Sox9 regulates cell state and activity of embryonic mouse mammary progenitor cells

Authors:

Naoko Kogata[1,2], Philip Bland[1], Mandy Tsang[1], Erik Oliemuller[1], Anne Lowe[1], and Beatrice A. Howard[1*]

Affiliations:

[1]The Breast Cancer Now Toby Robins Research Centre, Division of Breast Cancer Research, The Institute of Cancer Research, London, UK

[2]current address: Cellular Signalling and Cytoskeletal Function Lab, The Francis Crick Institute, London, UK

[*]Author for correspondence:

Beatrice A. Howard

e-mail: beatrice.howard@icr.ac.uk

phone: +44-203-437-7359

Keywords:

embryonic mammary gland, embryonic mammary progenitor cell, mammary lineage, multipotency, Sox9

Abstract

Embryonic mammary cells are a unique population comprised of undifferentiated, highly plastic progenitor cells that create normal mammary tissues. The mammary gland continues to develop after birth from descendants of embryonic mammary cells. Here, we establish novel cell lines from mouse mammary organs, immediately after they formed during prenatal development, to facilitate studies of primitive mammary cells, which are difficult to isolate in sufficient quantities for use in functional experiments. We show some lines can be induced to secrete milk, a distinguishing feature of mammary epithelial cells. Targeted deletion of *Sox9*, from one clone, leads to decreased ability to respond to lactogenic stimuli, consistent with a previously identified role for *Sox9* in regulating luminal progenitor function. Sox9 ablation also leads to alterations in 3D morphology and downregulation of Zeb1, a key epithelial-mesenchymal transition regulator. Prenatal mammary cell lines are an invaluable resource to study regulation of mammary progenitor cell biology and development.

Introduction

Embryonic breast epithelial cells are a unique cell population comprised of undifferentiated and highly plastic progenitor cells that ultimately give rise to all other postnatal breast epithelial cells. Lineage tracing studies have indicated that embryonic mammary cells are multipotent *in vivo*[1-3]. An important area of research in mammary gland biology is to determine the roles of genes and signalling pathways that regulate embryonic stages of mammary gland development, as many of these are also relevant to processes that are deregulated in cancer[4,5]. Despite their relevance to breast cancer research, the routine use of primary mid-gestation embryonic mammary cells for functional study is not currently feasible, due to the small size of the nascent organ.

In mice, mammary gland development commences at embryonic day 11 (E11) with the sequential appearance of five pairs of mammary primordia[6]. Local epithelial thickenings invaginate to the underlying tissue to form buds, which from E12.5 onwards are surrounded by a specialized condensed mammary mesenchyme. Mammary buds grow relatively slowly in size until E14 when the epithelial cells within the bud start to proliferate extensively and then invade into the underlying mesenchymal tissues[6]. These early stages of development are of particular interest as the cells have a number of unique properties. The epithelial cells within the E11-E13 stage mammary organ are largely quiescent[7,8]. At these stages of development, epithelial cells are thought to accrue within the mammary organ via localised cell movements[9,10]. Dissociated embryonic mammary cells from E12-E13-stage organs have minimal ability to efficiently repopulate cleared mammary fat pads, whilst cells from E16-18-stage organs have a much higher ability to engraft[11,12]. Intact mammary bud epithelium from E13-stage embryos can repopulate cleared fat pads suggesting a stem cell population has been delimited by mid-gestation[13]. Recent results from lineage tracing experiments indicate that embryonic mammary cells at E12-E13 stages are multipotent at the cellular level and become lineage-restricted prior to birth[1-3]. Although embryonic mammary progenitor cells from E12-13-stages harbour very distinct biological properties from embryonic mammary progenitor cells isolated from E16-17-stages and

from the postnatal mammary gland, a lack of appropriate *in vitro* models has limited their accessibility for many researchers.

Most studies of the embryonic mammary gland have relied on analyses of embryos from genetically modified mice or embryonic mammary organ explant cultures[14-17]. These methods require considerable training, expertise and the use of animals. However, using these novel embryonic mammary cell lines and standard two- and three-dimensional culture techniques, we model several key aspects of embryonic mammary gland development *in vitro.* Using CRISPR-Cas9 genome editing, we investigate the role of Sox9, an embryonic Sox transcription factor, which has been implicated in conferring stem cell state to differentiated postnatal mammary epithelial cells[18], in the regulation of stem cell activity and the differentiation potential of cells formed during early stages of embryonic mammary gland development. Our findings highlight the distinct biological features and context dependent regulation of embryonic mammary progenitor cells and are a novel resource for studying this unique cell population.

Results

Novel embryonic mammary progenitor cell lines established

To date, the establishment and maintenance of primary mammary embryonic epithelial cell cultures from mouse embryos at stages between E12-13 has not been possible. To overcome this issue, we have taken advantage of the Immortomouse, a genetically modified mouse, to introduce a temperature-sensitive Simian virus (SV) 40 antigen under control of an interferon regulated promoter that enables immortalisation of certain types of cells, including epithelial and mesenchymal cells[19]. Immortomice were bred with s-SHIP-GFP mice, which GFP expression marks epithelial progenitor cells, including those in the embryonic mammary primordia to facilitate and confirm dissection of the mammary organ when it is first morphologically distinct[20]. The majority of s-SHIP-GFP embryonic mammary epithelial cells are GFP+ and adjacent embryonic mammary mesenchymal cells are GFP- (Figure 1A and Supplementary Fig. 1A). Mammary primordia number three were microdissected from E12.0-stage embryos so that the mammary epithelium, mammary mesenchyme, and fat pad precursors were isolated and such that other GFP+ cell types marked in s-SHIP-GFP embryos[21] were excluded (Figure 1A). Excised mammary organs were embedded in thick Basement Membrane Extract (BME) for a month, so that embryonic mammary progenitor cells (eMPC) from the mammary organoid were able to proliferate (Figure 1B). This adaptation period appeared to be important for successful cell isolation since enzymatic digestion of Immorto:s-SHIP-GFP mammary organs immediately after microdissection did not give rise to cultivable eMPCs beyond one month. After proliferating for one month, eMPCs were harvested by enzymatically digesting the organoid culture grown in thick BME and expanding the cells in two-dimension culture. The bulk cells of this expanded eMPC are referred to subsequently as ePool (Figure 1B). To obtain clonal cell lines, single ePool cells were separated using fluorescence-activated cell sorter (FACS) according to s-SHIP-GFP expression status (GFP+ or GFP-) at the time of sorting (Figure 1B and Supplementary Fig. 1B). ~9% of GFP+ and ~2% of GFP- cells gave rise to viable clones. Sixteen clones were derived from the GFP+ fraction (designated eG1, eG2, eG2E9, etc). e1 and e2 were the only clones that survived from GFP- fraction (Supplementary Fig. 1C).

However, GFP expression is not necessarily indicative of tissue of origin since the activity of s-SHIP promoter is variable in cell populations expanded in 2D culture. eMPC clones expanded from single s-SHIP-GFP+ or single s-SHIP-GFP- cells contain both GFP+ and GFP- cells upon passage in 2D culture (Supplementary Fig. 1C).

Hierarchical cluster analysis of gene expression profiles obtained by RNA sequencing shows that eMPCs are distinct from other types of progenitor and stem cells, including embryonic stem cells (ESC), mouse embryonic fibroblasts (MEF), mesenchymal stem cells (MSC) and bone marrow mononuclear cells (BMMC) (Figure 2A). From 17204 significantly modulated genes, the 2059 most highly variable were selected when DESeq2 software was employed with the Likelihood ratio test (LRT), and using the intensity difference test (detailed selection strategy and gene lists are in Supplementary Dataset 1). Principal component analysis (PCA), based on 92 genes, produces a similar clustering of the eMPC clones with other stem cell types (Figure 2B and Supplementary Dataset 1).

Mammary organs from C57BL/6 E12.5-stage mouse embryos were microdissected so that mammary epithelium, mammary stroma (mammary mesenchyme, and adjacent fat pad precursor), and surface epithelium were separated and the tissues were used for gene expressing profiling using RNA-sequencing. Hierarchical clustering based on transcriptome profiles of the separated embryonic mammary primordia tissues and eMPCs and other stem cell types was performed (Supplementary Fig. 2 and Supplementary Dataset 2). To assess the global landscape of expression in this dataset, we performed PCA and 26 genes revealed clustering of eMPC lines into three distinct groups based on their expression of genes encoding Keratin and cell adhesion regulators (Figure 2C). Our previous results have shown that embryonic mammary cells are largely devoid of cells expressing differentiation markers, although a few cells present within the E12-stage embryonic mammary bud epithelium express stem cell/progenitor markers associated with mature postnatal mammary epithelial cell (MEC) lineages[22]. PCA identifies clusters based on expression of several genes, some of which control cell adhesion, including the protocadherin, *Pcdh7,* integrin, *Itga3*, and extracellular matrix (ECM) ligand, *Lama4,* that are highly expressed by a

subset of the cell lines (eG1, eG2C7, eG1C10), which also express Keratins (*Krt7*, *Krt18*, *Krt19*). These clones belonging to the cluster marked in green in Figure 2 may represent mammary epithelial progenitor cells. Clones belonging to another cluster, marked in yellow in Figure 2, including e1, which was selected further study, express high levels of a number of markers including, *Cd34, Cd74, Ebf2, Fabp4, Runx2,* and *Sox9* that are associated with a variety of stem cell types[23-28] and may represent less-differentiated progenitor/stem cells types. Two cells lines with less distinct features, (marked in orange in Figure 2), e2 and eG2 that cluster with the pool, were also selected for further study. eG1 was selected for further study since it expresses high levels of markers associated with an epithelial-state, including Keratins, the epithelial basement membrane protein *Lamc2,* and *Lgr4,* a regulator of mammary gland development and stem cell activity[29].

In total, two clones expanded from GFP+ cells (eG1 and eG2) and the two clones expanded from GFP- cells (e1 and e2) lines, along with the ePool, were used for further characterization based on their distinct morphologies, transcriptome profiles and cluster analysis results (Figure 3). The two GFP- cell lines (e1, e2) do not cluster together. However, single GFP+ and GFP- cells were expanded prior to other characterization including, RNA sequencing, and these cell populations contained a mixture of both GFP+ and GFP- cells (Supplementary Fig. 1C).

We selected a subset of markers that have direct links to stem cell biology and confirmed the RNA sequencing profiles using qRT-PCR (Figure 3B). We also used immunohistochemistry (IHC) to confirm expression of Sox9 and Twist1 protein in specific eMPC clones (Figure 3C). Sox9 and Twist1 protein expression were discordant with the RNA expression results, but mismatch between transcriptional and translational levels is a frequently occurring phenomena[30].

eMPCs harbour varying capacities for differentiation into the mesodermal lineage

We examined the capacity of eG1, eG2, e1, e2 and the ePool to differentiate into mesodermal lineages. Cells were exposed to differentiation media,

previously reported to induce MSCs toward adipocyte or endothelial cell lineages. Under adipogenic conditions, only e1 cells produced lipid-filled adipocytes in a substantial fraction of cells, indicating notable adipogenic ability, whereas eG1 and eG2 do not respond to adipogenic stimuli. (Figure 4A). Under endotheliogenic conditions, e1, eG2 and the ePool formed tubular network comparable to control Human umbilical vein endothelial cells (HUVECs) cultured in endothelium growth medium (Figure 4B). Both e1 and eG2 express a vascular progenitor cell marker *Cd34* prior to induction (Figure 3A). Overall, these functional studies suggest that clone e1 is the most responsive of the four cell lines to mesodermal lineage differentiation cues.

eMPCs are clonogenic and form distinct sphere morphologies

eG1, eG2, e1, e2, and the ePool were evaluated for mammary sphere-forming ability using the standard sphere formation technique. Each clone was cultured under sphere-forming conditions and formed spheres with highly distinct morphologies (Figure 5A). All eMPCs formed spheres, although e1 had the highest sphere-forming rate (3.93%) and eG1 the lowest (0.69%) (Figure 5B). Two clones (e1, eG1) were selected for more detailed functional assessments based on their distinct responses in functional assays and marker profiles. eG1 spheres show higher form factor, a measure of sphericity, when compared to those of e1, suggesting that eG1 spheres are relatively compact compared to the spheres formed from the other clones (Figure 5C-D). e1 cells had limited engraftment ability when injected into cleared mammary fat pads **(**Supplementary Figure 3**)**. No teratoma formation was observed, as would be expected to occur after xenografting of ESC cells.

eMPCs form differentiated mammary acini *in vitro*

eMPCs were evaluated for their ability to undergo functional alveolar differentiation by assessing cellular production of milk proteins by immunofluorescence staining with an antibody that detects milk specific proteins including, β-casein, free secretory component, and lactoferrin.

eMPCs were grown using conditions that induce lactogenic differentiation of mammary cells with prolactin. Under these conditions, e1, eG1, eG2 and ePool formed domes that produce milk protein, indicative of lactogenic differentiation (Figure 6A). In these assays, e1 and ePool produced the greatest number of alveoli-like structures, with an alveologenic capacity to produce milk similar to postnatal MECs. eG1, and eG2 also showed lactogenic ability, but it was not possible to directly compare with the others since it was not possible to maintain the confluent cell culture needed for assessment at identical timepoints with this assay (Figure 6B).

