Microfluidic single-cell transcriptional analysis rationally identifies novel surface marker profiles to enhance cell-based therapies

Robert C. Rennert1,*, Michael Januszyk1,2,*, Michael Sorkin1,*, Melanie Rodrigues1, Zeshaan N. Maan1, Dominik Duscher1,3,4, Alexander J. Whittam1, Revanth Kosaraju1, Michael T. Chung1, Kevin Paik1, Alexander Y. Li1, Michael Findlay1,5, Jason P. Glotzbach1, Atul J. Butte6 & Geoffrey C. Gurtner1

Current progenitor cell therapies have only modest efficacy, which has limited their clinical adoption. This may be the result of a cellular heterogeneity that decreases the number of functional progenitors delivered to diseased tissue, and prevents correction of underlying pathologic cell population disruptions. Here, we develop a high-resolution method of identifying phenotypically distinct progenitor cell subpopulations via single-cell transcriptional analysis and advanced bioinformatics. When combined with high-throughput cell surface marker screening, this approach facilitates the rational selection of surface markers for prospective isolation of cell subpopulations with desired transcriptional profiles. We establish the usefulness of this platform in costly and highly morbid diabetic wounds by identifying a subpopulation of progenitor cells that is dysfunctional in the diabetic state, and normalizes diabetic wound healing rates following allogeneic application. We believe this work presents a logical framework for the development of targeted cell therapies that can be customized to any clinical application.
Cell-based therapies have been proposed for regenerative medicine and wound healing applications. Progenitor cell therapies are being tested in clinical trials to either directly address diabetic pathophysiology, or to treat diabetic complications such as retinopathy, critical limb ischaemic and diabetic foot ulcers. However, existing cell-based approaches have been developed primarily empirically based on the ‘legacy’ surface markers (SMs) that were originally described for other cell types, making it difficult to decide how to proceed when trials fail. Recently, there has been an increased understanding of the heterogeneity of stem and progenitor cell populations, as well as a shift in the mechanistic hypothesis of cell therapies from direct tissue engraftment to enhancement of dysfunctional endogenous repair pathways. Thus, there is a need to rationally develop targeted cell-based approaches for specific clinical applications through the selection of cell subpopulations with desired transcriptional profiles.

Customized cell therapies require an in depth knowledge of both disrupted cellular pathways in diseased tissue and therapeutic cell SM profiles to isolate discrete cell pools for application. Progress has been made in understanding gross repair pathway disruptions in diseased tissues, which provides a basis for rationally replacing deficient growth factors and cytokines. While enrichment of progenitor cells has shown therapeutic promise, a more granular understanding of the subpopulation dynamics of diseased and therapeutic progenitor cell pools has proven challenging because the resolution afforded by traditional population-level assays is insufficient to capture the complex relationships in heterogeneous cell populations. Standard approaches rely on pooling RNA or protein from hundreds of thousands of cells to report aggregate gene expression, and are thus unable to detect differential distributions in gene expression among cell subgroups. Recent advances in high-throughput, microfluidic technology have enabled massively parallel single-cell gene expression analyses, with the resulting data providing insights into the relationships among cells in complex tissues. Leveraging this technique in previous work, we have combined single-cell transcriptional analysis with advanced mathematical modelling to characterize heterogeneity in putatively homogeneous populations, as well as identify critical perturbations in cell subpopulations in pathologic states. Most recently, we have utilized single-cell analysis to link defects in the neovascular potential of diabetic and aged progenitor cells to the selective depletion of specific cell subsets. These findings support the concept of functional heterogeneity within progenitor cell pools and highlight the potential of highly selected cell therapies to reverse specific cellular and pathophysiologic defects in diabetic and other impaired tissues.

In this work, we sought to create a rational framework to develop targeted cell therapies from heterogeneous progenitor populations for specific clinical diseases such as diabetes. Specifically, we hypothesized that single-cell transcriptional analyses could prospectively identify physiologically distinct progenitor cell subpopulations depleted in diabetes and with enhanced wound healing activity, based on the differences in individual cell gene expression distributions. Furthermore, the parallel assessment of intra-cellular and surface targets would enable subpopulation enrichment for therapeutic application by providing novel cell surface ‘recipes’. Importantly, this approach was designed to identify subpopulation-defining SMs comprehensively (by testing all 386 markers with commercially available antibodies) and blindly (assuming no mechanistic hypothesis). This comprehensive, blind approach greatly expands the potential SM pool and increases the likelihood of identifying subpopulations with robustly expressed markers to select cells.

