in aged cells that lose proliferation and differentiation capacity.

Previously, nine transcription factors expressed in SkeSCs isolated from BM (BM-SkeSCs) and from synovial tissue were identified (19). We investigated the mRNA levels of these transcription factors in fibroblasts, in undifferentiated ESCs, and iPSCs derived from the same pool of fibroblasts included in the analysis, in PSC-SkeSCs and in BM-SkeSCs (Fig. 4). From all nine transcription factors tested, *ETV1* was similarly expressed among fibroblasts, undifferentiated PSCs, and SkeSCs derived from them. However, *ETV1* mRNA levels were significantly lower in BM-SkeSCs. Other four transcription factors (*ETV5*, *FOXP1*, *HMGA* and *KLF12*) were equally expressed in fibroblasts in relation to SkeSCs. In contrast, the relative mRNA expression of *GATA6*, *PRDM16*, *SIM2* and *SOX11* was significantly upregulated in SkeSCs compared to fibroblasts.

We further investigated the mRNA level expression of these four transcription factors and integrin $\alpha 6$ in PSC-SkeSCs at early passages (passage 3), which have a increased capacity for osteogenic and adipogenic differentiation, and compared to late passage cells (passage 8) (Fig. 5a). In relation to early passage cells the relative mRNA expression of *GATA6*, *PRDM16* and *SIM2* decreased, while the relative mRNA levels of *SOX11* increased but not to significant levels in late passage cells. In contrast, it was observed that *ITGA6* mRNA levels had a four-fold change in late passage cells, which is reflective of the recovery of CD49f expression observed by flow cytometry at passage 8 (Fig. 3). Next, we investigated mRNA expression of *GATA6*, *PRDM16*, *SIM2*, *SOX11* and *ITGA6* in PSC-SkeSCs at passage 3 and in equivalent cells in early process of

differentiation into osteoblasts and adipocytes. Results indicated decreased expression of *ITGA6*, *GATA6*, *PRDM16* AND *SIM2* in cells undergoing adipogenic and osteogenic differentiation, while a non significant increase (<2 fold change) was observed for *SOX11* in cells in adipogenic differentiation.

Discussion

We identified CD49f and CD73 as cell surface biomarkers that more closely define SkeSCs, as their expression correlates with differentiation potential of these cells. While biomarkers CD29, CD105, CD146 and CD271 are present in SkeSCs, they are also expressed at similar levels in fibroblasts and in SkeSCs that have lost potential for differentiation. Our results validate a previous report (19) indicating the expression of transcription factors expressed in BM-SkeSCs. However, following our exclusion criteria, ETV1, ETV5, FOXP1, HMGA2 and KLF12 are not considered transcription factors that should define SkeSCs because are expressed at similar or lower levels than in fibroblasts. The common expression of GATA6, PRDM16, SIM2 and SOX11 among BM-SkeSCs and derived from several PSC lines identify them as representative transcription factors of this stem cell population. This is reinforced by the restricted expression of these transcription factors in SkeSCs with capacity for proliferation and differentiation.

Previously it was shown that knockdown of GATA6, SIM2 and SOX11 in SkeSCs suppresses their self-renewal, while the knockdown of PRDM16 reduces the adipogenic and osteogenic potential (19). Here, we demonstrate that cells expressing low

mRNA levels of these transcription factors have reduced proliferation levels and low capacity for differentiation into adipocytes and osteoblasts, indicating a functional expression in SkeSCs. The individual expression and function of these transcription factors in SkeSCs have been reported previously. For example, GATA has been found highly expressed in PSC-SkeSCs (22) and isolated from BM, adipose tissue, spleen, thymus and aorta (23). Interestingly, others found that unlike GATA6, which is expressed in all these populations of SkeSCs, genes involved in mesoderm segmentation and somite development were differently expressed in the same pool of SkeSCs populations (23). This suggests that GATA6 could be part of a core transcription factor signature of SkeSCs.

PRDM16 is expressed in neural stem/progenitor cells and hematopoietic stem cells, and promotes maintenance by modulating oxidative stress (24). We observed a decrease in mRNA levels of PRDM16 in cells that lost the capacity for proliferation and in early stage of differentiation, which could indirectly indicate a possible role in self-renewal as reported in hematopoietic stem cells. In contrast, PRDM16 has been implicated in the differentiation of adipocytes into brown fat (25), and it was found highly expressed in fetal SkeSCs that have a higher propensity to develop brown fat versus white fat (25).

