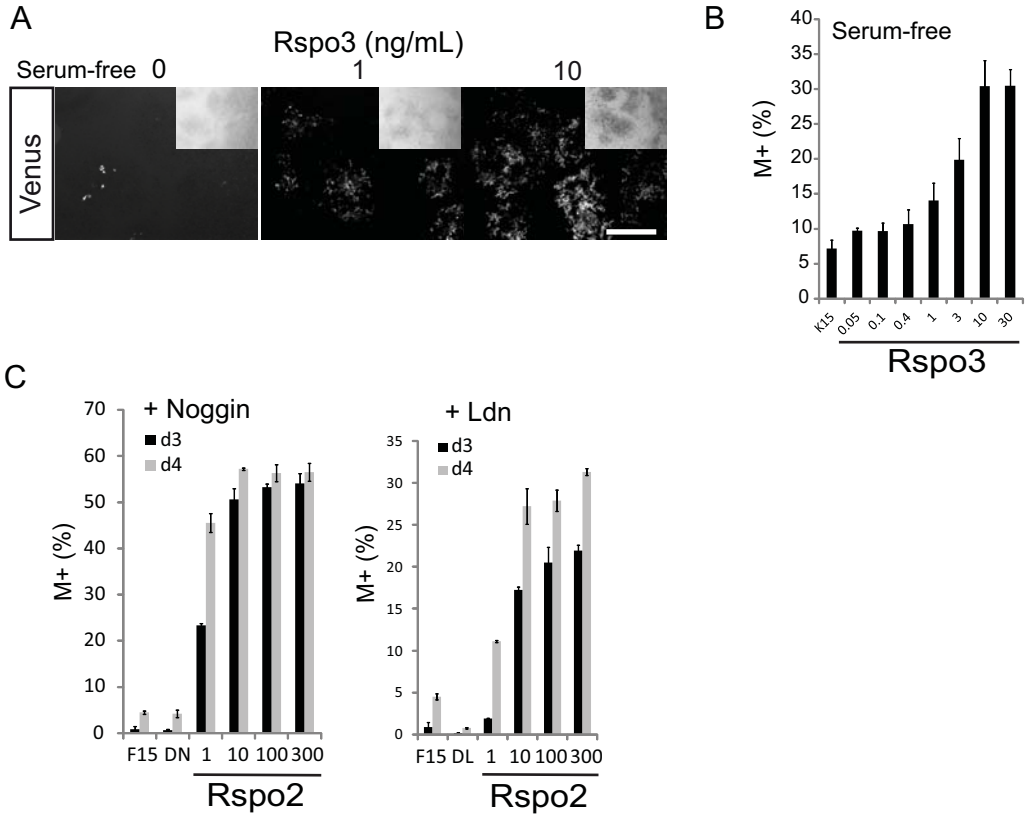


Figure S1: R-spondins induce the Msgn1-repV reporter in differentiating mouse ES cell cultures

(A) Activation of Msgn1-repV reporter with Rspo3 in serum-free conditions. Venus detection (white signal) of the reporter activation after 4 days of differentiation of Msgn1-repV ES cells, in a DMEM-based medium with 15% serum replacement (KSR) and supplemented or not with recombinant Rspo3. A transmitted light picture of each culture is shown in inset for each condition. Scale bar, 200µm

(B) Quantification of the proportion of Msgn1-repV -positive (M+) cells in cultures differentiated for 4 days in serum-free medium containing 15% KSR (K15) without or with various concentrations of Rspo3 (in ng/mL). Mean+/- s.d.

(C) Induction kinetic of the Msgn1-repV reporter (M+) in mouse ES cell culture in presence of Rspo2. Cultures were differentiated for 3-4 (d3-4) days in either control DMEM-based media containing 15% FBS alone (F15), or supplemented with DMSO 0.5% and Noggin (DN, Left graph), or with DMSO 0.5% and Ldn (DL, Right graph), with various concentrations of Rspo2 (in ng/mL). Mean+/- s.d.