**Results**

**Stem cell subpopulation and SM identification.** Utilizing human adipose-derived stem cells (hASCs) as a test progenitor cell pool, we first obtained a comprehensive profile of hASC SM expression through single-cell transcriptional analysis of ‘all’ known SMs with commercially available antibodies (Fig. 1a, Supplementary Data 1). This allowed us to cast the widest possible net in our search for novel subpopulation-defining markers without relying on a priori assumptions of gene expression. Using this approach, we identified over 200 markers that were expressed within hASCs. Focusing on the ~90 SMs with highest, non-uniform expression (which are most likely to distinguish biologically important cell subsets; Supplementary Fig. 1), we identified a distinct subpopulation of hASCs that was consistently present across multiple partial clustering permutations (Fig. 1b). This subpopulation could be defined with high sensitivity and specificity by two SM genes (DPP4 and CD55; Fig. 1c) and was present across multiple human patients (Supplementary Fig. 2a–d). To confirm that this was not an artefact, this transcriptionally defined subpopulation was recapitulated on a protein level (Supplementary Fig. 2a–b), and could be prospectively isolated based on the protein co-expression of DPP4 and CD55 using fluorescence-activated cell sorting (FACS) (Fig. 1d). To predict the function of this subpopulation, we first identified the targets overexpressed in our initial experiment (Supplementary Table 1) and evaluated their function based on the published literature. We found that these cells expressed increased levels of general stem cell markers (such as CD34 and CD73), as well as genes associated with cancer stem cells (CD99 and ITGB3) and embryonic stem cells (GGT1), suggesting that this subpopulation may have increased regenerative and wound healing potential.

**Stem cell subpopulation verification and in vitro characterization.** To confirm that these cells were functionally distinct from the parent pool, we performed repeat single-cell transcriptional analyses with a new gene list focused on intra-cellular targets thought to be important in wound healing (Supplementary Fig. 2e). We found that the identified cell subset, defined by and enriched via DPP4 and CD55 expression, displayed a transcriptional profile with increased expression of ‘intra-cellular’ genes related to cell survival, stemness, blood vessel growth and tissue remodelling (Fig. 1e–f, Supplementary Fig. 2f, 4b). To extend our results beyond the 96 genes explicitly analysed, canonical pathways overrepresented in these cells were determined using ingenuity pathway analysis (IPA) (Fig. 1g). These inferred pathways included multiple cardiovascular and wound healing processes, lending additional support to an enhanced wound healing potential for this discrete cell subset.

We next sought to confirm the functional importance of the DPP4/CD55 subpopulation. Starting in vitro we assessed cell survival, stemness, proliferation and colony-forming capacity, as these are idealized characteristics of effective cell therapies. As predicted, subpopulation enrichment led to enhancement in cell survival following exposure to an apoptotic stimulus (Fig. 2a,b, Supplementary Fig. 3), improved cell robustness as determined by proliferation and colony-forming capacity (Fig. 2c,d), and prolonged expression of cell stemness markers across extended passages (Fig. 2e).