Sox9 ablation in e1 leads to reduced expression of Zeb1

Mammary tissues originate from multipotent embryonic progenitors, which give rise to unipotent basal and luminal stem cells found and postnatal mammary tissues[31]. *Sox9* was selected for further study since it is amongst the earliest transcription factors expressed by cells within the embryonic mammary epithelium and its expression increases so that most cells express high levels of Sox9 by E14.5 (Supplementary Fig. 4)[22]. We assessed the role of *Sox9* in regulating embryonic mammary stem cell function using e1, a line that expresses a number of markers typically associated with stem/progenitor cells, including high levels of *Sox9*. CRISPR-Cas9 was used to create 3 independent subclones with gene deletions (or knockouts (KO)) of *Sox9* derived from e1, e1/*Sox9*-KO#1, e1/*Sox9*-KO#2, and e1-*Sox9*-KO#3 (Figure 7A).

We confirmed a reduction in *Sox9* levels in the 3 subclones with *Sox9* deletions compared to non-targeting guide control cells by qRT-PCR (Figure 7B), immunohistochemistry (Figure 7C), and western blotting (Supplementary Fig. 5B). We also detected a marked reduction of expression of *Zeb1,* a key epithelial-to-mesenchymal transition (EMT) regulator, and concomitant with the level of reduction of Sox9 (Figure 7B). Using IHC, we confirmed total loss of Sox9 expression in e1/*Sox9*-KO#1 subclone, but detected Sox9+ cells in 1-4% of e1/*Sox9*-KO#2 and e1/*Sox9*-KO#3 cells (Figure 7C). Reduced

detection of Sox9 in the 3 *Sox9*-KO subclones is correlated with reduction of Zeb1-expressing cells (Figure 7C).

Sox9 ablation increases e1 clonogenicity

In non-targeting guide control cells derived from e1, (e1/Co#1 and e1/Co#2), sphere-forming efficiency is similar in both the parental e1 clone and non-targeted e1 subclones (Figure 5B and 7D). In all three independently-derived e1 subclones with *Sox9* deletions, sphere-forming efficiencies increased by approximately two-fold, indicating that deletion of *Sox9* in e1 enhances clonogenic ability, a measure of stem cell activity (Figure 7D). When plated as multicellular aggregates on low-attachment plates, no change in spheroid morphology is observed (Supplementary Fig. 6). However, when spheroids are embedded with BME or matrigel, the spheroids from *Sox9*-KO cells are larger and exhibit a greater number of cellular protrusions when compared to controls indicative of altered morphogenetic and enhanced migratory capacity (Figure 7E-F and Supplementary Figure 6).

Sox9 ablation decreases response to lactogenic stimuli

We assessed expression of several stem/lineage markers in e1/*Sox9*-KO#1, which lacks detectable Sox9 when cells are stained by IHC. *Acta2 (α-SMA)*, a marker normally expressed by basal stem cells and differentiated myoepithelial cells, was increased when Sox9 was ablated. Levels of *Procr*, a marker for a rare basal mammary stem cell population and *Trp63*, a basal cell marker, were both decreased, suggesting profound alterations of the basal features of these cells (Figure 8A). We found other EMT regulators including, *Snail2/Slug* and *Twist* were expressed at lower levels in cells lacking Sox9, whilst *E-cadherin* levels, an epithelial marker, were slightly increased, consistent with a reduced EMT signature that would be expected from reduced Zeb1 levels.

RNA sequencing analysis confirmed that *Sox9* transcripts were not completely deleted from the three e1/*Sox9*-KO subclones. This was

anticipated since the targeting strategy does not remove the 5' transcript, including ATG start codon). e1/*Sox9*-KO#1 expresses the lowest *Sox9* levels compared to e1/*Sox9*-KO#2 and -KO#3 which show reduced *Sox9* levels compared to control cells (Supplementary Fig. 7). A small number of genes were consistently up- and down-regulated in all three e1/*Sox9*-KO lines when compared to e1/control subclones in transcriptomic analysis (Figure 8B). One is *Csf1,* which promotes postnatal mammary stem cell activity, and is expressed at higher levels in all *Sox9*-deficient e1 subclones. Two clusters of interest were identified after clustering of all hits identified based on either DESeq or Intensity Difference tests. One group includes *Sox9*, whilst the other group displays the opposite expression pattern. e1/*Sox9*-KO#1 and e1/*Sox9*-KO#3 clustered more closely together than with e1/S*ox9*-KO#2, and both lines displayed pronounced changes in morphologies when cultured in 3D, so these two lines were analysed together to identify genes consistently modulated in these e1/*Sox9*-KO cells compared to the control cells. Genes associated with innate immunity, cholesterol biosynthesis, cell fate are down-regulated, while genes regulating cell adhesion, Rap1 signalling and kinase activity are upregulated in e1/*Sox9*-KO#1 and e1/*Sox9*-KO#3 cells, identifying potential regulators of the morphological changes observed when cultured in 3D (Supplementary Dataset 3).

Results from lineage tracing studies have shown Sox9+ cells contribute to alveologenesis[32]. We therefore investigated the ability of e1/*Sox9*-KO cells to form fully differentiated mammary acini and secrete milk using an *in vitro* assay. Fewer alveoli-like structures derived from e1/*Sox9*-KO cells formed from all three e1/*Sox9*-KO lines and these were smaller and ill-defined structures compared to those formed from e1/control cells (Figure 8C-D). Our results demonstrate a diminished ability of e1/*Sox9*-KO embryonic mammary cells to respond to lactogenic stimuli and undergo functional alveologenesis. These results indicate Sox9 expression is required for embryonic mammary progenitor cells to attain a mature luminal progenitor cell phenotype capable of efficiently differentiating to form fully functional alveoli.

Discussion

In this study, we have established eighteen novel, cultivable, embryonic mammary progenitor cell lines (eMPCs). eMPCs will be a valuable resource for studying molecular regulators of normal breast development and understanding mammary cell fate and its plasticity. Based on gene expression and principal component analysis of the embryonic mammary progenitor cell lines, we identified three major clusters. One cluster (marked in yellow in Figure 2 and eleven out of the sixteen clones) is likely to represent mammary epithelial progenitor cells in an undifferentiated, plastic state since these cells express high levels of several stem cell makers; another cluster of 3 clones (marked in green in Figure 2) is likely to be representing more differentiated epithelial mammary progenitor cells since cells from this cluster express higher levels of Keratins. The third cluster of two clones and the pool (marked in magenta in Figure 2) appears to be composed of less plastic progenitor cells with restricted developmental potential. Four of the eighteen cell lines were selected for use in plasticity assays to represent the three major clusters and included both GFP- clones we had derived from single GFP- cells, as well as two of the sixteen GFP+ clones derived from single GFP+ cells. Although GFP+ expression was retained within the epithelial-appearing component after sustained organoid culture of the microdissected mammary organ, GFP expression is variable after expansion in 2D culture such that both GFP+ and GFP- cells are present in all eMPC lines, and we cannot assume GFP+ clones are derived from embryonic mammary epithelial cells or that GFP- cells represent mesenchymal cells. We cannot exclude the possibility that GFP+ clones could have originated from other GFP+ cells present within the embryo although our dissection protocol should exclude this possiblilty[21]. In addition, we observe a small percentage of cells that lack GFP expression within the mammary primordial epithelium.

Embryonic mammary lineage commitment is not completely understood[33]. It is still not clear how plastic or committed embryonic mammary cells are after removal from their native microenvironment. A subset of clones appeared to give rise to cell lines with mammary epithelial progenitor features and other clones appear less differentiated and more

plastic. Another small group clustered together with the ePool suggesting it consisted of a mixture of cell types**.** Heterogeneity exists between the clone profiles and their properties. This may reflect different levels of commitment of individual cells to the mammary lineage or loss of mammary epithelial phenotype. In addition, stem cells exist in a variety or spectrum of cell states that can interconvert[34,35]**,** so the observed heterogeneity could represent lines comprised of various components of a mammary progenitor spectrum.

During embryonic development, tissues reshape during a combination of morphogenetic processes. Tissue compaction is a morphogenetic process by which a tissue adopts a tighter structure. Both cell adhesion and cell contraction play roles in tissue compaction[36]. Of the four clones assayed, only eG1, a line that expresses Keratins, as well as other adhesion regulators (including protocadherins, integrins, ECM ligands) undergoes compaction. eMPC are clonogenic and form mammary acini *in vitro*. All four of the eMPC lines tested can be induced to secrete milk with the appropriate hormonal stimulation (although e2 and eG2 with low efficiency), which demonstrates these cells have been committed to the mammary lineage, an event that occurs during embryonic development[6]. Clone e1 displayed limited ability to engraft in cleared mammary fat pads; this is consistent with other reports that found dissociated mid-gestation stage mouse mammary cells have very low stem cell activity when assessed using the cleared fat pad assay[12]. Baseline sequencing permits monitoring for changes after passage and we have been able to recluster eMPC cells together after early and later passages when re-sequenced, suggesting these cells can be maintained without substantial changes.

Sox9 is amongst the earliest transcription factors expressed by cells within the prenatal mammary epithelium[22]. Sox9 is also expressed by basal and luminal Estrogen Receptor (ER)- postnatal mammary epithelial cells[32,37]. Sox9 is thought to act together with Slug to confer a stem cell state to postnatal mammary epithelial cells[18]. Regulation of prenatal and postnatal mammary stem cell function by *Sox9* produces distinct effects on clonogenic ability. In contrast to our finding of increased sphere-forming ability in e1 after *Sox9* deletion, knockdown of *Sox9* in primary postnatal mouse MECs causes

392 a marked reduction in organoid-forming ability, reduced *Slug* levels, and
393 diminishes gland-reconstituting activity[18]. Impaired ability to undergo EMT is
394 likely to be a common feature in all *Sox9*-deficient mammary cells. Mammary
395 stem cell function is highly impacted by *Sox9*, but cell context also appears to
396 be a significant factor, including developmental stage and state. A recent
397 study shows a role for SOX9 in the regulation of hormone resistance in breast
398 cancer[38], highlighting the relevance of embryonic mammary factor expression
399 in breast cancers and the need for further investigations into their ability to
400 modulate cell states.