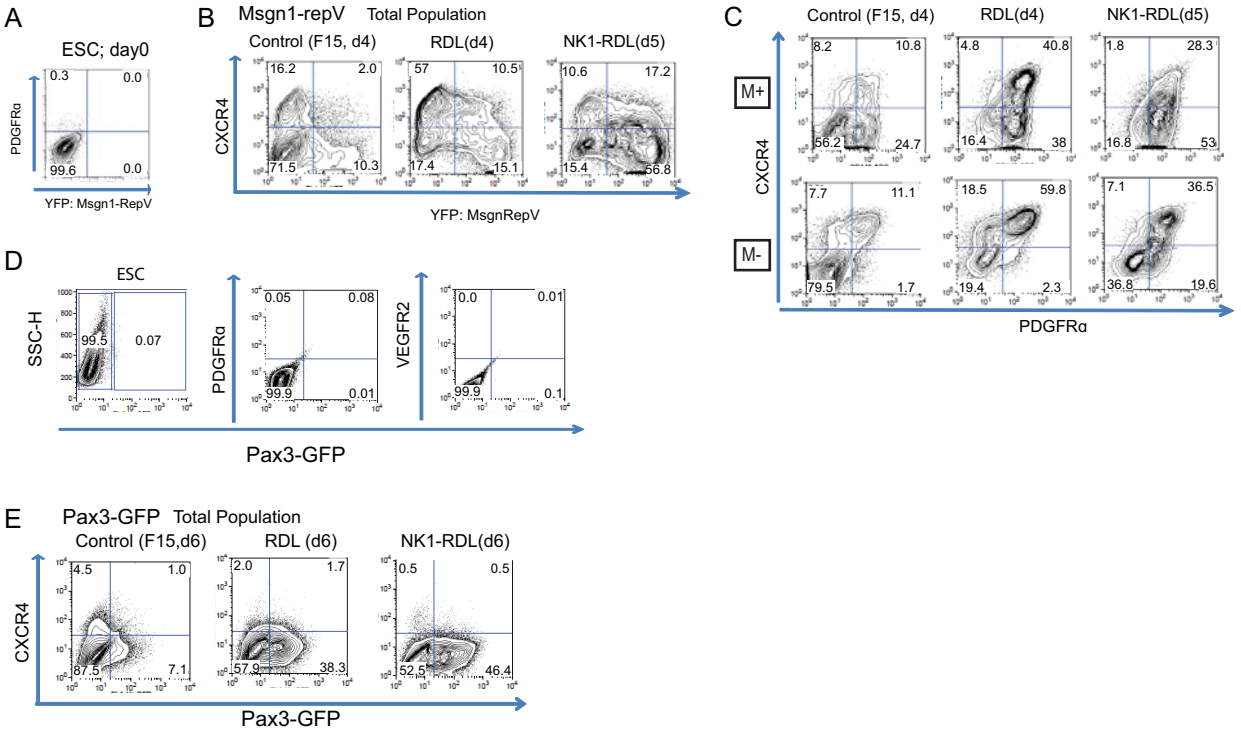


Figure S2: Immunophenotyping of mouse ES-derived PSM-like cells for PDGFRα, VEGFR2 and CXCR4

(A) Representative FACS analysis of undifferentiated (day0) mouse ES Msgn1-repV cells for PDGFRα expression.

(B) Representative FACS analysis of Msgn1-repV reporter cell cultures differentiated *in vitro* in control DMEM-based 15% FBS (F15, left), or in RDL medium without (center) or with pre-differentiation in NK1 medium (NK1- RDL, right) for 4-5 days and labeled with an anti-CXCR4 antibody.

(C) Representative FACS analysis of PDGFRα and CXCR4 expression in the Msgn1-RepV-positive (M+) and -negative (M-) populations, differentiated in (left) control DMEM-based 15% FBS (F15), in RDL medium without (center) or with pre-differentiation in NK1 medium (NK1- RDL, right) for 4-5 days.

(D) Representative FACS analysis of undifferentiated mouse ES Pax3-GFP reporter cell cultures for PDGFRα or forVEGFR2 expression. SSC-H: side scatter height.

(E) Representative FACS analysis of mouse ES Pax3-GFP reporter cell cultures differentiated *in vitro* in (left) control DMEM-based 15% FBS (F15), in RDL medium without (center) or with pre-differentiation in NK1 medium (NK1- RDL, right) for 6 days and labeled with an anti-CXCR4 antibody.

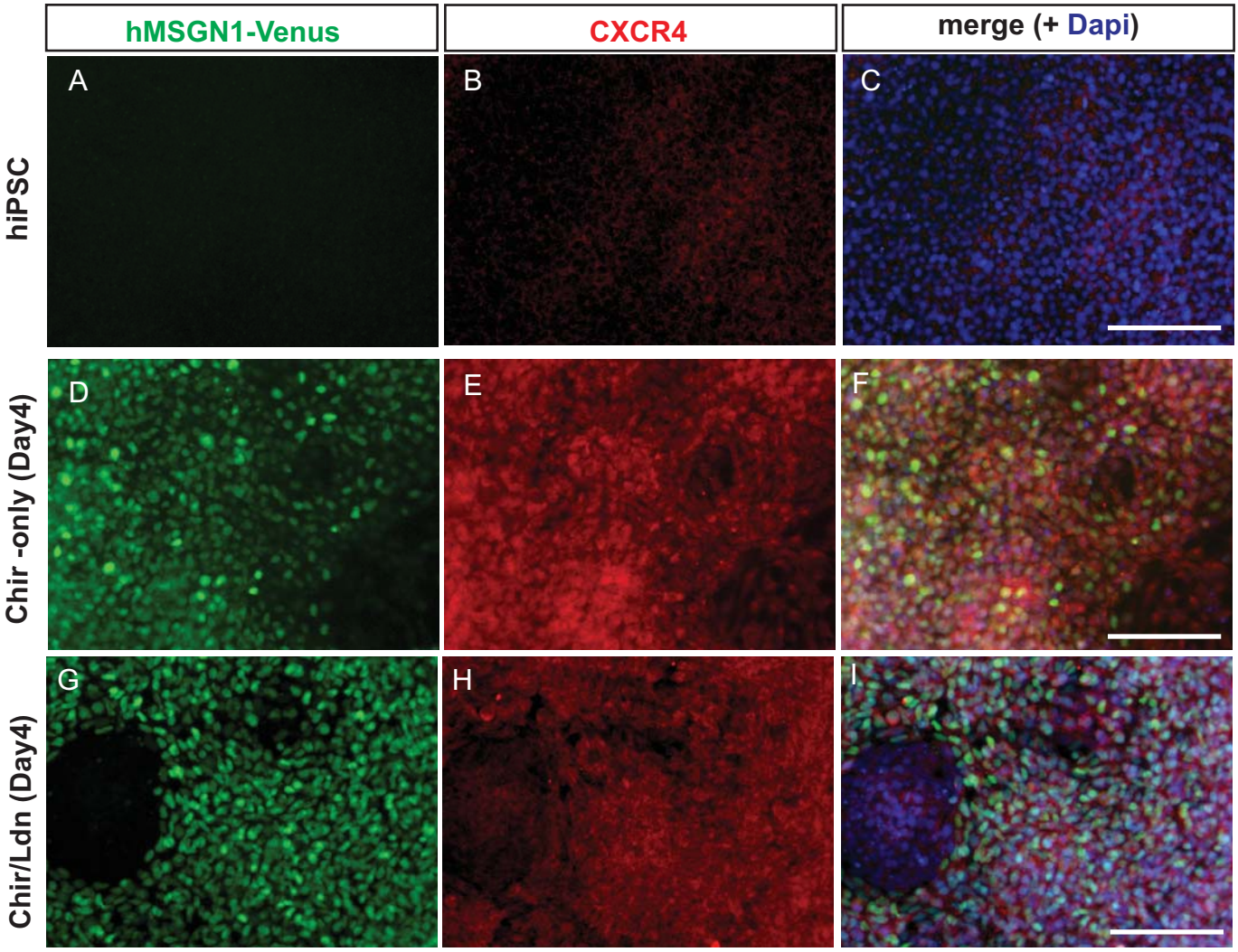


Figure S3: CXCR4 immunostaining on undifferentiated and differentiated human iPS cells
Undifferentiated hiPS hMSGN1-Venus cells (A-C), or differentiated for 4 days in Chir -only medium (D-F) or in Chir/Ldn medium (G-I) stained for Venus (hMSGN1-Venus, green) and CXCR4 (red) expression. Merge is shown with nuclear (Dapi, blue) counterstaining. Scale bar, 100µm.

