LETTER

https://doi.org/10.1038/s41586-018-0226-8

A stromal cell population that inhibits adipogenesis in mammalian fat depots

Petra C. Schwalie[1,5], Hua Dong[2,5], Magda Zachara[1,5], Julie Russeil[1], Daniel Alpern[1], Nassila Akchiche[2], Christian Caprara[3], Wenfei Sun[2], Kai-Uwe Schlaudraff[4], Gianni Soldati[3], Christian Wolfrum[2]* & Bart Deplancke[1]*

Adipocyte development and differentiation have an important role in the aetiology of obesity and its co-morbidities[1,2]. Although multiple studies have investigated the adipogenic stem and precursor cells that give rise to mature adipocytes[3–14], our understanding of their in vivo origin and properties is incomplete[2,15,16]. This is partially due to the highly heterogeneous and unstructured nature of adipose tissue depots[17], which has proven difficult to molecularly dissect using classical approaches such as fluorescence-activated cell sorting and Cre–lox lines based on candidate marker genes[16,18]. Here, using the resolving power of single-cell transcriptomics[19] in a mouse model, we reveal distinct subpopulations of adipose stem and precursor cells in the stromal vascular fraction of subcutaneous adipose tissue. We identify one of these subpopulations as CD142$^+$ adipogenesis-regulatory cells, which can suppress adipocyte formation in vivo and in vitro in a paracrine manner. We show that adipogenesis-regulatory cells are refractory to adipogenesis and that they are functionally conserved in humans. Our findings point to a potentially critical role for adipogenesis-regulatory cells in modulating adipose tissue plasticity, which is linked to metabolic control, differential insulin sensitivity and type 2 diabetes.

To study the molecular characteristics and the subpopulation structure of adipogenic stem and precursor cells (ASPCs), we performed single-cell RNA sequencing (scRNA-seq) on Lin$^-$(CD31$^-$CD45$^-$TER119$^-$)CD29$^+$CD34$^+$SCA1$^+$ cells from the subcutaneous stromal vascular fraction (SVF) of transgenic mice, in which red fluorescent protein is induced in cells that express *Dlk1* (Fig. 1a and Methods). CD29, CD34 and SCA1 are expected to enrich for stem cells, and DLK1 has previously been suggested to mark pre-adipocytes[20,21]. We obtained 208 high-quality cells (Fig. 1a, Extended Data Fig. 1a–d and Methods) and confirmed that the expression of all fluorescence-activated cell sorting (FACS) markers corresponded to the enrichment strategy that was used (Extended Data Fig. 1e). The cells grouped into three clusters, referred to as populations (P1, P2 and P3) (Fig. 1b, Supplementary Table 1 and Methods), each of which was characterized by the specific expression of hundreds of genes (Fig. 1c and Supplementary Table 2). Notably, only P2-specific genes enriched for adipogenesis-related functionality, including the adipogenic master regulator *Pparg* (Fig. 1c, Extended Data Fig. 1f, g and Supplementary Tables 2, 3). We further assessed the individual and combined expression of commonly used lineage marker genes (Supplementary Table 4 and Methods). We found that only *Fabp4*—which is associated with adipogenesis—was significantly differentially expressed across cells, showing high expression in P2 (Extended Data Fig. 1h and Supplementary Table 19). Consistently, P2 cells scored significantly higher on the adipogenic scale than did P1 or P3 cells (Extended Data Fig. 1i, Supplementary Table 19 and Methods). Virtually all cells scored high on the stem cell scale, as also exemplified by *Cd34* expression (Extended Data Fig. 1h, j and Supplementary Table 19), consistent with the selection of an ASPC-enriched cell pool. Finally, among genes previously used to mark pre-adipogenic populations[18], only *Pparg* and *Prdm16* expression was significantly correlated with *Fabp4* expression and enriched in P2 cells (Extended Data Fig. 1k, l and Supplementary Tables 5, 19).

To independently validate the three cell populations we detected, we assessed a larger number of Lin$^-$ SVF cells using the 10x Genomics Chromium platform. We considered 1,804 high-quality single cells (Fig. 1d, Extended Data Fig. 2a, b and Methods) and found that most of them reflected an ASPC state, on the basis of their high *Itgb1* (encoding CD29), *Cd34* and *Ly6a* (encoding SCA1) expression (Extended Data Fig. 2c). Consistently, mature adipocyte markers such as *Adipoq*, *Retn* and *Cidec* were virtually absent. We found that despite differences in cell stratification, capture and molecular assessment, these data largely recapitulated our initial findings (Fig. 1e, f, Supplementary Table 6 and Methods). The four clusters (G1–G4) that we detected could all be mapped back to the three previously described populations: G1 and G4 together approximately corresponded to P1, G2 to P2, and G3 to P3, with over 30% of the top 100 markers overlapping (Fig. 1f, Extended Data Fig. 2d, e, Supplementary Table 7 and Methods). This included the top-ranking transcription factors *Creb5* (in P1 and G1), *Peg3* (in P2 and G2) as well as *Meox2* (in P3 and G3). Consistently, G2-specific genes enriched for pathways related to adipogenesis (Supplementary Table 8), scored significantly higher on the adipogenic scale (Extended Data Fig. 2f and Supplementary Table 19) and included *Fabp4* (Fig. 1g, h and Extended Data Fig. 2g). *Pparg* was again strongly positively correlated with *Fabp4*, in contrast to most of the other (pre-)adipogenic markers with the notable exception of *Pdgfrb* (Fig. 1h, Extended Data Fig. 2g, h and Supplementary Tables 9, 19). *Cd34* and *Ly6a* were expressed at higher levels in G1 and G4 cells, and were significantly negatively correlated with *Pdgfrb*, *Fabp4* and *Pparg* expression (Fig. 1h and Supplementary Tables 9, 19). In summary, by using two distinct scRNA-seq-based approaches, we identified at least three subpopulations among Lin$^-$ SVF cells. The two major ones lie at opposite poles of stem cell-specific (*Cd34* and *Ly6a)* and pre-adipogenic (*Fabp4*, *Pparg* and *Cd36)* gene expression. These two subpopulations make up over 90% of the cells (Extended Data Fig. 2d) and can be further subclassified in increasingly homogenous cell groups, all of which are molecularly separated from a third and unrelated subpopulation (G3 or P3).

To assign adipogenic functionality to the three major subpopulations that we detected, we identified marker genes that encode surface proteins: *Cd55* and *Il13ra1* for P1 (and G1 and G4), *Aoc3* (encoding VAP1) and *Adam12* for P2 (and G2), and *F3* (encoding CD142) and *Abcg1* for P3 (and G3) (Fig. 2a and Extended Data Fig. 3a). Using FACS (Extended Data Fig. 3b), we isolated cellular fractions that were enriched for these respective markers and validated each subpopulation by quantitative PCR (Fig. 2b, Extended Data Fig. 3c and Supplementary Table 10). We then quantified lipid accumulation in response to a white fat differentiation cocktail for all ASPCs, marker-positive ASPCs and marker-negative ASPCs (Methods). We did not observe a difference in adipogenesis for P2 cells, whereas P1 cells showed an increased propensity and P3 cells a strongly decreased propensity to form adipocytes in

[1]Laboratory of Systems Biology and Genetics, Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL) and Swiss Institute of Bioinformatics, Lausanne, Switzerland. [2]Institute of Food Nutrition and Health, Eidgenössische Technische Hochschule Zürich, Schwerzenbach, Switzerland. [3]Swiss Stem Cell Foundation, Gentilino, Switzerland. [4]Concept-Clinic, Geneva, Switzerland. [5]These authors contributed equally: Petra C. Schwalie, Hua Dong, Magda Zachara. *e-mail: christian-wolfrum@ethz.ch; bart.deplancke@epfl.ch

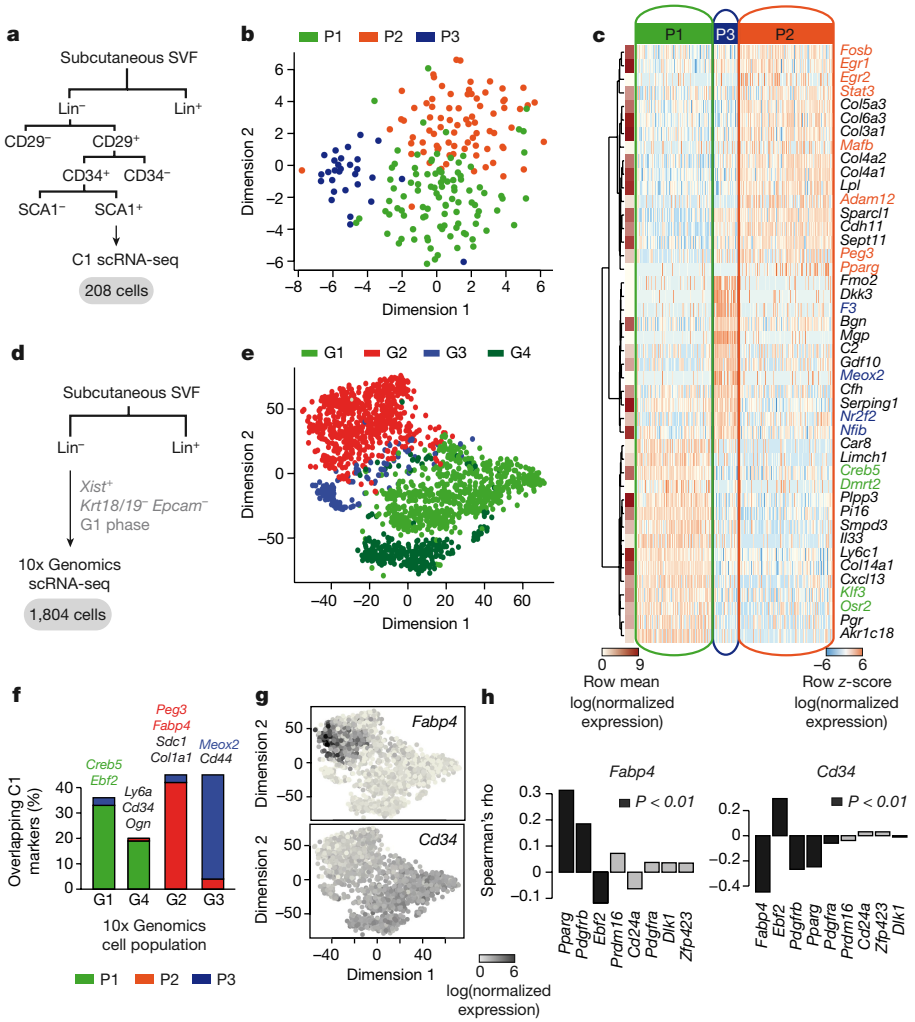


Fig. 1 | ScRNA-seq reveals the heterogeneity of subcutaneous ASPCs in the mouse. a, Schematic of FACS-based selection of ASPCs (Fluidigm C1). **b**, *t*-distributed stochastic neighbour embedding (*t*-SNE) 2D cell map (C1). $n = 208$ cells total, 3 biological replicates; populations P1, $n = 83$; P2, $n = 96$; P3, $n = 29$ cells. **c**, Heat map (blue-to-red) of expression of top 10 population markers and transcription factors among the top 100 population markers (in colour); hierarchically clustered rows, mean gene expression (white-to-red, left). $n = 208$ cells, 3 biological replicates. **d**, Schematic of FACS-based selection of Lin$^-$ SVF cells (10x Genomics). **e**, *t*-SNE 2D cell map (10x Genomics); populations G1, $n = 699$; G2, $n = 664$; G3, $n = 122$; and G4, $n = 319$ cells. **f**, Percentage of the top 100 10x Genomics population markers that overlap with the top 100 C1 population markers; transcription factors and lineage markers are highlighted. **g**, *t*-SNE 2D cell map (10x Genomics) highlighting the expression of an adipogenic (*Fabp4*) and a stem cell (*Cd34*) marker (black). **h**, Correlation (Spearman's rho) between the expression of (pre-)adipogenic markers and *Fabp4* (left) or *Cd34* (right). In **e**–**h**, $n = 1{,}804$ cells, 1 biological replicate. See Extended Data Figs. 1, 2 and 8 for related results.

response to the same stimulus (Fig. 2c, d and Extended Data Fig. 3d, e). Consistently, we measured significantly higher levels of expression of adipogenic markers in P1 cells compared to ASPCs depleted of P1 cells, and the reverse was true for P3 cells (Extended Data Fig. 3f and Supplementary Tables 10, 19). None of the populations exhibited differences in nuclei number, suggesting that the adipogenic differences we observed were not due to distinct proliferative properties (Extended Data Fig. 3g and Supplementary Table 19).

CD142$^-$ABCG1$^-$ and CD142$^-$ ASPCs (hereafter CD142$^-$(ABCG1$^-$) ASPCs) showed markedly higher differentiation compared to all ASPCs, which suggests that P3 cells inhibit adipogenesis (Fig. 2c, d and Extended Data Fig. 3d–f). To test this hypothesis, we performed a titration experiment in which CD142$^+$ABCG1$^+$ ASPCs were mixed with CD142$^-$ABCG1$^-$ ASPCs in ratios ranging from 0 to 100 per cent. We observed a nonlinear relation that showed a greater decrease in adipogenic differentiation than expected on the basis of a simple dilution effect (Fig. 2e, f). Again, we did not observe significant differences in nuclei numbers, which suggests that proliferation remained unaltered (Extended Data Fig. 4a and Supplementary Table 19). Thus, a fraction of ASPCs that is characterized by high CD142 and ABCG1 expression is refractory to adipogenesis and negatively regulates the adipogenic capacity of other ASPCs. A negatively modulatory function such as this is conceptually reminiscent of T regulatory cells or 'Tregs' that use their immunosuppressive capacity to maintain immune homeostasis and mediate peripheral tolerance[22]; therefore, we suggest that these 'adipogenesis-regulatory' cells be named 'Aregs'.

To characterize Aregs, we profiled the transcriptomes of all ASPCs, CD142$^+$ ASPCs and CD142$^-$ ASPCs at four time points: ex vivo, upon culture (after 5 and 24 h) and after adipogenic differentiation (Extended Data Fig. 4b–f, Supplementary Table 11 and Methods). This transcriptomic data further molecularly confirmed the non-adipogenic character of Aregs. Genes that have previously been associated[2] with adipogenesis—such as *Pparg*, *Fasn*, *Fabp4*, *Lpl* and *Fabp12*—were expressed at significantly lower levels in Aregs that were subjected to differentiation, when compared to both all ASPCs and to CD142$^-$ ASPCs (Extended Data Fig. 4g–i and Supplementary Tables 11, 12, 19). The majority (64% of the differentially expressed genes) of Areg-specific expression patterns were maintained upon plating and only a further 20% of the differentially expressed genes changed their expression after culture for another day (Extended Data Fig. 4j). The genes that maintained their expression included those encoding the two surface markers used for the isolation of Aregs (*F3* and *Abcg1*),

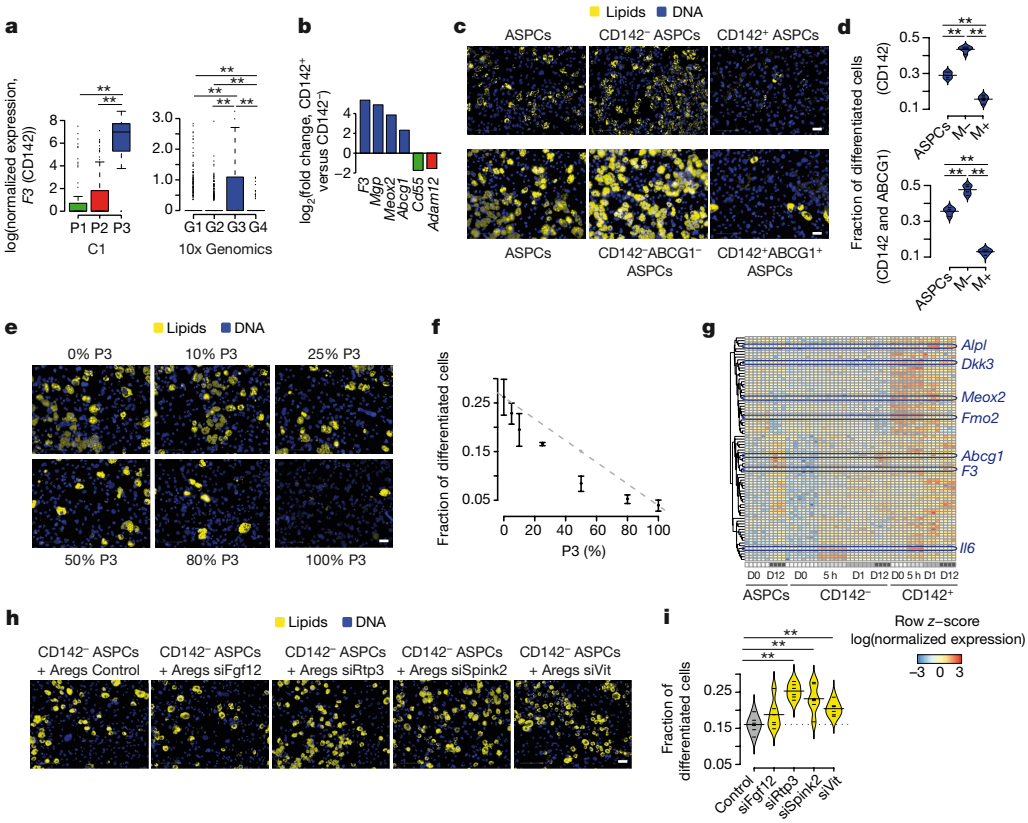


Fig. 2 | Aregs (CD142$^+$ABCG1$^+$ ASPCs) show a paracrine adipogenic inhibitory capacity. a, Expression of the P3 and G3 FACS marker *F3* (CD142) assessed using scRNA-seq; P1, $n = 83$ (green); P2, $n = 96$ (red); P3, $n = 29$ (blue) cells; G1, $n = 699$; G2, $n = 664$; G3, $n = 122$; and G4, $n = 319$ cells. $^{**}P \leq 0.01$, Wilcoxon rank-sum test. **b**, Fold-changes (CD142$^+$ versus CD142$^-$ ASPCs) of population-specific genes assessed using qPCR. **c**, Microscopy images of ASPC fractions after adipogenesis. **d**, Fraction of differentiated cells per ASPC type shown in **c**, $n = 4$–5 independent wells. M−, marker-negative, M+, marker-positive. In **c**, **d**, experiments were repeated at least three times, yielding similar results. **e**, Microscopy images of distinct ratios of Lin$^-$SCA1$^+$CD142$^+$ABCG1$^+$ (P3) and Lin$^-$SCA1$^+$CD142$^-$ABCG1$^-$ cells after adipogenesis. **f**, Mean and s.d. of the percentage of differentiated cells per titrated fraction shown in **e**; $n = 4$ independent wells. **g**, Heat map (blue-to-red) of expression of genes that show significantly higher expression levels in CD142$^+$ than in CD142$^-$ ASPCs after sorting (day 0, D0), plating (5 h) and culture (day 1, D1); expression levels upon adipogenesis (day 12, D12) are also included; lineage (osteogenic *Alpl* and inflammatory *Il6*) and genes among the top 10 scRNA-seq-derived markers (Fig. 1c) are highlighted; $n = 4$–8, 4 biological replicates, 1–3 independent wells each. **h**, Microscopy images after adipogenesis of Areg-depleted ASPCs co-cultured with Aregs with *Fgf12*, *Rtp3*, *Spink2*, *Vit* and control siRNA-mediated knockdown. **i**, Fraction of differentiated cells measured in the CD142$^-$ ASPCs underneath Aregs with the knockdowns shown in **h**. In **e**, **f**, **h**, **i**, experiments were repeated two times, yielding similar results. In **h**, **i**, $n = 6$, 2 biological replicates, 3 independent wells each. In all panels, nuclei are stained with DAPI (blue) and lipids are stained with LD540 (yellow). Scale bars, 50 μm (**d**, **i**). $^*P \leq 0.05$, $^{**}P \leq 0.01$, *t*-test. See Extended Data Figs. 3–5 for related results.

