ONLINE METHODS

Mouse lines

All animal protocols were approved by the Stanford University Administrative Panel on Laboratory Animal Care (APLAC) (protocol #10509 and #29398) and experiments were performed in compliance with the institutional guidelines of Stanford University. C57BL6 young adult mice were purchased from Jackson Laboratories (Bar Harbor, ME) and used at 2-4 months of age (median age 2 months). Mice ubiquitously expressing a green fluorescent protein (GFP) transgene and mice ubiquitously expressing a firefly luciferase (Fluc) transgene driven by the ACTB promoter (L2G85 strain) were obtained and genotyped as described previously[10]. Double transgenic GFP/Luciferase mice were generated by breeding the above strains and were confirmed by appropriate PCR-based strategies to validate the genotype. Cells from C57BL/6 mice were used for cell proliferation and protein expression experiments. Cells from GFP/Luciferase mice were isolated for transplantation experiments from young mice (median age 2 months). Pax7-ZsGreen transgenic mice were a kind gift from Dr. Kyba (University of Minnesota, Minneapolis, MN). Pax7-knock out mice were a kind gift from Dr. Rudnicki (University of Ottawa, Ontario, Canada). Double-transgenic *Pax7^{CreERT2};Rosa26-LSL-tdTomato* were generated by crossing *Pax7^{CreERT2}* mice obtained from Jackson Laboratory (Stock # 017763) and *Rosa26-LSL-tdTomato* obtained from Jackson Laboratory (Stock # 007914). We validated these genotypes by appropriate PCR-based strategies. We genotyped Pax7 knock mice using the following primer combinations: *wild type forward* 5'-GTGGGGTCTTCATCAACGGTC-3'; *mutant forward* 5' TCGTGCTTTACGGTATCGCCG-CTCCCG-3'; *reverse* 5'-GGGCTTGCTGCCTCCGATAGC-3'. Double-transgenic

Pax7^{CreERT2};Rosa26-LSL-tdTomato were genotyped using the following Tomato Primer Mix: MMW#314, 5'-AAG GGA GCT GCA GTG GAG TA-3'; MMW#315, 5'-CCG AAA ATC TGT GGG AAG TC-3'; MMW#316, 5'-GGC ATT AAA GCA GCG TAT CC-3'; MMW#317, 5'-CTG TTC CTG TAC GGC ATG G-3'. CreERT2 expression was validated using the following primers: *common forward* 5'-GCT GCT GTT GAT TAC CTG GC-3'; *wild type reverse* 5'-CTG CAC TGA GAC AGG ACC G-3'; *mutant reverse* 5'-CAA AAG ACG GCA ATA TGG TG-3'. The number of animals for each data set and all relevant details regarding the sample size are reported with each experiment.

Cell isolation and dissociation of muscle tissue

Tibialis Anterior (TA) and *Gastrocnemius* (GA) muscles were dissected and subjected to mechanical dissociation using the gentleMACS^{TM} dissociator (Miltenyi Biotech), followed by collagenase (0.25%) and dispase (0.04 U/ml; Roche, Indianapolis, IN) digestion at 37C for 2 hours. The resulting cell suspension was passed through a standard syringe needle and subsequently through a 70-μm nylon filter (BD Biosciences, San Jose, CA). No established cell lines were used in this study.

High-throughput flow cytometric surface screening

A high throughput flow cytometric screen of 176 cell surface antibodies was performed, using the Mouse Cell Surface Marker Screening Panel (BD Biosciences, San Jose, CA). Pax7-ZsGreen reporter mice were used, as Pax7 is the most faithful marker of the entire MuSC population throughout lifespan[9]. Briefly, mouse TA and GA muscles were isolated from Pax7-ZsGreen reporter mice[9], dissected and subjected to collagenase and dispase digestion. Non-muscle tissue was removed under a dissection microscope and muscle fibers were dissociated. After digestion, the remaining cell suspension was

passed through a nylon 70-μm filter. Single-cell suspension (500,000 cells/ well) was added to the antibody containing wells. Cells were stained first with the purified monoclonal antibodies, then with the respective biotinylated secondary antibodies and finally with AlexaFluor-647 conjugated streptavidin. Subsequently cells were analyzed on a BD-LSRII flow cytometer using FACS Diva Software. During analysis, ZsGreen positive cells were identified and candidate novel markers that were expressed by the MuSC population (ZsGreen positive cells) were identified. The same assay was performed on primary myoblasts. For myoblast preparation, adult murine hindlimb muscles were dissociated using a combination of mechanical and enzymatic digestion, as in the muscle stem cell isolation, and myoblasts were derived by differential plating as previously described[48].