Table S1. Comparison of myogenic protocols for differentiation of mouse ES cells

(A) Serum-containing and serum-free culture protocols comparison

	Serum-containing protocol (this study)	**Serum-free protocol (*Chal et al. 2015*)**	
mES cells maintenance	Lif medium on feeders (Gelatin -coated dish)	2i +Lif medium on feeders (Gelatin -coated dish)	
Prediff. passaging	Twice on gelatin-coated dish in 2i+Lif	Twice on gelatin-coated dish in 2i+Lif	
Seeding density	~15K/cm^2 on gelatin-coated dish	~15K/cm^2 on gelatin-coated dish	
Day0-2	**NK1** DMEM/F12:Neurobasal base + N2, B27 supplements + KSR 1% + BSA@ 25µg/mL + Insulin@ 10µg/mL (alt. ITS 1%)	**NK1 B4** DMEM/F12:Neurobasal base + N2, B27 supplements + KSR 1% + BSA@ 25µg/mL + Insulin@ 10µg/mL (alt. ITS 1%) **+ Bmp4@ 10ng/mL**	
Day2-4	**DF15 RDL/ CDL** DMEM (high glucose) base **+ FBS 15%** + b-ME@ 0.1mM + **R**spo3@ 10-30ng/mL or **C**HIR99021@ 1-5µM + **D**MSO@ 0.5% + **L**DN193189@ 0.1µM	**DK15 RDL/ CDL** DMEM (high glucose) base **+ KSR 15%** + b-ME@ 0.1mM + BSA@ 25µg/mL + **R**spo3@ 10-30ng/mL or **C**HIR99021@ 1-5µM + **D**MSO@ 0.5% + **L**DN193189@ 0.1µM	
Day4-6	**DK14F1 RDL/CDL** DMEM (high glucose) base **+ FBS 1%** **+ KSR 14%** + b-ME@ 0.1mM + "R,D,L or C,D,L"	**DK15 RDL/ CDL** DMEM (high glucose) base **+ KSR 15%** + b-ME@ 0.1mM + BSA@ 25µg/mL + "R,D,L or C,D,L"	
Day6-8	**DK14F1 PDL** DMEM (high glucose) base + FBS 1% + KSR 14% + b-ME@ 0.1mM + **PD173074@ 250nM** + **D**MSO@ 0.5% + **L**DN193189@ 0.1µM	**DK15 HIFL** DMEM (hi glucose) base + KSR 1% + b-ME@ 0.1mM + BSA@ 25µg/mL + **HGF@ 10ng/mL** + **IGF-1@ 2ng/mL** + **FGF-2@ 20ng/mL** + **L**DN193189 @ 0.1µM	
Day8+	**HS2%** DMEM (high glucose) base + b-ME@ 0.1mM + **HS 2%**	**DK15 HIF** Dmem (high glucose) base + KSR 1% + BSA@ 25µg/mL + b-ME@ 0.1mM + "H,I,F"	

* FBS: Fetal bovine serum; HS: Horse serum; BSA: Bovine serum albumin, KSR: Knock-Out ™ Serum Replacement

(B) Reagents references

B-27 supplement (50X) with Insulin and vitamin A, serum free	ThermoFisher	17504-044
N2 supplement (100X) with Insulin, serum free	ThermoFisher	17502-048
Knock-Out ™ Serum Replacement	ThermoFisher	10828028
DMEM/F12 glutamax	ThermoFisher	31331-093
Neurobasal	ThermoFisher	21103-049
DMEM high glucose w/o Sodium Pyruvate	ThermoFisher	41965-039
MEM-NEAA solution	ThermoFisher	11140-035
Pen/Step 10000 U/ml	ThermoFisher	15140-122
Beta-Mercaptoethanol (b-ME, 50 mM)	ThermoFisher	31350-010
DMEM high glucose with Sodium Pyruvate (for Feeders)	ThermoFisher	41966-029
L-glutamine	ThermoFisher	25030-123
Insulin–transferrin–selenium (ITS, 100X)	Life Tech.	41400-045
Bovine serum albumin (BSA, 0.1% wt/vol)	Sigma-Aldrich	A7906
Insulin, 2.5mg/mL	Sigma-Aldrich	91077C

Table S1. Comparison of myogenic protocols for differentiation of mouse ES cells

(A) Comparison serum-containing (this study) and serum-free (Chal et al. 2015) culture protocols.

(B) Product references for base media reagents

Table S2. Comparison of myogenic cultures obtained from differentiated mouse ES cells with serum-containing and serum-free protocols

	Serum-containing protocol (this study)	**Serum-free protocol (*Chal et al. 2015*)**	
PSM induction	up to 70% Msgn1^{+} (*) (with NK1 step) ~20-40% Pax3^{+}	~ 45-60% Msgn1^{+} ~ 30-50% Pax3^{+} (*) (* may incl. neural/neural crest cell type)	
Cultures viability	> 6 weeks	< 5 weeks	
Fiber length (µm)	1000+ (up to 3mm)	1087 ± 395 (up to 2mm)	
Fiber diameter (µm)	12.2 ± 3.79 (up to 20 µm)	11.5 ± 0.62	
Sarcomeric length (µm)	2.49 ± 0.16	2.39 ± 0.12	
Fusion index (nuclei/mm)	up to 40-45	26 ± 0.5	
Local Fiber density (fiber per mm^2)	up to 850 (in patches)	up to 300 (in patches)	
Yield (Fibers/Cell input)	0.5 - 0.75	0.3 - 0.6	