the early osteogenic marker *Alpl* and the inflammation-associated cytokine *Il6* (Fig. 2g). Forty-seven genes retained their specific expression across plating, culture and adipogenic differentiation, including those genes that encoded the transcription factor MEOX2, the secreted factor DKK3 and the enzyme FMO2, all among the top genes associated with P3 in the scRNA-seq data (Figs 1c, 2g). Consistent with the scRNA-seq results (Extended Data Fig. 1f, g), we found that genes that were expressed at high levels in Aregs immediately after sorting are characteristic of blood vessels and enrich for the 'Hedgehog signalling' pathway (Extended Data Fig. 4k–m and Supplementary Tables 3, 12). However, we did not observe consistently higher expression of endothelial marker genes in Aregs (Extended Data Fig. 4n).

We further investigated the mechanism by which Aregs exert their adipogenesis-suppressive effect by testing whether the inhibition requires direct cell–cell contact. Transwell experiments in which ASPCs, Aregs or Areg-depleted ASPCs were inserted above ASPCs revealed that adipogenesis differences were maintained in the absence of direct contact (Extended Data Fig. 5a, b), which suggests a paracrine signalling mechanism. To identify genes that may mediate this signal, we performed a loss-of-function screen targeting 23 genes that showed high expression in Aregs compared to Areg-depleted ASPCs and were expressed only at low levels in mature adipocytes (Supplementary Table 13 and Methods). We subjected plated SVF cells to short interfering RNA (siRNA)-mediated reduction in expression of each of these genes, reasoning that knockdown would be effective only in Aregs. Thus, any changes in differentiation would be attributable to alterations in Areg activity. We found that the knockdown of *Rtp3*, *Spink2*, *Fgf12* and *Vit* significantly increased the differentiation of plated SVF cells, but not cell number (Extended Data Fig. 5c–f and Supplementary Table 19), which suggests that these genes modulate the inhibitory potential of Aregs. To validate these findings, we used transwells comparing the anti-adipogenic effect of wild-type Aregs to the anti-adipogenic effect in Aregs in which each of these genes was knocked down. Loss of *Rtp3* and, to a lesser extent, *Spink2* and *Vit* lowered the paracrine inhibitory potential of Aregs (Fig. 2h, i and Extended Data Fig. 5g–j). Our findings therefore support a paracrine mechanism of action, in which these genes may act upstream of one or several factors secreted by Aregs or, in the case of *Spink2*, as a direct signalling molecule.

To investigate whether Aregs constitute a conserved cell population that functions similarly in humans, we used a FACS isolation strategy based on CD142 that was analogous to the strategy we used for the mouse model (Extended Data Fig. 6a and Methods). Across three individuals, CD142$^+$ cells exhibited higher levels of *F3* gene expression when compared to CD142$^-$ cells (Extended Data Fig. 6b and

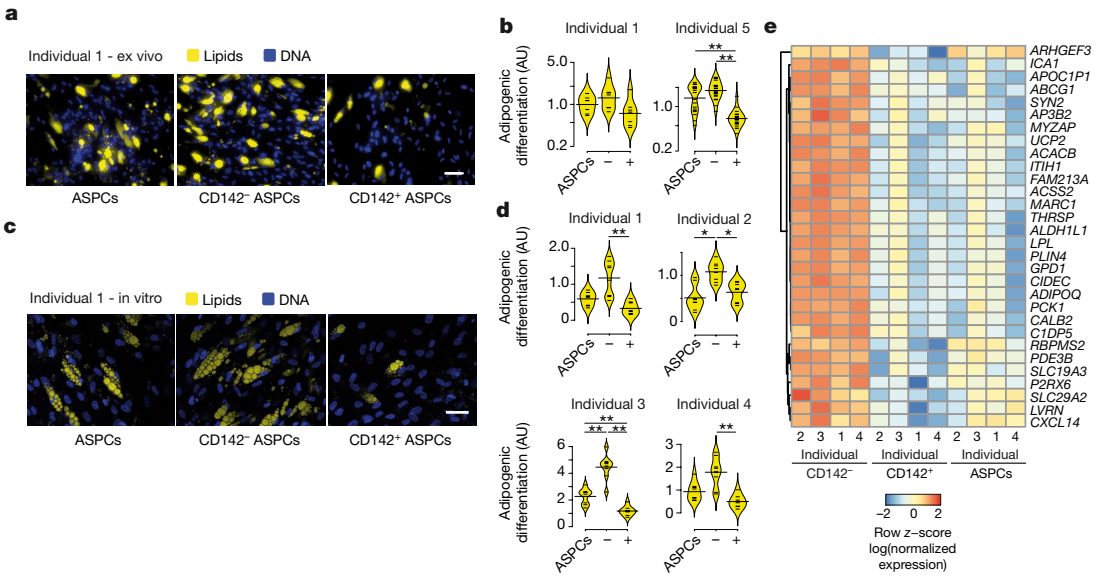


Fig. 3 | Aregs and their inhibitory capacity are conserved in humans. **a**, Microscopy images of ex vivo human ASPCs, CD142$^-$ ASPCs and CD142$^+$ ASPCs after adipogenesis. **b**, Adipogenic differentiation (arbitrary units, AU) per each fraction shown in **a** (individual 1, $n = 7$ fields of view) and another individual (individual 5, $n = 20$ fields of view). −, CD142$^-$; +, CD142$^+$. **c**, Microscopy images of in vitro human (individual 1) ASPCs, CD142$^-$ ASPCs and CD142$^+$ ASPCs after adipogenesis. In **a**, **c**, experiments were repeated at least three times (distinct individuals), yielding similar results. **d**, Adipogenic differentiation per each fraction shown in **c**, and in three other individuals (individuals 1 to 4, $n = 8$ fields of view). −, CD142$^-$; +, CD142$^+$. **e**, Expression of genes that were most-highly expressed in CD142$^-$ versus CD142$^+$ ASPCs upon adipogenesis (day 12, D12); $n = 4$ biological replicates (distinct individuals). In all panels, nuclei are stained with Hoechst (blue) and lipids are stained with Bodipy (yellow). Scale bars, 50 μm, $*P \leq 0.05$, $**P \leq 0.01$, Wilcoxon rank-sum test. See Extended Data Fig. 6 for related results.

Supplementary Tables 14, 19), validating our approach. We observed only a mild enrichment of *ABCG1* expression (Extended Data Fig. 6b), which may point to species-specific gene expression differences. Similar to mouse CD142$^+$ ASPCs, human CD142$^+$ ASPCs were refractory to adipogenic differentiation, as assessed by lipid accumulation (Fig. 3a, b and Methods) and marker gene expression (Extended Data Fig. 6b and Supplementary Table 14). Consistent with the mouse data, *F3* expression increased during adipogenesis, which negated the difference in expression between the CD142$^+$ and CD142$^-$ populations (Extended Data Fig. 6b). We further examined whether the adipogenic inhibitory property is maintained after culture by separating the CD142$^+$ population from the CD142$^-$ population in cultured (passage 1) ASPCs (Extended Data Fig. 6c), and then inducing adipogenesis. Population differences were consistent across four individuals (Fig. 3c, d). We observed significantly larger numbers of mature adipocytes among CD142$^-$ cells than among all ASPCs, which suggests that human CD142$^+$ ASPCs also have an inhibitory capacity (Fig. 3c, d and Supplementary Table 19). Transcriptomic profiling of induced cells revealed hundreds of differentially expressed genes among the three fractions, with CD142$^-$ ASPCs exhibiting the highest number of differences (Extended Data Fig. 6d and Supplementary Table 15). The adipocyte-like phenotype of CD142$^-$ cells was globally confirmed by the high expression of genes associated with adipogenic functionality (Fig. 3e, Extended Data Fig. 6e–i and Supplementary Table 16), including the adipogenic markers *ADIPOQ*, *CIDEC* and *LPL*. Finally, we performed a cross-species experiment in which human Lin$^-$ SVF cells were co-cultured with mouse Aregs or Areg-depleted ASPCs. We observed a reduction in adipogenesis in the transwells containing Aregs (Extended Data Fig. 6j, k), which suggests that the paracrine signal is conserved across species. Taken together, our data suggest that Aregs are functionally conserved in humans.

Having characterized Aregs based on cell-culture experiments, we next aimed to study their distribution and effect in vivo. Immunofluorescence staining of cryosections of the subcutaneous adipose tissue of mice revealed a CD142 signal that indicated a perivascular sheath, which featured small individual cells and was located within the adipose parenchyma but not the lymph nodes (Fig. 4a and Extended Data Fig. 7a–d, arrows). By combining staining for CD142 with staining for the stem cell marker SCA1, we identified a small fraction of CD142$^+$SCA1$^+$ cells that are also located perivascularly (Fig. 4b–d and Extended Data Fig. 7e–g, arrows). This is consistent with our gene expression data, which suggest that Aregs constitute a small fraction of all ASPCs as well as a link to blood vessels (Extended Data Figs 1g, 2d and 4l). We therefore hypothesize that the CD142$^+$SCA1$^+$ cells we detected correspond to Aregs.

Using a FACS strategy analogous to the one used above, we found that Aregs were also present in visceral fat depots (Fig. 4e and Methods). The visceral SVF contained a significantly higher proportion of CD142$^+$ABCG1$^+$ cells, when compared to the subcutaneous SVF (Fig. 4e and Supplementary Table 19). Furthermore, we found that obese *ob/ob* mice have significantly more Aregs than do lean mice, in both the subcutaneous and the visceral adipose depots (Fig. 4e and Supplementary Table 19). To assess whether the presence of Aregs exerts an anti-adipogenic effect in vivo, we embedded (1) total SVF or CD142$^-$ABCG1$^-$ SVF cells and (2) Lin$^-$SCA1$^+$ or CD142$^-$Lin$^-$SCA1$^+$ SVF cells in Matrigel, injecting them into the left and right subcutaneous layer of the mouse abdomen (Methods). After implantation, the pad-bearing mice were fed a high-fat diet for three weeks to induce adipogenesis. Consistent with our in vitro data, we observed a notable difference between the Matrigel pads for all mice (Fig. 4f–h and Extended Data Fig. 7h–j). The absence of Aregs either from total SVF cells or from ASPCs led to a significantly higher number of mature adipocytes being formed in the implants, as also revealed by the much whiter appearance of Matrigel pads containing CD142$^-$(ABCG1$^-$) cells (Supplementary Table 19). The staining of implant cryosections for endothelial cells using isolectin IB4 revealed a similar degree of vascularisation (Extended Data Fig. 7k, l). These findings strongly suggest that Aregs control the formation of mature adipocytes in vivo.

Our single-cell transcriptomics-based approach using a mouse model revealed at least three cell populations in the subcutaneous ASPC-enriched SVF. The largest one was CD55$^+$, and showed enrichment of the adipogenic stem cell markers *Cd34* and *Ly6a* as well as a high adipogenic differentiation capacity in vitro. A recent study has

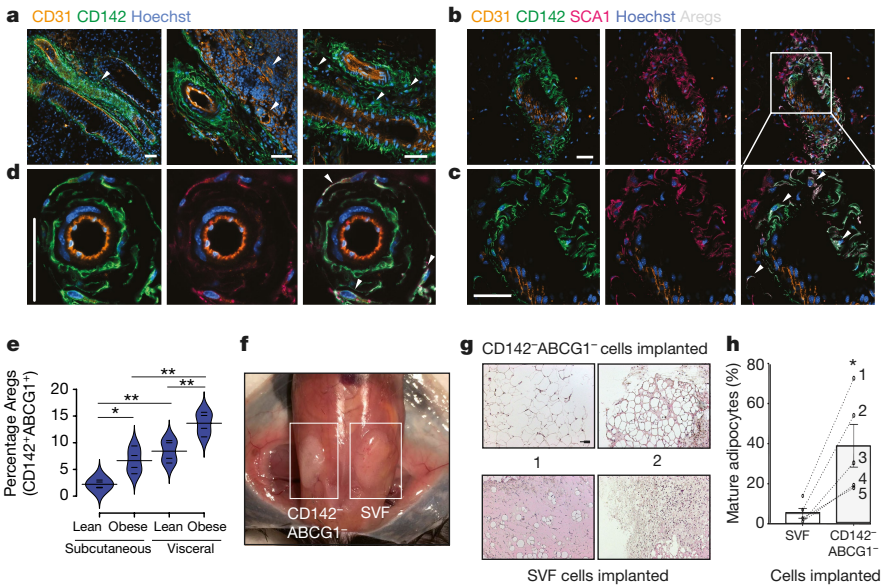


Fig. 4 | Aregs locate proximal to the vasculature and inhibit adipogenesis in vivo. a–**d**, Microscopy images of the in vivo localization of CD31 (orange), CD142 (green) and SCA1 (pink) in subcutaneous white adipose tissue in mice. Nuclei are stained with Hoechst (blue). Scale bars, 50 μm. Experiments were repeated at least three times, yielding similar results. Arrows in **a** indicate CD142$^+$ around big vessels within the adipose parenchyma (left), CD142$^+$ absence in the lymph nodes (middle) and CD142$^+$ in individual cells (right); arrows in **c, d** indicate individual cells presenting both CD142 and SCA1 staining (white). **e**, Fraction of Aregs (CD142$^+$ABCG1$^+$ ASPCs) in the subcutaneous and visceral SVF from lean and obese mice; $n = 4$ biological replicates. **f**–**h**, Total SVF and Areg-depleted (CD142$^-$ABCG1$^-$) SVF cells were implanted into the subcutaneous adipose depot in Matrigel. **f**, Image of implantation sites after three weeks of high-fat diet. **g**, Microscopy images of Matrigel plugs with total SVF or CD142$^-$ABCG1$^-$ SVF cells. Scale bar, 100 μm. **h**, Mean and s.d. of the percentage of mature adipocytes that developed within the plugs ($n = 5$ biological replicates) analysed using CellProfiler[25]. *$P \leq 0.05$, **$P \leq 0.01$ paired t-test. See Extended Data Fig. 7 for related results.

identified a homologous subpopulation in the human SVF, which exhibited an enhanced regenerative potential and was dysfunctional in diabetes[23]. These CD55$^+$ cells may therefore exhibit the highest plasticity of all Lin$^-$ SVF cells in both mice and humans, a hypothesis that warrants further investigation. We determined that a small (less than 10% of Lin$^-$ SVF cells) CD142$^+$ subpopulation suppresses adipocyte formation in vitro and in vivo. To our knowledge, we provide the first molecular characterization and functional follow-up of this CD142$^+$ subpopulation and show that the mechanism of action of this subpopulation possibly involves the genes *Spink2*, *Rtp3*, *Vit* and/or *Fgf12*. Our findings suggest that depot- and condition-specific differences in hyperplasia might not be solely attributable to alterations in the number of ASPCs, but instead might also be influenced by the number of Aregs, as these latter cells may control the de novo adipogenic capacity of adipose tissue. As adipocytes can also arise in tissues such as bone marrow and muscle[24], it is tempting to speculate that de novo adipogenesis in these tissues is also controlled through the presence of Areg-like cells. Together, our findings point to a potentially critical role for Aregs in modulating the plasticity and metabolic signature of distinct fat-cell-containing systems, within which they may constitute essential components of the elusive adipogenic precursor niche.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0226-8.