Isolation of muscle stem cells and novel progenitor populations

Muscle tissues was isolated and digested to a single-cell suspension as described above. Cells were incubated with biotinylated antibodies reactive to CD45, CD11b, CD31, and Sca1 (BD Biosciences, San Jose, CA) for 30 minutes at 4C and washed. Cells were then incubated with streptavidin magnetic beads (Miltenyi Biotech, Auburn, CA), streptavidin-APC-Cy7 (Invitrogen, Carlsbad, CA), α7 integrin-PE antibody (AbLab, Vancouver, Canada), and CD34-eFluor660 antibody (eBioscience, San Diego, CA) or CD9-APC antibody and CD104-FITC antibody (eBioscience, San Diego, CA). After magnetic depletion of the biotin-positive cells, the lineage-negative cells were stained with DAPI and sorted on a FACSAria cell sorter in purity mode, using FACS Diva software (BD Biosciences, San Jose, CA). To isolate MuSCs, we first gated for viable cells (DAPI negative), then for cells negative for the lineage markers (CD45, CD31,

CD11b, and Sca1), finally for cells positive for both CD34 and α7 integrin, which represent the muscle stem cell fraction. To simultaneously isolate stem cells and the novel progenitor populations using the new sorting strategy, we first gated for viable cells (DAPI negative), then for cells negative for the lineage markers (CD45, CD31, CD11b, and Sca1) and positive for α7 integrin and CD9. Finally we sorted SC-P2 based on their unique expression pattern of CD9 and CD104. Flow cytometry scatter plots were generated using FlowJo v8.7 (Treestar, Ashland, OR).

Transplantation of stem cells and novel progenitor cell populations

Sorted cell populations were transplanted immediately following FACS isolation into the TA muscle of recipient mice as previously described[10,30]. Cells from GFP/Luciferase mice were transplanted into gender-matched, hindlimb-irradiated NOD/SCID mice (Jackson Laboratories, Bar Harbor, ME). NOD/SCID mice (2-4 months of age, median 2 months) were anesthetized with ketamine (2.4 mg per mouse) and xylazine (240 µg per mouse) by intraperitoneal injection and irradiated by a single dose of 18 Gy administered to the hindlimbs, with the rest of the body shielded in a lead jig. Transplantation was performed within three days post-irradiation. Freshly isolated cells were counted by hemocytometer and resuspended at desired cell concentrations in PBS with 2.5% goat serum and 1 mM EDTA. Cells were transplanted by intramuscular injection into the tibialis anterior muscle in a 10 µl volume.

In vivo IdU labeling

Mice were weighed, anesthetized with isofluorane and injected intraperitoneally with IdU eight hours prior to sacrificing the animal (150 mg/Kg body weight per injection)[32].

Bioluminescence imaging

Bioluminescence imaging (BLI) was performed using a Xenogen-100 system, as previously described[10], at 5 weeks post-transplant. The system is comprised of a light-tight imaging chamber, a charge-coupled device (CCD) camera with a cryogenic refrigeration unit and the appropriate computer system (Living-Image Software, Caliper LifeSciences). Briefly, the animals were anesthetized under isofluorane and injected intraperitoneally with a 100ul volume of luciferin diluted in PBS (0.1mmol/Kg body weight, Caliper LifeSciences). Immediately after injection, images were acquired each minute for a total of 15 min and data were stored for subsequent analysis, using Living Image software (Caliper LifeSciences). Bioluminescence images acquired at 12 min post-luciferin injection were used for analysis. Bioluminescence signal was calculated by drawing a consistent region-of-interest (ROI) over each hindlimb and quantifying the resulting signal. A bioluminescence signal value of 80,000 photons/s represented our detecion threshold, as this level was previously determined to represent the presence of one or more GFP-positive muscle fiber in transplanted tissue[10,30]

Muscle injury

Mice (8-10 weeks) were acutely injured by a single 10 µl intramuscular injection of notexin (10 µg/ml; Latoxan, France) into the TA muscle and two injections in the GA muscle at the indicated time points.