401 Cell state is thought to influence epithelial cell phenotype and
402 plasticity[39]. Mature postnatal MECs express Keratins, indicative of being in an
403 epithelial state (E). Loss of Sox9 and ability to undergo EMT in postnatal
404 MECs should push them towards a more highly epithelial state, which, in
405 theory, would have less stem cell activity according to several recent
406 models[40]. Stem cell activity is thought to be highest when cells are in an
407 intermediate epithelial/mesenchymal state[35]. Deleting *Sox9* from e1 leads to
408 reduced Zeb1 levels and reduced propensity to undergo EMT. As a result, e1/
409 *Sox9*-KO cells would be expected to shift from the embryonic mesenchymal
410 (M) state towards epithelial to an intermediate hybrid M/E state, which, in
411 theory, should have more stem cell activity and mixed epithelial (adhesive)
412 and mesenchymal (migratory) properties that confer ability to move
413 collectively as clusters (Supplementary Fig. 8)[41,42]. Our results clearly show
414 that there are profound differences between the effects from loss of *Sox9*
415 function on clonogenic ability in embryonic mammary progenitor cells
416 compared to postnatal MECs, which exist in relatively high epithelial cell
417 states. Despite these differences, these findings are consistent with current
418 models of EMT/MET mediation of cell state in regulating stem cell function.
419 Analysis from lineage tracing data indicated that Sox9 marks luminal
420 progenitors that give rise to alveolar progenitors[32]. Fewer alveoli-like
421 structures formed from e1/*Sox9*-KO cells and were smaller than control
422 alveoli, which is consistent with a requirement for active Sox9 signallling in
423 luminal progenitors to produce functional alveoli. *Sox9* is required for luminal
424 progenitor cell survival, but basal/myoepithelial cells can withstand loss of

Sox9 in studies using a conditional deletion of *Sox9* from the mammary gland using Mouse mammary tumour virus (MMTV)-Cre[37]. Mice with *Sox9*-deficient mammary glands could lactate normally, but mosaic MMTV-Cre mediated gene deletion was reported, so it is likely that sufficient numbers of Sox9+ luminal progenitors were retained to produce functional alveoli that had undergone terminal differentiation[37]. Our results support a model in which Sox9 is required for the specification of mammary progenitor cells from early embryonic stages of through postnatal development as a key regulator of mammary gland development and luminal progenitor homeostasis and maintenance.

Our study reveals complex interactions that underlie mammary lineage regulation, cell state and progenitor cell activity. Plasticity, defined as the ability of cells to dynamically change cell state, is crucial for many processes during mammary gland development and tissue homeostasis. Reversible EMT is central to tissue development, epithelial stemness, and cancer metastasis. SOX9 has been linked to all three of these processes[18,37,43].

Cell context is key in studies of mammary cell function and embryonic cells are highly plastic and multipotent, giving rise to both basal and luminal postnatal mammary cells. Cancer progression involves the loss of differentiated phenotype and acquisition of progenitor/stem cell features, which have recently been put forward as a hallmark of cancer[44]. These cell lines will be an invaluable resource to explore unresolved questions related to biology of mammary progenitor and stem cells, including those relevant to the origin of breast cancer.

Methods

Animal experiments

All animal work was carried out under UK Home Office project and personal licenses following local ethical approval from The Institute of Cancer Research Ethics Committee and in accordance with local and national guidelines. Female mice (*Mus musculus*) were used for all experiments except breeding. Immortomouse and SCID/Beige mice were purchased from Charles River (Harlow, UK)**.** *s-SHIP-GFP* mice were a gift from the late Professor Larry Rohrschneider (Fred Hutchinson Cancer Research Center, Seattle, WA, USA). Mice were housed in IVC cages on a 12h light/dark cycle and received food and water *ad libitum*. Mouse genotyping for detecting *Immorto* and *s-SHIP-GFP* transgenes was performed at Transnetyx (Cordova, TN, USA) and the results were used for colony maintenance.

Isolation and culture of eMPC

Immortomouse were backcrossed with C57BL6/J (B6) for four generations. The Immortomouse and *s-SHIP-GFP* mice were bred to obtain *Immorto:s-SHIP-GFP* embryo. E12.0-stage *Immorto:s-SHIP-GFP* embryo was dissected to isolate only the organ of mammary primordium number three as a micro-tissue, using Dumont #5 forceps and Tübingen Spring Scissors (Fine Science Tools GmbH, Heidelberg, Germany) using a Leica M205FA fluorescent stereomicroscope (Leica, Wetzlar, Germany). The remaining embryonic tissue was used for genotyping the *Immorto* transgene. Micro-tissue was immediately embedded into 200 μl of ice-cold Cultrex BME type 1, stem cell qualified, growth factor reduced (3434-005-02, AMS Biotechnology Ltd, Abingdon, Oxford, UK) within a well of a 48-well plate. The matrix-embedded tissue was transferred into CO_2 incubator equilibrated to 33ºC, 5% CO_2 atmosphere for 3 hours. Culture medium was overlaid on the solidified BME during the pre-incubation. Culture medium contains Mesencult MSC basal medium for mouse (05501), Mesencult Stimulatory Supplements for Mouse (05502), 2 mM L-Glutamine (17100; all from STEMCELL Technologies UK Ltd., Cambridge, UK), 5 U/ml Mouse Interferon gamma (IFNγ) (315-05, Pepro Tech EC Ltd., London, UK), and 0.1 mg/ml Primocin (ant-pm1, Source

Bioscience, Nottigham, UK). eMPCs migrated and proliferated within BME from the *Immorto:s-SHIP-GFP* micro-tissue for 1 month. Cells were harvested by trypsinisation and further expanded on two-dimensional dishes pre-coated thinly with BME, and referred to henceforth, as eMPC pool, abbreviated as ePool.

Generation of eMPC clones

Fluorescence activated cell sorting

ePool cells were harvested from the culture dish and adjusted to the concentration of 2 x 10^6 cells/ml with PBS. To stain apoptotic and necrotic cells, cells were stained with DAPI at the final concentration of 1 μg/ml for 5 min at room temperature prior to cell sorting. ePool were passed through a 40 μm filter and sorted using BD Aria fluorescence-activated cell sorter (BD Biosciences, Oxford, UK). Live single ePool cells were positively gated using their forward and side scatter profiles (Supplementary Fig. 1B). ePool population with high s-SHIP-GFP expression and ePool GFP- population were subsequently gated based on green channel profile.

Clone derivation from single cells

Live single ePool GFP+ and ePool GFP- cells were individually sorted into 96 multiwell plates. The multiwell plates were precoated with either Cultrex BME or Matrigel, growth factor reduced (734-0268, VWR International, Lutterworth, Leicestershire, UK) and filled with eMPC culture medium. The single cells on multiwell plates were incubated at CO_2 incubator equilibrated to 33ºC, 5% CO_2 atmosphere without changing medium for a month. eMPC clones proliferated for a month after which they were trypsinised and transferred to dishes precoated with BME for further expansion. Only two GFP- eMPC clones were grown from a Matrigel-precoated multiwell plate and named as e1 and e2 (Figure 1B). Fourteen GFP+ eMPC clones were recovered from two Matrigel-precoated multiwell plates, which are named as eG1, eG2, eG2E9…, depending on well position. Only one GFP+ eMPC clone, eG1B3, was recovered from a BME-precoated plate, suggesting that Matrigel is preferable for initial clonal cell expansion.

Cell culture

The following mouse C57BL/6 (B6) cell lines were purchased and cultured for RNA isolation according to manufacturer's instructions; Bone Marrow Mononuclear cells, (c57-6271F, Caltag Medsystems Buckingham, UK); Embryonic Stem Cells, (SCRC-1002, LGC standards, Middlesex, UK); Mouse Embryonic Fibroblasts, untreated, (GSC-6002, AMS Biotechnology Ltd); Mesenchymal Stem cells, (MUBMX-01001 Cambridge Bioscience, Cambridge, UK). B6 postnatal mammary epithelial cells were isolated as described[45]. The mammary epithelial fraction was further purified using EasySep Mouse Epithelial Cell Enrichment Kit according to manufacturer's instructions (19758, STEMCELL Technologies UK). Human Umbilical Vein Endothelial Cells were cultured according to manufacturer's instruction (191027, Lonza, Basel, Switzerland).

Total RNA isolation and quality control

Embryonic mammary tissues (embryonic mammary primordial epithelium (MPE), embryonic mammary stroma (includes mammary mesenchyme (MM) and fat pad precursor (FPP)), and surface epithelium (Epi), were micro-dissected from E12.5 B6 mice following the protocol described in[46]. Total RNA from cells and mammary tissue were extracted using RNeasy Plus Micro kit and RNeasy Mini kit (74034 and 74104, respectively, Qiagen, Manchester, UK), RNA concentration and gDNA contamination were determined using Qubit 2.0 Fluorometer and accompanying kits (Thermo Fisher Scientific, Waltham, MA, USA). RNA samples highly contaminated with gDNA for more than 10% of RNA were treated with TURBO DNA-free Kit (AM1907, Thermo Fisher Scientific) and concentrated using RNA Clean and Concentrator-5 (R1015, Zymo Research, Irvine, CA, USA). All samples were assessed for RNA quality using Agilent 2100 Bioanalyzer system (Agilent Technologies, Cheshire, UK) and confirmed that RNA Integrity Number was greater than 9.0.

RNA sequencing

cDNA library preparation was carried out at Oxford Genomics Centre, The Wellcome Trust Centre for Human Genetics using PolyA+ RNA enrichment method for total RNA from cultured cells and SMARTer method for total RNA

551 from embryonic mammary tissue, respectively. mRNA fraction was selected
552 from the total RNA before conversion to cDNA. Second strand cDNA
553 synthesis incorporated dUTP. The cDNA was end-repaired, A-tailed and
554 adapter-ligated. Prior to amplification, samples underwent uridine digestion.
555 The prepared libraries were size selected, multiplexed and quality checked
556 before paired end sequencing over three lanes of a flow cell. Amplified cDNA
557 from embryonic mammary tissues were generated by the SMARTer
558 amplification kit. The cDNA was end-repaired, A-tailed, adapter-ligated and
559 amplified. The prepared libraries were size selected, multiplexed and quality
560 checked before paired end sequencing on four lanes of a flow cell. Data was
561 aligned to the reference genome, mm10, and quality checked.
562 RNA sequencing files were submitted to ArrayExpress as accession E-
563 MTAB-6846, MTAB-6856, E-MTAB-6859. FastQ files were truncated to a
564 consistent length of 75bp using trim galore v0.4.3 and were then aligned
565 against the mouse GRCm38 genome assembly using hisat2 v2.0.5 using
566 options --no-mixed and --no-discordant. Mapped positions with MAPQ values
567 of <20 were discarded. Gene expression was quantitated using the RNA-Seq
568 quantitation pipeline in SeqMonk software v1.37.0
569 (https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/) in opposing
570 strand specific library mode. For count based statistics, raw read counts over
571 exons in each gene were used. For visualisation and other statistics
572 log2RPM (reads per million reads of library) expression values were used.
573 Differentially expressed genes were selected based on passing two statistical
574 filters -the DESeq2 Likelihood Ratio Test with a cutoff of $p<0.05$ following
575 multiple testing correction and the SeqMonk Intensity Difference filter on
576 log2RPM values with a sample size of 1% of all genes and a cutoff of $p<0.05$
577 after multiple testing correction. Hierarchical clustering was performed on per-
578 gene median centred log2RPM expression values using Pearson's
579 correlation. Gene cluster separation was performed by segmenting the tree at
580 an R value of 0.5. Clusters containing <50 genes were discarded. PCA was
581 performed on column centred log2RPM values without additional scaling.
582 The intensity difference test used a locally matched subset of 1% of
583 genes based on average expression. From these a local standard deviation in
584 log2RPM difference values was calculated and used to calculate the

probability of the cumulative distribution function for a normal distribution with this standard deviation, using the observed difference in the gene being tested. P-values were corrected for multiple testing using Benjamini and Hochberg multiple testing correction.

cDNA synthesis and real-time quantitative reverse transcription PCR

cDNA library from total RNA was synthesised using QuantiTect Reverse Transcription Kit (205310, Qiagen). Comparative C_T ($\Delta\Delta C_T$) real-time **quantitative** PCR was performed using Taqman Gene Expression assays as previously described[17] on either ABI Prism 7900HT sequence detection system (Applied Biosystems, Foster City, CA, USA) or QuantStudio 6 Flex system (Thermo Fisher Scientific). *Actb* was used as an endogenous control and fold change normalised to a comparator sample was calculated. The assay probes are listed in Supplementary Dataset 4.