Received: 3 May 2017; Accepted: 16 May 2018;
Published online: 20 June 2018

1. Müller, S., Kulenkampff, E. & Wolfrum, C. in *Metabolic Control* (ed. Herzig, S.) 251–263 (Springer International Publishing, Cham, 2015).
2. Rosen, E. D. & Spiegelman, B. M. What we talk about when we talk about fat. *Cell* **156**, 20–44 (2014).
3. Crisan, M. et al. A perivascular origin for mesenchymal stem cells in multiple human organs. *Cell Stem Cell* **3**, 301–313 (2008).
4. Tang, W. et al. White fat progenitor cells reside in the adipose vasculature. *Science* **322**, 583–586 (2008).
5. Vishvanath, L. et al. Pdgfrβ$^+$ mural preadipocytes contribute to adipocyte hyperplasia induced by high-fat-diet feeding and prolonged cold exposure in adult mice. *Cell Metab.* **23**, 350–359 (2016).
6. Zannettino, A. C. W. et al. Multipotential human adipose-derived stromal stem cells exhibit a perivascular phenotype *in vitro* and *in vivo*. *J. Cell. Physiol.* **214**, 413–421 (2008).
7. Cai, X., Lin, Y., Hauschka, P. V. & Grottkau, B. E. Adipose stem cells originate from perivascular cells. *Biol. Cell* **103**, 435–447 (2011).
8. Gupta, R. K. et al. *Zfp423* expression identifies committed preadipocytes and localizes to adipose endothelial and perivascular cells. *Cell Metab.* **15**, 230–239 (2012).
9. Jiang, Y., Berry, D. C., Tang, W. & Graff, J. M. Independent stem cell lineages regulate adipose organogenesis and adipose homeostasis. *Cell Reports* **9**, 1007–1022 (2014).
10. Tran, K.-V. et al. The vascular endothelium of the adipose tissue gives rise to both white and brown fat cells. *Cell Metab.* **15**, 222–229 (2012).
11. Majka, S. M. et al. De novo generation of white adipocytes from the myeloid lineage via mesenchymal intermediates is age, adipose depot, and gender specific. *Proc. Natl Acad. Sci. USA* **107**, 14781–14786 (2010).
12. Billon, N. et al. The generation of adipocytes by the neural crest. *Development* **134**, 2283–2292 (2007).
13. Sowa, Y. et al. Adipose stromal cells contain phenotypically distinct adipogenic progenitors derived from neural crest. *PLoS ONE* **8**, e84206 (2013).
14. Su, X. et al. Fascia origin of adipose cells. *Stem Cells* **34**, 1407–1419 (2016).
15. Berry, R., Jeffery, E. & Rodeheffer, M. S. Weighing in on adipocyte precursors. *Cell Metab.* **19**, 8–20 (2014).
16. Sanchez-Gurmaches, J. & Guertin, D. A. Adipocyte lineages: tracing back the origins of fat. *Biochim. Biophys. Acta* **1842**, 340–351 (2014).
17. Cristancho, A. G. & Lazar, M. A. Forming functional fat: a growing understanding of adipocyte differentiation. *Nat. Rev. Mol. Cell Biol.* **12**, 722–734 (2011).
18. Jeffery, E. et al. Characterization of Cre recombinase models for the study of adipose tissue. *Adipocyte* **3**, 206–211 (2014).
19. Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C. & Teichmann, S. A. The technology and biology of single-cell RNA sequencing. *Mol. Cell* **58**, 610–620 (2015).
20. Rodeheffer, M. S., Birsoy, K. & Friedman, J. M. Identification of white adipocyte progenitor cells *in vivo*. *Cell* **135**, 240–249 (2008).
21. Hudak, C. S. et al. Pref-1 marks very early mesenchymal precursors required for adipose tissue development and expansion. *Cell Reports* **8**, 678–687 (2014).

22. Vignali, D. A. A., Collison, L. W. & Workman, C. J. How regulatory T cells work. *Nat. Rev. Immunol.* **8**, 523–532 (2008).
23. 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 (2016).
24. Shook, B. et al. The role of adipocytes in tissue regeneration and stem cell niches. *Annu. Rev. Cell Dev. Biol.* **32**, 609–631 (2016).
25. Gubelmann, C. et al. Identification of the transcription factor ZEB1 as a central component of the adipogenic gene regulatory network. *eLife* **3**, e03346 (2014).

Acknowledgements We thank J. Auwerx, W. Chen, E. Dorcey, N. Gheldof, O. Naveiras and K. Schoonjans for constructive discussions and carefully reading the manuscript. This research was supported by the Human Frontier Science Program LT001032/2013 (to P.C.S.), the Swiss National Science Foundation Grants (#31003A_162735 and 31003A_162887), the Kristian Gerhard Jebsen Foundation for Metabolic Research (to B.D.) and by institutional support from the Swiss Federal Institute of Technology in Lausanne (EPFL) and Zürich (ETH). We thank the EPFL FCCF, BIOP (especially R. Guiet and O. Burri), HCF and GECF for cell sorting, imaging, histology and sequencing support, as well as the VITAL-IT platform (University of Lausanne) for computational support.

Reviewer information *Nature* thanks O. Stegle and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions B.D., C.W., P.C.S., M.Z. and H.D. designed the study and wrote the manuscript. P.C.S. conducted all analyses related to transcriptomics. N.A., J.R., H.D. and M.Z. performed the single-cell experiments. H.D. and M.Z. conducted all FACS, qPCR, cell culture and related imaging assays. H.D. and W.S. performed all experiments related to transplants. H.D. performed all siRNA-based knockdown experiments. M.Z. and J.R. conducted all histological assays. C.C., K.-U.S. and G.S. provided access to human samples and helped with processing them. M.Z. and D.A. performed all experiments related to human cells. All authors read and approved the final manuscript.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0226-8.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0226-8.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to C.W. or B.D.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment, except for the quantification of in vivo implantation experiments (Fig. 4f–h and Extended Data Fig. 7h–l).

Bioethics. All mouse experiments were conducted in strict accordance with the Swiss law and all experiments were approved by the ethics commission of the state veterinary office (VD2984/2015, 60/2012, 43/2011, 80/2014). The work on human ASPC cultures derived from human lipoaspirate samples was approved by the ethical commission of Canton Ticino (CE 2961 from 22.10.2015) and conforms to the guidelines of the 2000 Helsinki declaration. The anonymized samples were collected under signed informed consent.

Generation of the DLK1 (also known as PREF1)–RFP mouse. The new mouse strain Tg(Pref1-CreER)426 Biat (Pref1-CreER) was generated using the bacterial artificial chromosome (BAC) RP24-334E14 (BACPAC Resources Center) according to previously published methods[26,27].

Isolation of the mouse SVF. Subcutaneous and visceral adipose tissue depots were dissected from male and female 8–11-week-old wild-type, DLK1–RFP or male *ob/ob* C57BL/6J mice into ice-cold PBS. The tissue was finely minced using scissors, transferred into collagenase (Sigma-Aldrich #C6885-1G, 2 mg/ml of collagenase buffer (25 mM $NaHCO_3$, 12 mM KH_2PO_4, 1.2 m $MgSO_4$, 4.8 mM KCl, 120 mM NaCl, 1.4 mM $CaCl_2$, 5 mM Glucose, 2.5% BSA, pH = 7.4)) and incubated for 1 h at 37 °C under agitation. Following a 10-min centrifugation at 400 *g* at room temperature, floating mature lipid-filled adipocytes were aspirated and the cell suspension was filtered through a 40-μm cell strainer to ensure a single cell preparation. Next, a red blood cell lysis was performed by incubating the pelleted cells in the red blood cell lysis buffer (154 mM NH_4Cl, 10 mM $KHCO_3$, 0.1 mM EDTA) for 5 min, followed by two washes (a 5-min centrifugation at 400*g*, room temperature) with FACS buffer (PBS with 3% fetal bovine serum (FBS) (Gibco #10270-106), 1 mM EDTA, 1% penicillin–streptavidin (Gibco #15140122)).

Isolation of human SVF. Fresh lipoaspirates were washed twice with DPBS with calcium and magnesium (Gibco #14040091) in 100-ml syringes (VWR International #720-2528) (50 ml lipoaspirate per 40 ml DPBS with calcium and magnesium) and incubated with 0.28 U/ml of liberase TM (Roche #05401119001) for 45 min at 37 °C under agitation. The digested tissue was mixed with 1% human albumin (CSL Behring) in DPBS −/− (Gibco #14190094) (50 ml lipoaspirate per 40 ml 1% human albumin in DPBS −/−) and shaken vigorously to liberate the stromal cells (procedure performed twice). The aqueous phase was recovered and centrifuged (all centrifugations performed at 400 *g* for 5 min at room temperature). The cell pellet was resuspended in 15 ml remaining buffer and filtered through a 100-μm and then a 40-μm cell strainer to ensure a single-cell preparation, centrifuged and resuspended in 5 ml of 5% human albumin (CSL Behring). The viability and the number of nucleated cells in the cell suspension obtained was determined using a Nucleostainer, after which a red blood cell lysis was performed using VersaLyse solution (Beckman Coulter #A09777) according to the manufacturer's recommendations.

scRNA-seq. For the initial (Fluidigm C1) scRNA-seq experiments, we used the DLK1–RFP mouse strain, as recent results have suggested that adipogenic progenitors may be marked by the expression of *Dlk1*[21]. Specifically, in this strain, RFP is induced in *Dlk1*-expressing cells upon feeding with tamoxifen. Thus, all cells in which the *Dlk1* promoter has been active become RFP^+ and can be easily isolated by FACS. We investigated both RFP^+ and RFP^- cells to explore the differences (if any) in cell composition and molecular characteristics between the two cell populations. However, as no differences were observed, we treated and discussed the data irrespective of RFP status. Cells (550 RFP^+ and 550 RFP^-, 1,100 in total) at a concentration of 200 cells/μl were sorted by FACS directly into the C1 Single-Cell Auto Prep Array IFC 10–17 μm chip (Fluidigm #100-5760, C1) for 96 cells with the efficiency of ~10%. The presence of the cells was validated by microscopy visualization (Olympus Cell Xcellence, objective: UPLFLN 10×/0.30). Single-cell cDNA synthesis was performed on the C1 Single-Cell Auto Prep system (Fluidigm) according to the manufacturer's specifications. Single-cell libraries were multiplexed and sequenced across 3 lanes of HiSeq 2000 (Illumina) using 100-bp single-end sequencing. All results relating to the Fluidigm C1 single-cell experiment (Figs 1a–c, 2a and Extended Data Figs 1, 2a and 8) are based on 3 biological replicates performed on 3 different days, each stemming from 3–6 male and female 10–11-week-old mice.

A second scRNA-seq experiment was performed using the Chromium Single Cell Gene Expression Solution (10x Genomics), following the manufacturer's protocol. The SVF was isolated from six 8-week-old mice (3 males and 3 females) as described above. Cells were stained with the anti-mouse antibodies CD31-AF488, CD45-AF488 and TER119 AF488 (Biolegend #303110, #304017 and #116215, respectively), and 280,000 Lin^- cells were isolated using a Becton Dickinson FACS Aria III sorter. Cells were washed and resuspended in 250 μl FACS buffer (PBS, 2% FBS, 1 mM EDTA), targeting the required 1,000 cells/μl concentration, accounting for a 10–20% loss. We pipetted 9.7 μl cell suspension (concentration of 913 cells/μl, ~8,800 cells), targeting the recovery of ~5,000 cells. Single-cell RNA-seq libraries were obtained following the 10x Genomics recommended protocol, using the reagents included in the Chromium Single Cell 3′ v2 Reagent Kit. Libraries were sequenced on the NextSeq 500 v2 (Illumina) instrument using 150 cycles (18 bp barcode + UMI, and 132-bp transcript 3′ end), obtaining ~5×10^8 raw reads.

scRNA-seq data analysis. For the C1 scRNA-seq experiments, FASTQ files containing 100-bp-long single-end sequenced tags (reads) from 3 replicates of 96 cells each were trimmed and filtered using prinseq 0.20.3[28] with the parameters '-custom_params 'A 70%;T 70%;G 70%;C 70%' -trim_tail_left 36 -trim_tail_right 36 -lc_method dust -lc_threshold 45 -min_gc 1 -out_format 3' and cutadapt 1.5[29] with the parameters '-m 36 -q 20' and the Nextera adaptor sequence. The retained tags were evaluated using FastQC v.0.11.2 and aligned to the Ensembl 84[30] gene annotation of the NCBI38/mm10 mouse genome using STAR 2.4.0g[31] with the parameters '–runThreadN 4 –runMode alignReads –outFilterType BySJout –outFilterMultimapNmax 20 –alignSJoverhangMin 8 –alignSJDB overhangMin 1 –outFilterMismatchNmax 999 –outFilterMismatchNoverLmax 0.04 –alignIntronMin 20 –alignIntronMax 1000000 –alignMatesGapMax 1000000 g –outSAMtype BAM SortedByCoordinate'. The number of tags per gene was calculated using htseq-count 0.6.0[32] with the parameters 'htseq-count -m intersection-nonempty -s no -a 10 -t exon -i gene_id'. Cells that appeared as doublets in the microscopy images (33) and those that had <40% or <400,000 aligned reads were excluded, which resulted in a total of 208 cells, of which 63 were RFP^- according to the microscopy images. Genomic alignment rates, number of detected genes per cell, RNA spike-in recovery as well as correlations between replicate experiments and those with bulk population (control) samples suggested that our data were of high quality (Extended Data Fig. 1a–d). RFP^+ cells showed significantly higher expression of *RFP* than RFP^- cells ($P < 0.01$, Wilcoxon rank-sum test, data not shown). For each gene, expression estimates per gene were expressed as log-transformed counts per million (cpm) (log_cpm/log(norm_expr) in Fig. 1c and Extended Data Fig. 1), by dividing total tags per gene by the total number of gene-aligned reads per cell and taking the $\log(x+1)$ of this value. Using the marker genes listed in Supplementary Table 4, we calculated 'scores'—a single numeric value representative of the expression of multiple marker genes—as the sum of log_cpms across all markers in a category. We further filtered and normalized the data using the package M3Drop_1.0.1[33] and the function M3DropCleanData() with the parameters 'is.counts=T, min_detected_genes=3000', obtaining 17,287 genes and 208 cells. We used the function M3DropDropoutModels from the M3Drop package to fit the modified Michaelis–Menten equation, a logistic regression (logistic) or a double exponential (ZIFA) function to the relationship between mean expression and dropout-rate (proportion of zero values). After visual inspection of the three fits and examination of the sum of squared residuals and sum of absolute residuals, we determined that the Michaelis–Menten method produced the best fit to our data. We therefore used it to estimate genes that had a significantly higher number of drop-outs than expected by chance (referred to as differentially expressed genes above, and informative genes below) using the M3Drop function M3DropDifferentialExpression() at a false discovery rate (FDR) of 0.05, which resulted in 527 genes (listed in Supplementary Table 1). We used these genes to obtain a 2D representation of the cells, while maintaining the similarity relationships between them using *t*-SNE[34] as implemented in the package Rt-SNE_0.11[34]. Cluster analysis, including silhouette analysis, was performed using the SC3_1.3.6 package[35] on the 527 differentially expressed genes, using $k = 3$ or 4 (Fig. 1b and Extended Data Fig. 8a). P1 and P3 were stable across different cluster number choices, whereas P2 could be further subdivided. The cell grouping we obtained did not correspond to biological replicates, which suggests that the main signal captured by the selected genes is biological and not technical (Extended Data Fig. 8b, c). Silhouette analysis supported a three-cluster partitioning (Extended Data Fig. 8d), which also corresponded closely to the distribution of *Fabp4* expression. Marker genes per cluster were obtained using M3DropGetMarkers() function, with the top 200 included in Supplementary Table 2 and used for functional enrichment analysis (Supplementary Table 3). To test the robustness of our results to methodological choices, we alternatively used a set of 1,827 biologically highly variable (>0.5, FDR 0.05) genes calculated using the scran_1.2.2 package[36] instead of the drop-out-based gene selection. We found that 520 of these genes were among the 527 significantly differentially expressed genes described above (Extended Data Fig. 8e). The clustering was highly similar to that previously described, with P3 being largely unchanged and ~20% of P1 cells attributed as P2 cells (Extended Data Fig. 8f). Over 60% of the top 100 markers identified for each of the three populations were identical (Extended Data Fig. 8g) and the *t*-SNE projection placed P3 outside the P1–P2 space (Extended Data Fig. 8h). In Fig. 1c, we displayed in colour the top 10 markers per cluster, additional transcription-factor coding genes that were among the top 100 markers per cluster and highlighted genes that code for transcription factors and cell surface proteins. Heat maps were generated using gplots_3.0.1 and the function heatmap.2() with the parameters 'scale="row",

Colv = F, Rowv = as.dendrogram (cluster)' (per row *z*-score transformed log(normalized expression), blue-to-red). Row means per gene are displayed on the left, in white-to-dark-red. All correlations were calculated based on log_cpm values, with the function cor() and the parameters 'method="Spearman'. Multiple testing correction using the function 'p.adjust and the parameters 'method = Bonferroni') was applied for *Fabp4* and *Cd34* expression correlations, respectively.

The 10x Genomics scRNA-seq data was processed using cellranger-2.1.0, default parameters and the mouse NCBI38/mm10 genome. Molecular counts were obtained for 2,919 cells (filtered matrix), with 174,362 mean reads/cell, an average of 60.5% reads mapping to the transcriptome and 3,404 median genes detected per cell. We also filtered outlier cells using the median absolute deviation from the median total library size (logarithmic scale) as well as total gene numbers (logarithmic scale), as implemented in scran[36], using a cutoff of 3 (isOutlier, nmads = 3). log(normalized expression) values were obtained using size factors per cell, estimated with scran. The majority of the cells expressed high levels of the housekeeping gene *Actb*, as well as the ASPC marker genes *Cd34*, *Ly6a* and *Itgb1* (Extended Data Fig. 8i). An initial analysis using all cells revealed that one of the major sources of variation was related to the expression of the gene *Xist*, as exemplified by the strong clustering of cells expressing *Xist* in the 2D *t*-SNE projection (Extended Data Fig. 8j). As *Xist* was expressed (>2 reads/cell) in 93% of the C1 cells, for comparative consistency we included only cells that expressed *Xist* (≥2 molecular counts per cell) in the reported results. We also noted a strong cell division signal in a small (27) number of cells, as quantified by Cyclone[37], and exemplified by *Mki67* expression (Extended Data Fig. 8j). As the C1 data contained cells that were predicted only to be in G1 phase (with a single exception), we also excluded the cells predicted to be in S or G2–M phase. Finally, as we noticed high levels of expression of epithelial-specific genes in a small subset of cells (again, dissimilar to the C1 experiment), we excluded all cells showing any (>0) *Epcam* and *Krt19* or *Krt18* expression. The final reported dataset consists of 1,804 cells (Fig. 1d). Genomic alignment rates and number of detected genes/cell suggested that our data were of high quality (Extended Data Fig. 2a, b). We further used both methods detailed above (M3Drop and highly variable genes) to determine a set of informative genes, and obtained a set of 550 genes with a significantly higher number of drop-outs than expected by chance (FDR = 0.0001, a stringent cutoff designed to exclude the selection of many genes expressed at very low levels, given that the 10x Genomics data showed a different drop-out distribution as compared to the C1 data) and 223 highly variable genes (FDR 0.1 and biological variability >0.3). Combining these two sets of genes gave 631 unique informative genes (Supplementary Table 6), which we used to perform consensus clustering using SC3_1.3.6 and to display the cells in 2D using *t*-SNE, similar to as described above. Although the sum of the squared differences decreased rapidly from using two to using three or four clusters (Extended Data Fig. 8k), silhouette analysis suggested a local maximum at four clusters (Extended Data Fig. 8l). We therefore used $k = 4$ in the manuscript. We note that a larger choice of k could provide further insight into additional subgroupings of the reported populations. Marker genes per cluster were again obtained using M3DropGetMarkers() function; the top 200 are included in Supplementary Table 7, and were used for functional enrichment analysis. The top 100 marker genes were used to compare the clusters (groups, G) with the previous ones (populations, P) (Fig. 1f and Extended Data Figs 2e, 8m), revealing >30% overlap for each population with its respective group.