Lineage tracing

Pax7^{CreERT2}*; Rosa26-LSL-Tomato* mice were treated with five consecutive daily intraperitoneal injections at a tamoxifen dose of 75mg/kg body weight to activate tdTomato expression following *Pax7*-dependent Cre mediated recombination. Ten days after the last tamoxifen injection, mice were subjected to intramuscular injection of 10 µl

of Notexin (Latoxan, France). Flow cytometry analysis was perfomed at 6 days post-injury.

Intracellular flow cytometry

Muscle stem cells were isolated from young mice by FACS sorting, fixed in 1.6% paraformaldehyde, permeabilized in methanol and stored at -80°C. Thawed cells were washed two times with staining buffer (PBS with 0.5% BSA and 2.5mM EDTA), and blocked in staining buffer for 30 minutes at room temperature. Cells were stained with the following primary antibodies, mouse monoclonal anti-Pax7 antibody (sc-81648, Santa Cruz Biotechnology, Santa Cruz, CA), rabbit polyclonal anti-Myf5 antibody (Clone C20, Santa Cruz Biotechnology, Santa Cruz, CA), mouse monoclonal anti-MyoD antibody (clone 5.8A, BD Biosciences, San Jose, CA), mouse monoclonal anti-Myogenin antibody (Clone F5D, BD Biosciences, San Jose, CA), rabbit polyclonal anti-ZsGreen antibody (632474, Clontech, Mountain View, CA) for 1 hour at room temperature. Subsequently, cells were washed and incubated in the presence of secondary donkey anti rabbit-AlexaFluor-488 and goat anti-mouse IgG1-AlexaFluor-647 antibodies (Jackson ImmunoResearch Laboratories) for 30 minutes at room temperature. Finally, cells were washed twice with staining buffer and analyzed on a BD-FACSCalibur flow cytometer using CellQuest software (BD Biosciences, San Jose, CA).

Antibody conjugation with metal isotopes

Purified antibodies were conjugated to the indicated metals for mass cytometry analysis using the MaxPAR antibody conjugation kit (DVS Sciences, Toronto, Canada) according to the manufacturer's instruction. Following labeling, antibodies were diluted

in Candor PBS Antibody Stabilization solution (Candor Bioscience GmbH, Wangen, Germany) to 0.2 mg/mL and stored long-term at 4°C. Each antibody clone and lot was titrated to optimal staining concentrations using murine myoblasts, as well as murine muscle and spleen cell suspensions. Antibodies against the intracellular myogenic transcription factors were first optimized by flow cytometry performed on murine muscle stem cells sorted from Pax7 Zs-Green mice and murine myoblasts. A list of all mass cytometry antibodies used for analysis can be found in Table S1.

Mass cytometry staining

Dead cells were stained using cisplatin (WR International, Radnor, PA, Cat# 89150-634) as previously described[49]. In particular, cells were resuspended in 1ml serum-free DMEM at $2x10^6$ cells/ml and cisplatin was added at a final concentration of 25 μM for 1 min at room temperature. The reaction was quenched with 3ml of DMEM/15% FBS. Samples were centrifuged at 1400 rpm for 10 min and cell pellets were resuspended in cell staining media (CSM: PBS with 0.5% bovine serum albumine (BSA) and 0.02% sodium azide) and fixed with 1.6% paraformaldehyde (PFA) for 10 minutes on ice. Cells were washed with CSM and stained with antibodies against surface markers included in the mass cytometry panel for 1h at room temperature. Cells were washed twice with CSM and permeabilized with methanol for 10 minutes on ice. Cells were washed twice with CSM and stained with antibodies against intracellular markers included in the mass cytometry panel for 1h at room temperature. Cells were washed twice with CSM and stained with 1 mL of 1:5000 191/193Ir DNA intercalator[50] (DVS Sciences, Richmond Hill, Ontario, Canada) diluted in PBS with 1.6% PFA for 20 mins at room temperature.