Table S2. Comparison of myogenic cultures obtained from differentiated mouse ES cells with serum-containing and serum-free protocols

Comparison of myogenic culture obtained from mouse ES cells with serum-based protocol (this study) and serum-free protocol (Chal et al. 2015). While early premyogenic differentiation is comparable between the protocols, long term culture and maturation of the resulting muscle fibers is greatly enhanced in serum-containing protocol, as serum-free cultures progressively loose viability after 4weeks of culture. Values for serum-free, 3-4 week old cultures are from Chal et al. 2015. Analysis and quantifications of serum-containing, 5-6 week old cultures were done as described previously (Chal et al. 2015), see also Extended Data section. Values are shown as mean ± standard deviation or a range of values.

Table S3: List of genes differentially expressed between posterior and anterior PSM transcriptional domains. Expression intensity of each expressed probeset (RMA annotation) was compared between the posterior PSM domain (fragments 2 and 3 of PSM array series) and the anterior PSM domain (fragments 4, 5 and 6 of PSM array series). Comparison was done by Significance Analysis of Microarrays (SAM) analysis using MeV4.5 (TM4 Microarray Software suite). For each probeset, the corresponding gene and the relative expression fold change between the posterior and anterior PSM domains are shown. The False Discovery Rate was fixed at 0.525. Genes validated by *in situ* hybrization in this study are highlighted (yellow), or validated genes in the literature are also highlighted (orange).

Click here to Download Table S3

Table S4: List of signature genes for posterior and anterior PSM domains used for expression heatmaps. For each marker gene, Affymetrix probeset identity is shown. Lists were assembled from this study and published literature.

Click here to Download Table S4

Table S5: Complete lists of signature genes for each Paraxial mesoderm and Neural tube domains (GSL method). Signature gene lists were generated by the GSL method (see Materials and Methods for details) for each paraxial mesoderm domains in the PSM array serie, namely Tail bud, Posterior PSM, Anterior PSM and Somite. The neural tube array series was arbitrarily subdivided into a posterior and anterior domains and signature gene lists were generated for both.

Click here to Download Table S5

Table S6: Venn diagram comparing the gene signatures list (GSL) of PSM versus Neural tube (NT) array series. (left) Comparison between the core Neural tube GSL genes and the core PSM GSL genes. Core signature genes are genes found significantly upregulated in domain of a serie for a given tissue. (right) Comparison between all the signature Neural tube genes with all the signature PSM genes. For each subgroup, the corresponding list of signature genes is provided.

Click here to Download Table S6

Table S7: Lists of signature genes for mouse ES cells and differentiated FACS-sorted Msgn1-RepV-positive and Pax3-GFP-positive populations (GSL method). Signature gene lists were generated by the GSL method (see Materials and Methods for details) for undifferentiated mouse ES cells and differentiated populations (Msgn1-repV^{+}, Pax3-GFP^{+}) for 4 to 5 days of differentiation in Rspo3/DMSO/Ldn (RDL) or Chir/DMSO/Ldn (RDL) medium. The top 400 probe sets are provided for each condition. For each probe set, the corresponding gene and the relative expression fold change compared to the reference median value are shown.

Click here to Download Table S7

Table S8: List of genes differentially expressed in in 4 day-old hMSGN1-GFP-positive cells cultured in Chir/Ldn or Chir only media versus undifferentiated hiPSC. (TabA) List of genes upregulated >2 folds in MSGN1-GFP+ cells in Chir/Ldn versus undifferentiated hiPSC. (TabB) List of genes upregulated >2 folds in MSGN1-GFP+ cells in Chir only versus undifferentiated hiPSC

Click here to Download Table S8

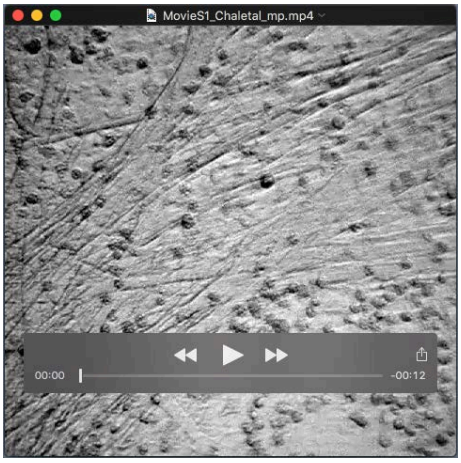
Table S9: List of genes differentially expressed in 4 day-old hMSGN1-GFP-positive cells cultured in Chir/Ldn versus Chir only media. (TabA) List of genes upregulated >2 folds in hMSGN1-GFP+ cells in Chir/LdnL vs Chir only conditions. (TabB) List of genes upregulated >2 folds in MSGN1-GFP+ cells in Chir only vs Chir/Ldn conditions. For each probeset/gene Fold change (red) and False Discovery Rate (FDR(BH), yellow) are highlighted.

Click here to Download Table S9

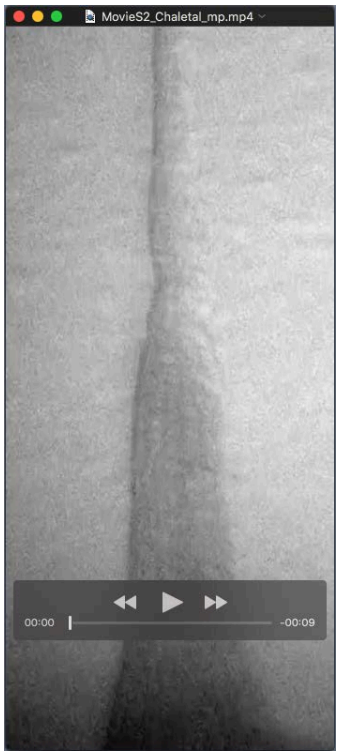
Table S10: *In situ* probes information

For each gene/probe, the ENSEMBL reference sequence number is provided. The primer couples sequences (F: forward, R: reverse) used to amplify the region of interest and the corresponding probe template size are provided.