SIM2 is expressed at early fetal life and has been involved in development of the forebrain, ribs, vertebrae, limb skeletal muscles and kidneys (26). SIM2^(-/-) mice develop rib, vertebral and craniofacial abnormalities and die soon after birth due to breathing failure (27). This implicates SIM2 in skeletal tissue development by SkeSCs, and supports our findings of its expression in BM-SkeSCs and PSC-SkeSCs..

Interestingly, SIM2 is associated with KLF4, NANOG, OCT4 (POU5F1), SOX2 in mouse ESCs (28), and results also show significantly higher expression of SIM2 in human PSCs compared to fibroblasts (Fig. 4), suggesting a possible role for SIM2 in human pluripotency.

SOX11 was found to be significantly expressed in SkeSCs compared to fibroblasts, and its expression remains without significant change after differentiation, or in late passage cells. Other reports indicate that SOX11 expression decreases with expansion of SkeSCs and its knockdown with siRNA reduces the proliferation and capacity for osteogenic differentiation (19, 29). Indeed, SOX11 has been identified as a key transcription factor in the osteogenesis. SOX11-deficient mice develop multiple craniofacial and skeletal malformations, such as cleft palate or cleft lips, duplication in vertebrae L4,5 and kinked tails (30). SOX11^(-/-) died around E8.5 with growth arrest at this stage and rudimentary somites (31).

Finally, we observed that integrin $\alpha 6$ (CD49f) expression decreases in cells during expansion and differentiation. Similar findings were reported in SkeSCs isolated from the fetal and adult BM (17). Knockdown of CD49f reduces differentiation potential of SkeSCs (17). Furthermore, CD49f plays an important role in the self-renewal of PSCs (21), skeletal muscle stem cells (32), glioblastoma (33) and breast cancer stem cells (34); and its expression has been identified in 35 population of stem cells (35). Our flow cytometry data indicates that although close to 100% of cells were CD29 and CD105 positive during all 8 passages, only about 70% of PSC-SkeSCs were CD49f positive at passage 3, and its expression decreased in subsequent passages and in cells that lost potential for proliferation and differentiation. This indicates the presence of two

subpopulations based on CD49f expression. We propose that the CD49f positive subpopulation contains SkeSCs, while the CD49f negative population may contain differentiated cells. Future experiments will be required to verify this and elucidate its implications.

Conclusions

Our results identify four transcription factors (GATA6, PRDM16, SIM2 and SOX11) and a cell surface marker (CD49f) as reliable biomarkers that better indicate the functional status of SkeSCs. This has significant implications in regenerative medicine, in particular regarding craniofacial related diseases in which SkeSCs are used for treatment. Currently there are close to 900 clinical trials using stem cells defined as mesenchymal stem cells. Most likely these stem cells are defined using the standard characterization, which we proved identify functional and non-functional cells. Thus, we proposed the inclusion of further bio-markers that better identify functional SkeSCs to obtain consistent outcomes.

Acknowledgements

LGVD acknowledge funding from the Center of Biomedical Research and the Honors College of Oakland University.