Barcoding

Barcoding was performed on about 1 million fixed cells, as described previously[51,52]. In particular, fixed cells were washed once in CSM, once in PBS, and finally in the permeabilization solution (PBS with 0.02% saponin). Cells were then resuspended in the residual permeabilization solution and incubated with the barcoding dyes for 15 minutes at room temperature. After barcoding, cells were washed twice with CSM, pooled in a single tube and stained with isotope-chelated antibodies. Upon data collection, individual samples were recovered using the single-cell deconvolution algorithm[52].

Mass cytometry measurement

Cells were acquired on the CyTOF TM mass cytometer (DVS Sciences, Richmond Hill, Ontario, Canada) at an event rate of approximately 500 cells per second as previously described[4]. The instrument was run in high-resolution mode (Mass resolution ~700) with internally calibrated dual-count detection. Noise reduction and cell extraction parameters were: cell length 10-65, lower convolution threshold 10. Samples were normalized using beta beads[53].

X-Shift analysis and graphic display of single-cell mass cytometry data

Events were clustered based on a combination of surface markers and myogenic transcription factors (α7 integrin, CD34, CD9, CD104, Pax7, Myf5, MyoD, Myogenin) using the X-Shift algorithm and clusters were validated[13]. In order to visualize the spatial relationships between the cell types within these X-shift clusters, 2000 randomly sampled cells from each cluster were subjected to a force-directed layout[13,15]. All conditions from a single experiment were processed simultaneously so that the resulting map would capture all populations present in the entire dataset.

Principal Component Analysis

Principal Component Analysis (PCA) plots were generated using the PCA function in the FactoMineR package in R. Dendrograms were fitted using hclust function in R. The distance matrix was calculated through the dist function using Euclidean parameters.

Statistics and Reproducibility

All animal experiments were randomized. No statistical method was used to predetermine sample size. The investigators were not blinded to allocation during experiments and outcome assessment. Transplantation experiments were performed with at least n=12 mice, two independent experiments. CyTOF analysis of muscle tissue during homeostasis and acute injury was performed with at last n=8 mice, 3 independent experiments, unless otherwise noted. The line indicated in the scatter plots represents the median (Fig. 3c), mean ± SEM (Fig. 2c, 4d, e) or mean ± SD (Fig. 3f). Bar graphs represent mean ± SEM. ANOVA or multiple t-test analysis was performed for experiments with more than 3 groups, as indicated in the legend. Bonferroni post hoc analysis was performed to determine significant differences between groups. Statistical significance was determined by $p<0.05$. *, **, *** and **** represent statistical significance at $p<0.05$, $p<0.01$, $p<0.001$ and $p<0.0001$ respectively. NS represents statistically non-significant. Source data of all representative experiments are provided in Supplementary Table 2.

Code Availability

The computational code used in the study can be obtained at https://github.com/nolanlab/vortex/releases.

Data Availability

CyTOF data that support the findings of this study have been deposited in Flowrepository.org under accession code FR-FCM-ZY3E, FR-FCM-ZY3F, FR-FCM-ZY3G, FR-FCM-ZY3H, FR-FCM-ZY3J, FR-FCM-ZY3K, FR-FCM-ZY3W). Source data for Fig. 1-6 and Supplementary Fig. 1-3 have been provided as Statistics Source Data Supplementary Table. All other data supporting the findings of this study are available from the corresponding author upon reasonable request.

Supplementary References

48. Rando, T. A. & Blau, H. M. Primary mouse myoblast purification, characterization, and transplantation for cell-mediated gene therapy. *The Journal of Cell Biology* **125,** 1275–1287 (1994).
49. Fienberg, H. G., Simonds, E. F., Fantl, W. J., Nolan, G. P. & Bodenmiller, B. A platinum-based covalent viability reagent for single-cell mass cytometry. *Cytometry* **81A,** 467–475 (2012).
50. Ornatsky, O. I. *et al.* Study of cell antigens and intracellular DNA by identification of element-containing labels and metallointercalators using inductively coupled plasma mass spectrometry. *Anal. Chem.* **80,** 2539–2547 (2008).
51. Behbehani, G. K. *et al.* Transient partial permeabilization with saponin enables cellular barcoding prior to surface marker staining. *Cytometry A* **85,** 1011–1019 (2014).
52. Zunder, E. R. *et al.* Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm. *Nat Protoc* **10,** 316–333 (2015).
53. Finck, R. *et al.* Normalization of mass cytometry data with bead standards. *Cytometry A* **83,** 483–494 (2013).