Gene	ENSEMBL Ref. Seq	Primer sequences (5'>3')	Amp Size (bp)
Chst7	ENSMUSG00000037347	F- CCATCGATGGTGGTGCAACTCCTCCGAG R- CTTGGTCTCCAGTGGTGTCTC	628
Greb1	ENSMUSG00000036523	F- CCATCGATCAGTGACCACAGCAACAGGTC R- CCTGCTGGTACTGCCTTATCTC	680
Tpm1	ENSMUSG00000032366	F- ATCGATGGACCAGGAGCGGAAGCT R- CAGGGCCAGCTTTAGCTCG	622
Eogt	ENSMUSG00000035245	F- ATCGATTCTGAGGCTGACGATGCGC R- GCTGAGTGAGGTAGAGATTCAG	711
Tnfrsf19	ENSMUSG00000060548	F- ATCGATCAGTGCGGACCTGGCATG R- GAAGGCTGGCATCGTGTCC	725
Greb1L	ENSMUSG00000042942	F- CCATCGATGTCTGACTCAAACAGCCCACC R- CTCAGTTCTTGAAAGAGGACGAAC	800
Ifitm1	ENSMUSG00000025491	F- CCATCGATGAGAGATGCCTAAGGAGCAGC R- CCAGAATCTGTTATCTACTGTGAGATG	350
Add3	ENSMUSG00000025026	F- CCATCGATCATGTCTCCTGATCTACGACAG R- CTGAGGACGAGCACCTTACA	700
Lhfpl2	ENSMUSG00000045312	F- ATCGATGATGCTCTGGACCCTCCTG R- GGACTTTGTCACTGGAGGTTG	596
Fbn2	ENSMUSG00000024598	F- ATCGATCCTATTGCTGCCCAGGATGG R- CCAGCAGGGCATCTGCAC	634
Abca1	ENSMUSG00000015243	F- ATCGATTGAGCTACCCACCCTAGCAA R- TGTCCATGTTGTAGCGCAGT	573
Fam101a	ENSMUSG00000037962	F- CCATGCATCAGTCATGGTGGGTCACCTG R- CCTCCTCAGAGACTGGACTC	640
Myocd	ENSMUSG00000020542	F- ATCGATGCCTCCACGGCAGAAAGG R- GGTGCATCCCAGGGTAGC	733
Shroom3	ENSMUSG00000029381	F- ATCGATGAGTTGTGGGTTACTGGGGG R- TGCCTACATCCACCACAAAA	452
Vtn	ENSMUSG00000017344	F- CCATCGATCTACTTGTTCAAGGGTAGTCAGTA R- GAGGATTCACAGAGTCCACTC	727
Arg1	ENSMUSG00000019987	F- CCATCGATCACATGCAGCAGCAGCAGC R- CATCACCTTGCCAATCCCCAG	670
Pgm5	ENSMUSG00000041731	F- ATCGATGGTACTTCAGCAGGACGGC R- GGTTGGCAGCGATGATGGC	759
Ism1	ENSMUSG00000074766	F- CCATCGATGCTCACGCTGCACATCACG R- GGCAGTTTGGACGGTCACAG	350



Movie 1: Spontaneous contractile activity of differentiated mouse ES-derived myofibers. Mouse ES cells were differentiated in to PSM-like cells according to the RDL medium protocol and subsequently cultured in HS2% for 3 weeks and imaged with transillumination. Real time speed.



Movie 2: Spontaneous muscle bundle formation and contractile activity in adherent differentiated ES cultures. Mouse ES cells were differentiated in to PSM-like cells according to the CDL medium protocol and subsequently cultured in HS2%. Large fiber bundle formed spontaneously *in vitro* anchoring only by extremities (not shown)**.** Note synchronous twitching of the bundle of aligned myofibers. Real time speed.

Supplementary Materials and Methods

Mouse ES cell culture and differentiation

Maintenance

Mouse ES lines Msgn1-repV, Pax3-GFP, Myog-repV and Pax7-GFP have been described previously (Chal et al., 2015). Undifferentiated mouse ES cells were cultured on feeders (mitomycin-C inactivated mouse embryonic fibroblasts) at 37 °C in 5% CO2, in a maintenance medium composed of DMEM supplemented with 15% fetal bovine serum (FBS, Millipore), penicillin, streptomycin, 2 mM L-glutamine, 0.1 mM nonessential amino acids, 0.1% β-mercaptoethanol and 1,500 U/mL LIF. Prior to the NK1 pre-differentiation step, cells were first passaged twice onto gelatin-coated, feeder-free culture plates in 2i+LIF maintenance medium (Ying et al., 2008).

Analysis of R-spondin3 and Noggin action on mouse ES cells differentiation

ES cells were trypsinized and plated at approximately 10-20,000 cells/ cm^2 in gelatin-coated 24 well plate in a DMEM-based medium, with 15% FBS (DF15), supplemented or not with Rspo3 (Peprotech, R&D Biosystems), CHIRON99021 (Chir; Tocris, Stemgent), Dkk1 (R&D Biosystems), 0.5% DMSO (Sigma), Noggin (R&D Biosystems) and LDN-193189 (Ldn, Tocris, Stemgent) at various concentrations. Media was typically changed every second day. Experiments were done in biological triplicates.