**Disease effects on stem cell subpopulation dynamics.** We have previously catalogued the depletion of progenitor cell subsets from therapeutic cell populations in murine disease models, consistent with the hypothesis that cellular perturbations underlie disease-specific sequelae. Building on this work, we verified that...
Figure 1 | Single-cell transcriptional analysis identifies a subpopulation of human ASCs with putatively enhanced regenerative potential. (a) Single-cell transcriptional screening of all known cell SMs to identify those with differential expression (most useful for cell subtyping). Gene expression presented as fold change from median (yellow—high expression, 32-fold above median to blue—low expression, 32-fold below median; grey—no expression). (b) Single-cell analysis focused on high copy number, differentially distributed SM genes identified a cell subpopulation present across repeated k-means clusterings. (c) Linear discriminate analysis (LDA) identified SMs for prospective subpopulation isolation, with ROC analysis of cluster sensitivity and specificity utilizing the ‘best’ individual or groups of genes determined using forward feature selection. (d) Single-cell confirmation of prospective hASC subpopulation isolation via FACS using two LDA-defined SMs (DPP4 and CD55). (e) Positive hASC subpopulation enrichment enhances gene expression distributions for multiple genes related to tissue regeneration (selected significantly affected genes displayed as determined via Kolmogorov–Smirnov testing). (f) Single-cell whisker plots and pooled cell RT-PCR demonstrating a confirmation of selected single-cell gene distribution findings on a population level. (g) Top scoring IPA-constructed transcriptome network based on the genes significantly increased following positive hASC selection. Significant ‘seed’ genes are coloured in red to distinguish them from the remaining ‘inferred’ entities in the network. * indicates P ≤ 0.05 for positive selection versus hASCs or negative selection, via one-way ANOVA. Error bars represent s.e.m.
Figure 2 | Effect of prospective hASC selection and co-morbidities on ASC subpopulation dynamics to inform cell source decisions. (a,b) Enrichment for the transcriptionally identified hASC subpopulation enhances cell survival following exposure to an in vitro apoptotic stimulus (Fas ligand; measuring caspase activation (red)), (c,d) increases cell proliferation and clonogenicity and (e) prolongs stemness marker (CD34) expression. (f) The transcriptionally identified ASC subpopulation is significantly depleted and possesses deregulation of critical signalling pathways visible on single-cell analysis in the setting of both diabetes and aging. Gene expression presented as fold change from median (yellow—high expression, 32-fold above median to blue—low expression, 32-fold below median; grey—no expression). (h) Principal component projections of individual cells (left) and genes (right) demonstrating considerable segregation among phenotypes, driven largely by vascular/tissue remodelling genes. (i) Single-cell transcriptional analysis of healthy, aged and diabetic mASCs reveals that the depletion/dysfunction of cluster 1 cells in these states is not a result of cell SM loss and redistribution to other clusters (expression profiles of subpopulation-defining SMs and tissue remodelling genes highlighted). (j) Flow cytometric analysis demonstrating dynamic DPP4/CD55 subpopulation increases in wild-type wounds, supporting their role in the wound healing process. The DPP4/CD55 subpopulation was also elevated in diabetic and aged wounds as compared with uninjured skin, with a trend toward compensatory overrecruitment consistent with an impaired cellular functionality. * indicates $P \leq 0.05$ via one-way ANOVA or Student’s t-test (healthy versus aged or diabetic in f, day 7 versus respective controls in j). † indicates $P \leq 0.05$ for positive versus negative selection via Student’s t-test. Error bars represent s.e.m. Scale bar, 50 μm.
this DPP4/CD55 subpopulation was abundantly present in the adipose tissue of healthy mice (Supplementary Fig. 4). However, we found that both diabetes and aging were associated with decreases in the number of these cells in adipose tissue (Fig. 2f), as well as deregulation of critical cellular wound healing signalling pathways (Fig. 2g, h). This cellular dysfunction was not explained by SM drift (Fig. 2i), and was confirmed in humans in the setting of diabetes (Supplementary Fig. 5a). These data support the predicted based on the transcriptional and physiologic wound healing and supporting the clinical therapeutic potential of subpopulation-enriched, autologous or allogeneic ASC-based therapies in the setting of impaired murine diabetic wounds. As guided by these findings, we evaluated the in vivo therapeutic potential of subpopulation-enriched, allogeneic ASC-based therapies in the setting of impaired murine diabetic wounds. As predicted based on the transcriptional and in vitro enhancements of the identified ASC subpopulation, and its dysregulation in the diabetic state, a single application of FACS-enriched healthy ASCs accelerated wound closure rates and improved dermal regeneration (Fig. 2). However, trends toward compensatory overrecruitment were observed in these pathologic states, consistent with an impaired or deleterious cellular functionality. These data support a diminished therapeutic potential of allogeneic cells in these pathologic settings, even following subpopulation enrichment, and suggest that allogeneic cell pools may be more beneficial in patients with underlying disease.

**Therapeutic efficacy of stem cell subpopulation enrichment.** Guided by these findings, we evaluated the in vivo therapeutic potential of subpopulation-enriched, allogeneic ASC-based therapies in the setting of impaired murine diabetic wounds. As predicted based on the transcriptional and in vitro enhancements of the identified ASC subpopulation, and its dysregulation in the diabetic state, a single application of FACS-enriched healthy ASCs accelerated wound closure rates and improved dermal regeneration (Fig. 2). However, trends toward compensatory overrecruitment were observed in these pathologic states, consistent with an impaired or deleterious cellular functionality. These data support a diminished therapeutic potential of allogeneic cells in these pathologic settings, even following subpopulation enrichment, and suggest that allogeneic cell pools may be more beneficial in patients with underlying disease.