FACS-based cell isolation of mouse cells. The isolated single-cell suspension was diluted to 0.75 or 1×10^7 cells/ml with FACS buffer (PBS with 3% FBS, 1 mM EDTA, 1% penicillin–streptavidin) and the following fluorophore-conjugated antibodies were added (in titration-determined quantities, Supplementary Table 17): anti-mouse CD31-AF488, anti-mouse CD45-AF488, anti-mouse TER119-AF488 (BioLegend #303110, #304017 and #116215, respectively) for selecting the Lin^- population; anti-mouse SCA1-PE-Cy7 (BioLegend #122513), anti-mouse CD34-BV421 (BioLegend #119321) and anti-mouse CD29-PerCP-efluor710 (eBiosciences #46-0291) to enrich the Lin^- population with ASPCs; anti-mouse CD55-APC (BioLegend #131802), anti-mouse VAP1 (Abcam #ab81673) conjugated with allophycocyanin (APC) (Lightning-Link Allophycocyanin (APC) Conjugation Kit (Innova Biosciences #705-0030), anti-mouse CD142-PE (SinoBiological #50413-R001) and anti-ABCG1 antibody (Invitrogen #PA5-13462) conjugated with APC for separating populations negative and positive for the given marker. The cells were incubated with the cocktail of antibodies on ice for 20 min protected from light, after which they were washed and stained with DAPI (Sigma #D9542) or propidium iodide (Molecular Probes #P3566) for assessing viability, and subjected to FACS using a Becton Dickinson FACSAria II sorter. Compensation measurements were performed for single stains using compensation beads (eBiosciences #01-2222-42).

The following gating strategy was applied while sorting the cells: first, the cells were selected based on their size and granulosity or complexity (side and forward scatter), and then any events that could represent more than one cell were eliminated. Next, the Lin^-($CD31^-CD45^-TER119^-$) population was selected, followed by $Lin^-SCA1^+CD34^+CD29^+$ (C1), $Lin^-SCA1^+CD34^+$, Lin^-SCA1^+, or Lin^- (10x) selection, which were used as controls for the further analysed populations that were negative or positive for a given marker: $Lin^-SCA1^+CD55^-$ ($CD55^-$) and $Lin^-SCA1^+CD55^+$ ($CD55^+$), $Lin^-SCA1^+VAP1^-$ ($VAP1^-$) and $Lin^-SCA1^+VAP1^+$ ($VAP1^+$), $Lin^-SCA1^+ABCG1^-$ ($ABCG1^-$) and $Lin^-SCA1^+ABCG1^+$ ($ABCG1^+$), $Lin^-SCA1^+CD34^+CD142^-$ ($CD142^-$) and $Lin^-SCA1^+CD34^+CD142^+$ ($CD142^+$), and $Lin^-SCA1^+CD142^-ABCG1^-$ ($CD142^-ABCG1^-$) and $Lin^-SCA1^+CD142^+ABCG1^+$ ($CD142^+ABCG1^+$). The measurements were acquired using Diva software supplemented on the Becton Dickinson FACS Aria II sorter and analysed using Kaluza analysis software. The mouse sorting experiments were performed two to four times (Extended Data Fig. 3b: $CD55^+$ ASPCs = 8.7% ± 1.9%, $VAP1^+$ ASPCs = 37.0% ± 4.1%, $CD142^+$ $ABCG1^+$ ASPCs = 2.2% ± 0.7%, $CD142^+$ ASPCs = 9.6% ± 1.3%; Fig. 4e: lean subcutaneous $CD142^+ABCG1^+$ ASPCs = 2.2% ± 0.7%, obese subcutaneous $CD142^+ABCG1^+$ ASPCs = 6.6% ± 2.3%, lean visceral $CD142^+ABCG1^+$ ASPCs = 8.4% ± 2.1%, obese visceral $CD142^+ABCG1^+$ ASPCs = 13.7% ± 2.1%); all stemming from male mice.

FACS-based cell isolation of human cells. The isolated single cell suspension was diluted to 1×10^7 cells/ml with FACS buffer (DPBS −/−) with 1% human platelet serum (Cook Medical #G34936 PL-NH-500) and the following fluorophore-conjugated antibodies (Supplementary Table 17) were added: anti-human CD31-AF488, anti-human CD45-AF488 for selecting the Lin^- population, and anti-human CD142-PE (BioLegend #303110, #304017 and #365204, respectively) for separating the $CD142^+$ and $CD142^-$ populations. 7-Aminoactinomycin D (7-AAD) (Beckman Coulter #A07704) was used for assessing viability and Syto40 (Molecular Probes #S11351) was used for discerning nucleated cells. The cells were then processed as detailed above. The human ex vivo experiments were replicated for three different individuals (individuals 1 and 5, individual 4 not shown) (Extended Data Fig. 6a and Fig. 3a, b, $CD142^+$ ASPCs = 3.10% ± 0.85%). The human in vitro experiments were performed for four different individuals (individuals 1 to 4) (Extended Data Fig. 6c and Fig. 3c, d, $CD142^+$ ASPCs = 3.38% ± 0.64%). The qPCR gene expression assay was performed on ex vivo and differentiated human $CD142^+$ ASPCs and $CD142^-$ ASPCs from three different individuals (individual 1, individual 4 and individual 5) (Extended Data Fig. 6b). The 3′ RNA-seq gene expression assay (see below) was performed on in vitro and differentiated human ASPCs, $CD142^+$ and $CD142^-$ ASPCs from four distinct individuals (individuals 1 to 4).

Ex vivo adipogenic differentiation of mouse cells. The same number of cells was sorted directly into flat-bottom microscopy-adapted cell culture plates (Corning #353219), cultured to confluence in high glucose DMEM medium (Gibco #61965026) supplemented with 10% FBS and 1% penicillin–streptavidin, and treated with a single dose of white adipocyte differentiation induction cocktail (0.5 mM 3-isobutyl-1-methylxanthine (IBMX, Sigma #15879), 1 μM dexamethasone (Sigma #D2915), 170 nM insulin (Sigma #19278)), followed by maintenance treatment (1 μg/ml insulin) every 48 h. The culture was carried out until day 8, 9 or 10 after induction, at which point cells were stained for imaging or collected for RNA extraction.

Ex vivo adipogenic differentiation of human cells. The same number of cells was sorted directly into flat-bottom microscopy-adapted cell culture plates in high glucose MEMalpha medium (Gibco #32561037) supplemented with 5% human platelet serum and 50 μg/ml Primocin (InvivoGen #ant-pm-1). The cells at confluence were treated with induction cocktail (high glucose DMEM, 10% FBS, 50 μg/ml Primocin, 0.5 mM IBMX, 1 μM dexamethasone, 1.7 μM insulin, 0.2 mM indomethacin) for 7 days, followed by maintenance cocktail treatment (high glucose DMEM, 10% FBS, 50 μg/ml Primocin, 1.7 μM insulin) for another 7 days. For expanded human ASPCs, TrypLE Select reagent (Gibco #12563011) was used to collect the cells from the cell culture plates.

Quantification of in vitro differentiation (mouse and human). For experiments shown in Fig. 2c–f, h, i and Extended Data Figs 3d, e, 5d, e, h, differentiated cells were fixed with 4% formaldehyde before staining with Hoechst (nuclei) and LD540 (lipid droplets). Images were taken per well with an automatic imaging system (Operetta, Perkin Elmer) and analysed for lipid droplet content using the Harmony software. The differentiation quantification experiments were repeated at least 3 times, with multiple independent wells per experiment. At least 15 images per well (96-well plate) were acquired. One well is represented (Fig. 2c and Extended Data Fig. 3d), and quantification of 4 or 5 independent wells is shown (Fig. 2d and Extended Data Fig. 3e).

For experiments shown in Fig. 3a–d and Extended Data Figs 5a, b, 6j, k, once the cells differentiated fully, they were stained with live fluorescence dyes: Bodipy (boron-dipyrromethene, Invitrogen #D3922) for lipids and Hoechst for nuclei. Cells were incubated with the dyes in FluoroBrite phenol red-free DMEM medium (Gibco #A1896701) supplemented with 10% FBS and 1% penicillin–streptavidin for 30 min at 37 °C in the dark, washed twice with warm PBS and imaged in FluoroBrite medium using a Leica DMI4000 wide filed microscope (objectives: PL-S-APO 5×/0.15, PL-S-APO 10×/0.30, HC-PL-APO 20×/0.70) or a Zeiss

LSM700 confocal inverted microscope (objectives: EC Plan-Neofluar 10×/0.30, Plan-Apochromat 20×/0.80). To accurately estimate and represent differences in adipose differentiation, a quantification algorithm for image treatment was developed in collaboration with the EPFL BIOP imaging facility. In brief, image analysis was performed in ImageJ/Fiji, lipid droplets (yellow) and nuclei (blue) images were filtered using a Gaussian blur (sigma equal to 2 and 3, respectively) before an automatic thresholding. The automatic thresholding algorithm selections were chosen on the basis of visual inspection of output images. At least 7 images per well (96-well plate) were acquired. Images of one technical replicate (field of view) are represented in Fig. 3a, and quantification for 2 individuals is represented in Fig. 3b for ex vivo experiments. Images of one technical replicate are represented in Fig. 3c, and quantification for 4 individuals is represented in Fig. 3d for in vitro experiments. Data points shown in Fig. 3b, d correspond to technical replicates (fields of view). The experiments were performed for at least three individuals.

Titration experiment. An equal number of $Lin^-SCA1^+CD142^-ABCG1^-$ and $Lin^-SCA1^+CD142^+ABCG1^+$ cells (see 'FACS-based cell isolation of mouse cells') were plated in 10-cm collagen I (Corning #354249)-coated cell culture plates and cultured for expansion in high glucose DMEM medium supplemented with 10% fetal bovine serum and 1% penicillin–streptavidin, refreshed every 48 h. At 90% confluence, the $CD142^-ABCG1^-$ and $CD142^+ABCG1^+$ populations were mixed at different ratios (ranging from 0 to 100%) in a 96-well plate at a density of 25,000 cells per well. Cells were differentiated as described in 'Ex vivo adipogenic differentiation of mouse cells'. The experiment was replicated twice, with $n = 4$ independent wells each. At least 15 images per well (96-well plate) were acquired. One independent well is represented in Fig. 2e, and quantification of four independent wells is represented in Fig. 2f.

siRNA-mediated knockdown experiments. For each gene, a pool of 2–4 siRNA probes (Supplementary Table 18) was reverse-transfected to total SVF (Extended Data Fig. 5c–f) or $CD142^+$ ASPCs (Aregs) (Fig. 2h, i). Seventy-five thousand cells/cm^2 were plated with 100 nM of a given siRNA dissolved in 1.5% Lipofectamine RNAiMAX (Invitrogen #13778150) in Opti-MEM I reduced serum medium (Invitrogen #31985062), and high glucose DMEM medium supplemented with 2.5% FBS (w/o penicillin–streptavidin). After 24 h, the medium was changed to high glucose DMEM medium supplemented with 10% FBS and 1% penicillin–streptavidin and after 48 h the cells were collected for determining knockdown efficiency. The experiment was replicated twice, with $n = 6$–8, 2 biological replicates and 3–4 independent wells each. At least 15 images per well (96-well plate) were acquired. One independent well is represented in Extended Data Fig. 5d, and quantification of $n = 6$–8 independent wells from two biological replicates and two independent experiments (S1 and S2) is represented in Extended Data Fig. 5e, f.

Transwell experiments. The corresponding genes were knocked down in Aregs on the transwell inserts (Corning #3381) as described in 'siRNA-mediated knockdown experiments'. 24 h after transfection, the inserts were washed with PBS and changed to high glucose DMEM medium supplemented with 10% FBS and 1% penicillin–streptavidin. Afterwards, the inserts were co-cultured with $CD142^-$ ASPCs and at 48 h, both cell populations were treated with white adipocyte differentiation cocktail (see 'Ex vivo adipogenic differentiation of mouse cells') and imaged at day 6 after induction (see 'Quantification of in vitro differentiation'). The experiment was replicated twice, with $n = 6$, 2 biological replicates and 3 independent wells each. At least 15 images per well (96-well plate) were acquired. One independent well is represented in Fig. 2h, and quantification of $n = 6$–8 independent wells from two biological replicates and two independent experiments (S1 and S2) is represented in Fig. 2i.

The experiment shown in Extended Data Fig. 5a, b was performed three times, a technical replicate (two fields of views of one well) is shown in Extended Data Fig. 5a and the quantification of two biological replicates is represented in Extended Data Fig. 5b; data points correspond to technical replicates (fields of view of one well). The experiments were imaged at day 8, 9 or 10.

In vivo differentiation. In Fig. 4f–h, 10^6 SVF cells and 10^6 $CD142^-ABCG1^-$ SVF cells were re-suspended in 150 µl of Matrigel (Corning #356234) and injected subcutaneously in each flank of the same 5–6-week-old mouse. This experimental set-up minimizes inter-individual variation, and enables a direct comparison of the effect of $CD142^+ABCG1^+$ cell removal on adipocyte differentiation capacity in vivo. In Extended Data Fig. 7i, j, 10^6 Lin^-SCA1^+ cells and 10^6 $Lin^-SCA1^+CD142^-$ cells were injected as described for Fig. 4f–h. After 3 weeks of high-fat diet, all Matrigel plugs were excised and fixed in 4% paraformaldehyde (PFA) overnight, dehydrated and embedded in paraffin. Sections of 4 µm were stained with haematoxylin and eosin. From each plug, images of at least 3 full sections were taken and adipocyte numbers, as well as the number of nuclei, were determined with the software CellProfiler[38], as previously described[25]. Five (Fig. 4f–h and Extended Data Fig. 7h) and seven (Extended Data Fig. 7i, j) biological replicates are represented. The experiments were performed once each.

Vascularisation. The Matrigel plugs were obtained as described in 'In vivo differentiation'. After excision, the Matrigel plugs were frozen in Neg-50 Frozen Section Medium (ThermoFisher #6502). Sections of 100 µm were prepared using a Reichert cryostat microtome. Primary antibody anti-isolectin GS-IB4 Alexa Fluor-488 (Invitrogen #121411) was applied in 5% goat serum and incubated 12 h at 4 °C. The sections were then washed with PBS 3 times for 10 min, at room temperature. Finally, the sections were incubated with Hoechst (5 µg/ml), washed twice for 10 min with PBS and mounted with ProLong Diamond Antifade Mountant (Thermo Fisher #P36965). The slides were then imaged with a Leica TCS SP8 microscope. One replicate is represented in Extended Data Fig. 7k. The experiment was performed once, and included 6 biological replicates (Extended Data Fig. 7l).

RNA isolation and quantitative PCR. *RNA isolation of sorted cells.* Live cells were collected in RLT+ lysis buffer (Qiagen #1053393) and flash-frozen on dry ice. Cell lysates were homogenized with QIAshredders before RNA isolation using the RNeasy Plus Micro Kit (Qiagen #74034). Reverse transcription was performed using the QuantiTect whole transcriptome kit (Qiagen #207045), following the manufacturer's recommendations for a standard-yield reaction (2 h of amplification time), or SuperScript VILO cDNA Synthesis Kit (Invitrogen #11754050). mRNA expression was normalized to *36b4* (also known as *Rplp0*) (for all mouse experiments, if not otherwise specified) or to *Hprt1* (for the human experiments and the experiments represented in Fig. 2b and Extended Data Fig. 3c ($CD142^+$ versus $CD142^-$)).

RNA isolation of differentiated cells. mRNA was isolated and transcribed into cDNA using the MultiMACS cDNA Synthesis Kit (Miltenyi #130-094-410) or collected into Tri-Reagent (Molecular Research Center #TR118). Direct-zol RNA kit (Zymo Research #R2052) was used to extract RNA, followed by reverse transcription using the SuperScript VILO cDNA Synthesis Kit (Invitrogen). Expression levels of mRNA were assessed by real-time PCR using the PowerUp SYBR Green Master Mix (Thermo Fisher Scientific #A25743). mRNA expression was normalized to *36b4* (for mouse experiments) or to *Hprt1* or *HPRT1* (for mouse or human experiments presented in Fig. 2b (third panel) and Extended Data Fig. 3c (third panel)).

Gene expression assays in Extended Data Fig. 3c show mean values over $n = 4$ or 5 biological replicates; the experiment was performed once. Figure 2b and Extended Data Fig. 3c ($CD142^+$ versus $CD142^-$) show $n = 3$ technical qPCR replicates, from 1 sorting tube. The experiment was performed twice. Extended Data Fig. 3f shows $n = 4$ or 5 independent wells, the experiment was replicated four times. Extended Data Fig. 5c, g represents 2–3 independent wells from 2 biological replicates each ($n = 4$–6 independent wells). The experiment was performed once. Extended Data Fig. 5j represents 3 independent wells from 2 biological replicates each ($n = 6$). The experiment was replicated two times. Extended Data Fig. 6b represents three biological replicates, from three individuals (individuals 1, 4 and 5).