Antibodies

Antibodies used for whole-mount immunofluorescence, immunohistochemistry, immunofluorescence, and western blotting in this study are listed in Supplementary Dataset 4.

Adipogenesis assay

eMPC and MSCs (positive control) were plated at 100% confluency in Osteogenic/Adipogenic Base Media (CCM007, Bio-Techne Ltd, Oxford, UK) containing 1 µg/ml Insulin, 0.5 mM 3-Isobutyl-1-methylxanthine and 0.25 mM Dexamethasone (19278, I5879 and D2915 from Sigma, Dorset, UK, respectively). After 6 days, cells were fixed with 4% Paraformaldehyde/PBS and lipid droplets were detected by HCS LipidTOX™ Red Neutral Lipid Stain (H34476, Thermo Fisher Scientific) at 1:200 in PBS. Lipid-labelled cells were photographed using EVOS FL fluorescent microscope equipped with 10x Plan LWD FL 0.30NA objective (Thermo Fisher Scientific).

***In vitro* tube formation assay**

Tube formation ability of eMPCs was examined following a published protocol[47]. Briefly, eMPCs, MSCs (negative control) and HUVECs (positive control) were plated at ~80% confluency using endothelial growth medium (EGM) (EBM-2 medium containing EGM-2 SingleQuot Kit, CC-3156 and CC-3162 from Lonza, respectively). On the next day, 96 well plates filled with 50-80 µl Cultrex BME were prepared on ice without forming air bubbles and pre-incubated at CO_2 incubator equilibrated to 37ºC, 5% CO_2 atmosphere to allow matrix gelation. Cells were harvested by trypsinisation and re-suspended at the concentration of 1.5 x 10^5 cells /ml in either endothelial basal medium 2 (EBM-2) or EGM-2 medium. 100 µl of single cell suspension (containing total 15,000 cells) was gently added on the top of BME gel. After incubation for 24 hours, cells were labelled with Calcein AM solution at the final concentration of 2 µM in EBM-2 medium and incubated in CO_2 incubator for 30 min. Calcein AM-labelled cells were photographed using EVOS FL fluorescent microscope equipped with 2x Plan LWD 0.06NA objective (Thermo Fisher Scientific). The fluorescent images were analysed to detect mesh structure using a developed package of ImageJ, Fiji[48] (http://imagej.net/Fiji/Downloads), and its Plugin, Angiogenesis Analyzer (Carpentier G., Angiogenesis Analyzer for ImageJ (2012); https://imagej.nih.gov/ij/macros/toolsets/Angiogenesis%20Analyzer.txt). Cells were also tested for tube formation assay with EBM-2 medium containing 10 mM Sulforaphane (angiogenic inhibitor, S4441, Merck KGaA, Darmstadt, Germany) as a control to observe its inhibitory effect on the assay.

Mammary alveologenesis assay

Lactogenic ability of eMPCs was examined following an established protocol[49]. Briefly, eMPCs and MECs (positive control) were plated at 100% confluent in RPMI 1640 medium (21875, Thermo Fisher Scientific) containing 10% horse serum (16050-122, Thermo Fisher Scientific), 5 µg/ml Insulin, 10 ng/ml EGF (236-EG-200, Bio-Techne Ltd), 10 mM HEPES (H0887, Sigma), and Primocin (Bioscience). On the following day, in order to withdraw the EGF and enhance differentiation, rather than cell proliferation, cell medium was changed to EGF- medium containing 5% horse serum, 10 mM HEPES, and

Primocin in RPMI 1640 medium. After 3 days, mammary alveologenesis was induced by replacement with priming medium (EGF- medium containing 5 µg/ml mouse Prolactin (757908, Biolegend, London, UK), 5 µg/ml Insulin, and 0.1 mM Dexamethasone). Cells were fixed with 4% Paraformaldehyde/PBS 6 days after induction and stained with anti-Milk antibody and DAPI. Alveoli were photographed using EVOS FL microscope and the number of alveoli larger than 100 µm per 24 well was counted manually.

For studies of e1 with reduced *Sox9* expression, the assay was modified since e1/*Sox9*-knockout cells detached upon induction of lactogenic differentiation using the standard assay. Cells were plated around 80-90% confluence in MesenCult Expansion medium with MesenCult supplement, 200 mM L-Glutamine, 5 U/ml IFNγ (Pepro Tech) and 0.1mg/ml Primocin. Once cells reached 100% confluence (on day 4), cell medium was changed to RPMI 1640 medium containing 5% horse serum, 10 mM HEPES, and Primocin. On day 5, mammary alveologenesis was induced with medium containing 5 µg/ml mouse Prolactin, 5 µg/ml Insulin, and 0.1 mM Dexamethasone. Induction medium was replenished every 3 days. On day 11, cells were fixed with 4% Paraformaldehyde/PBS and stained with anti-Milk antibody and DAPI. Alveoli were photographed using EVOS FL microscope and the number of alveoli in different size ranges were counted manually.

Sphere-forming assays

A clonogenic mammosphere protocol and media were adapted for eMPCs[50]. Mouse EpiCult-B complete medium (05610, STEMCELL Technologies) was supplemented with 0.6% Methylcellulose (HSC011, Bio-Techne Ltd), 20 ng/ml basic FGF (3139-FB-025/CF, Bio-Techne Ltd), 20 ng/ml EGF, 2% NeuroCult™ SM1 Without Vitamin A, 10 µg/ml Heparin, 1 µg/ml Hydrocortisone (05731, 07980, and 07926, respectively, from STEMCELL Technologies UK Ltd), 10 ug/ml Insulin, 5 U/ml Mouse IFNγ, and 0.1 mg/ml Primocin. Trypsinised eMPCs and MSCs were passed through 40 µm filter to make absolute single cell suspension and the cell number was adjusted to 2-4 x 10^5 cells/ml. 10,000 cells were added into 2 ml of the supplemented Epicult-B medium. Cells were pipetted thoroughly to be distributed homogenously

within the viscous Methylcellulose-containing medium and transferred into Corning ultra-low attachment 6 well (10154431, Fisher Scientific, Loughborough, UK). 6 well plates were incubated within a humidified box at 33ºC incubator for eMPCs and 37ºC for MSCs, respectively. At day 7, sphere images were photographed and automatically scored using Single Colony Verification program in Celigo cytometer (Nexcelom Bioscience LLC, Lawrence, MA, USA).

Spheroid growth in hydrogel

For analysing 3D morphology, e1/control and e1/*Sox9*-KO cells were seeded with 5000 cells/well onto Corning Costar Ultra-Low Attachment 96 well plates (Corning Kennebunk ME, USA) filled with eMPC culture medium. On day 3, 50 µl of Corning Matrigel Growth Factor Reduced Basement Membrane Matrix (356231) or Cultrex PathClear 3-D Culture Matrix RGF BME I (3434-005) was added into each well (except control wells). Medium was replenished every 3 days. On day 9, 50 µl of 1% agarose was added to the control wells. All spheroids were subsequently fixed with 4% Paraformaldehyde (PFA)/PBS and were photographed using EVOS FL microscope.

CRISPR/Cas9-mediated gene targeting of *Sox9*

We took advantage of lentiviral transduction of *Cas9* gene within eMPCs followed by transfection of pre-made CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA) for genetic ablation of *Sox9*, which were purchased from GE Healthcare (Little Chalfont, UK). Edit-R Lentiviral Blast-Cas9 Nuclease Particles (VCAS10128, Dharmacon, Lafayette, Colorado, USA) were transduced into e1 and Cas9-expressing e1 population was expanded with 10 µg/ml Blasticidin (ant-bl-1, Source Bioscience Nottingham, UK) from day 4 onward. 2 x 10^5 cells of Blasticidin-resistant e1 were seeded on 6 well and on the next day treated with 5 µl DharmaFECT 1 Transfection Reagent (T-2001-02, Dharmacon) together with 5 µl of 10 µM Edit-R CRISPR-Cas9 Synthetic tracrRNA (U-002000-20, Dharmacon) and 5 µl of 10 µM *Sox9*-targeting crRNA or non-targeting crRNA. Cell populations with *Sox9*

mutations were screened by Heteroduplex formation assay using QuickExtract DNA Extraction Solution (QE0905T, Qiagen, Peterborough, UK) and EnGen™ Mutation Detection Kit (E3321S, New England Biolabs, Herts, UK) according to manufacturer's instructions. We tested 4 different crRNA and the primer sets to detect mutation within each crRNA-binding region; details are provided in Supplementary Dataset 4C. 2 cell populations (Sox9-01 and Sox9-02 in Supplementary Fig. 5A) that were verified for *Sox9* mutation were further sorted using BD Aria to develop single cell-derived subclones. 3 e1/*Sox9*-knockout (KO) subclones, KO#1, KO#2, and KO#3, produced no Sox9 expression detectable by Western blot (Supplementary Fig. 5B). These three e1/*Sox9*-KO subclones as well as two e1/control clones, Co#1 and Co#2, were used in this study.

Mammary fat pad injection of e1

To allow bioluminescence imaging of e1 cells, cells were labelled with carrying red-shifted *Luciola Italica* luciferase transgene using lentivirus particles, RediFect Red-FLuc-Puromycin (CLS960002, Perkin Elmer, Buckinghamshire, UK). The transduced e1/Red-FLuc cells were selected with 5 μg/ml Puromycin for 3 weeks. e1/Red-FLuc cells were harvested and resuspended at a concentration of 1 x 10^6 cells/ml in PBS. 100 μl of cells were injected into both the left and right inguinal mammary fat pad number 4 of 10-week-old SCID/Beige female mice. Mice were injected 100 μl of 15 mg/ml D-Luciferin (119222, Perkin Elmer) in PBS intraperitoneally and imaged after 5 min using IVIS Illumina II (Perkin Elmer). Engrafted mammary glands were harvested 1-2 weeks after mammary fat pad injections and fixed in 4% PFA in PBS for immunohistochemistry.

Statistical Analysis

The data in the graphs are presented as mean and the standard error of the mean. The data was analysed by two-tailed ANOVA, Student's t-test, or two-tailed, paired t-test using GraphPad Prism 7 software. *P-value* ≤ 0.0001 is considered as extremely significant (****), $P \leq 0.001$ as highly significant (***), $P \leq 0.01$ as very significant (**), $P \leq 0.05$ as significant (*), and $P > 0.05$ as not significant (ns), respectively.

Data availability statement

The authors declare that the data supporting the findings of this study are available within the article and its supplementary information files. All source data underlying the graphs and charts presented in the main figures is available in Supplementary Dataset 5.

References

1 Van Keymeulen, A. & Blanpain, C. Tracing epithelial stem cells during development, homeostasis, and repair. *The Journal of cell biology* **197**, 575-584, doi:10.1083/jcb.201201041 (2012).

2 Wuidart, A. *et al.* Early lineage segregation of multipotent embryonic mammary gland progenitors. *Nat Cell Biol* **20**, 666-676, doi:10.1038/s41556-018-0095-2 (2018).

3 Lilja, A. M. *et al.* Clonal analysis of Notch1-expressing cells reveals the existence of unipotent stem cells that retain long-term plasticity in the embryonic mammary gland. *Nat Cell Biol* **20**, 677-687, doi:10.1038/s41556-018-0108-1 (2018).

4 Zvelebil, M. *et al.* Embryonic mammary signature subsets are activated in Brca1-/- and basal-like breast cancers. *Breast cancer research : BCR* **15**, R25, doi:10.1186/bcr3403 (2013).

5 Howard, B. & Ashworth, A. Signalling pathways implicated in early mammary gland morphogenesis and breast cancer. *PLoS Genet* **2**, e112, doi:10.1371/journal.pgen.0020112 (2006).