**Therapeutic efficacy of stem cell subpopulation enrichment.** Guided by these findings, we evaluated the in vivo therapeutic potential of subpopulation-enriched, allogeneic ASC-based therapies in the setting of impaired murine diabetic wounds. As predicted based on the transcriptional and in vitro enhancements of the identified ASC subpopulation, and its dysregulation in the diabetic state, a single application of FACS-enriched healthy ASCs accelerated wound closure rates and improved dermal regeneration (Fig. 2). However, trends toward compensatory overrecruitment were observed in these pathologic states, consistent with an impaired or deleterious cellular functionality. These data support a diminished therapeutic potential of allogeneic cells in these pathologic settings, even following subpopulation enrichment, and suggest that allogeneic cell pools may be more beneficial in patients with underlying disease.
regeneration compared with application of sorted or depleted cells, effectively ‘normalizing’ diabetic healing to wild-type kinetics (Fig. 3a–d). These cells persisted for up to 16 days in situ (Supplementary Fig. 5b), likely acting via IPA-predicted cytokine-mediated improvements in local wound healing pathways, and the upregulation of fibroblast collagen production (Fig. 3e,g)21. In the field of wound healing, very few interventions are capable of ‘normalizing’ the wound healing timeline, demonstrating the power of this approach. Importantly, depletion of this ASC subpopulation completely abrogated the beneficial effects of cell therapy (Fig. 3a–e), and enriched medium did not have the same beneficial wound healing effect as a single direct cell subpopulation application (Fig. 3f). Consistent with our single-cell data demonstrating impairment of diabetic ASCs across subpopulations, enrichment of diabetic ASCs did not restore their wound healing capacity (Fig. 3g, Supplementary Fig. 6a). These data suggest that healthy, enriched ASCs are both necessary and sufficient to maximally improve wound healing using this approach.

Regarding the safety of ASC application, one-time dosing was used to limit allogeneicity, and cell applications were tolerated without local or systemic signs of rejection. Moreover, treated wounds demonstrated considerable vascularity on closure (Supplementary Fig. 6b), with no evidence of tumour formation or wound breakdown 3 months after application. Further increasing the clinical potential of these findings, rapid subpopulation enrichment of hASCs was also possible via magnetic-assisted cell sorting (MACS) for the cluster-discriminating SmDs PFP4 and CD55 (Supplementary Fig. 6c), enabling a total processing time (from tissue collection to cell application) on the scale of an hour. These data demonstrate the potential utility of targeted cell subpopulation enrichment for cell-based therapeutics utilizing a process that is adaptable to virtually any pathologic state.

Discussion

This work establishes the efficacy of single-cell analysis for the rational enhancement of cell-based therapeutics by addressing pathologic alterations in progenitor cell biology. Importantly, this methodology overcomes the inherent challenges progenitor cell heterogeneity imposes on our understanding of pathologic cellular perturbations and the standardization of cell-based approaches, and enables the logical assessment and development of targeted cell therapies addressing an underlying cause of specific clinical defects. Importantly, by making no a priori assumptions of cell SM expression, this approach is adaptable to any cell population. It also definitively informs decisions regarding the utility of autologous versus allogeneic cell sources.

hASCs were the target population assessed here due to their therapeutic potential, limited immunogenecity and safety profile33–37. Our findings build on our previous work demonstrating ASC subpopulation perturbations in diabetes37, critically identifying novel SmDs for prospective subpopulation isolation, testing and therapeutic application. While there are currently no FDA-approved progenitor cell treatments for superficial wounds, multiple products containing mesenchymal progenitors are in early stages of clinical testing8. Our data support the safety and effectiveness of hASC therapies for the treatment of diabetic wounds, and suggest that sub-fractionation of other progenitor cell populations may similarly enhance their therapeutic potential in this setting. On the basis of these findings, future direct comparisons with other treatment modalities are warranted. This experimental framework did not seek to address the underlying defects of diabetes. We nonetheless envision that a similar methodology would be beneficial to more curative cell therapies, such as pancreatic islet cell transplantation. Given the adaptability of this approach to any cell type, the authors feel this technique has the potential to standardize and improve cell-based therapies for any disease state.