Serum-based differentiation of mouse ES cells toward a PSM-like fate

ES cells cultured on Feeders (+LIF) were trypsinized and plated as single cells at approximately 10-20,000 cells/ cm^2 in gelatin-coated, feeder-free, 96, 24 or 6-well plates directly DMEM-based medium, with 15% FBS (DF15), supplemented with recombinant 10 ng/ml Rspo3 (Peprotech, R&D Biosystems), 0.5% DMSO (Sigma), and 0.1 µM LDN-193189 (Ldn, Tocris, Stemgent) for 2 d. Alternatively, Rspo3 was replaced with the GSK3-β inhibitor CHIRON99021 (Chir; Tocris, Stemgent) at 1-3 µM. At 2 days of differentiation, medium was changed to a DMEM-based medium, with reduced serum, typically 1% FBS, 14% KSR (DK14F1) supplemented with Rspo3 (or Chir), DMSO and Ldn (RDL or CDL media) as indicated above. Alternatively, ES cells cultured on gelatin (+2i + LIF) were trypsinized and plated at 10- 20,000 cells /cm^2 on gelatin-coated, feeder-free, 96, 24 and 6-well plates and pre-differentiated in serum-free N2B27 medium supplemented with 1% Knock-out Serum Replacement (KSR, Gibco), thereafter (NK1 medium) for 2 days (Table S4). Cells were then changed to the RDL (or CDL) medium described above.

Media were refreshed every 2 days until day 6. While both methods generated PSM-like cells, the NK1 pre-differentiation protocol was however more robust in generating large amount of PSM-like progenitors and was used for most of the long term *in vitro* myogenic differentiation (Table S4). PSM differentiation experiments were performed at least 30 times independently, on 4 different mouse ES cell lines.

Serum-based skeletal muscle differentiation of the mouse PSM-like cells

Following 6 days of differentiation into PSM-like cells, cultures were changed to a DMEM-based reduced serum medium (DK14F1) supplemented with 0.25 µM of the MEK inhibitor PD173074 (Stemgent) and 0.1 µM Ldn (PdL medium) for 2 days. After day 8 of differentiation, cultures were changed to a DMEM-based medium with 2% Horse serum (HS2%) (Table S4). Media was changed every other day. Myogenic differentiation experiments were performed at least 20 times independently, on 4 different mouse ES cell lines.

Neural differentiation

Protocol to generate $Pax3^{+}$ neural rosette was based on "dual-smad" inhibition method described previously (Chambers et al., 2009). Pax3-GFP mouse ES cells were seeded as single cells as for mesodermal differentiation but cultured instead in a DMEM-based medium with 15% FBS and supplemented with 10µM SB431542 (Tocris) and 0.1 µM LDN-193189 (Tocris).

Human iPS cell culture and differentiation

Maintenance

Undifferentiated human PS cell lines H9 (WA09) and hiPS11a (HSCI) cells were cultured on Matrigel (hESC-qualified, Corning)-coated dishes in mTeSR1 media (StemCell Technologies). Cells were passaged as aggregates. Lines were regularly confirmed to be mycoplasma-free using a VenorGEM detection kit (Sigma).

Serum-free and transgene-free PSM-like differentiation of the human iPS cells

Human iPS cells were differentiated essentially according to (Chal et al., 2016). Briefly, hPS colonies were dissociated with Accutase (StemCell Technologies) and plated as single cells on Matrigel -coated 24 and 12-well plates (approximately 15,000–30,000 cells/cm^2) in mTeSR1 supplemented with 10µM ROCK inhibitor (Y-27632, Sigma) for one day. The medium was changed to a DMEM-based medium supplemented with Insulin-Transferrin-Selenium (ITS, Gibco), 3 µM CHIRON99021 (Axon MedChem, Tocris) and 0.5 µM LDN-193189 (Axon

MedChem, Stemgent) (CL medium). At day 3, 20 ng/ml FGF-2 (R&D Systems) was added for additional 3 d. Differentiation experiments were performed at least 15 times independently, on 2 unrelated human iPS lines. All human iPS and ES cell experiments were done according to local regulations (IGBMC and Brigham and Women's Hospital) and in agreement with national and international guidelines.

Human hMSGN1-Venus hiPS reporter line generation

The *MSGN1* locus of undifferentiated hiPS11a was targeted by the CRISPR-Cas9 method (Cong et al., 2013; Ran et al., 2013) using the pSpCas9 (BB) vector (pX330; Addgene plasmid ID: 42230) to clone sgRNAs targeting the stop codon of *hMSGN1* exon1, and a corresponding targeting vector, *hMSGN1-2A-nls-Venus*, containing a ~1.5kb 5'- homology arm (HA) followed by a 2A-peptide sequence (for translational cleavage) and a nuclearly localized (nls) Venus (YFP), followed by a Neomycin selection cassette and a ~1kb 3'HA. Both heterozygous and homozygous integrations were selected and coding regions were sequenced to verify that they do not contain indels. Three clones were further validated for correct expression and differentiation.

Flow cytometry analysis

Cells were trypsinized and analyzed by flow cytometry on a FACS calibur (BD Biosciences) according to reporter expression. Gating was determined for each reporter line using corresponding undifferentiated culture as a baseline control. Data are represented as % of GFP (or YFP)$^+$ cells in the culture. **Sorting.** Msgn1-repV$^+$ (M$^+$) and Pax3-GFP$^+$ (P$^+$) cells were isolated by FACS from Msgn1-repV and Pax3-GFP cultures respectively differentiated for 4-6 days in RDL (or CDL) medium. Biological triplicates were generated. Gated fractions were sorted either on FACS Aria (BD Biosciences) or S3 cell sorter (Bio-Rad). hMSGN1-Venus$^+$ were sorted at 4-5 days of differentiation. Data were further analyzed with FlowJo software. Sorted populations were either processed for microarray analysis, RT-qPCR or for transplantation experiments.

Immunophenotyping. Cells were trypsinized, resuspended in blocking solution PBS- 5% FBS supplemented with 5μM Rock inhibitor (Y-27632, Sigma) and incubated for 10 min at 4C. Conjugated primary antibodies, mouse anti-PDGFRα (CD140a, clone APA5)-APC and mouse anti-VEGFR2 (Flk-1/KDR/CD309, clone Avas12a)-PE (eBioscience) were added at 1μg and 0.5μg per million of cells, respectively. Corresponding isotype staining controls were also run in parallel. Cells were incubated for 20 mins at 4°C. Next, cells were washed with PBS and

resuspended in PBS-2% FBS and analyzed or sorted for RT-qPCR analysis. Flow cytometry analyses were performed at least three times independently.