Supplementary Figure 1 (related to Figure 1). CyTOF analysis of skeletal muscle tissue.

(a) Intracellular staining of Pax7 by flow cytometry in sorted muscle stem cells (left panel) and myoblasts (right panel). **(b)** Muscle cells isolated from Pax-Zs green reporter mice were fixed and permeabilized for intracellular staining. Cells were simultaneously stained with antibodies against Zs-Green and Pax7. Cells that were positive for Pax7 were gated and the fraction of Zs-Green^{+} cells was quantified to be 90%. **(c)** CyTOF antibody titration. Isotope-chelated anti-mouse antibodies against the surface marker of muscle stem cells, α7 integrin, and intracellular myogenic transcription factors Pax7, MyoD, Myogenin have been titrated using positive (muscle) and negative (spleen) controls to optimize signal, achieve saturation and minimize background. **(d)** Gating strategy on CyTOF samples as described in Fig. 2a. Individual contour plots are shown. **(e)** Histogram plot of CD9 (left panels) and CD104 (right panels) expression in MuSCs (upper panels) and myoblasts (lower panels) compared to the respective isotype control. **(f)** Screening data were analyzed using the Bland-Altman method to measure significant differences in signal intensity (also known as Median Fluorescence Intensity (MFI)) of individual markers in myoblasts compared to MuSCs. The percentage difference from the average MFI (100 x (Myoblast MFI - MuSC MFI)/ Average MFI) is plotted (y axis) as a function of the average MFI ((Myoblast MFI + MuSC MFI)/2) (x axis). **(g)** PCA plot of Live/Lineage^{-}/α7 integrin^{+}/CD9^{+} cells by population in uninjured (day 0) samples. Protein expression levels were clustered by their $\log_2$ median intensities (representative experiment, n=3 mice)**.**

Supplementary Figure 2 (related to Figure 2 and 3). **Functional characterization of the newly identified progenitor population in skeletal muscle.**

(a) Representative biaxial dot plots of CD9 (y axis) by CD104 (x axis) colored by channel, showing MAPKAPK2 phosphorylation in populations SC, P1, P2 and P3 in $Pax7^{-/-}$ muscle (right) and WT control (left), isolated from neonates (upper panels) (n=3 mice, 2 independent experiments) and 3 weeks old mice (lower panels) (n=1 $Pax7^{-/-}$; mean ± SEM from n=10 WT, 2 independent experiments). **(b)** Representative biaxial dot plots of CD9 by CD104 as in (a) colored by Myogenin expression. **(c)** Representative biaxial dot plots of CD9 by CD104 as in (a) colored by Pax7 expression. **(d)** Individual populations were sorted by FACS and cultured in differentiation media for one week. Images were acquired with an AxioPlan2 epifluorescent microscope (Carl Zeiss) with ORCA-ER digital camera (Hamamatsu Photonics). Each population was differentiated to yield fusion competent cells (n=4, 2 independent experiments). Scale bar, 50 μm. **(e)** Representative images showing the gating strategy on samples analyzed by flow cytometry at day 0 (upper panels) and day 6 (lower panels) (Figure 3d-f). Live cells are identified based on lack of DAPI staining. $Lineage^+$ cells ($CD45^+/CD11b^+/CD31^+/Sca1^+$) are excluded from the analysis and myogenic cells are enriched by gating on the $\alpha 7integrin^+/CD9^+$ fraction. A biaxial plot of CD9 (y axis) by CD104 (x axis) (far right) shows populations SC, P1 and P2 (representative images, n= 3 mice per condition).

Supplementary Figure 3 (related to Figure 5). **Molecular characterization of stem and progenitor cells during acute muscle injury identifies cell state transitions.** **(a)** Supervised clustering analysis enforcing a 2-cluster solution in the stem (SC) and progenitor (P1, P2) cell populations during the time course of recovery from acute injury (Day 0 = D0; Day 3 = D3; Day 6 = D6), (representative experiment, n=3 mice per condition). Dendrograms were fitted using hclust function in R. The distance matrix was calculated through the dist function using euclidean parameters. **(b)** Representative biaxial dot plot of CD9 by CD104 colored by channel, showing expression of CD9 (upper panel) and CD82 (lower panel), in adult mice (n= 6 mice, 2 independent experiments). CD82 is highly expressed only in a small subset of populations P1 and P2.