Methods

hASC isolation. Human abdominoplasty specimens were obtained after acquiring informed consent from patients, in accordance with the Stanford University Institutional Review Board guidelines. For initial experiments, ASCs were collected from the tissue samples of multiple adult female patients without major medical conditions who were undergoing elective abdominoplasty procedures. For experiments on the effect of diabetes on hASC subpopulations, ASCs were isolated from consecutive patients (n = 3) undergoing elective bariatric surgery who met predefined criteria for diabetes mellitus (haemoglobin-A1c (HbA1c) > 6.5). Controls for these experiments were adult patients without major medical conditions undergoing elective abdominoplasty (n = 4) without a history of diabetes mellitus or HbA1c ≥ 6.0. Human ASCs from all groups were isolated based on an established protocol38. Raw human abdominoplasty specimens were manually minced, washed and treated with 0.075% collagenase type I (Sigma-Aldrich, St Louis, MO) in Hank’s balanced salt solution (Life Technologies, Grand Island, NY) for 1 h at 37 °C with gentle agitation. The reaction was stopped with the addition of fetal bovine serum (FBS), and after centrifugation, the pelleted stromal vascular fraction was prepared for FACS as described below.

Animals. Young (3 months) and aged (21 months) wild-type mice (C57BL/6), and young diabetic mice (db/db; BKS.Cg-Dock7m / Leprdb/l) were obtained from Jackson Laboratories (Bar Harbor, ME) and the National Institute on Aging (NIA, Bethesda, MD). Luciferase / GFP mice (FVB-Tg(CAG-luc-GFP)P2G8Scilo/J) were also obtained Jackson laboratories. All protocols were approved by the Stanford Administrative Panel on Laboratory Animal Care.

mASC isolation. Murine ASCs (mASCs) were isolated from young, aged, diabetic and luciferase / GFP mice (FVB-Tg(CAG-luc-GFP)P2G8Scilo/J) were also obtained Jackson laboratories. All protocols were approved by the Stanford Administrative Panel on Laboratory Animal Care.

mFACS. Fresh human or mASCs were sorted on a FACS Aria II instrument (BD Biosciences, San Jose, CA) with the use of a 100-μm nozzle. Cells were isolated as described above, and incubated for 20 min in FACS buffer (phosphate-buffered saline (PBS) supplemented with 2% FBS) containing one of the following antibody combinations: (1) anti-human eGFP- conjugated CDA5 (Bioscience, San Diego, CA), APC- or PE-conjugated CD34 (BD Biosciences), FITC-conjugated CD31 (BD Biosciences), PE- or APC-conjugated DPP4 (BD Biosciences) and PE-Cy7-conjugated CD55 (Bioscience, San Diego, CA); (2) anti-mouse eGFP-conjugated CD45 (eBioscience), APC-conjugated CD34 (Bioscience), PE-Cy7-conjugated CD55 (Bioscience), FITC-conjugated DPP4 (BD Biosciences) and PE-conjugated CD31 (Bioscience); (3) anti-mouse PE-Cy7-conjugated CD55 (Bioscience), APC-conjugated CD34 (Bioscience), eGFP-conjugated CD31 (Bioscience) and PE-Cy7-conjugated CD55 (Bioscience). Using a Becton Dickinson flow cytometric cell sorter, cells were either sorted as single cells into 6 μl of lysis buffer for single-cell transcriptional analysis, or as populations for subsequent culture. ASCs were defined with the SM profile CDA5- /CD31- /CD55- (to exclude contaminating hematopoietic and endothelial cells found within the SVF), and CDA5 and DPP4 expression within this ASC population was used for positive and negative subpopulation selection (see Supplementary Fig. 2ab for gating scheme). Cells sorted for culture were plated onto conventional tissue culture plates in the DMEM (Life Technologies) supplemented with 10% FBS and 1% P/S. Plated cells were cultured under standard conditions (37 °C, 5% CO2) and used at or before passage two. In vitro assays were run in triplicate unless otherwise stated.

In vitro survival assays. hASC survival was assessed using two methods following exposure to the apoptotic stimulus Fas ligand (1, 10 or 100 ng per ml for 5 h; Human Recombinant Super FasL, Enzo Life Sciences, Farmingdale, NY). Apoptotic cell cascades were assessed using the CaspaTag Caspase 3–7 kit (Millipore, Billerica, MA) according to the manufacturer’s instructions. Briefly, cells were seeded on a chamber slide, and following exposure to FasL, cells were washed and incubated for 1 h at 37 °C with the CaspaTag reagent solution. After washing, red fluorescence was captured using a Leica DM5000 microscope (Leica Microsystems, Heidelberg, Germany) equipped with a DFC500FX camera. Hoechst staining was used to label nuclei, and untreated cells were used as control. Fluorescence intensity (595 nm) was analysed using Image J software (NIH, Bethesda, MD). Brightfield images were also obtained for assessment of apoptotic cell morphology.