RT-qPCR

Total RNA was extracted from PSC cultures using Trizol (Invitrogen) or with the RNeasy micro-kit (Qiagen). RT-qPCR was performed on 5 ng total RNA using QuantiFast SYBR Green RT-qPCR Kit (Qiagen) and gene-specific primers (Primerbank) and run on a LightCycler 480II (Roche). β-actin was used as an internal control for mouse system, and GAPDH for the human system. For mouse tail/PSM reference, mouse E9.5 tails were microdissected posterior to the level of the forming somite (S0) and pooled for total RNA extraction. For embryonic and fetal muscle references, CD1 mouse E11.5 trunk muscles and mouse E17.5 back muscles were micro-dissected for total RNA extraction.

Microarrays generation

Generation of the mouse neural tube (NT) microarray series was done as described previously for the PSM microarray series (Chal et al., 2015). Briefly, CD1 mouse E9.5 embryos were micro-dissected into caudo-rostral series of consecutive ~100 µm fragments. Two series of 6 fragments were generated from 2 different embryos, from the tail bud to the level of the newly formed somite (S0) level (Figure 1B). RNA from each fragment was extracted with Trizol (Invitrogen) and used to generate probes hybridized on GeneChip Mouse Genome 430 2.0 arrays, as described previously (Chal et al., 2015). Differentiated mouse ES and human iPS cultures were dissociated, FACS-sorted for reporter expression and processed as described previously (Chal et al., 2015). Mouse ES cells samples were hybridized on GeneChip Mouse Genome 430 2.0 arrays, while hiPS samples were hybridized on GeneChip Human Genome U133 2.0 arrays (Affymetrix).

Microarray data analysis

Initial filtering and preprocessing, including background correction, quantile normalization and summarization, was performed using both RMA and MAS with the R Bioconductor package (R version 2.12.1, Bioconductor version 2.8). Expression sets were then filtered according to Calls information. Probe sets expression fold changes between conditions (biological duplicates or triplicates) were calculated using the “Comparative Marker Selection” module of GenePattern (Reich et al., 2006). Histogram expression profiles of gene probesets were generated from MAS values. Further analysis was performed using the Manteia database (Tassy and Pourquie, 2014). Hierarchical clustering was performed on Microarray RMA data. Clustering was computed with

an Average linkage method and Euclidean distances. Both an Approximately unbiased (AU) and Bootstrap probability (BP) P values were calculated (pvclust, R package). Clusters with AU and BP P values > 0.95 are highlighted by color-coded boxes. Expression heatmaps were generated with TM4-MeV (Saeed et al., 2006).

Gene signature lists (GSL) method

A gene expression reference was created by using all microarrays of wild-type mouse tissues deposited in GEO corresponding to Affymetrix Mouse Genome 430 2.0 Arrays (GEO platform id GPL1261) as of 08-2011. 1320 microarrays data sets from 255 distinct experiments were downloaded using a Perl script (Bioperl). Normalization was done by calculating the mean values for each microarray. The median values for the distribution of those mean values across all microarrays were determined. This median was then used as a scaling factor for each value on each microarray. Once all microarrays had been normalized, the median expression value for each probeset was defined as the reference value for that particular probeset/gene. Gene signature list specific to one experimental condition was generated by normalizing the corresponding microarray data as done for the reference dataset. Signature gene for a given conditions were any probeset whose normalized expression value was 10 times higher than the corresponding reference value.

PSM-like cells preparation for transplantation into injured *tibialis anterior* (TA) muscle

Msgn1-repV or Pax3-GFP mouse ES cells were differentiated for 4-6 days and pretreated with 10μM ROCK inhibitor (Y-27632, Tocris) one day before being trypsinized. One day prior to injection, the *tibialis anterior* muscles of a cohort of $Rag2^{-/-}$: $\gamma c^{-/-}$ mice (see below) were injected with 20μL of 10μM Cardiotoxin (Latoxan). Cell preparation was filtered through a 30μm mesh and reporter-positive cells were sorted using a FACS Aria II or a Moflow Astrios (BD Biosciences). Sorted cells were replated and permanently labeled by transduction overnight with a CAG-GFP lentivirus (MOI of 20-30). Cells were then washed several time and re-incubated for overnight in medium at 37°C before preparation for transplantation (Gayraud-Morel et al., 2012). For control satellite cells transplantation, Pax7-GFP^{+} were freshly isolated from Tg:Pax7-nGFP mouse hindlimbs and labeled overnight in suspension with a CAG-GFP lentivirus (Gayraud-Morel et al., 2012). 50,000 to 100,000 cells, resuspended in 10μl of culture medium, were injected in the pre-injured *Tibialis anterior* (TA) muscles of 3-4 month -old $Rag2^{-/-}$: $\gamma c^{-/-}$ male mice. Intramuscular injections were done under Ketamine/ Xylasin anesthesia. Grafted *Tibialis anterior* muscles were collected 1-2 months post transplantation and processed for cryosection and

immunohistochemistry (Gayraud-Morel et al., 2012). Transplantation experiments were done at least 3 times independently on cohorts of 3-5 animals. Experiments on mice were done according to European and international ethical regulations (Institut Pasteur, IGBMC).

In situ hybridization

Whole mount *in situ* hybridization was carried out as described (Henrique et al., 1995). Antisense mRNA probes for novel PSM markers identified by microarray and gene signature analysis were synthesized from mouse E.9.5 embryonic tail cDNA with sequence specific primers (Sigma, Table S8). Images were acquired on a Leica M125 stereomicroscope. For each probe, at least 3 embryos were hybridized.