References

1. Bianco P. "Mesenchymal" stem cells. *Annu Rev Cell Dev Biol.* 2014;30:677-704.

2. Bianco P, Robey PG. Skeletal stem cells. *Development*. 2015;142(6):1023-7.
3. Dominici M, Le Blanc K, Mueller I, Slaper-Cortenbach I, Marini F, Krause D, et al. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy*. 2006;8(4):315-7.
4. Zuk PA, Zhu M, Ashjian P, De Ugarte DA, Huang JI, Mizuno H, et al. Human adipose tissue is a source of multipotent stem cells. *Mol Biol Cell*. 2002;13(12):4279-95.
5. Shi S, Gronthos S. Perivascular niche of postnatal mesenchymal stem cells in human bone marrow and dental pulp. *J Bone Miner Res*. 2003;18(4):696-704.
6. Schwab KE, Gargett CE. Co-expression of two perivascular cell markers isolates mesenchymal stem-like cells from human endometrium. *Hum Reprod*. 2007;22(11):2903-11.
7. Ryu KH, Cho KA, Park HS, Kim JY, Woo SY, Jo I, et al. Tonsil-derived mesenchymal stromal cells: evaluation of biologic, immunologic and genetic factors for successful banking. *Cytotherapy*. 2012;14(10):1193-202.
8. Romanov YA, Svintsitskaya VA, Smirnov VN. Searching for alternative sources of postnatal human mesenchymal stem cells: candidate MSC-like cells from umbilical cord. *Stem Cells*. 2003;21(1):105-10.
9. Brown SE, Tong W, Krebsbach PH. The derivation of mesenchymal stem cells from human embryonic stem cells. *Cells Tissues Organs*. 2009;189(1-4):256-60.
10. Villa-Diaz LG, Brown SE, Liu Y, Ross AM, Lahann J, Parent JM, et al. Derivation of mesenchymal stem cells from human induced pluripotent stem cells cultured on synthetic substrates. *Stem Cells*. 2012;30(6):1174-81.
11. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 2007;131(5):861-72.
12. Lorenz K, Sicker M, Schmelzer E, Rupf T, Salvetter J, Schulz-Siegmund M, et al. Multilineage differentiation potential of human dermal skin-derived fibroblasts. *Exp Dermatol*. 2008;17(11):925-32.
13. Sacchetti B, Funari A, Michienzi S, Di Cesare S, Piersanti S, Saggio I, et al. Self-renewing osteoprogenitors in bone marrow sinusoids can organize a hematopoietic microenvironment. *Cell*. 2007;131(2):324-36.
14. Tormin A, Li O, Brune JC, Walsh S, Schutz B, Ehinger M, et al. CD146 expression on primary nonhematopoietic bone marrow stem cells is correlated with in situ localization. *Blood*. 2011;117(19):5067-77.
15. Cattoretti G, Schiro R, Orazi A, Soligo D, Colombo MP. Bone marrow stroma in humans: anti-nerve growth factor receptor antibodies selectively stain reticular cells in vivo and in vitro. *Blood*. 1993;81(7):1726-38.
16. Lee RH, Seo MJ, Pulin AA, Gregory CA, Ylostalo J, Prockop DJ. The CD34-like protein PODXL and alpha6-integrin (CD49f) identify early progenitor MSCs with increased clonogenicity and migration to infarcted heart in mice. *Blood*. 2009;113(4):816-26.
17. Yang Z, Dong P, Fu X, Li Q, Ma S, Wu D, et al. CD49f Acts as an Inflammation Sensor to Regulate Differentiation, Adhesion, and Migration of Human Mesenchymal Stem Cells. *Stem Cells*. 2015;33(9):2798-810.

18. Wagner W, Wein F, Seckinger A, Frankhauser M, Wirkner U, Krause U, et al. Comparative characteristics of mesenchymal stem cells from human bone marrow, adipose tissue, and umbilical cord blood. *Exp Hematol*. 2005;33(11):1402-16.
19. Kubo H, Shimizu M, Taya Y, Kawamoto T, Michida M, Kaneko E, et al. Identification of mesenchymal stem cell (MSC)-transcription factors by microarray and knockdown analyses, and signature molecule-marked MSC in bone marrow by immunohistochemistry. *Genes Cells*. 2009;14(3):407-24.
20. Villa-Diaz LG, Kim JK, Lahann J, Krebsbach PH. Derivation and long-term culture of transgene-free human induced pluripotent stem cells on synthetic substrates. *Stem Cells Transl Med*. 2014;3(12):1410-7.
21. Krebsbach PH, Kuznetsov SA, Satomura K, Emmons RV, Rowe DW, Robey PG. Bone formation in vivo: comparison of osteogenesis by transplanted mouse and human marrow stromal fibroblasts. *Transplantation*. 1997;63(8):1059-69.
22. Zhang L, Wang H, Liu C, Wu Q, Su P, Wu D, et al. MSX2 Initiates and Accelerates Mesenchymal Stem/Stromal Cell Specification of hPSCs by Regulating TWIST1 and PRAME. *Stem Cell Reports*. 2018.
23. Sagi B, Maraghechi P, Urban VS, Hegyi B, Szigeti A, Fajka-Boja R, et al. Positional identity of murine mesenchymal stem cells resident in different organs is determined in the postsegmentation mesoderm. *Stem Cells Dev*. 2012;21(5):814-28.
24. Chuikov S, Levi BP, Smith ML, Morrison SJ. Prdm16 promotes stem cell maintenance in multiple tissues, partly by regulating oxidative stress. *Nat Cell Biol*. 2010;12(10):999-1006.
25. Morganstein DL, Wu P, Mane MR, Fisk NM, White R, Parker MG. Human fetal mesenchymal stem cells differentiate into brown and white adipocytes: a role for ERRalpha in human UCP1 expression. *Cell Res*. 2010;20(4):434-44.
26. Shaw L, Johnson PA, Kimber SJ. Gene expression profiling of the developing mouse kidney and embryo. *In Vitro Cell Dev Biol Anim*. 2010;46(2):155-65.
27. Goshu E, Jin H, Fasnacht R, Sepenski M, Michaud JL, Fan CM. Sim2 mutants have developmental defects not overlapping with those of Sim1 mutants. *Mol Cell Biol*. 2002;22(12):4147-57.
28. Letourneau A, Cobellis G, Fort A, Santoni F, Garieri M, Falconnet E, et al. HSA21 Single-Minded 2 (Sim2) Binding Sites Co-Localize with Super-Enhancers and Pioneer Transcription Factors in Pluripotent Mouse ES Cells. *PLoS One*. 2015;10(5):e0126475.
29. Larson BL, Ylostalo J, Lee RH, Gregory C, Prockop DJ. Sox11 is expressed in early progenitor human multipotent stromal cells and decreases with extensive expansion of the cells. *Tissue Eng Part A*. 2010;16(11):3385-94.
30. Sock E, Rettig SD, Enderich J, Bosl MR, Tamm ER, Wegner M. Gene targeting reveals a widespread role for the high-mobility-group transcription factor Sox11 in tissue remodeling. *Mol Cell Biol*. 2004;24(15):6635-44.
31. Bhattaram P, Penzo-Mendez A, Sock E, Colmenares C, Kaneko KJ, Vassilev A, et al. Organogenesis relies on SoxC transcription factors for the survival of neural and mesenchymal progenitors. *Nat Commun*. 2010;1:9.
32. Rayagiri SS, Ranaldi D, Raven A, Mohamad Azhar NIF, Lefebvre O, Zammit PS, et al. Basal lamina remodeling at the skeletal muscle stem cell niche mediates stem cell self-renewal. *Nat Commun*. 2018;9(1):1075.