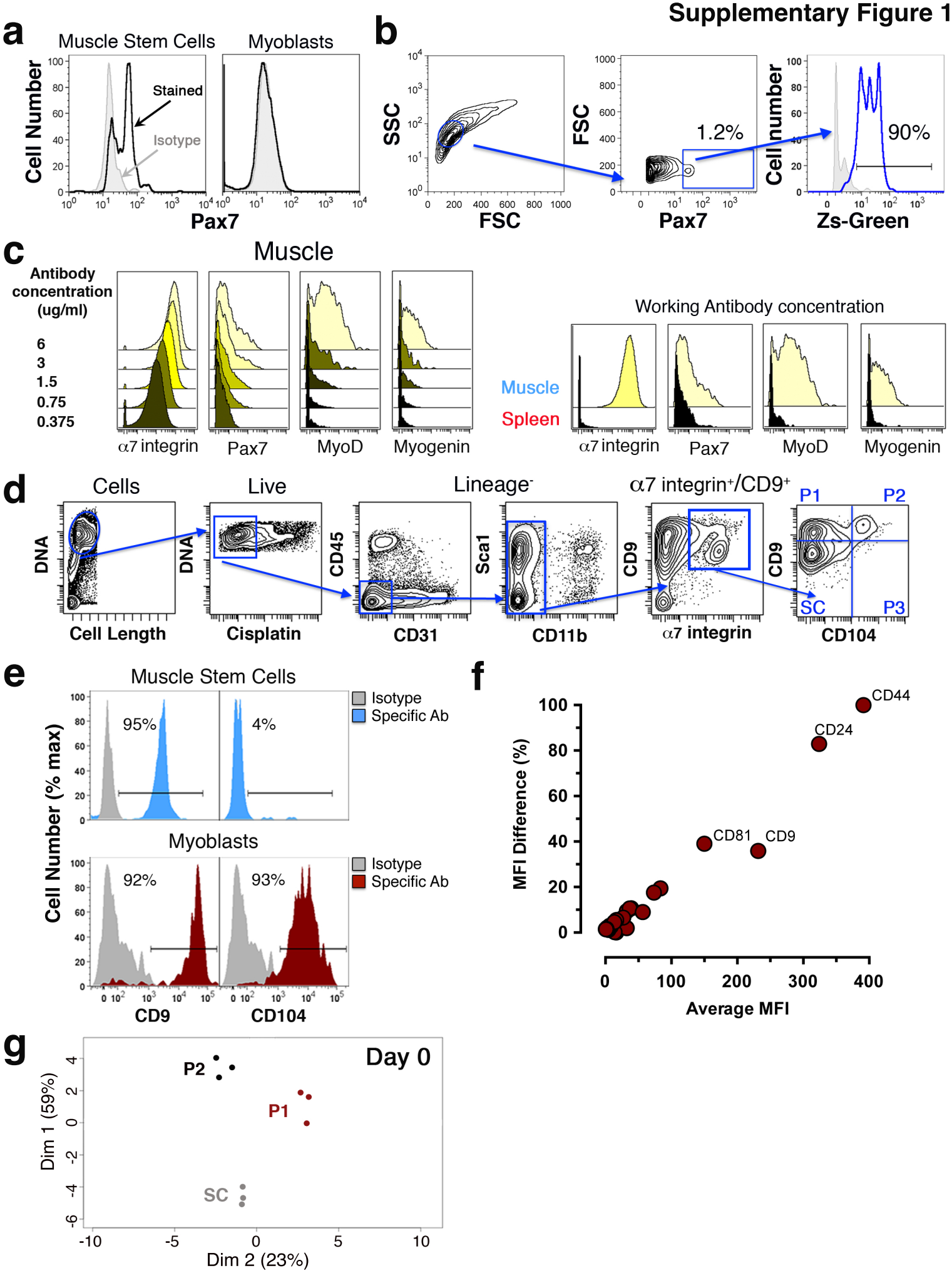
Supplementary Table 1 (related to Figure 1). Markers measured by Single-Cell Mass Cytometry.