6 Propper, A. Y., Howard, B. A. & Veltmaat, J. M. Prenatal morphogenesis of mammary glands in mouse and rabbit. *Journal of mammary gland biology and neoplasia* **18**, 93-104, doi:10.1007/s10911-013-9298-0 (2013).

7 Balinsky, B. On the developmental processes in mammary glands and other epidermal structures. *Transactions of the Royal Society Edinburgh* **62**, 1-31 (1949-1950).

8 Lee, M. Y. *et al.* Ectodermal influx and cell hypertrophy provide early growth for all murine mammary rudiments, and are differentially regulated among them by Gli3. *PloS one* **6**, e26242, doi:10.1371/journal.pone.0026242 (2011).

9 Propper, A. Y. Wandering epithelial cells in the rabbit embryo milk line. A preliminary scanning electron microscope study. *Developmental biology* **67**, 225-231 (1978).

10 Propper, A. Y., Howard, B. A. & Veltmaat, J. M. Prenatal Morphogenesis of Mammary Glands in Mouse and Rabbit. *Journal of mammary gland biology and neoplasia*, doi:10.1007/s10911-013-9298-0 (2013).

11 Makarem, M. *et al.* Developmental changes in the in vitro activated regenerative activity of primitive mammary epithelial cells. *PLoS Biol* **11**, e1001630, doi:10.1371/journal.pbio.1001630 (2013).

12 Spike, B. T. *et al.* A mammary stem cell population identified and characterized in late embryogenesis reveals similarities to human breast cancer. *Cell stem cell* **10**, 183-197, doi:10.1016/j.stem.2011.12.018 (2012).

13 Sakakura, T., Nishizuka, Y. & Dawe, C. J. Capacity of mammary fat pads of adult C3H/HeMs mice to interact morphogenetically with fetal mammary epithelium. *J Natl Cancer Inst* **63**, 733-736 (1979).

14 Voutilainen, M. *et al.* Ectodysplasin regulates hormone-independent mammary ductal morphogenesis via NF-kappaB. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 5744-5749, doi:10.1073/pnas.1110627109 (2012).

15 Hens, J. R. *et al.* BMP4 and PTHrP interact to stimulate ductal outgrowth during embryonic mammary development and to inhibit hair follicle

induction. *Development* **134**, 1221-1230, doi:10.1242/dev.000182 (2007).
16 Veltmaat, J. M. *et al.* Gli3-mediated somitic Fgf10 expression gradients are required for the induction and patterning of mammary epithelium along the embryonic axes. *Development* **133**, 2325-2335, doi:10.1242/dev.02394 (2006).
17 Kogata, N., Oliemuller, E., Wansbury, O. & Howard, B. A. Neuregulin-3 regulates epithelial progenitor cell positioning and specifies mammary phenotype. *Stem Cells Dev* **23**, 2758-2770, doi:10.1089/scd.2014.0082 (2014).
18 Guo, W. *et al.* Slug and Sox9 cooperatively determine the mammary stem cell state. *Cell* **148**, 1015-1028, doi:10.1016/j.cell.2012.02.008 (2012).
19 Jat, P. S. *et al.* Direct derivation of conditionally immortal cell lines from an H-2Kb-tsA58 transgenic mouse. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 5096-5100 (1991).
20 Rohrschneider, L. R., Custodio, J. M., Anderson, T. A., Miller, C. P. & Gu, H. The intron 5/6 promoter region of the ship1 gene regulates expression in stem/progenitor cells of the mouse embryo. *Dev Biol* **283**, 503-521, doi:S0012-1606(05)00270-8 [pii] 10.1016/j.ydbio.2005.04.032 (2005).
21 Bai, L. & Rohrschneider, L. R. s-SHIP promoter expression marks activated stem cells in developing mouse mammary tissue. *Genes Dev* **24**, 1882-1892, doi:10.1101/gad.1932810 (2010).
22 Wansbury, O. *et al.* Transcriptome analysis of embryonic mammary cells reveals insights into mammary lineage establishment. *Breast Cancer Research* **13**, doi:Artn R79 Doi 10.1186/Bcr2928 (2011).
23 Sidney, L. E., Branch, M. J., Dunphy, S. E., Dua, H. S. & Hopkinson, A. Concise review: evidence for CD34 as a common marker for diverse progenitors. *Stem Cells* **32**, 1380-1389, doi:10.1002/stem.1661 (2014).
24 dos Santos, C. O. *et al.* Molecular hierarchy of mammary differentiation yields refined markers of mammary stem cells. *Proc Natl Acad Sci U S A* **110**, 7123-7130, doi:10.1073/pnas.1303919110 (2013).
25 Qian, H. *et al.* Molecular characterization of prospectively isolated multipotent mesenchymal progenitors provides new insight into the cellular identity of mesenchymal stem cells in mouse bone marrow. *Mol Cell Biol* **33**, 661-677, doi:10.1128/MCB.01287-12 (2013).
26 Shan, T., Liu, W. & Kuang, S. Fatty acid binding protein 4 expression marks a population of adipocyte progenitors in white and brown adipose tissues. *FASEB J* **27**, 277-287, doi:10.1096/fj.12-211516 (2013).
27 Wang, C. Q., Jacob, B., Nah, G. S. & Osato, M. Runx family genes, niche, and stem cell quiescence. *Blood Cells Mol Dis* **44**, 275-286, doi:10.1016/j.bcmd.2010.01.006 (2010).
28 Jo, A. *et al.* The versatile functions of Sox9 in development, stem cells, and human diseases. *Genes Dis* **1**, 149-161, doi:10.1016/j.gendis.2014.09.004 (2014).
29 Wang, Y. *et al.* Lgr4 regulates mammary gland development and stem cell activity through the pluripotency transcription factor Sox2. *Stem Cells* **31**, 1921-1931, doi:10.1002/stem.1438 (2013).

30 Liu, Y., Beyer, A. & Aebersold, R. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**, 535-550, doi:10.1016/j.cell.2016.03.014 (2016).
31 Rodilla, V. *et al.* Luminal progenitors restrict their lineage potential during mammary gland development. *PLoS Biol* **13**, e1002069, doi:10.1371/journal.pbio.1002069 (2015).
32 Wang, C., Christin, J. R., Oktay, M. H. & Guo, W. Lineage-Biased Stem Cells Maintain Estrogen-Receptor-Positive and -Negative Mouse Mammary Luminal Lineages. *Cell Rep* **18**, 2825-2835, doi:10.1016/j.celrep.2017.02.071 (2017).
33 Howard, B. A. In the beginning: the establishment of the mammary lineage during embryogenesis. *Seminars in cell & developmental biology* **23**, 574-582, doi:10.1016/j.semcdb.2012.03.011 (2012).
34 Liu, S. *et al.* Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem Cell Reports* **2**, 78-91, doi:10.1016/j.stemcr.2013.11.009 (2014).
35 Wahl, G. M. & Spike, B. T. Cell state plasticity, stem cells, EMT, and the generation of intra-tumoral heterogeneity. *NPJ Breast Cancer* **3**, 14, doi:10.1038/s41523-017-0012-z (2017).
36 Turlier, H. & Maitre, J. L. Mechanics of tissue compaction. *Semin Cell Dev Biol* **47-48**, 110-117, doi:10.1016/j.semcdb.2015.08.001 (2015).
37 Malhotra, G. K. *et al.* The role of Sox9 in mouse mammary gland development and maintenance of mammary stem and luminal progenitor cells. *BMC Dev Biol* **14**, 47, doi:10.1186/s12861-014-0047-4 (2014).
38 Jeselsohn, R. *et al.* Embryonic transcription factor SOX9 drives breast cancer endocrine resistance. *Proc Natl Acad Sci U S A* **114**, E4482-E4491, doi:10.1073/pnas.1620993114 (2017).
39 Varga, J. & Greten, F. R. Cell plasticity in epithelial homeostasis and tumorigenesis. *Nat Cell Biol* **19**, 1133-1141, doi:10.1038/ncb3611 (2017).
40 Hong, T. *et al.* An Ovol2-Zeb1 Mutual Inhibitory Circuit Governs Bidirectional and Multi-step Transition between Epithelial and Mesenchymal States. *PLoS Comput Biol* **11**, e1004569, doi:10.1371/journal.pcbi.1004569 (2015).
41 Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. Emt: 2016. *Cell* **166**, 21-45, doi:10.1016/j.cell.2016.06.028 (2016).
42 Revenu, C. & Gilmour, D. EMT 2.0: shaping epithelia through collective migration. *Curr Opin Genet Dev* **19**, 338-342, doi:10.1016/j.gde.2009.04.007 (2009).
43 Malladi, S. *et al.* Metastatic Latency and Immune Evasion through Autocrine Inhibition of WNT. *Cell* **165**, 45-60, doi:10.1016/j.cell.2016.02.025 (2016).
44 Malta, T. M. *et al.* Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell* **173**, 338-354 e315, doi:10.1016/j.cell.2018.03.034 (2018).
45 Smalley, M. J. Isolation, culture and analysis of mouse mammary epithelial cells. *Methods Mol Biol* **633**, 139-170, doi:10.1007/978-1-59745-019-5_11 (2010).
46 Sun, L., Lee, M. Y. & Veltmaat, J. M. A non-enzymatic microsurgical dissection technique of mouse embryonic tissues for gene expression

profiling applications. *Int J Dev Biol* **55**, 969-974, doi:10.1387/ijdb.113424ls (2011).
47 Arnaoutova, I. & Kleinman, H. K. In vitro angiogenesis: endothelial cell tube formation on gelled basement membrane extract. *Nat Protoc* **5**, 628-635, doi:10.1038/nprot.2010.6 (2010).
48 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).
49 Morrison, B. & Cutler, M. L. Mouse Mammary Epithelial Cells form Mammospheres During Lactogenic Differentiation. *J Vis Exp*, doi:10.3791/1265 (2009).
50 Boras-Granic, K., Dann, P. & Wysolmerski, J. J. Embryonic cells contribute directly to the quiescent stem cell population in the adult mouse mammary gland. *Breast Cancer Res* **16**, 487, doi:10.1186/s13058-014-0487-6 (2014).

Figure legends

Figure 1. Establishment of novel embryonic mammary progenitor cell lines.

(A) Schematic illustration of experimental scheme of eMPC derivation from E12.0-stage Immorto;s-SHIP-GFP mammary organ number three. After microdissection of MP3 to include three tissues: GFP+ MPE (GFP staining in green), GFP-, ERα+ MM (ERα staining in magenta) and adjacent ERα- FPP (DAPI staining in blue), MP3 was cultured on thick BME for 4 weeks.

(B) Brightfield and GFP images of cultured MP3 at Day 1, Day 11, Day 25. After culture, a variety of cell types including lipid-containing adipocyte-like cells (area shown in black box and yellow arrow), as well as contractile muscle-like cells (black arrowhead) were observed within the cell population. After enzymatic dissociation, eMPC were plated and expanded in 2D culture to obtain a pool of eMPCs (ePool). eMPC cells were subject to single cell sorting into GFP+ and GFP- cell populations using fluorescence-activated cell sorting. Clones were expanded as single-cell derived clones to create the eighteen single cell-derived clones from GFP- cells (e1 and e2) or GFP+ cells (eG1, eG2E9, etc.) described in this study. scale bar, 200 μm.

MP3, mammary primordium 3, MPE, mammary primordial epithelium, Epi, epithelium, MM, mammary mesenchyme, FPP, fat pad precursor. eMPC, embryonic mammary progenitor cells (eMPC), BF, brightfield.

Figure 2. Global gene expression landscape of novel embryonic mammary progenitor cell lines.

(A) Dendrograms from hierarchical clustering of RNA sequencing data using 2059 most significantly-changing genes of eMPC and other types of progenitor/stem cells. Eighteen eMPC clones are classified into three main clusters, which are highlighted: e1 cluster (yellow box), e2/eG2/ePool cluster (magenta box), and eG1 cluster (green box).