Bulk mRNA-seq. Bulk mRNA sequencing (Figs 2, 3 and Extended Data Figs 4, 6) was performed as previously described[39,40]. In brief, 50 ng of total RNA from each sample was reverse transcribed in a 96-well plate using Maxima H Minus Reverse Transcriptase (Thermo Fisher Scientific #EP0753) with individual oligo-dT primers, featuring a 6-nt-long multiplexing barcode, and template switch oligo (Microsynth Custom made). Specifically, each oligo-dT primer is biotinylated and has the following structure: 5′-ACACTCTTTCCCTACACGACGCTCTTCCGATCT[BC6] [N15] [T30] VN-3′, in which [BC6] represents a 6-nt barcode that is specific to each well and N15 represents a stretch of random nucleotides forming a unique molecular identifier (UMI). Thus, for each well, we have generated a unique combination of barcodes and UMIs to identify each well (sample) and transcript. Next, all the samples were pooled together, purified using the DNA Clean and Concentrator kit (Zymo Research #D4014), and treated with exonuclease I (NEB or New England BioLabs #M0293S). The full-length cDNA library was amplified using a single primer and purified with Agencourt AMPure XP beads (Beckman Coulter A63881). The sequencing library was prepared by tagmentation of 10 ng full-length cDNA with a Tn5 transposase made in house, at 55 °C for 9 min[41]. Tagmented DNA was purified with the DNA Clean and Concentrator kit and PCR amplified using NEBNext High-Fidelity 2X PCR Master Mix (NEB or New England BioLabs #M0541S) with an i7 adaptor identical to Illumina Nextera and custom i5 (Microsynth custom made). The PCR reaction was then purified twice with AMPure beads or Agencourt AMPure XP beads and the average fragment size of the library was evaluated using a Fragment Analyzer (Advanced Analytical) before paired-end sequencing with NextSeq 500 (Illumina).

Analysis of bulk mRNA-seq data. For conventional bulk mRNA sequencing analysis, FASTQ files containing 100-bp-long single-end sequenced tags (reads) from four replicates of $Lin^-CD29^+CD34^+SCA1^+RFP^+$ and $Lin^-CD29^+CD34^+SCA1^+RPF^-$ each were analysed. Reads from each sample were trimmed and filtered using prinseq 0.20.3[28] with the parameters '-custom_params 'A 70%;T 70%;G 70%;C 70%' -trim_tail_left 36 -trim_tail_right 36 -lc_method dust -lc_threshold 45 -min_gc 1 -out_format 3' and cutadapt 1.5[29] with the parameters '-m 36 -q 20' and the Nextera adaptor sequence. The retained tags were evaluated using FastQC v.0.11.2 and aligned to the Ensembl 84[30] gene annotation of the NCBI38/mm10 mouse genome using STAR 2.4.0g[31] with the parameters

'–runThreadN 4–runMode alignReads–outFilterType BySJout–outFilterMultimapNmax 20–alignSJoverhangMin 8 –alignSJDB overhangMin 1–outFilterMismatchNmax 999–outFilterMismatchNoverLmax 0.04–alignIntronMin 20–alignIntronMax 1000000–alignMatesGapMax 1000000 g–outSAMtype BAM SortedByCoordinate'. The number of tags per gene was calculated using htseq-count 0.6.0[32] with the parameters 'htseq-count -m intersection-nonempty -s no -a 10 -t exon -i gene_id'. For each gene, expression estimates per gene were expressed as log-transformed counts per million (log_cpm), by dividing total tags per gene by the total number of gene-aligned reads per cell and taking the $\log(x+1)$ of this value.

To analyse the data from barcoded mRNA sequencing (Areg-related experiments in mouse and human, Figs 2, 3 and Extended Data Figs 4, 6)[40], the following workflow was used. Reads from barcoded mRNA-seq experiments have two barcodes, corresponding to the two levels of multiplexing. The first one is common to standard protocols and is used to separate the libraries. The second is specific to the barcoded mRNA-seq protocol and is used to separate the multiplexed samples from the bulk data. The first demultiplexing step was performed with the Illumina BaseSpace platform, and the second was performed using custom scripts. FASTQ files containing 62-bp-long single-end sequenced tags (reads) from at least four biological replicates (4 biological replicates, each with 1–3 independent wells) per ASPC population and sample were analysed: 6 ASPC replicates D0, 4 ASPC samples D12, 8 CD142$^-$ ASPC replicates D0, 7 CD142$^-$ ASPC 5 h, 7 CD142$^-$ ASPC 24 h, 4 CD142$^-$ ASPC D12, and 4 replicates each for CD142$^+$ ASPC D0, 5 h, 24 h and D12 (Fig. 2g and Extended Data Fig. 4). Reads from each sample were trimmed and filtered using prinseq 0.20.3[28] with the parameters '-custom_params 'A 70%;T 70%;G 70%;C 70%' -trim_tail_left 36 -trim_tail_right 36 -lc_method dust -lc_threshold 45 -min_gc 1 -out_format 3' and cutadapt 1.5[29] with the parameters '-m 36 -q 20' and the Nextera adaptor sequence. The retained tags were evaluated using FastQC v.0.11.2 and aligned to the Ensembl 84[30] gene annotation of the NCBI38/mm10 mouse genome using STAR 2.4.0g[31] with the parameters '–runThreadN 4–runMode alignReads–outFilterType BySJout–outFilterMultimapNmax 20–alignSJoverhangMin 8 –alignSJDB overhangMin 1–outFilterMismatchNmax 999–outFilterMismatchNoverLmax 0.04–alignIntronMin 20–alignIntronMax 1000000–alignMatesGapMax 1000000 g–outSAMtype BAM SortedByCoordinate'. The number of tags per gene was calculated using htseq-count 0.6.0[32] with the parameters 'htseq-count -m intersection-nonempty -s no -a 10 -t exon -i gene_id'.

Genes with a count per million greater than 2 in at least 4 samples were retained, providing a filtered dataset of 26,159 expressed genes across 52 samples. The samples were of high quality, as assessed by the number of aligned reads and detected genes per cell and by correlations with the P3 single-cell data (Extended Data Fig. 4c, d). Raw counts were normalized using mean-variance modelling at the observational level, as implemented in the voom() function in limma_3.30.4[42] and further using combat() in sva_3.22.0[43] to adjust for batch effects. Differential expression was computed on the normalized values using the limma_3.30.4 pipeline at an FDR of 0.05 and fold-change cutoff of 2. We detected a large number of changes that were induced upon adipogenic differentiation in all three cell fractions (CD142$^+$, CD142$^-$ and ASPCs) (Extended Data Fig. 4e, Supplementary Table 11), each of which had over 3,500 significantly differentially expressed genes (FDR 0.05, fold-change 2). Across time points, the highest number of differences was observed between Aregs and CD142$^-$ cells or ASPCs (Extended Data Fig. 4e). The high transcriptional similarity between CD142$^-$ cells and ASPCs (Extended Data Fig. 4e, f) was expected based on the low fraction of Aregs (generally $<10\%$) contained in ASPCs. Most differences arose early upon plating (after 5 h and 24 h), but dampened after adipogenic differentiation (Extended Data Fig. 4e). Heat maps displaying row-normalized (z-score transformation) expression (log_cpm) values were generated using pheatmap_1.0.8 and the parameters 'clustering_distance_rows = "correlation",cluster_cols = F, clustering_method = "average", scale = "row"'. Extended Data Fig. 4g contains all genes that were expressed at significantly lower levels after adipogenic differentiation (day 12, D12) in CD142$^+$ ASPCs compared to both CD142$^-$ ASPCs and all ASPCs; Fig. 2g contains all genes that were expressed at significantly higher levels in CD142$^+$ ASPCs compared to CD142$^-$ ASPCs after sorting (day 0, D0) and after culturing (5 h and day 1, D1); Extended Data Fig. 4k contains all genes that were expressed at significantly higher levels in CD142$^+$ ASPCs compared to CD142$^-$ ASPCs after sorting (day 0, D0); Extended Data Fig. 4n shows the expression of selected endothelial marker genes. All correlations were calculated based on log_cpm values, with the function cor() and the parameters 'method="Spearman"'.

The analysis of the human barcoded mRNA-seq data (displayed in Fig. 3 and Extended Data Fig. 6) was performed in a manner analogous to that described above for mouse data, but reads were aligned to the Ensembl 84[30] gene annotation of the GRCh38/hg20 human genome instead. Figure 3e contains only the top 30 genes that were expressed at significantly higher levels (fold-change = 2, FDR = 0.1) in CD142$^-$ ASPCs compared to CD142$^+$ ASPCs, and Extended Data Fig. 6g contains all genes that were expressed at significantly higher levels (fold-change = 2, FDR = 0.1) in CD142$^+$ ASPCs compared to CD142$^-$ ASPCs.

Immunofluorescence. Mice were anaesthetized with isoflurane and perfused with PBS (5 min) followed by 4% PFA (paraformaldehyde, electron microscopy grade (VWR #100504-858) (5 min). The subcutaneous fat pads were dissected and post-fixed in 4% PFA for 2 h at 4 °C upon gentle shaking. Next, the tissue was washed with PBS and incubated with 30% sucrose for 24 h at 4 °C upon gentle shaking. Cryoblocks were prepared using a Cryomatrix (Thermo Fisher Scientific #6769006) and 30-μm sections were generated using a Leica CM3050S cryostat at −30 °C. The sections were deposited onto glass slides and incubated in −20 °C acetone for 10 min, after which they were dried at room temperature for 10 min, rehydrated in PBS and blocked with 10% goat serum (Invitrogen #31872) supplemented with 0.3% TritonX100 (Sigma #T9284) for 1 h. Primary antibodies (Supplementary Table 17, hamster anti-mouse CD31, BioRad #MCA1370A, rabbit anti-mouse CD142, SinoBiological #50413-R001, rat anti-mouse SCA1, BioLegend #122501) were applied in 1% goat serum with 0.3% TritonX100 and incubated for 24 h at 4 °C. The sections were then washed with PBS three times for 10 min at room temperature. Secondary antibodies (Supplementary Table 17, goat anti-hamster AF-546, Molecular Probes #A21111, donkey anti-rabbit AF-488, Molecular Probes #A21247, goat anti-rat AF-647, Molecular Probes #A21206) were applied in PBS with 0.3% TritonX100 and incubated in the dark for 1 h at room temperature, followed by three 10-min washes with PBS at room temperature in the dark. Finally, the sections were incubated with Hoechst (5 μg/ml), washed two times for 10 min with PBS and mounted with Fluoromount G (Southern Biotech #0100-01). The slides were then imaged with a Zeiss confocal LSM700 microscope (objectives: EC Plan-Neofluar 10×/0.30, Plan-Apochromat 20×/0.80, Plan-Apochromat 40×/1.30). The results presented in Fig. 4a–d were replicated in at least three independent experiments. We note that we also verified that the signal we detected is not the result of autofluorescence of the adipose tissue (Extended Data Fig. 7a) or from unspecific binding of secondary antibodies (Extended Data Fig. 7b). We also showed that the perivascular staining is not an artefact caused by perfusing the animals with 4% PFA (Extended Data Fig. 7c).

Annotation and functional enrichment analyses. Gene annotations were obtained from Ensembl 84 (mar2016.archive.ensembl.org)[30], either by direct download or through biomaRt_2.30.0[44]. Enrichment analysis was performed using the webserver Enrichr with default parameters[45] and annotations provided by KEGG[46], Wikipathways[47], or gene expression data provided by the GTEx project[48]. For Extended Data Fig. 1g, the top 10 enriched terms were considered. Only highlights are displayed in the figures, and full results are available in Supplementary Tables 3, 8, 12 and 16.

Other computational analyses and data processing remarks. All computational analyses were performed using R version 3.3.2 and Bioconductor version 3.4. All t-tests and Wilcoxon rank-sum tests were unpaired and two-sided, if not otherwise specified. All box plots were generated and displayed in R, using the boxplot() function with default parameters. The median value is indicated with a black line, and a coloured box (hinges) is drawn between the 1st and 3rd quartiles (interquartile range, IQR). The whiskers correspond to approximately 1.5× interquartile range (±1.58 interquartile range divided by the square root of n) and outliers are drawn as individual points. All bean plots are generated and displayed in R, using the beanplot() (package beanplot) function with default parameters ('bw = "SJ-dpi", kernel = "gaussian", cut = 3, cutmin = -Inf, cutmax = Inf'). The mean value is indicated with the widest black line, and individual values are indicated with narrower lines. Kernel density estimates are displayed in colours. All bar plots display mean values as centres and the standard deviation as error bars. All included microscopy images and macroscopic images are representative. Exact P values are provided in the Supplementary Table 19; only the P value ranges are provided in the figure legends.

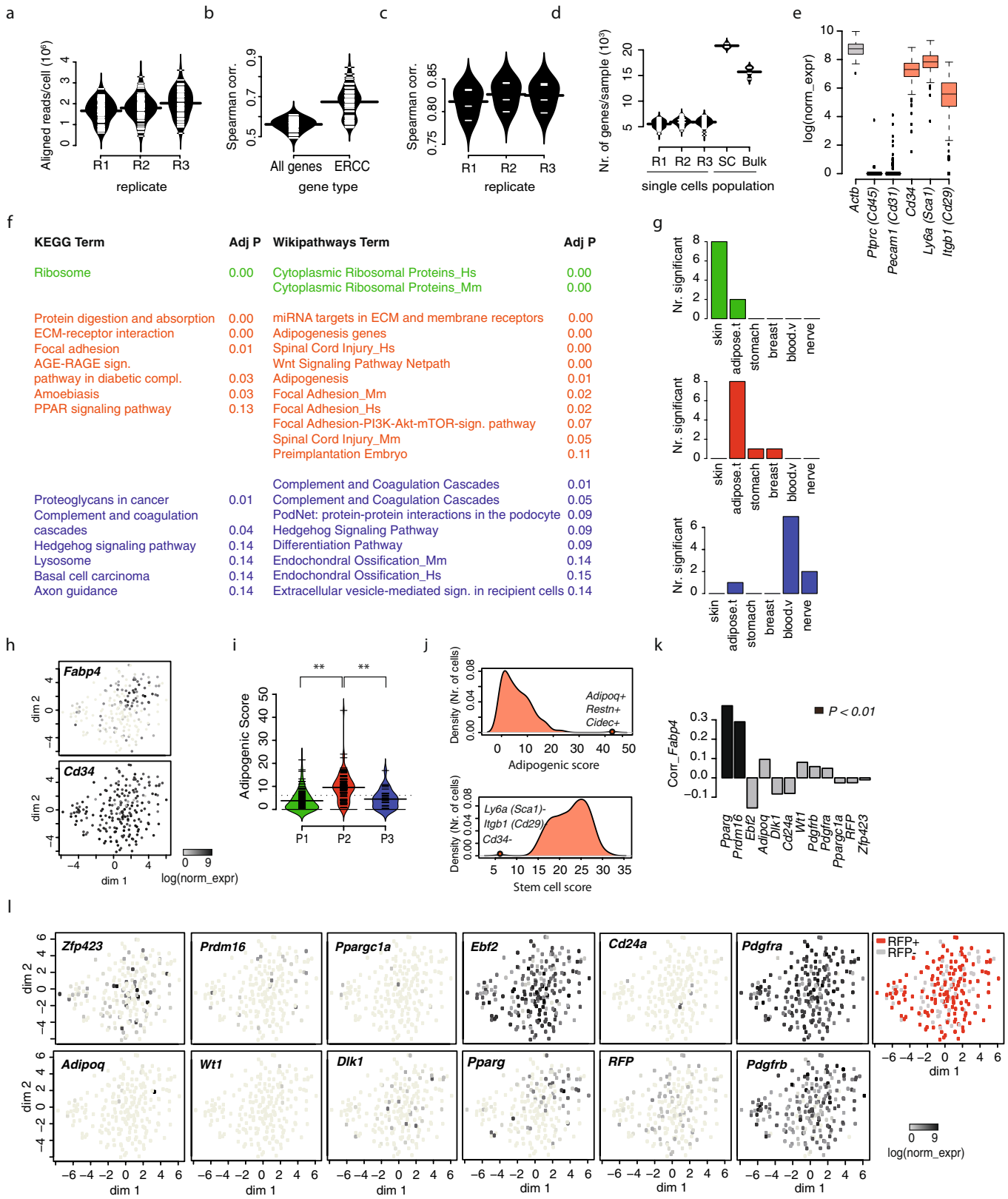
Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Code availability. Sample scripts used to process the data are available at https://github.com/DeplanckeLab/Areg.

Data availability. All raw and processed RNA-seq data have been uploaded in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) with the accession numbers E-MTAB-5785, E-MTAB-5818, E-MTAB-5802, E-MTAB-5787 and E-MTAB-6677. Microscopy images are available from the corresponding authors upon reasonable request.

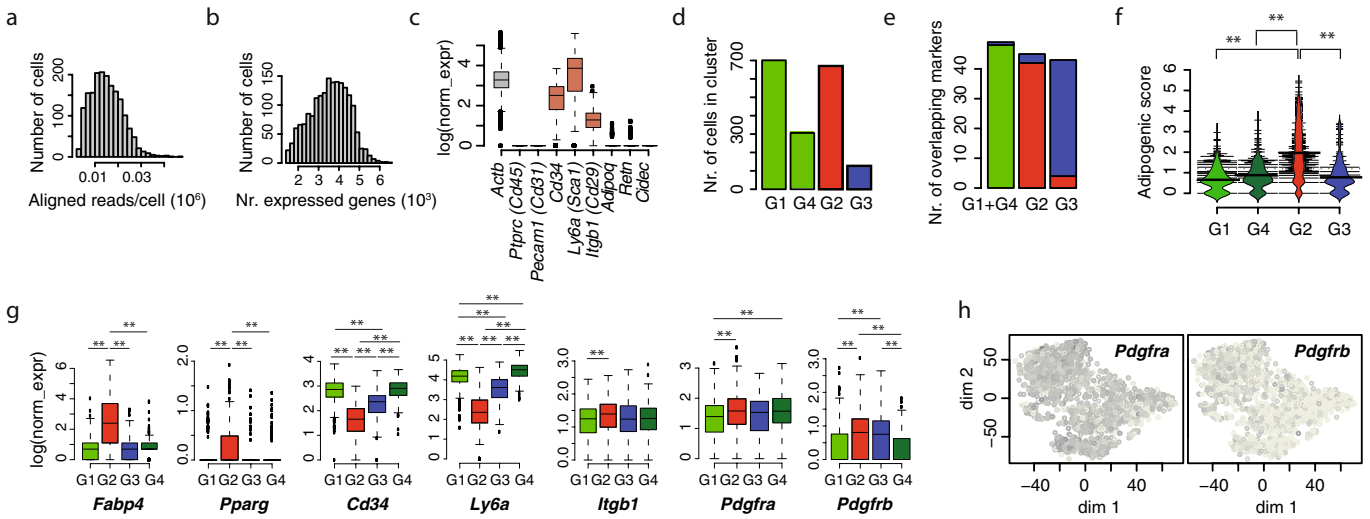
26. Johansson, T. et al. Building a zoo of mice for genetic analyses: A comprehensive protocol for the rapid generation of BAC transgenic mice. *genesis* **48**, 264–280 (2010).
27. Rosenwald, M., Perdikari, A., Rülicke, T. & Wolfrum, C. Bi-directional interconversion of brite and white adipocytes. *Nat. Cell Biol.* **15**, 659–667 (2013).
28. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863–864 (2011).

29. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
30. Yates, A. et al. Ensembl 2016. *Nucleic Acids Res.* **44**, D710–D716 (2016).
31. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
32. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
33. Andrews, T. S. & Hemberg, M. Modelling dropouts allows for unbiased identification of marker genes in scRNASeq experiments. Preprint at https://www.biorxiv.org/content/early/2016/07/21/065094 (2016).
34. Van Der Maaten, L. J. P. & Hinton, G. E. Visualizing data using *t*-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
35. Kiselev, V. Y. et al. SC3: consensus clustering of single-cell RNA-seq data. *Nat. Methods* **14**, 483–486 (2017).
36. Lun, A. T. L., McCarthy, D. J. & Marioni, J. C. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res.* **5**, 2122 (2016).
37. Scialdone, A. et al. Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods* **85**, 54–61 (2015).
38. Carpenter, A. E. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).
39. Pradhan, R. N. et al. Dissecting the brown adipogenic regulatory network using integrative genomics. *Sci. Rep.* **7**, 42130 (2017).
40. Alpern, D., Gardeux, V., Russeil, J. & Deplancke, B. Time- and cost-efficient high-throughput transcriptomics enabled by Bulk RNA Barcoding and sequencing. Preprint at https://www.biorxiv.org/content/early/2018/01/30/256594 (2018).
41. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protocols* **9**, 171–181 (2014).
42. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
43. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882–883 (2012).
44. Drost, H.-G. & Paszkowski, J. Biomartr: genomic data retrieval with R. *Bioinformatics* **33**, 1216–1217 (2017).
45. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
46. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
47. Kelder, T. et al. WikiPathways: building research communities on biological pathways. *Nucleic Acids Res.* **40**, D1301–D1307 (2012).
48. The GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).

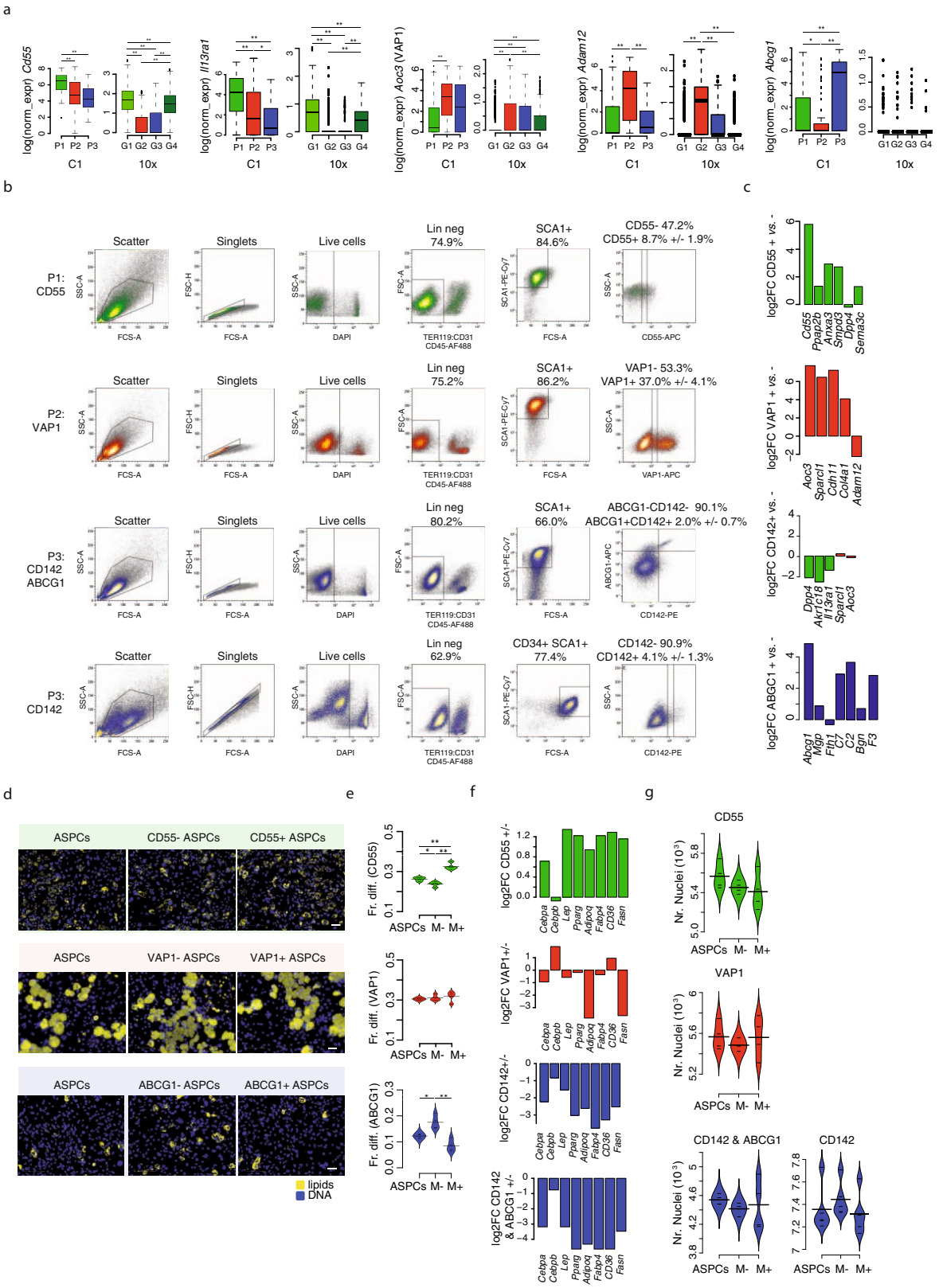


Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | ScRNA-seq reveals the heterogeneity of ASPCs (Fluidigm C1). a, Number of aligned reads per cell for each of the three Fluidigm C1 (C1) scRNA-seq experiments. R1, $n = 74$; R2, $n = 71$; R3, $n = 63$ single cells from three independent biological replicates, each derived from a pool of mice. **b**, Correlations (corr, Spearman's rho) between expression in single cells across all genes or only the artificial RNA spike-ins (ERCC). **c**, Correlation (Spearman's rho) between expression in merged single cell (per individual biological replicate, R1–R3) and bulk population ($n = 4$ biological replicates) samples. **d**, Number of expressed genes/sample in each of the three biological scRNA-seq replicates (R1–R3), as well as the merged individual replicates (SC, $n = 3$) or the bulk population control samples (bulk, $n = 4$). **e**, Expression of ubiquitous *Actb*, negative (*Ptprc* and *Pecam1*) and positive (*Cd34*, *Ly6a* and *Itgb1*) FACS markers. **f**, **g**, Pathways[47] significantly enriched (**f**) and top 10 significantly enriched GTEx[48] tissue samples (**g**) among the top 200 genes specific to one of the three populations (P1, green; P2, red; P3, blue). Full enrichment results are listed in Supplementary Table 3. **h**, *t*-SNE 2D maps of all analysed cells (C1), highlighting the expression of the stem cell marker *Cd34* and the adipogenic marker *Fabp4* (black). **i**, Bean plots showing the distribution of adipogenic scores across cells belonging to one of the three populations (P1, $n = 83$; P2, $n = 96$; P3, $n = 29$ cells). $^{**}P \leq 0.01$, Wilcoxon rank-sum test. **j**, Stem cell (lower panel) and adipogenic (upper panel) score distributions across all analysed cells, highlighting a single outlier cell not expressing—or expressing very low levels of—*Itgb1* (CD29), *Cd34* and *Ly6a* (SCA1) as well as very high levels of the mature adipogenic markers *Adipoq*, *Restn* and *Cidec*. **k**, Correlation (Spearman's rho) between the expression of 12 (pre-)adipogenesis-related genes and the expression of *Fabp4*; significant (Bonferroni multiple testing adjusted $P \leq 0.01$) correlations in black. **l**, *t*-SNE 2D maps of all analysed cells, highlighting the expression of *RFP* and various genes previously used to mark adipogenic precursors or pre-adipocytes (black); microscopic RFP status also displayed. In **e**, **h**, **j**, **l**, $n = 208$ cells from three independent biological experiments, each based on a pool of mice.

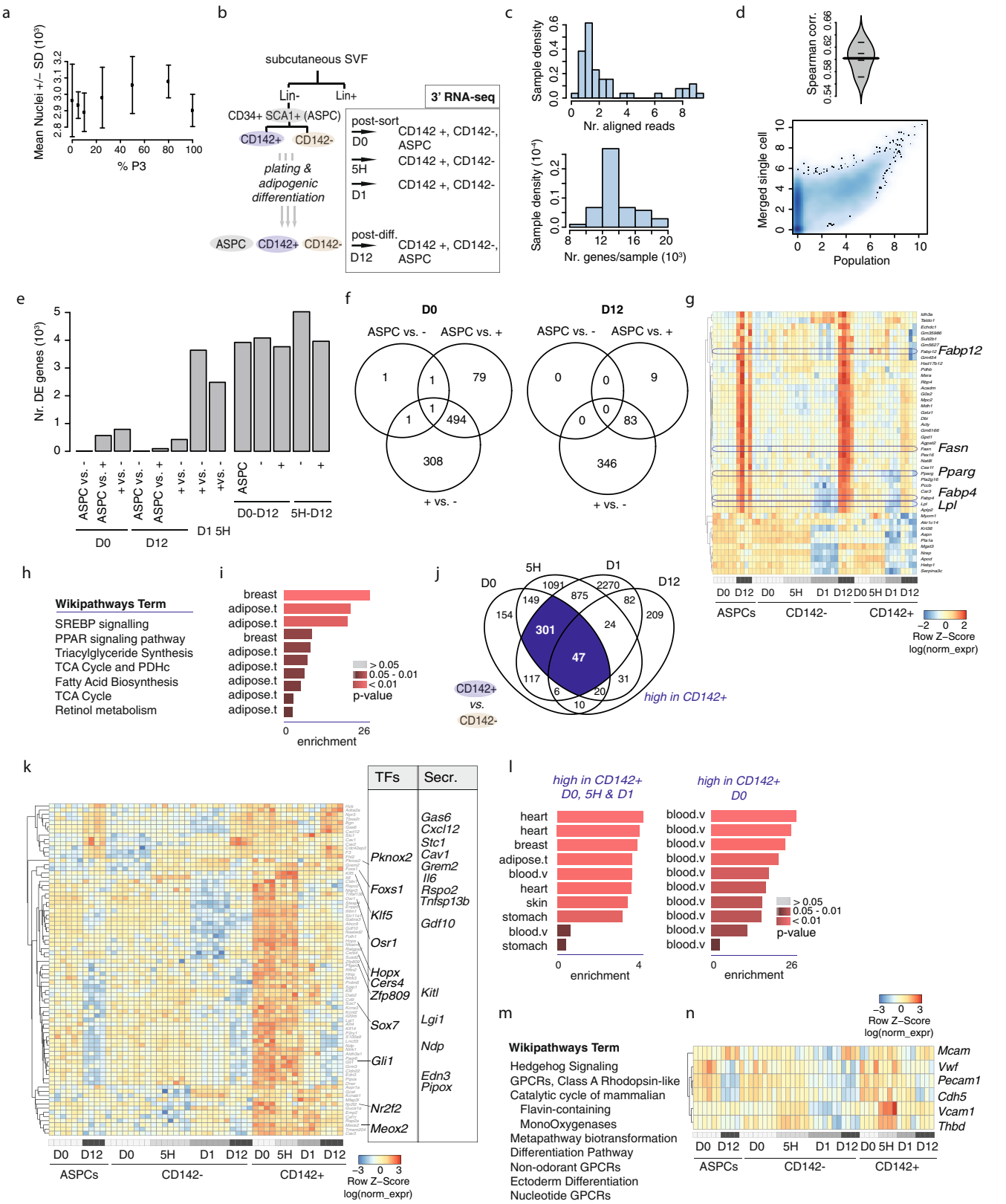


Extended Data Fig. 2 | ScRNA-seq reveals the heterogeneity of ASPCs (10x Genomics). **a**, **b**, Number of aligned reads per cell (**a**) and number of expressed genes per cell (**b**) for the scRNA-seq experiment performed using the 10x Genomics Chromium instrument. **c**, Expression of ubiquitous *Actb*, negative (*Ptprc* and *Pecam1*), positive (*Cd34*, *Ly6a* and *Itgb1*) FACS markers and adipocyte markers (*Adipoq*, *Retn* and *Cidec*). **d**, Number of cells in each of the four 10x Genomics populations (G1–G4). **e**, Percentage of the top 100 marker genes of the 10x Genomics populations (G1 and G4, G2, and G3) that overlap the top 100 marker genes of the C1 cell populations (P1–P3). **f**, Bean plots showing the distribution of adipogenic scores across cells that belong to one of the four 10x Genomics populations (G1–G4). Only *P* values for G2 comparisons are marked. **g**, Box plots showing the distribution of log(normalized expression) values of pre-adipogenic and ASPC markers across the four 10x Genomics populations (G1–G4). $*P \leq 0.05$, $**P \leq 0.01$, Wilcoxon rank-sum test. In **e**–**g**, G1, $n = 699$; G2, $n = 664$; G3, $n = 122$; and G4, $n = 319$ cells. **h**, *t*-SNE 2D maps of all analysed cells (10x Genomics), highlighting the expression of the pre-adipogenic markers *Pdgfra* and *Pdgfrb* (black). In **c**, **h**, $n = 1{,}804$ cells from one biological experiment based on a pool of mice.



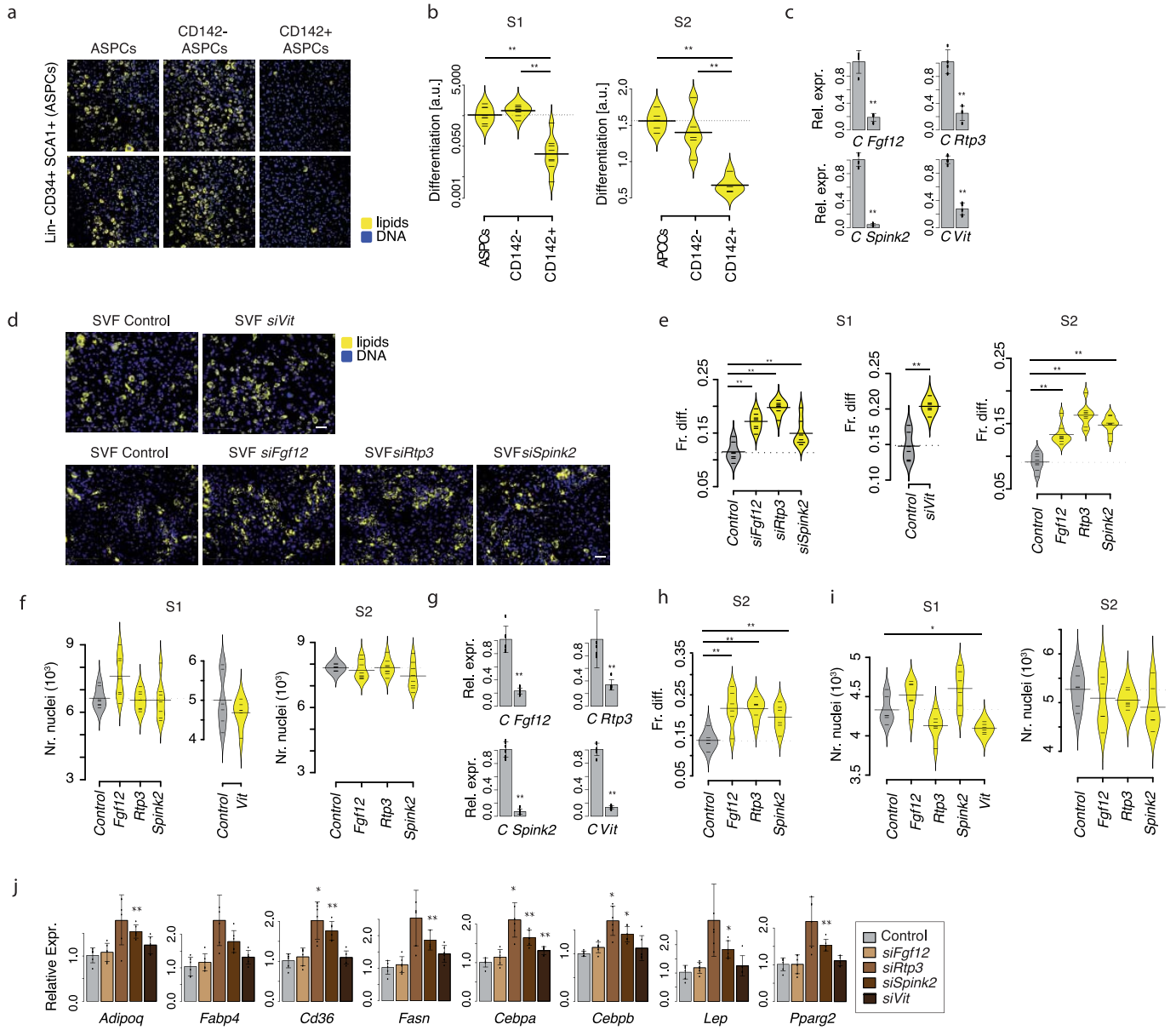
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Three ASPC subpopulations with distinct adipogenic differentiation capacity. a, Box plots showing the distribution of log(normalized expression) values for surface markers with available FACS-grade antibodies selected for subpopulation follow-up: CD55 and IL13RA1 (P1, G1 and G4), VAP1 and ADAM12 (P2 and G2) and ABCG1 (P3 and G3). $*P \leq 0.05$, $**P \leq 0.01$, Wilcoxon rank-sum test. P1, $n = 83$; P2, $n = 96$; P3, $n = 29$; G1, $n = 699$; G2, $n = 664$; G3, $n = 122$; and G4, $n = 319$ cells. **b**, FACS-based sorting strategy to isolate the three identified populations. **c**, qPCR-based expression fold-changes ($\log_2$(fold-change (FC) marker-positive versus marker-negative ASPCs)) for a panel of P1-, P2- and P3-specific genes. **d**, Microscopy images of distinct ASPC fractions after adipogenic differentiation; $CD55^+$ (P1 or G1 and G4), $VAP1^+$ (P2 or G2), and $ABCG1^+$ (P3 or G3). Nuclei are stained with DAPI (blue) and lipids with LD540 (yellow). Scale bar, 50 μm. **e**, Bean plots showing the distribution of the fraction of differentiated cells (Fr. diff.) per each ASPC subpopulation shown in **d**, $n = 4$ independent wells. M−, marker negative; M+, marker positive. $*P \leq 0.05$, $**P \leq 0.01$, *t*-test. **f**, qPCR-based expression fold-changes ($\log_2$(fold-change, marker-positive versus marker-negative ASPCs)) upon adipogenic differentiation for a panel of adipogenic marker genes. **g**, Bean plots showing the distribution of the mean nuclei number for the four differentiated ASPC fractions shown in **d**, **e** and Fig. 2c, d. $n = 4$ or 5 independent wells. In **d**, **e**, **g**, the experiments were repeated independently three times, yielding similar results; representative images are shown. All panels, population 1 (P1, green), population 2 (P2, red), population 3 (P3, blue).