Single-cell mass cytometry antibodies and reagents panel, with description of clone number and supplier.

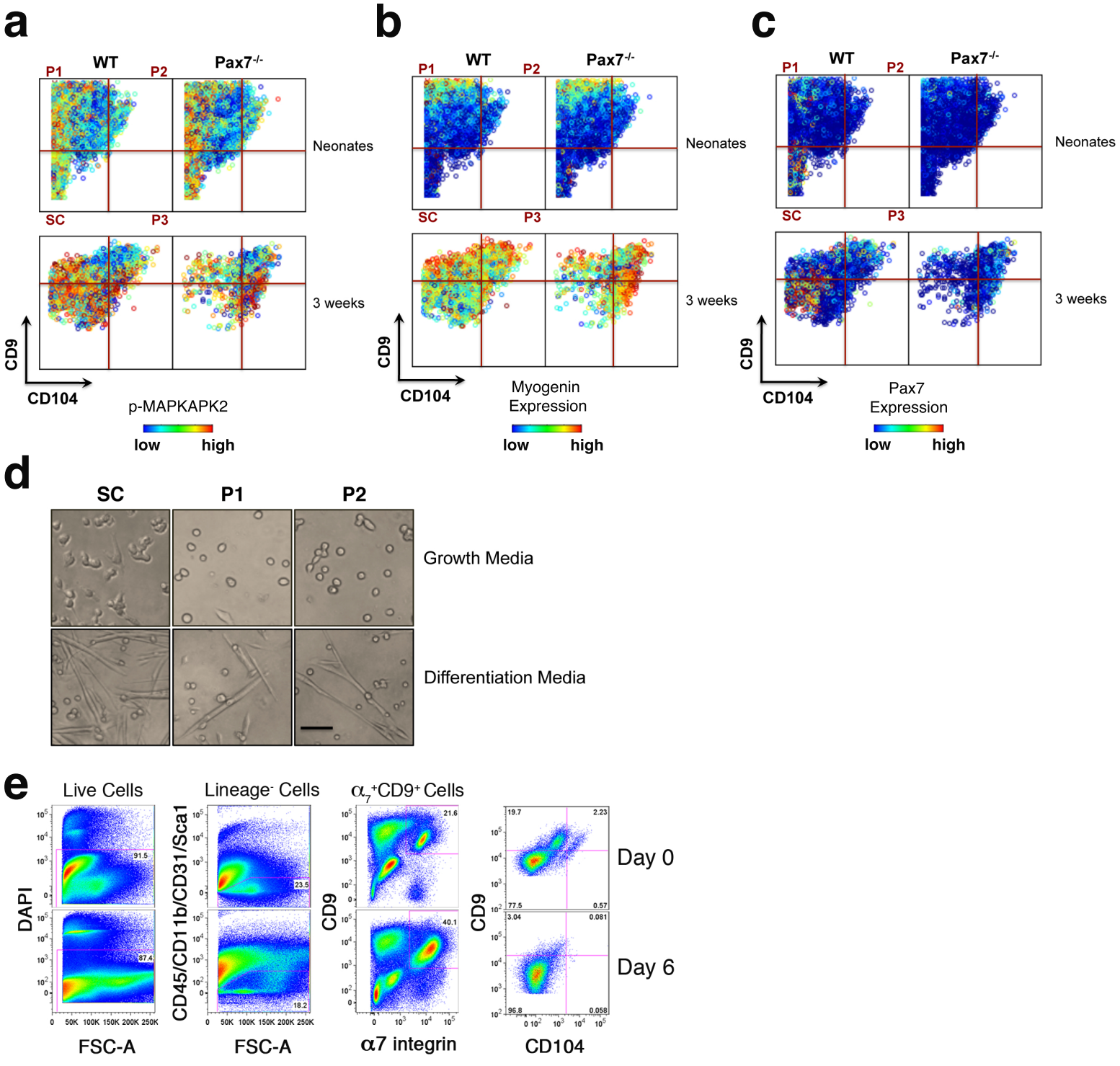
Supplementary Table 2. Statistics Source Data.

Source data and statistics of all experimental replicates relative to representative images presented in the main figures are compiled.

Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3