(B) PCA of eMPC and other types of progenitor/stem cells. Loading plot using principal component (PC) 2 (18%) and PC3 (14%) of the 17,204 most variable genes. Three main eMPC clusters (yellow, magenta, and green) recapitulate the hierarchical clustering results shown in (B). The heatmap shows hierarchical clustering based on top 0.5% of PC2 (92 genes).

(C) PC1 and PC2 loading plot using the most highly significant genes in the dataset of eMPC and other progenitor/stem cells combined with embryonic mammary tissues. Three eMPC clusters (yellow, magenta, and green) and an embryonic mammary tissue cluster (purple) from PCA are highlighted based PC1. The heatmap shows results of hierarchical clustering based on 26 genes from PC1 (top high and low rotation genes analysed using the Likelihood ratio test (LRT) as well as the intensity difference test).

eMPC, embryonic mammary progenitor cell; MP3, mammary primordium 3; BME, basement membrane extract; PC, principal component, MPE, mammary primordial epithelium, Epi, epithelium, MM, mammary mesenchyme, FPP, future fat pad precursor. ESC, embryonic stem cells, MEF, mouse embryonic fibroblasts, MSC, mesenchymal stem cells and BMMC, bone marrow mononuclear cells. Early and late indicate clones that were sequenced at early (5th) and later (13th) passage.

Figure 3. Progenitor and lineage-associated marker expression in select eMPCs.

(A) Heatmap depicting stem cell and epithelial marker expression in four eMPC selected for further characterisation.

(B) qRT-PCR for *Aldh1a1*, *Krt19*, *Krt18, Sox9*, *Snai2*, *Twist1*, *Procr*, *Trp63*, *Fabp4* in eMPC clones. Relative quantification (RQ) of each sample is normalised to that of ePool and shown in log2 plot with error bar of triplicates. (n = 3, mean ± s.e.m.).

(C) Immunohistochemistry staining of Sox9 and Twist expression in eMPC. Scale bar, 200 μm.

eMPC, embryonic mammary progenitor cell; MEF, mouse embryonic fibroblast; MSC, mesenchymal stem cell; ESC, embryonic stem cell.

Figure 4. eMPCs harbour varying potential for differentiation into mesodermal lineage.

(A) *In vitro* differentiation of eMPC clones and positive control (MSC) with adipogenic stimuli. Neutral lipid staining merged with bright-field image shows accumulation of lipid droplets (red) within cells after 6 days of induction. Scale bar, 100 μm.

(B) Vasculogenesis assay results of eMPCs and positive control (HUVEC). Tubes formed in medium with growth factors (EGM) and without growth factors were stained with Calcein AM (white) after 24 hours incubation. The total area of networks (blue, yellow, and green lines) was analysed and presented in box plots with whiskers denoting minimum and maximum values (n = 4 mean ± s.e.m.). Statistical significance was computed using one-way ANOVA and Dunnett's multiple comparisons test as ****$P \leq 0.0001$, P*** ≤ 0.001, ns = no significance. Two-tailed P values are ePool EGM 0.0001, e1 EGM 0.0001, e2 EGM 0.1607, eG1 EGM 0.9994, eG2 EGM 0.0005, MSC EGM 0.9999, and HUVEC EGM 0.0001, when compared to control HUVEC EBM. Scale bar, 200 μm.

eMPC, embryonic mammary progenitor cell; MSC, mesenchymal stem cell; HUVEC, Human umbilical vein endothelial cells; EGM, endothelial cell growth media EBM, endothelial basal growth media.

Figure 5. eMPC lines are clonogenic and display distinct acinar morphologies when grown as mammopheres in 3-D culture.

(A) Bright-field images of single cell-derived spheres. Scale, 40 μm.

(B) Colony-formation ability presented in box plots with whiskers denoting minimum and maximum values showing number of cells (out of 10,000 cells plated) that gave rise to spheres in anchorage-independent cultures.

Statistical significance was computed for each cell line compared to control (MSC) using one-way ANOVA and Dunnett's multiple comparisons test, where ****$P \leq 0.0001$, ** $P \leq 0.01$, *$P \leq 0.05$, ns = not significant. Two-tailed P values are ePool 0.0152, e1 0.0001, e2 0.3582, eG1 0.0092, and eG2 0.1107 when compared to control MSC. (n = 3 mean ± s.e.m.).

(C) Compactness of individual spheres presented where a circle with Form factor = 1 is most condensed. eG1 forms more compact spheres compared to those of e1, which is also evident in images in (A). Statistical significance was computed using unpaired t-test, where two-tailed ****P = 0.0001. Bar indicates mean value ± s.e.m. n = 3.

(D) Images of e1 and eG1 spheres stained by immunofluorescence with CD44 (red), an adhesion receptor expressed by basal mammary cells which mediates epithelial-stromal and cell-cell interactions and DAPI (blue).

Scale bar, 50 μm.

eMPC, embryonic mammary progenitor cell; MSC, mesenchymal stem cell.

Figure 6. eMPCs have lactogenic potential.

(A) Alveologenesis assay results for eMPC clones and positive control (MEC). Formation of alveolar-like structures and milk production was assessed by morphological changes using bright-field microscopy, as well as IF using anti-milk antibody (red). Scale bar, 40 μm.

(B) Quantification of number of alveolar-like structures greater than 100 μm. The plot is shown as count per 24 well. (n = 3, mean ± s.e.m.).

eMPC, embryonic mammary progenitor cell; MEC (mammary epithelial cell).

Figure 7. Loss of *Sox9* in e1 enhances colony formation ability and alters sphere morphology.

(A) Schematic illustration for establishment of *Sox9*-targeted subclones from Cas9-expressing e1.

(B) qRT-PCR for *Sox9* and EMT regulator, *Zeb1*. Each RQ is normalised to Co#1. Statistical significance between control group and *Sox9* knockout group was computed using unpaired t-test. $^{**}P \leq 0.01$, $^{*}P \leq 0.05$. Two-tailed *P* values are 0.0306 for *Sox9* and 0.0036 for *Zeb1*, respectively ($n = 3$, mean with 95% confidence interval).

(C) Immunohistochemistry for Sox9 and Zeb1 showing Zeb1 expression is decreased by *Sox9* reduction. Scale, 200 μm.

(D) Representative images of spheres grown in methylcellulose (left) and colony-formation ability presented in box plots (right) with minimum and maximum values showing number of control or *Sox9*-targeted cells (out of 10,000 cells plated) that gave rise to spheres in anchorage-independent cultures. Statistical significance was calculated using unpaired t-test. Two-tailed $^{****}P = 0.0001$. ($n = 3$, mean ± s.e.m.). Scale bar , 100 μm.

(E) Representative images of spheroids grown in BME from control and *Sox9*-targeted cells. Scale bar, 400 μm.

(F) Quantification of area of spheroids and (G) increase in area of protrusions from spheroids grown in BME from e1/control and e1*Sox9*-KO cells. ($n = 8$, mean + S.D.) Statistical significance was calculated using one-way ANOVA and multiple comparisons, *** P=0.0002 in F and G.

e1, embryonic mammary progenitor cell 1; Co, control, non-targeted cells; KO, knockout; *Sox9*-targeted cells; MEC, mammary epithelial cell.

Figure 8. Effect of *Sox9* ablation on embryonic mammary progenitor cell fate and function.

(A) qRT-PCR for *Sox9*, *Zeb1, Snai2*, *Twist*, *Procr*, *Trp63*, *Acta2*, and *Ecad* in e1/Co#1 and e1/*Sox9*-KO#1 cells. Relative quantification of each sample is normalised to that of e1/Co#1 and shown in Log2 plot with error bar of triplicates. ($n = 3$, mean ± s.e.m.).

(B) Heat map from RNA-seq of e1/control and e1/*Sox9*-KO cells of genes that were significantly up- or down-regulated in all three e1/*Sox9*-KO subclones using both DESeq and Intensity Difference test.

(C) *In vitro* alveologenesis assay results. Formation of alveoli-like structure and milk production was assessed by morphological changes using bright-field microscopy, as well as IF using anti-milk antibody (red) and DAPI. Scale bar, 200 μm.

(D) *In vitro* alveologenesis assay results. Quantification of number of alveoli less than 200 μm, 200-400 μm, and greater than 400 μm per well. Statistical significance between control group and *Sox9* knockout group was computed using two-tailed paired t-test. $P = 0.0044$. ($n = 3$).

e1, embryonic mammary progenitor cell; Co#1/#2, control #1/#2, non-targeted e1 cells; KO#1/#2/#3, knockout #1, #2, #3; *Sox9*-targeted e1 cells.

Acknowledgements

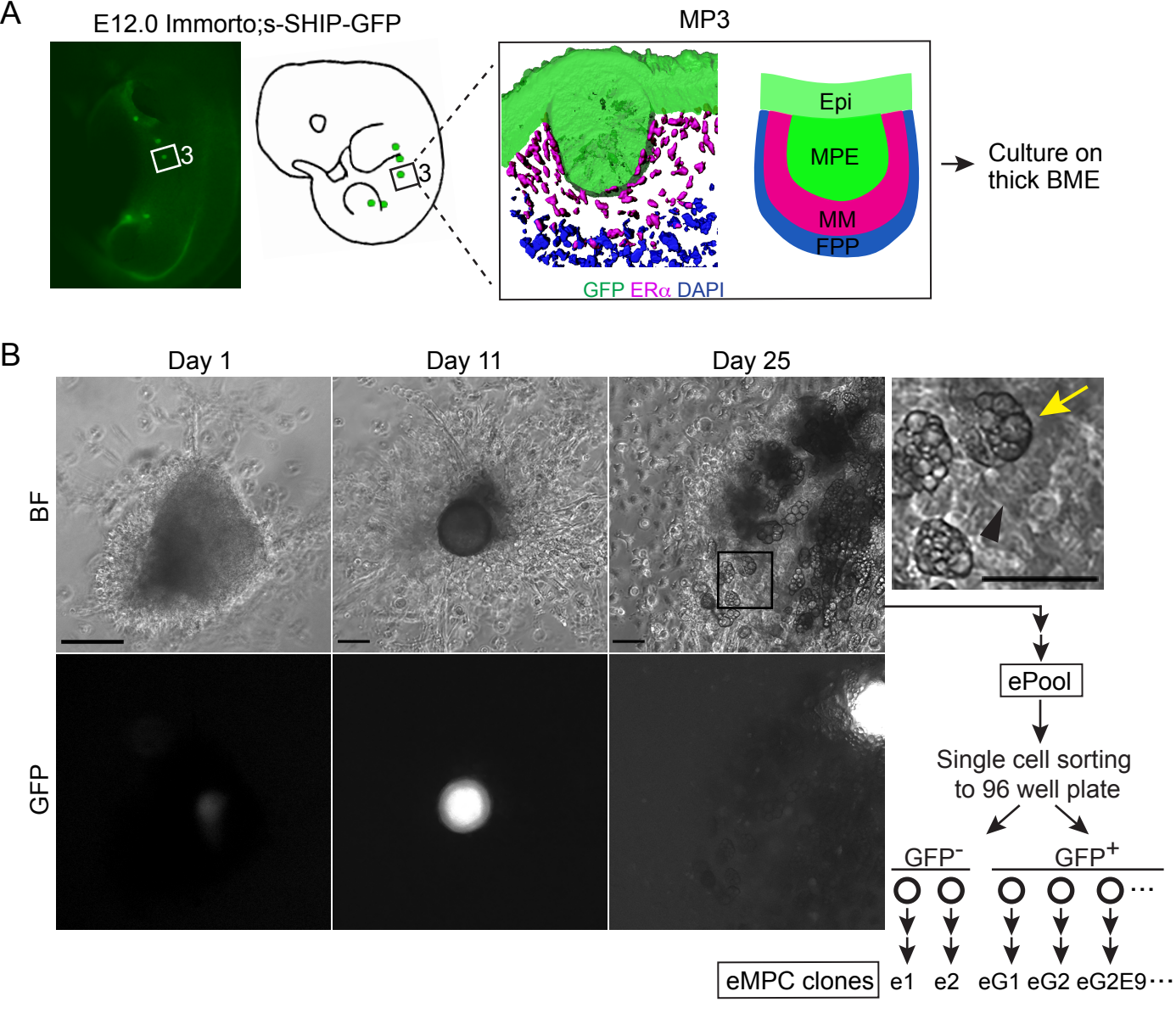
This work was funded by Programme Grants from Breast Cancer Now as part of Programme Funding to the Breast Cancer Now Toby Robins Research Centre. We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics (funded by Wellcome Trust grant reference 090532/Z/09/Z) for the generation of sequencing data and Dr Simon Andrews, Babraham Insititute, for performing bioinformatics analysis.