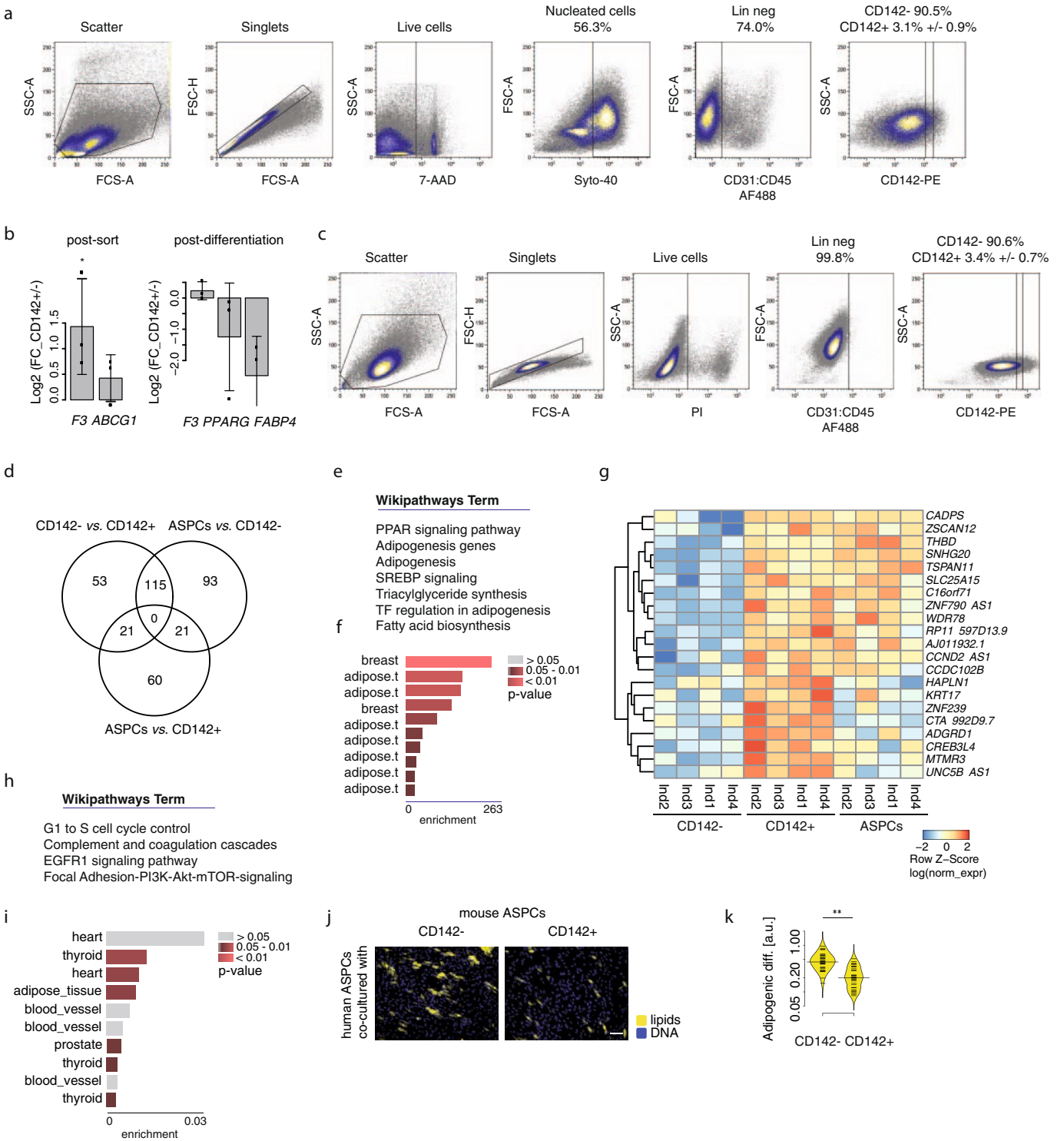


Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Aregs (CD142$^+$ABCG1$^+$ ASPCs) show adipogenic inhibitory capacity. a, Mean and s.d. of number of nuclei per experiment; *x* axis represents the percentage of Lin$^-$SCA1$^+$CD142$^+$ABCG1$^+$ cells (P3, Aregs) mixed in with Lin$^-$SCA1$^+$CD142$^-$ABCG1$^-$ cells; $n = 4$ independent wells. **b**, Schematic of sample collection aiming at a detailed characterization of CD142$^+$ ASPCs versus CD142$^-$ ASPCs and all ASPCs, after sorting (D0), upon plating (5 h), after culturing (D1) and after adipogenic differentiation (D12). **c**, Number of aligned reads (top) and number of expressed genes per sample (bottom) across barcoded RNA-seq samples. **d**, Correlation (Spearman's rho) between the barcoded mRNA-seq samples after sorting, and the merged scRNA-seq P3 cells; $n = 4$ biological replicates (top). Example of the correlation (Spearman's rho) between the log(normalized expression) estimates across merged scRNA-seq P3 cells and one barcoded mRNA-seq replicate (bottom). **e**, Number of genes ($\times 10^3$) that were significantly differentially expressed (FDR 0.05, fold-change 2) in all comparisons between all ASPCs, CD142$^+$ ASPCs and CD142$^-$ ASPCs, after sorting (D0), upon plating (5 h), after culture (D1) and after adipogenic differentiation (D12) (left) as well as upon adipogenic induction (D0 versus D12, D0–D12, and 5 h versus D12, 5 h–D12), in all ASPCs, CD142$^+$ ASPCs and CD142$^-$ ASPCs (right). **f**, Venn diagram showing overlaps between significantly differentially expressed genes (FDR 0.05, fold-change 2) in CD142$^+$ ASPCs (+) versus CD142$^-$ ASPCs (−) and all ASPCs (ASPC), after sorting (D0) and after adipogenic differentiation (D12). **g**, Heat map of the expression (row-wise *z*-scores of log(normalized expression), blue-to-red) of genes significantly differentially expressed (FDR 0.05, fold-change 2) in CD142$^+$ ASPCs versus CD142$^-$ ASPCs and versus all ASPCs upon adipogenic differentiation (D12). Adipogenic marker genes are highlighted. **h**, **i**, Pathways[47] (**h**) and top 10 significantly enriched GTEx tissue samples (**i**) that are significantly enriched among the genes displayed in **f**. **j**, Venn diagram showing overlaps between significantly differentially expressed genes (FDR 0.05, fold-change 2) in CD142$^+$ ASPCs versus CD142$^-$ ASPCs at all four time points that were assessed. **k**, Heat map of the expression (row-wise *z*-scores of log(normalized expression), blue-to-red) of genes that were expressed at significantly higher levels (FDR 0.05, fold-change 2) in CD142$^+$ ASPCs versus CD142$^-$ ASPCs after sorting (D0). Genes encoding transcription factors (TFs) and secreted factors (Secr.) are highlighted. **l**, Top 10 significantly enriched GTEx tissue samples among the genes that were expressed at significantly higher levels (FDR 0.05, fold-change 2) in CD142$^+$ ASPCs versus CD142$^-$ ASPCs after sorting, plating and culture (D0, 5 h and D1, respectively) (left) as well as after sorting only (D0) (right). **m**, Pathways[47] significantly enriched among the genes displayed in **k**. For **h**, **i**, **l**, **m**, full enrichment results are listed in Supplementary Table 12. **n**, Heat map of the expression (row-wise *z*-scores of log(normalized expression), blue-to-red) of a panel of endothelial marker genes. In **g**, **k**, **n**, $n = 4$–8, 4 biological replicates, 1–3 independent wells each.

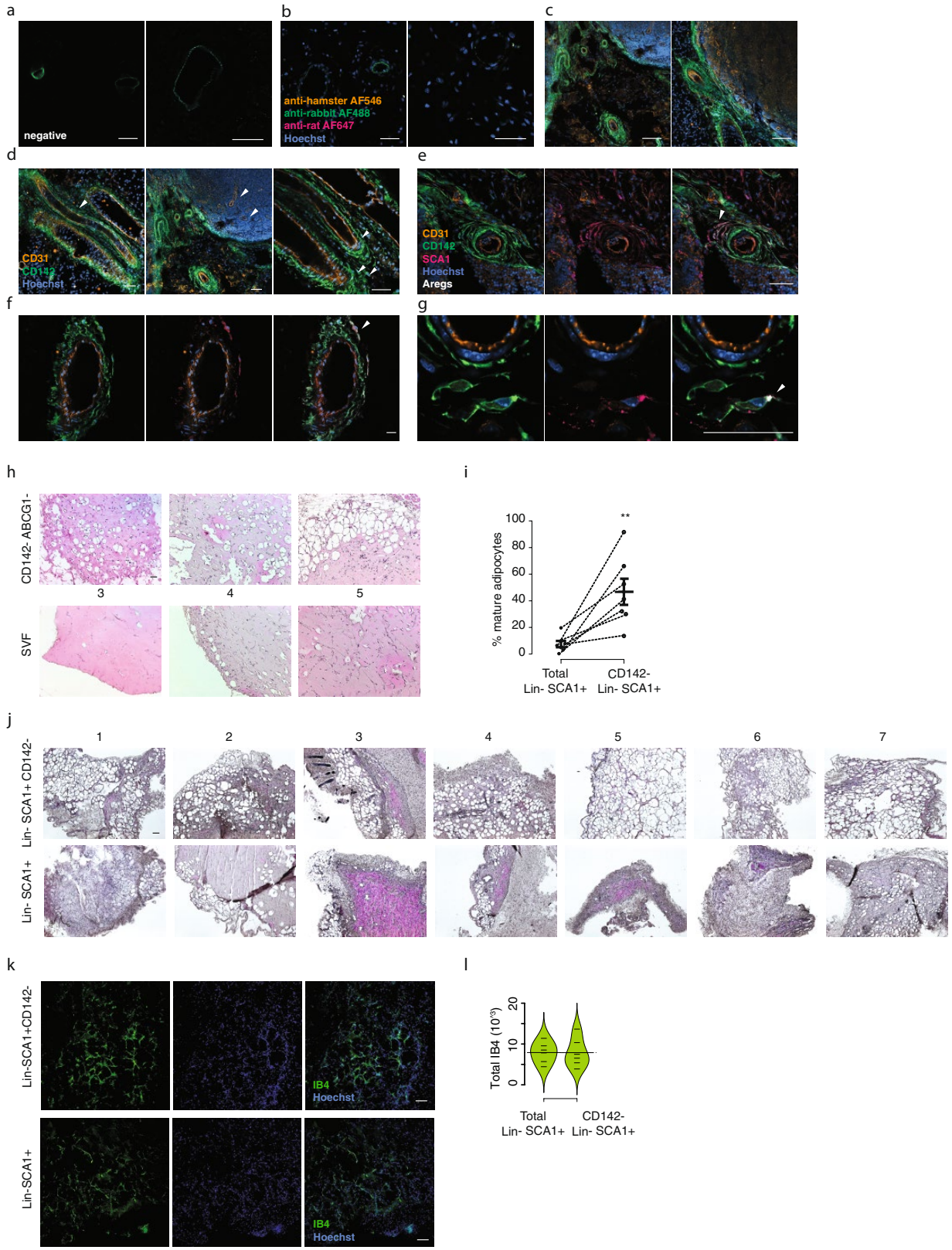


Extended Data Fig. 5 | The adipogenic inhibitory capacity of Aregs is paracrine. a, Microscopy images (after adipogenic differentiation) of mouse ASPCs co-cultured in transwells with all ASPCs, CD142$^-$ ASPCs or CD142$^+$ ASPCs. Nuclei are stained with Hoechst (blue) and lipids with Bodipy (yellow). Scale bar, 50 μm. The experiments were repeated independently two times, yielding similar results; representative images are shown. **b**, Bean plots showing the extent of differentiation (arbitrary units, a.u.) per each ASPC fraction shown in **a**, as well as an additional independent biological replicate. $^{**}P \leq 0.01$, Wilcoxon rank-sum test. Experiment S1: ASPCs and CD142$^-$ ASPCs, $n = 8$; CD142$^+$ ASPCs, $n = 7$ fields of view; experiment S2: ASPCs and CD142$^-$ ASPCs, $n = 5$; CD142$^+$ ASPCs, $n = 7$ fields of view. **c**, qPCR-measured gene expression (Rel. expr.) of knockdowns of *Fgf12*, *Rtp3*, *Spink2* ($n = 4$) and *Vit* ($n = 6$) in total SVF. Two biological replicates, 2 or 3 independent wells each. **d**, Microscopy images of plated SVF cells with knockdowns of *Fgf12*, *Rtp3*, *Spink2* and *Vit*, as well as control (scramble siRNA) knockdowns, after adipogenesis. Nuclei are stained with DAPI (blue) and lipids are stained with LD540 (yellow). Scale bars, 50 μm. **e**, Bean plots showing the distribution of the fraction of differentiated SVF cells with knockdowns of *Fgf12*, *Rtp3*, *Spink2* and *Vit*, as well as control knockdowns. **f**, Bean plots showing the distribution of the mean number of nuclei for the differentiated fractions shown in **e**. In **e**, **f**, for *Fgf12*, *Rtp3* and *Spink2* in experiment S1: $n = 7$, 2 biological replicates, 3 or 4 independent wells each; in experiment S2: $n = 8$, 2 biological replicates, 4 independent wells each; for *Vit*: $n = 6$, 2 biological replicates, 3 independent wells each. **g**, qPCR-measured gene expression of *Fgf12*, *Rtp3*, *Spink2* and *Vit* in Aregs of the experiments shown in Fig. 2h, i; $n = 6$, 2 biological replicates, 3 independent wells each. **h**, Bean plots showing the distribution of the fraction of differentiated cells, as measured in the CD142- ASPCs underneath Aregs with the knockdowns of *Fgf12*, *Rtp3*, *Spink2*, and control; $n = 6$, 2 biological replicates, 3 independent wells each (independent replication of the experiment shown in Fig. 2h, i). **i**, Bean plots showing the distribution of the mean number of nuclei for the differentiated fractions shown in Fig. 2h, i (experiment S1) and Extended Data Fig. 5h (experiment S2). $n = 6$, 2 biological replicates, 3 independent wells. **j**, Expression of adipogenic markers measured using qPCR in the CD142$^-$ ASPCs underneath Aregs with knockdown of genes shown in Fig. 2h, i; $n = 6$: 2 biological replicates, 3 independent wells each; Experiments were repeated two times, yielding similar results. In **c**–**j**, $^*P \leq 0.05$, $^{**}P \leq 0.01$, *t*-test.