Downstream annexin V activation was also assessed following FasL exposure. Using the Mitochondrial Membrane Potential/Apoptosis Kit (Invitrogen, Carlsbad, CA), according to the manufacturer’s instructions. Following exposure to the FasL, cells were washed, lifted using trypsin and...
incubated with MitoTracker Red dye for 30 min at 37 °C and 5% CO2. Cells were then washed, incubated in Alexa Fluor 488 annexin V for 15 min, and analysed via flow cytometry (FACS Aria II instrument; BD Biosciences). Untreated cells were used as negative controls.

**Cell proliferation.** HASC proliferation was assessed following FACS sorting using a BrdU proliferation assay. Sorted hASCs (2 × 10^5) were seeded in 96-well plates under standard culture conditions (DMEM with 10% FBS, 37 °C and 5% CO2). After overnight adherence, the cells were visualised using a differential interference contrast microscope. The effect of DPP4/CD55 conditioned medium on fibroblast collagen production was assessed following FACS sorting of healthy and diabetic ASCs. Sorted mASCs (2 × 10^5) were seeded in 24-well plates under standard culture conditions (DMEM with 10% FBS, 37 °C and 5% CO2). After reaching 90% confluence, the media was switched to DMEM containing 1% FBS. The effect of DPP4/CD55 ASC conditioned medium on diabetic wounds was assessed using a MultiSort Kit (Miltenyi Biotec, San Diego, CA). Freshly isolated human SVF was stained with Biotin-conjugated anti-human CD55 (Biolegend) for 20 min on ice, washed, stained with anti-Biotin MicroBeads and enriched/depleted for DPP4/CD55 positive, negative and control mASCs using a MACS Separation Column (Miltenyi Biotec). Sorted cells and unsorted controls were stained with FACS antibodies as above for subsequent analysis.

**Gene list generation.** The approach to selection of ~ 90 SM genes for initial subpopulation analysis was one of exclusion, with a focus on SMs with the highest potential for protein based sub-fractionation. Specifically, roughly 400 candidate gene targets with known commercially available antibodies were initially screened via single-cell analysis (Supplementary Data 1). Increasingly restrictive filters were then used to remove targets that would be unlikely to yield FACS-separable subpopulations (Supplemental Fig. 1). The approach excluded genes not to be expressed in our target population, followed by those expressed in > 95% or < 5% of these cells. Of the remaining genes, those with the lowest median channel number were selected. To exclude genes with identical median channel number, we included the initial SM gene list (Supplemental Table 2, Supplemental Fig. 7a). Subsequent human and murine gene lists were then generated using cluster
defining genes (Fig. 1b) and intra-cellular targets thought to be important in wound healing (Supplementary Figs 2e,7c,e; Supplementary Tables 3,4).

**Statistical analysis.** Results are presented as mean ± s.e.m. Standard data analysis was performed using a Student’s t-test or one-way analysis of variance (ANOVA), with subsequent comparisons between individual methods completed using a Tukey’s post hoc analysis. Wound healing curves were assessed at each time point using one- or two-way ANOVA (historical wild-type murine wound healing curve overlaid for visual comparison). Results were considered significant for P < 0.05.

Analysis of single-cell data was performed using an established protocol. Expression data from all chips for each experiment were normalized to the median expression of each gene in the pooled sample, before being converted to base 2 logarithmically. Absolute threshold was defined as ± 5 cycle thresholds from the median (or 32-fold increases/decreases in expression) in the non-treated and non-expressing data points were allocated to this floor. To aid data visualization, colour-coded clustergrams were then produced using hierarchical clustering, with a ‘complete’ linkage function and Euclidean distance metric (MATLAB, R2011b, MathWorks, Natick, MA).

To identify subpopulations within this single-cell transcriptional data, k-means clustering was used with a standard Euclidean distance metric, with each cell assigned membership to each cluster as dictated by similarities in gene expression distributions (decreasing the within-cluster sum of square distances in the 96-dimensional gene hyperspace) in MATLAB, and clustering was repeated for k = 2 through 5. Optimally partitioned clusters were then sub-grouped using hierarchical clustering to facilitate visualization of data patterning within and across these clusters.