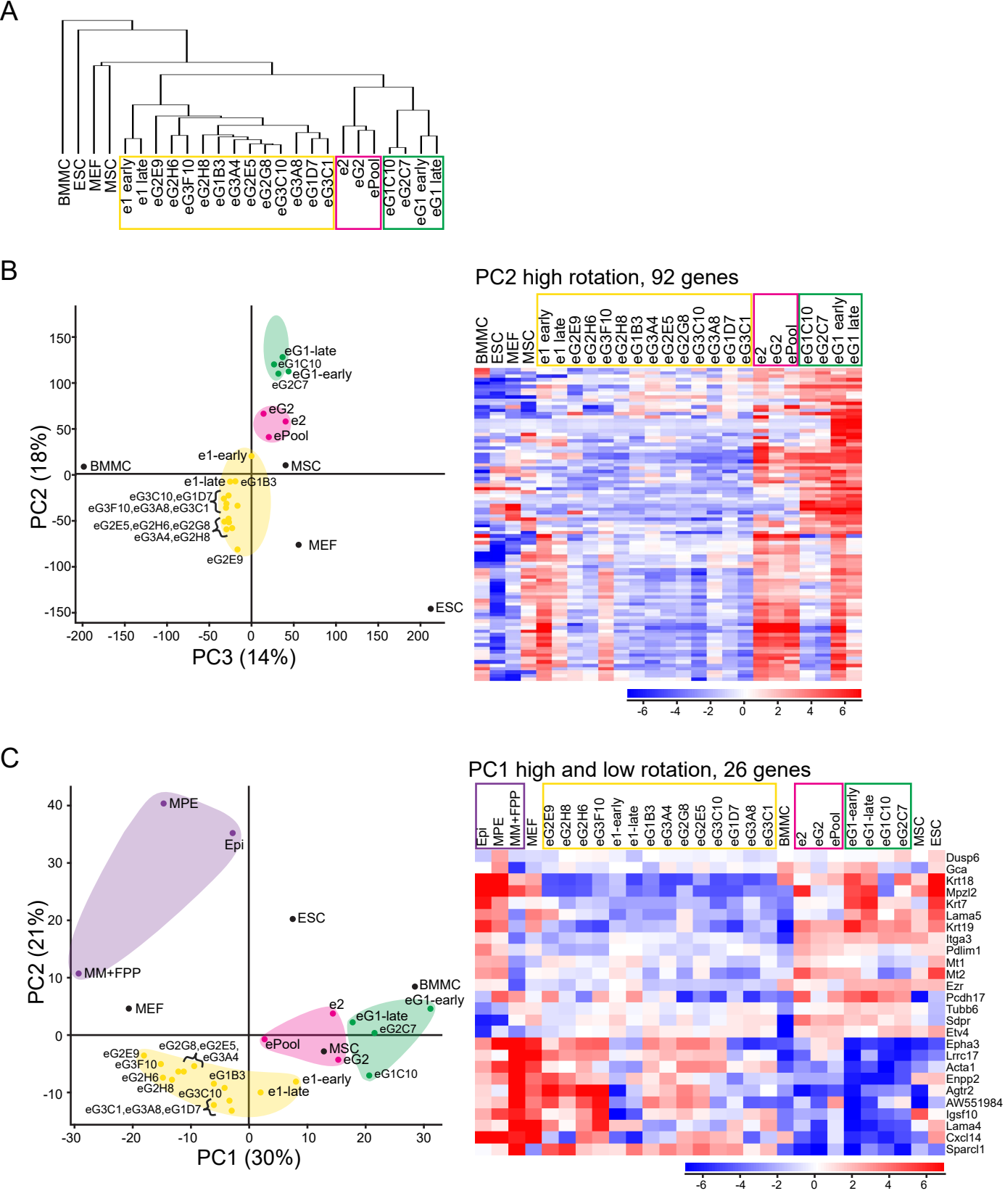
Conflict of interest

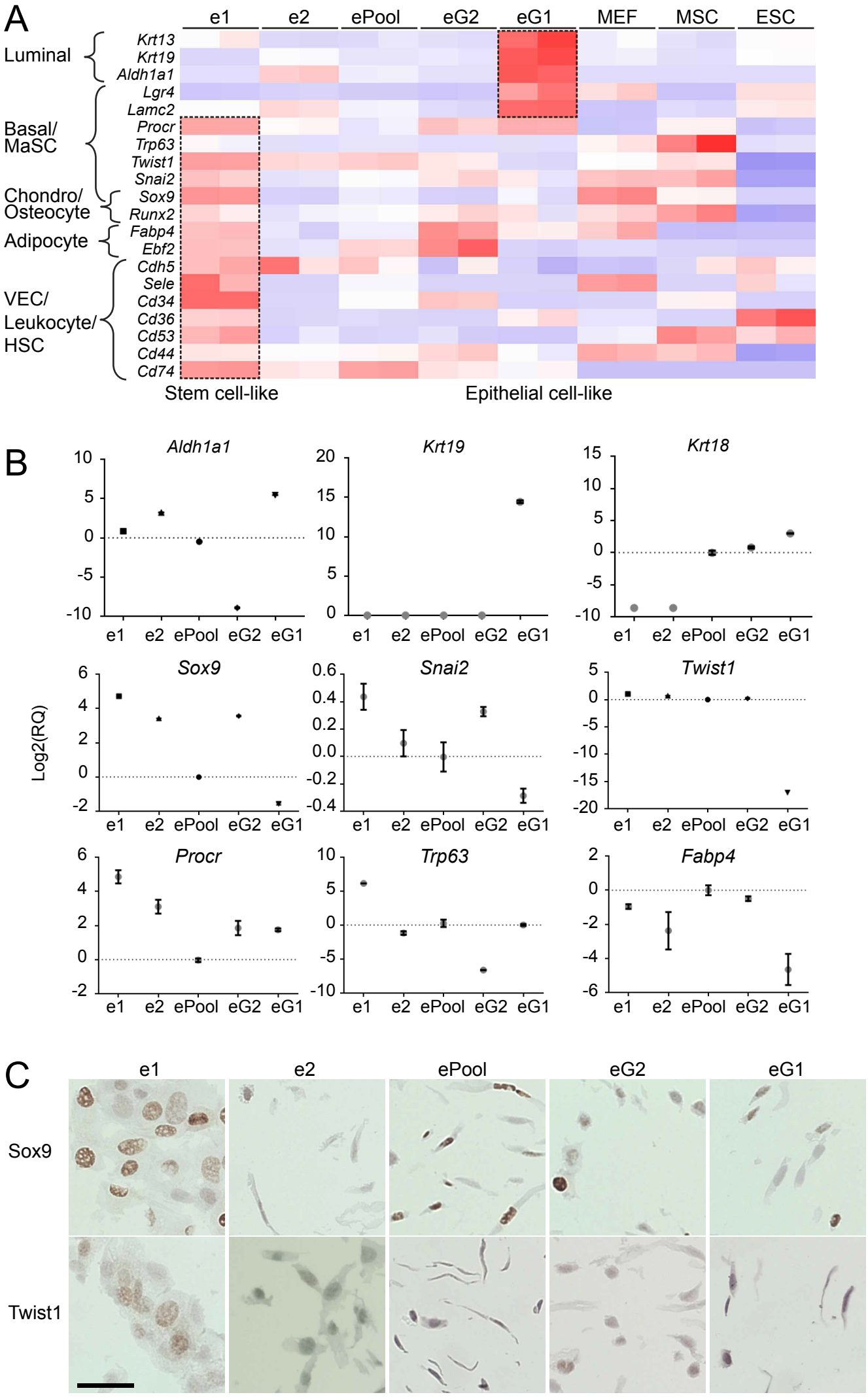
The authors declare no conflict of interest.

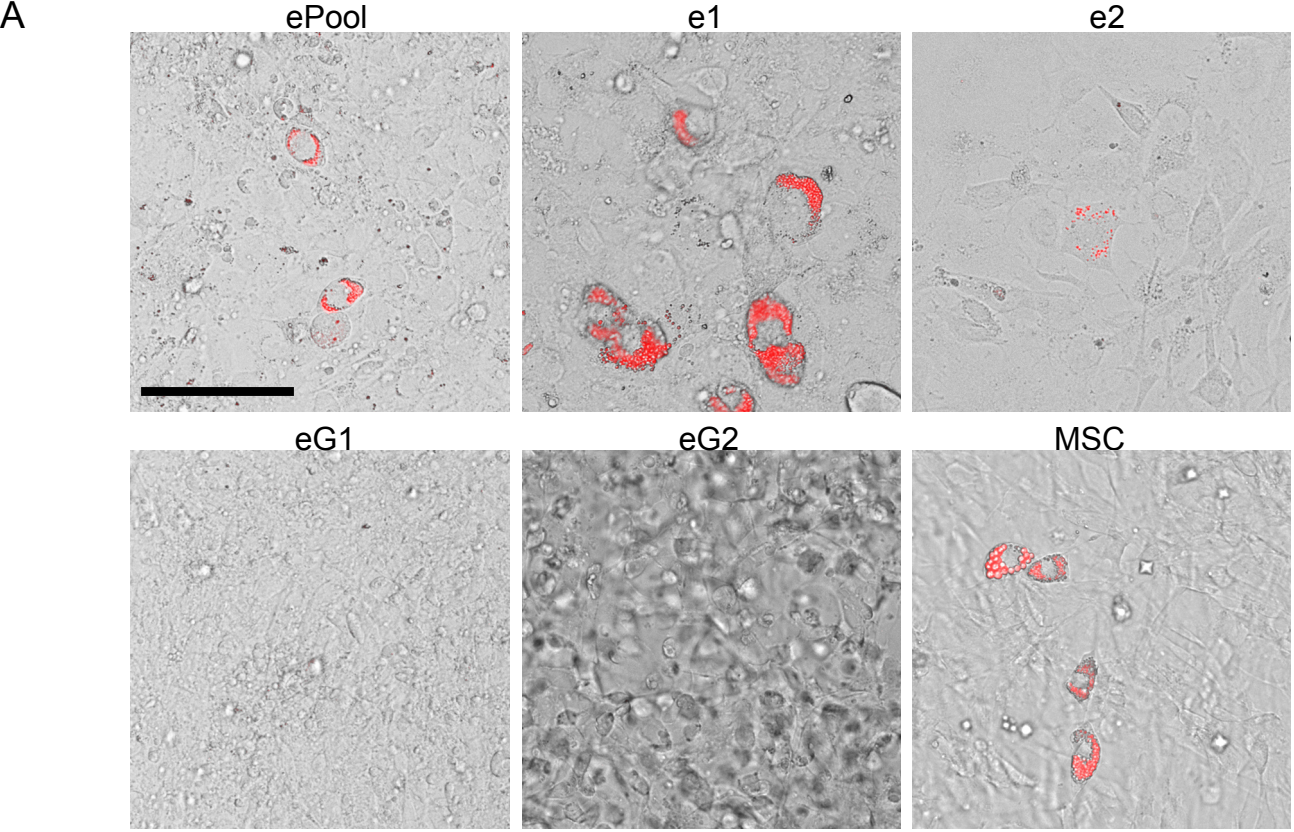
Author contributions

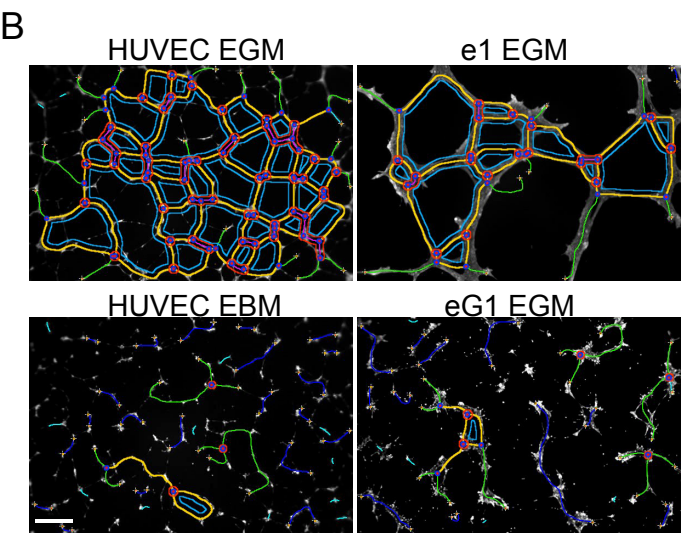
NK and BAH were responsible for study design, experimental work, data analysis and manuscript preparation; NK, PB, MT, EO, AL performed the experiments and analysed the data; NK manuscript editing, experimental details; BAH wrote the manuscript; all authors commented on the manuscript.