33. Lathia JD, Gallagher J, Heddleston JM, Wang J, Eyler CE, Macsworlds J, et al. Integrin alpha 6 regulates glioblastoma stem cells. *Cell Stem Cell*. 2010;6(5):421-32.
34. Goel HL, Gritsko T, Pursell B, Chang C, Shultz LD, Greiner DL, et al. Regulated splicing of the alpha6 integrin cytoplasmic domain determines the fate of breast cancer stem cells. *Cell Rep*. 2014;7(3):747-61.
35. Krebsbach PH, Villa-Diaz LG. The Role of Integrin alpha6 (CD49f) in Stem Cells: More than a Conserved Biomarker. *Stem Cells Dev*. 2017;26(15):1090-9.

FIGURE LEGENDS

Fig. 1. Skeletal stem cells lost potential for differentiation over time. Human pluripotent stem cells were differentiated into skeletal stem cells (SkeSCs) and subcultured for eight consecutive passages. Top panel shows representative micrographs of SkeSCs at passages 3, 6 and 8 to highlight that the cell spindle morphology of these cells remains through all passages. The middle and bottom panel show representative micrographs of osteoblasts and adipocytes from SkeSCs at passage 3, 6 and 8 and after staining for alkaline phosphatase and Oil Red O, respectively.

Fig. 2. Skeletal stem cells reduce proliferation capacity over time. A) Representative image of RT-PCR gel showing reduce in KI67 expression in skeletal stem cells (SkeSCs) derived from human pluripotent stem cells at passage 6 compared to cells at passage 3. TBP was used as loading control and to normalize the expression of KI67 for the densitometry analysis of bands. B) Graph indicating the cell doubling time in hours (h) of SkeSCs at passage 3 and 6.

Fig. 3. Expression of cell surface biomarkers in skeletal stem cells. Skeletal stem cells (SkeSCs) derived from human pluripotent stem cells were subcultured for eight

consecutive passages and flow cytometry analysis was performed at passages 3, 6 and 8 to quantify the percentage of cells with positive expression of CD29, CD49f, CD73, CD104, CD105, CD146, CD271 and CD45. The asterisk (*) denotes statistical significant differences ($p < 0.050$) detected by student *t*-test.

Fig. 4. Identification of transcription factors expressed in skeletal stem cells. Graph showing relative mRNA level expression of transcription factors (*ETV1*, *ETV5*, *FOXP1*, *GATA6*, *HMG2*, *KLF12*, *PRDM16*, *SIM2* and *SOX11*) in human pluripotent stem cells (hPSCs), and skeletal stem cells derived from the pluripotent stem cells (hPSC-SkeSCs) and isolated from bone marrow extracts (hBM-SkeSCs). A fold change equal or bigger to 2 was used to identify significant differences in relation to relative mRNA expression levels on fibroblasts. The relative mRNA levels of each gene were normalized to house keeping gene *TBP*.

Fig. 5. Validation of biomarkers expressed skeletal stem cells. Graphs comparing the relative mRNA levels of integrin $\alpha 6$ (ITGA6) and transcription factors between skeletal stem cells (SkeSCs) at passage 3 to cells at passage 8 (A), and cells in early stage of adipogenic and osteogenic differentiation (B). The relative mRNA levels of each gene were normalized to house keeping gene *TBP*.