Two-sample Kolmogorov–Smirnov tests were used to identify genes with significantly different distribution patterns between population clusters and/or groups, using a cutoff value of P < 0.05 with Bonferroni correction for multiple samples. For comparisons among subgroups, the empirical distribution of cells from any given cluster was evaluated against the distribution of all the remaining cells in the experiment.

To identify those SM gene combinations best able to distinguish each cluster, forward feature selection was used using linear discriminate analysis with fivefold cross-validation. The resulting receiver operating characteristic curves were constructed to compare the sensitivity and specificity of each gene set in discriminating the cluster of interest and areas under the curve (AUC) calculated. Selection of the n = 2 gene model (CD55 and DPP4) was made subjectively based on the breakdowns in the distribution of associated AUC values.

To construct transcriptome networks based on the genes significantly increased following positive selection for the SM proteins CD55 and DPP4, IPA (Ingenuity Systems, Redwood City, CA) was used. In this analysis, the 96 genes from the corresponding single-cell analysis (as opposed to the entire transcriptome) were defined as the reference set (that is, possibility space). This approach was taken to prevent bias of enrichment calculations in IPA’s internal algorithm.

**Data availability.** The data that support the findings of this study are available from the corresponding author on request.

**References**

1. Rennert, R. C. et al. Biological therapies for the treatment of cutaneous wounds: phase III and launched therapies. Expert Opin. Biol. Ther. 13, 1523–1541 (2013).
2. Calafiore, R., Montanucci, P. & Basta, G. Stem cells for pancreatic beta-cell repair. Stem Cells Transl. Med. 3, 30–34 (2014).
3. ClinicalTrials.gov. Bethesda: US National Library of Medicine. Available from: https://clinicaltrials.gov. Accessed March 4th, 2016.
4. Lin, C. S., Xin, Z. C., Dui, J. & Lue, T. F. Commonly used mesenchymal stem cell markers and tracking label limitations and challenges. Histol. Histopathol. 28, 1109–1116 (2013).
5. Graf, T. & Stadtfeld, M. Heterogeneity of embryonic and adult stem cells. Stem Cell Rev. 7, 309–322 (2011).
6. Rennert, R. C. et al. Isolation of human adipose-derived stromal cells using laser-assisted liposuction and their therapeutic potential in regenerative medicine. Stem Cell Transl. Med. 2, 808–817 (2013).
7. Vilaíta, M. et al. Biodistribution, long-term survival, and safety of human adipose tissue-derived mesenchymal stem cells transplanted in nude mice by high sensitivity non-invasive bioluminescence imaging. Stem Cell Dev. 17, 993–1003 (2008).
8. Rennert, R. C. et al. Enhancement of mesenchymal stem cell angiogenic capacity and stemness by a biomimetic hydrogel scaffold. Biomaterials 33, 80–90 (2012).
9. Garg, R. K. et al. Capillary force seeding of hydrogels for adipose-derived stem cell delivery in wounds. Stem Cell Transl. Med. 3, 1079–1089 (2014).
10. Wang, V. W. et al. Focal adhesion and microtubule mechanical force to skin fibrosis via inflammatory signaling. Nat. Med. 18, 148–152 (2012).
Acknowledgements
We thank Yujin Park for her assistance with tissue processing and staining. Flow cytometric analyses for this manuscript were completed at the Stanford Shared FACS Facility. Funding for the stem cell research performed in our laboratory has been provided by the Hagey Family Endowed Fund in Stem Cell Research and Regenerative Medicine, the Armed Forces Institute of Regenerative Medicine (United States Department of Defense), the National Institutes of Health (R01-DK074095, R01-EB005718, R01-AG025016 and #LM0077033) and the Oak Foundation.

Author contributions
All authors contributed to the idea generation, design and completion of this work. R.C.R., M.J. and M.S. contributed equally to the idea generation, experimental work and manuscript preparation. M.R., Z.N.M., D.D., A.J.W., R.K., M.T.C., K.P., A.Y.L., M.F., J.P.G. and A.J.B. contributed to the experimental work and manuscript preparation. G.C.G. guided the idea generation, experimental work and manuscript preparation.

Additional information
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Rennert, R. C. et al. Microfluidic single-cell transcriptional analysis rationally identifies novel surface marker profiles to enhance cell-based therapies. Nat. Commun. 7:11945 doi: 10.1038/ncomms11945 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/