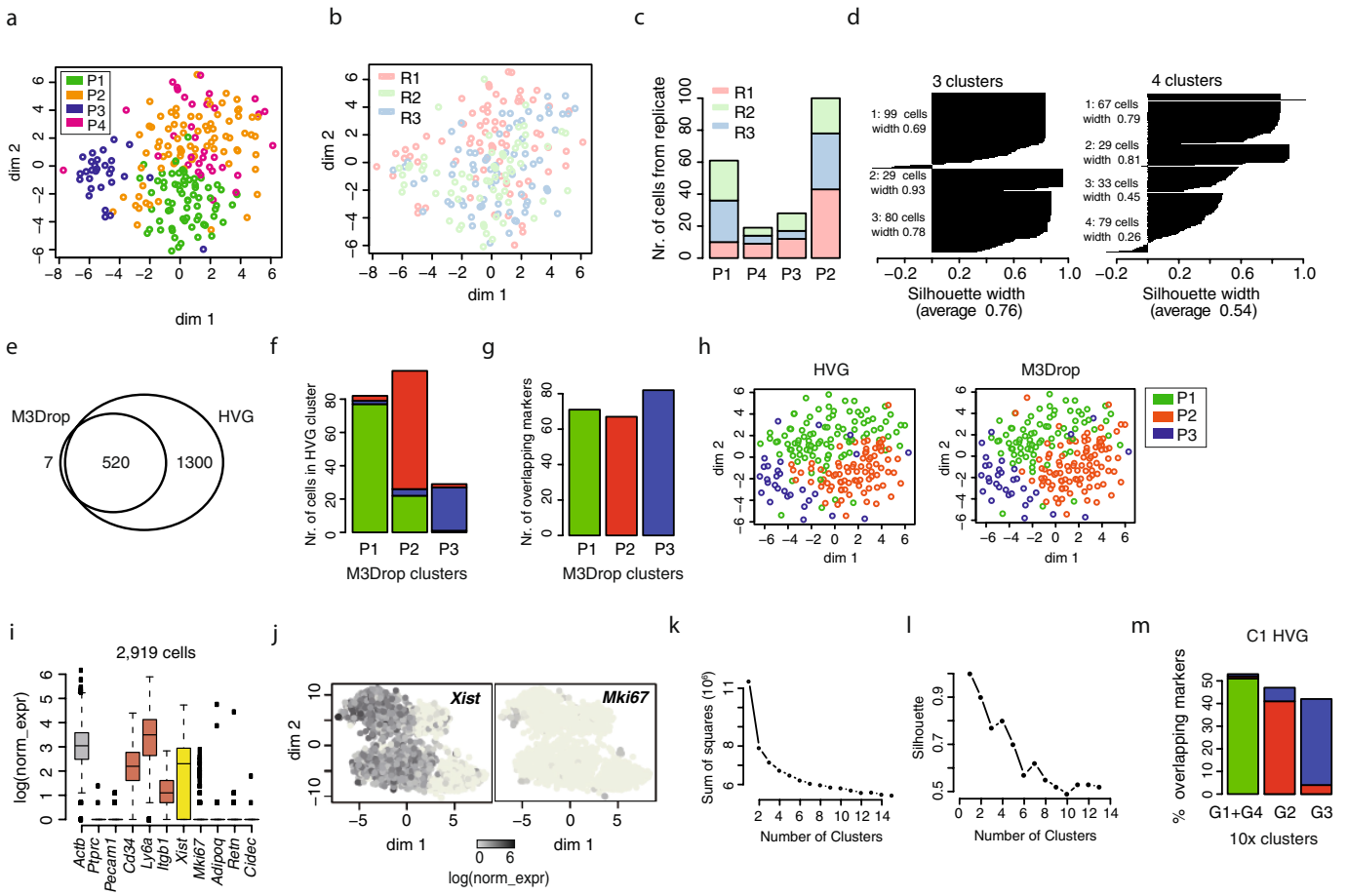
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Aregs and their inhibitory capacity are conserved in humans. a, FACS-based strategy to isolate human (one individual is shown) ex vivo $CD142^+$ ASPCs and $CD142^-$ ASPCs. **b**, qPCR-based expression fold-changes ($\log_2$(fold change)) of $CD142^+$ ASPCs versus $CD142^-$ ASPCs for *F3* and *ABCG1* after sorting (left) and *F3*, *PPARG* and *FABP4* after differentiation (right). $n = 3$ biological replicates (distinct individuals). $^*P \leq 0.05$, one-sided paired *t*-test. **c**, FACS-based strategy to isolate the human (one individual is shown) in vitro $CD142^+$ ASPCs and $CD142^-$ ASPCs. **d**, Venn diagram showing overlaps between significantly differentially expressed genes (FDR 0.1, fold-change 2) in $CD142^+$ ASPCs versus $CD142^-$ ASPCs and all ASPCs after adipogenic differentiation (D12). **e**, **f**, Pathways[47] (**e**) and top 10 GTEx tissue samples (**f**) significantly enriched among the genes that were expressed at significantly higher levels (FDR 0.1, fold-change 2) in $CD142^-$ ASPCs versus $CD142^+$ ASPCs after adipogenic differentiation (Fig. 3e). **g**, Heat map of the expression (row-wise *z*-scores of log(normalized expression), blue-to-red) of genes that were expressed at significantly higher levels (FDR 0.1, fold-change 2) in $CD142^+$ ASPCs versus $CD142^-$ ASPCs after adipogenic differentiation (D12). **h**, **i**, Pathways[47] (**h**) and top 10 GTEx tissue samples (**i**) significantly enriched among the genes that were expressed at significantly higher levels (FDR 0.1, fold change 2) in $CD142^+$ ASPCs versus $CD142^-$ ASPCs after adipogenic differentiation (D12) and displayed in **g**. For **e**, **f**, **h**, **i**, full enrichment results are provided in Supplementary Table 16. In **d**–**i**, $n = 4$ biological replicates (distinct individuals). **j**, Microscopy images (after adipogenic differentiation) of human ex vivo ASPCs co-cultured with mouse Areg- ($CD142^+$) or Areg-depleted ($CD142^-$) ASPCs. Nuclei are stained with Hoechst (blue) and lipids with Bodipy (yellow). Scale bar, 50 μm. The experiment was performed once. **k**, Bean plots showing the extent of adipogenic differentiation (arbitrary units, a.u.) per each ASPC fraction shown in **j**. $n = 35$ fields of view. All panels except b: $^*P \leq 0.05$, $^{**}P \leq 0.01$, Wilcoxon rank-sum test.



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Aregs locate proximal to the vasculature and inhibit adipogenesis in vivo. a–g, Microscopy images of the in vivo localization of CD31 (orange), CD142 (green) and SCA1 (pink) markers in mouse subcutaneous white adipose tissue. Nuclei are stained with Hoechst (blue). Scale bar, 50 μm. The experiments were repeated independently at least three times, yielding similar results; representative images are shown. Negative control (**a**) or tissue cryosections stained with the indicated secondary antibodies only (**b**). In **c**, perfused (left) versus non-perfused (right) tissue stained for the indicated markers. In **d**, arrows indicate $CD142^+$ around big vessels within the adipose parenchyma (left), outside the lymph nodes (middle) and within individual cells (right); in **e–g**, arrows indicate individual cells presenting both CD142 and SCA1 staining (white). **h**, Total SVF and Areg-depleted SVF cells ($CD142^-ABCG1^-$) were implanted into the subcutaneous adipose depot in Matrigel. Microscopy images of Matrigel plugs with total SVF or $CD142^-ABCG1^-$ SVF cells. Scale bar, 100 μm. Individuals 3 to 5 are shown, and individuals 1 and 2 are presented in Fig. 4g. **i**, Mean and s.d. of the percentage of mature adipocytes developed within the plugs composed of either Lin^-SCA1^+ or $Lin^-SCA1^+CD142^-$ ASPCs ($n = 7$ biological replicates) analysed using CellProfiler[25]. $^*P \leq 0.01$, paired *t*-test. Mean and s.d. are shown. **j**, Microscopy images of Matrigel plugs corresponding to **i**. Scale bar, 100 μm. **k**, Microscopy images of Matrigel plugs stained with isolectin GS-IB4 (green) and Hoechst (blue) corresponding to **i**. Scale bar, 100 μm. **l**, Quantification of vascularisation within the plugs, composed of either Lin^-SCA1^+ or $Lin^-SCA1^+CD142^-$ ASPCs ($n = 6$ biological replicates) analysed using CellProfiler[25].



Extended Data Fig. 8 | Supplementary methods for the scRNA-seq analyses. a, **b**, *t*-SNE 2D maps of all analysed cells (Fluidigm C1), highlighting the four subpopulations (P1–P4) obtained when clustering analysis was performed with $k = 4$ (**a**) and cells belonging to one of three biological replicates (**b**) (pink, R1; green, R2; blue, R3). **c**, For each of the four clusters shown in **a**, the number of cells stemming from each biological replicate is highlighted. **d**, Silhouette analysis results for $k = 3$ and $k = 4$. **e**, Venn diagram showing the overlap between the 527 significantly differentially expressed genes identified using M3Drop in the Letter, and the 1,827 genes showing high variability (HVG). **f**, For each cluster shown in Fig. 1b, we determined the number of cells contained in the alternative clustering obtained by considering 1,827 highly variable genes for SC3 analysis. The vast majority of cells are similarly attributed, irrespective of gene set choice. **g**, Percentage of the top 100 marker genes of the three subpopulations described in the Letter that overlap with the top 100 marker genes of the subpopulations derived from the analysis based on highly variable genes. **h**, *t*-SNE 2D maps of all analysed cells (C1) considering only the highly variable genes, highlighting the three subpopulations (P1–P3) identified by cluster analysis in the Letter (M3Drop, right) and using the highly variable genes (HVG, left). In **a**, **b**, **h**, $n = 208$ cells from three independent biological experiments, each based on a pool of mice. **i**, Marker gene expression across all 10x Genomics cells, including a housekeeping gene (*Actb*), the negative FACS markers *Ptprc* (*Cd45*) and *Pecam1* (*Cd31*), the positive FACS markers *Cd34*, *Ly6a* (*Sca1*) and *Itgb1* (*Cd29*), and the mature adipogenic markers *Adipoq*, *Retn* and *Cidec*. **j**, *t*-SNE 2D maps of all 10x Genomics cells highlighting the expression of *Xist* and the cell division marker *Mki67* (black). **k**, Sum of squares for distinct number of clusters (2 to 15) in the 10x Genomics data. In **i**, **j**, $n = 2{,}919$ cells from one biological experiment based on a pool of mice, before G1, *Xist*, *Krt18* or *Krt19* and *Epcam* filtering. **l**, Silhouette width for distinct number of clusters (2 to 13) in the 10x Genomics data. **m**, Percentage of the top 100 marker genes of the 10x Genomics populations (G1–G4) that overlaps with the top 100 marker genes of the populations (P1–P3) derived from the C1 data based on the highly variable genes.

nature research

2017-05-05724B
Corresponding author(s): Christian Wolfrum, Bart Deplancke

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	FACS data acquisition: Becton Dickinson FACSDiva Sc-transcritomics: C1 System Software, 10X Genomics Chromium software, Sequencing: NextSeq 500 software Image acquisition: Olympus cellSens, Zeiss ZEN 2009, Leica LAS-AF, LAS 2009, PerkinElmer Harmony 4.5 qPCR acquisition: QuantStudio 6 and 7 Flex Real-Time PCR System software
Data analysis	FACS data analysis: Beckman Coulter Kaluza Analysis (sc)RNA-seq: STAR 2.4.0g, htseq-count 0.6.0, cellranger-2.1.0 (10x data), M3Drop_1.0.1, Rtsne_0.11 , SC3_1.3.6, scran_1.2.2, limma_3.30.4, sva_3.22.0, pheatmap_1.0.8, biomaRt_2.30.0, Enrichr (Kuleshov et al., 2016) Image processing and quantification: Fiji (ImageJ 2.0.0-rc-64/1.51s), CellProfiler 2.0, PerkinElmer Harmony 4.5

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All raw and processed RNA-seq data have been uploaded in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) with the accession numbers E-MTAB-5785, E-MTAB-5818, E-MTAB-5802, E-MTAB-5787, and E-MTAB-6677.
Fig. 1, Fig. 2, Fig. 3, Extended Data Fig. 1, Extended Data Fig. 2, Extended Data Fig. 3, Extended Data Fig. 4, Extended Data Fig. 6, Extended Data Fig. 8 have associated raw data.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The employed SVF isolation protocol results in ~2 million cells/mouse. Given cell concentration and input numbers required for the scRNA-seq experiments (per manufacturer's specification), we used 3-6 mice for the Fluidigm C1 (between the three biological replicates) and 6 mice for the 10x experiments. For the validation and functional follow-up experiments, we aimed at plating 13.5 k/cm2 (ex vivo) and 83.5 k/cm2 (in vitro), experimentally determined optimal cell densities. We thus inferred the number of mice required for each of the performed experiments, given distinct cell fractionations and replicate numbers. For subsequent bulk RNA sequencing experiments validating and characterizing Aregs, we performed four biological replicate (experiments conducted on different days), as commonly employed in the field.
Data exclusions	All animals were included in the experiments. In the single-cell analysis, we excluded cells that appeared as doublets in the microscopic images (33) and those that had <40% or <400.000 aligned reads, resulting in a total of 208 cells, in the Fluidigm C1 experiments. In the 10x experiments, we filtered outlier cells using the median absolute deviation from the median total library size and total gene numbers, as well as based on Xist, Epcam, Krt19/Krt18 expression, and predicted cell cycle phase, such that the data is conceptually comparable to the Fluidigm C1 one. The final reported dataset consists of 1,804 cells.
Replication	The scRNA-seq experiment was performed in 3 biological replicates (Fluidigm C1) and independently validated on a larger cell population, with a distinct methodology (10x). Functional follow-ups of the CD142+ population were replicated across two distinct laboratories typically ranging from two to five independent biological replicates. All attempted replicates were successful except the mouse-human transwell experiment, where a replicate was not successfully conducted.
Randomization	No randomization was applied. All the animals used in the study were wild-type, male and female, age- and weight-matched.
Blinding	The in vivo cell implantation and subsequent mouse surgery was performed by the same person for experimental reasons therefore implantation was not blinded. Computational analysis of the section was performed blinded.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Primary antibodies
Antibody, Application, Conjugate, Supplier, Ref # ,Clone, Lot #
anti-ABCG1 FACS Unconjugated Invitrogen PA5-13462 - SF2410263
anti-CD29 FACS PerCP-efluor710 eBiosciences 46-0291 HMb1-1 E1287-104
anti-CD31 FACS AF488 BioLegend 303110 WM59 B213987
anti-CD31 FACS AF488 BioLegend 102514 MEC13.3 B191232, B190518, B161677
anti-CD31 Histology Unconjugated Bio-Rad MCA1370Z 2H8 980601W
anti-CD34 FACS BV421 BioLegend 119321 MEC14.7 B208699, B202457
anti-CD45 FACS AF488 BioLegend 368512 2D1 B182155, B221064
anti-CD45 FACS AF488 BioLegend 103122 30-F11 B202059, B146208
anti-CD55 FACS Unconjugated BioLegend 131802 RIKO-3 B157653
anti-CD142 FACS PE BioLegend 365204 NY2 B192204, B192205
anti-CD142 FACS PE SinoBiological 50413-R001-P 001 HG08OC1605, HA10OC2402-B
anti-CD142 Histology Unconjugated SinoBiological 50413-R001 001 HA10MA0201
anti-SCA1 FACS PE-Cy7 BioLegend 122514 E13-161.7 B194434, B154904
anti-SCA1 Histology Unconjugated Abcam ab51317 E13-161.7 GR285475-2
anti-TER119 FACS AF488 BioLegend 116215 TER-119 B175180, B202457, B149788
anti-VAP1 FACS Unconjugated Abcam ab81673 7-88 GR216611-1

Secondary antibodies
Antibody Application Conjugate Company Ref # Lot #
anti-hamster Histology AF546 Molecular Probes A21111 1729639
anti-rabbit Histology AF488 Molecular Probes A21247 - 1796375
anti-rat Histology AF647 Molecular Probes A21206 - 1780352, 1750833

Validation

Primary antibodies
Antibody ,Tested Application, Publication DOI
anti-ABCG1 FC, ICC, IF, IHC, WB N/A
anti-CD29 FC 10.3389/fendo.2015.00129
anti-CD31 FC 10.1038/nbt.3576
anti-CD31 FC 10.1038/ncomms11302
anti-CD31 FC, IF 10.1038/nature08878
anti-CD34 FC, IF 10.4049/jimmunol.1100558
anti-CD45 FC 10.1038/ncb3354
anti-CD45 FC 10.1038/ncomms11533
anti-CD55 FC 10.1152/ajpcell.00213.2011
anti-CD142 FC 10.1038/srep40329
anti-CD142 FC N/A
anti-CD142 FC N/A
anti-SCA1 FC 10.1371/journal.pbio.1002562
anti-SCA1 ICC, IF 10.1038/nm.3866
anti-TER119 FC 10.1016/j.bbalip.2016.06.010
anti-VAP1 IHC, FC, ICC, IF, IP 10.1016/S0002-9440(10)62300-0

Secondary antibodies
Antibody, Publication DOI
anti-hamster https://www.ncbi.nlm.nih.gov/pubmed/27239212
anti-rabbit https://www.ncbi.nlm.nih.gov/pubmed/28445462
anti-rat https://www.ncbi.nlm.nih.gov/pubmed/28448484

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

The new mouse strain (C57BL/6J) Tg(Pref1-CreER)426 Biat as Pref1-CreER was generated and used (10-11-week-old, males and females) in the Fluidigm C1 scRNA-seq experiment.
All mice used in the remaining experiments were wild-type C57BL/6J males and females (median of age: 8.7 weeks ranging 7-11 weeks, median of weight: 22.5 g ranging 20.4-23.4 g).

Wild animals

No wild animals were used in the study.

Field-collected samples

The study did not involve field-collected samples.

Human research participants

Policy information about studies involving human research participants

Population characteristics

Ind1: Female, 36 yo, European, non-diabetic
Ind2: Female, 71 yo, European, non-diabetic

	Ind3: Male, 38 yo, European, non-diabetic Ind4: Male, 47 yo, American, non-diabetic Ind5: Female, 51 yo, Yorubian, non-diabetic
Recruitment	The participants voluntarily underwent liposuction and donated the material for the study.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Metods: Mouse SVF isolation, Human SVF isolation and Mouse FACS-based cell isolation, Human FACS-based cell isolation
Instrument	Becton Dickinson FACSAria II
Software	Acquisition: Becton Dickinson FACSDiva software Analsyis: Beckman Coulter Kaluza Analysis software
Cell population abundance	Mouse (Extended Data Fig. 3b): CD55+ ASPCs = 8.7% +/- 1.9%, VAP1+ ASPCs = 36.5% +/- 4.1%, CD142+ ABCG1+ ASPCs = 2.2% +/- 0.7%, CD142+ ASPCs = 9.6% +/- 1.3%; Human (Extended Data Fig. 6a,c): CD142+ ASPCs: ex-vivo: 3.10% +/- 0.85%, in-vitro: 3.38% +/- 0.64%
Gating strategy	Gating strategy description: Methods: Mouse FACS-based cell isolation, Human FACS-based cell isolation Gating strategy graphical representation: Mouse: Extended Data Fig. 3b, Human: Extended Data Fig. 6a,c.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.