Immunohistochemistry

For cell cultures, cultures plates were fixed in 4% formaldehyde for 20 mins at 25 °C or overnight at 4°C. Cultures were rinsed three times in PBS, followed by incubation in blocking buffer composed of Tris-buffered saline (TBS) supplemented with 1% FBS and 0.1% Triton X-100 for 30 mins. Primary antibodies were then diluted in blocking buffer and incubated overnight at 4 °C. The next day, cultures were washed three times with TBST (TBS supplemented with 0.5% Tween-20) and incubated with secondary antibodies (1:500) and counterstained with DAPI or Hoechst (1:1000) for at least 6 hours. Cultures were finally washed with TBST followed by PBS, before analysis.

Dissected TA muscles were prepared for cryosection according to (Gayraud-Morel et al., 2012). 12 µm transversal serial sections were prepared and incubated overnight with primary antibodies in blocking buffer on a Sequenza rack system (Shandon). Following TBST washes, slides were incubated with secondary antibodies conjugated with AlexaFluor (Molecular probes) at 1:500 in blocking buffer, and DAPI or Hoechst (1:1000) were used as counterstain. Following TBST and PBS washes, slides were mounted with Prolong Gold Antifade (Molecular Probes). Primary antibodies used in this study were anti-GFP (Abcam; 1:300), anti-Tbx6 (Abcam; 1:400), anti-MyoD (5.8A, Santa Cruz; 1:200), anti-Myogenin (F5D, DSHB; 1:250), anti-Pax3 (DSHB; 1:250), anti-Pax7 (DSHB; 1:250), anti-Slow MyHC (NOQ7.5.4D, Sigma; 1:250), anti-Embryonic MyHC (F1.652, DSHB; 1:250), anti-Fast MyHC (MY-32, Sigma; 1:300), anti-Dystrophin (Mandra1, Sigma; 1:400), anti-Dystrophin (DY4/6D3, Leica; 1:200), anti-Laminin (Abcam; 1:800), anti-CXCR4 (Abcam; 1:100) and anti-Ki67 (Abcam; 1:300). AchR was detected with conjugated Bungarotoxin (Molecular Probes).

Quantifications

Myofiber cross sectional area (CSA) was measured from immunostained (Dystrophin) transverse section of engrafted TA using Fiji (Schindelin et al., 2012). Two independent engrafted muscles were analyzed for each condition and the CSA of >150 individual fibers were measured per muscle. *In vitro* mouse ES-derived myofiber width was measured from immunostained (Fast MyHC) fibers in differentiated cultures using Fiji. For each condition, three fields of fibers were analyzed and the width of >70 individual fibers were measured per cultures. *In vitro* sarcomeric length was measured from immunostained (Fast MyHC) fibers in differentiated cultures using Fiji. Measurements were done along the main axis of the fiber, perpendicular to the striations. Fifteen individual mature fibers were identified and for each of them >20 consecutive sarcomeric units were measured. *In vitro* myogenic marker expression was quantified from immunostained cultures. For each condition, >5 myogenic fields per well were quantified. Results are expressed as a % of positive nuclei.

Image acquisition and processing

Live or fixed samples were acquired either on a Zeiss Axiovert or Evos FL for systems for transmitted light and/or fluorescent images. Images were processed with Adobe Photoshop and measurements were done with Fiji. For muscle contractions, movies were taken with a Leica DMRB microscope using a photometric FX camera and a 20X objective. Images were taken at 10Hz. Image sequences were processed in Fiji and saved in AVI format at real time speed.

Statistical analysis

For array data hierarchical clustering, Approximately unbiased (AU) and BP P values were both calculated (pvclust, R package). Clusters with AU and BP P > 0.95 were considered significant. For differential array expression data comparison, unpaired t-test P values were calculated using the Comparative Marker Selection module of GenePattern (Reich et al., 2006). Differences were considered significant for P values ≤ 0.05. For qRT-PCR data, unpaired Student t-test and two-tailed P values were calculated. Differences were considered significant for P values ≤ 0.05.

Supplementary material References

Chal, J., Al Tanoury, Z., Hestin, M., Gobert, B., Aivio, S., Hick, A., Cherrier, T., Nesmith, A. P., Parker, K. K. and Pourquie, O. (2016). Generation of human muscle fibers and satellite-like cells from human pluripotent stem cells in vitro. *Nature protocols* **11**, 1833-1850.

Chal, J., Oginuma, M., Al Tanoury, Z., Gobert, B., Sumara, O., Hick, A., Bousson, F., Zidouni, Y., Mursch, C., Moncuquet, P., et al. (2015). Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. *Nat Biotechnol* **33**, 962-969.

Chambers, S. M., Fasano, C. A., Papapetrou, E. P., Tomishima, M., Sadelain, M. and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol* **27**, 275-280.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823.

Gayraud-Morel, B., Chretien, F., Jory, A., Sambasivan, R., Negroni, E., Flamant, P., Soubigou, G., Coppee, J. Y., Di Santo, J., Cumano, A., et al. (2012). Myf5 haploinsufficiency reveals distinct cell fate potentials for adult skeletal muscle stem cells. *J Cell Sci* **125**, 1738-1749.

Henrique, D., Adam, J., Myat, A., Chitnis, A., Lewis, J. and Ish-Horowicz, D. (1995). Expression of a Delta homologue in prospective neurons in the chick. *Nature* **375**, 787-790.

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A. and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature protocols* **8**, 2281-2308.

Reich, M., Liefeld, T., Gould, J., Lerner, J., Tamayo, P. and Mesirov, J. P. (2006). GenePattern 2.0. *Nat Genet* **38**, 500-501.

Saeed, A. I., Bhagabati, N. K., Braisted, J. C., Liang, W., Sharov, V., Howe, E. A., Li, J., Thiagarajan, M., White, J. A. and Quackenbush, J. (2006). TM4 microarray software suite. *Methods Enzymol* **411**, 134-193.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods* **9**, 676-682.

Tassy, O. and Pourquie, O. (2014). Manteia, a predictive data mining system for vertebrate genes and its applications to human genetic diseases. *Nucleic Acids Res* **42**, D882-891.

Ying, Q. L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P. and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-523.