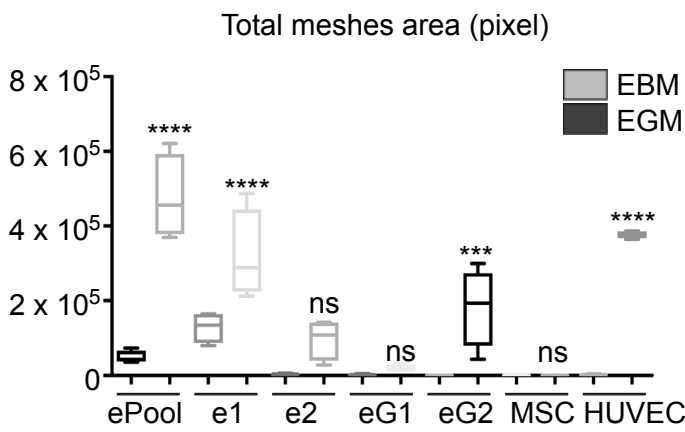

A
E12.0 Immorto;s-SHIP-GFP
3
3
MP3
Epi
MPE
MM
FPP
GFP ERα DAPI
Culture on
thick BME
B
Day 1
Day 11
Day 25
BF
GFP
ePool
Single cell sorting
to 96 well plate
GFP-
GFP+
eMPC clones
e1
e2
eG1
eG2
eG2E9

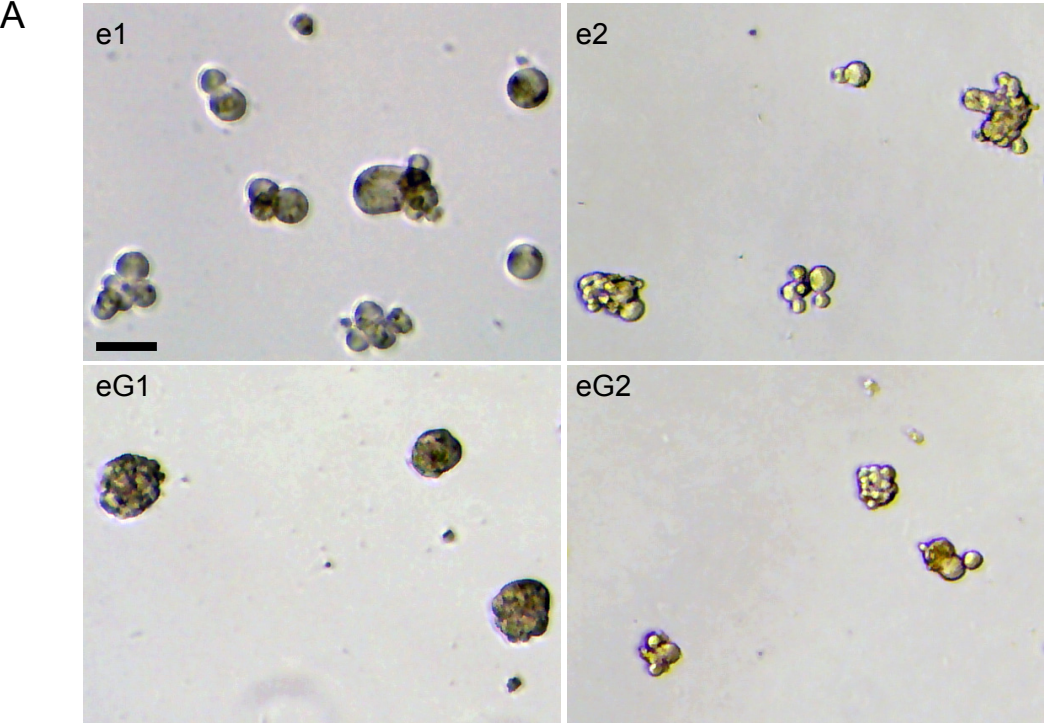

A
BMMC
ESC
MEF
MSC
e1 early
e1 late
eG2E9
eG2H6
eG3F10
eG2H8
eG1B3
eG3A4
eG2E5
eG2G8
eG3C10
eG3A8
eG1D7
eG3C1
e2
eG2
ePool
eG1C10
eG2C7
eG1 early
eG1 late
B
PC2 (18%)
PC3 (14%)
eG1-late
eG1C10
eG1-early
eG2C7
eG2
e2
ePool
e1-early
BMMC
MSC
e1-late
eG1B3
eG3C10,eG1D7
eG3F10,eG3A8,eG3C1
eG2E5,eG2H6,eG2G8
eG3A4,eG2H8
eG2E9
MEF
ESC
PC2 high rotation, 92 genes
C
PC2 (21%)
PC1 (30%)
MPE
Epi
MM+FPP
MEF
ESC
BMMC
eG1-early
e2
eG1-late
eG2C7
ePool
MSC
eG2
eG1C10
e1-early
e1-late
eG2E9
eG3F10
eG2H6
eG2H8
eG1B3
eG3C10
eG3C1,eG3A8,eG1D7
eG2G8,eG2E5,
eG3A4
PC1 high and low rotation, 26 genes
Dusp6
Gca
Krt18
Mpzl2
Krt7
Lama5
Krt19
Itga3
Pdlim1
Mt1
Mt2
Ezr
Pcdh17
Tubb6
Sdpr
Etv4
Epha3
Lrrc17
Acta1
Enpp2
Agtr2
AW551984
Igsf10
Lama4
Cxcl14
Sparcl1

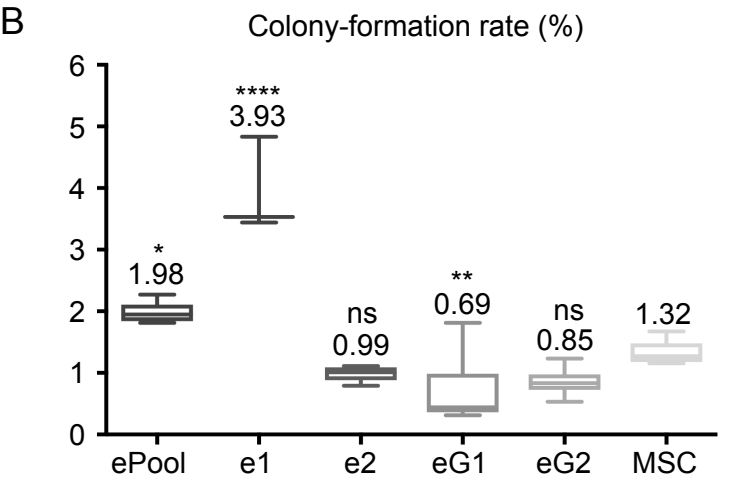

A
e1
e2
ePool
eG2
eG1
MEF
MSC
ESC
Luminal
Krt13
Krt19
Aldh1a1
Basal/
MaSC
Lgr4
Lamc2
Procr
Trp63
Twist1
Snai2
Chondro/
Osteocyte
Sox9
Runx2
Adipocyte
Fabp4
Ebf2
VEC/
Leukocyte/
HSC
Cdh5
Sele
Cd34
Cd36
Cd53
Cd44
Cd74
Stem cell-like
Epithelial cell-like
B
Aldh1a1
Krt19
Krt18
Sox9
Snai2
Twist1
Procr
Trp63
Fabp4
Log2(RQ)
C
Sox9
Twist1

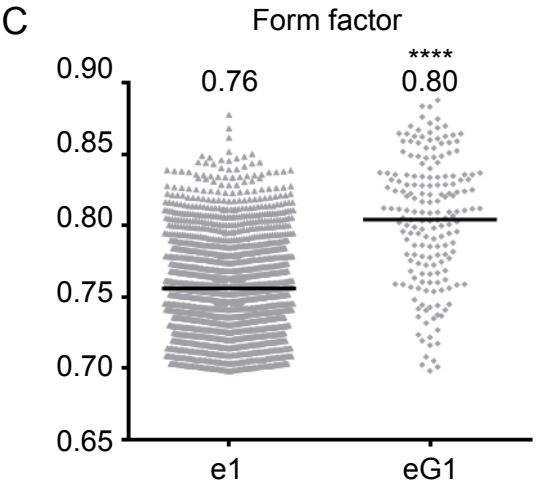

A
ePool
e1
e2
eG1
eG2
MSC

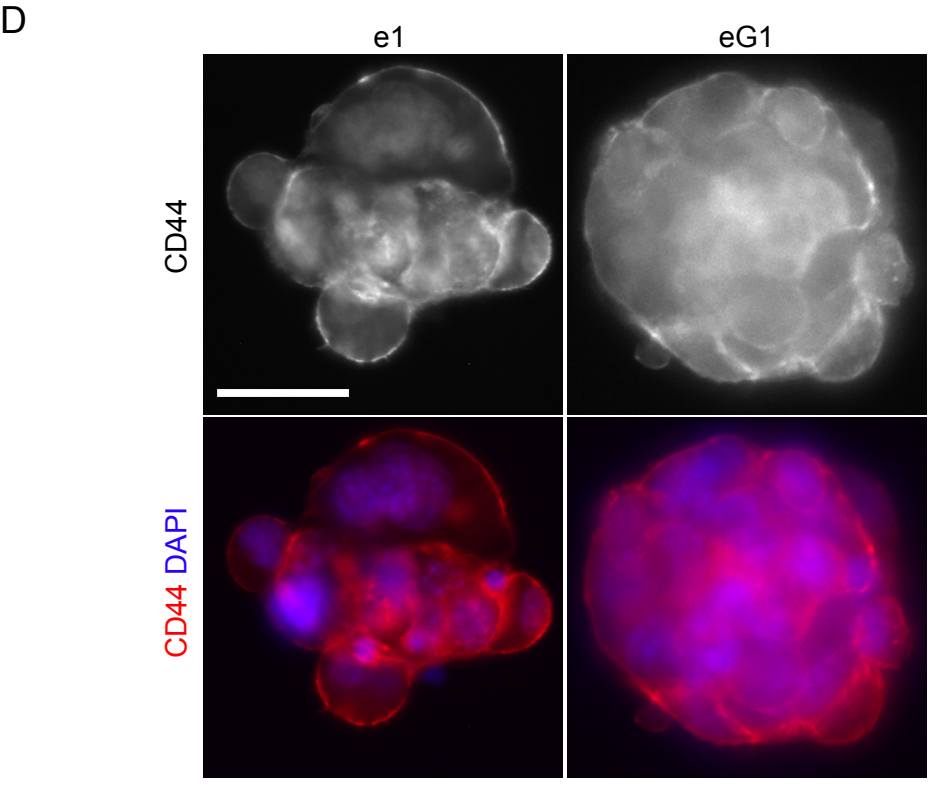

B
HUVEC EGM
e1 EGM
HUVEC EBM
eG1 EGM

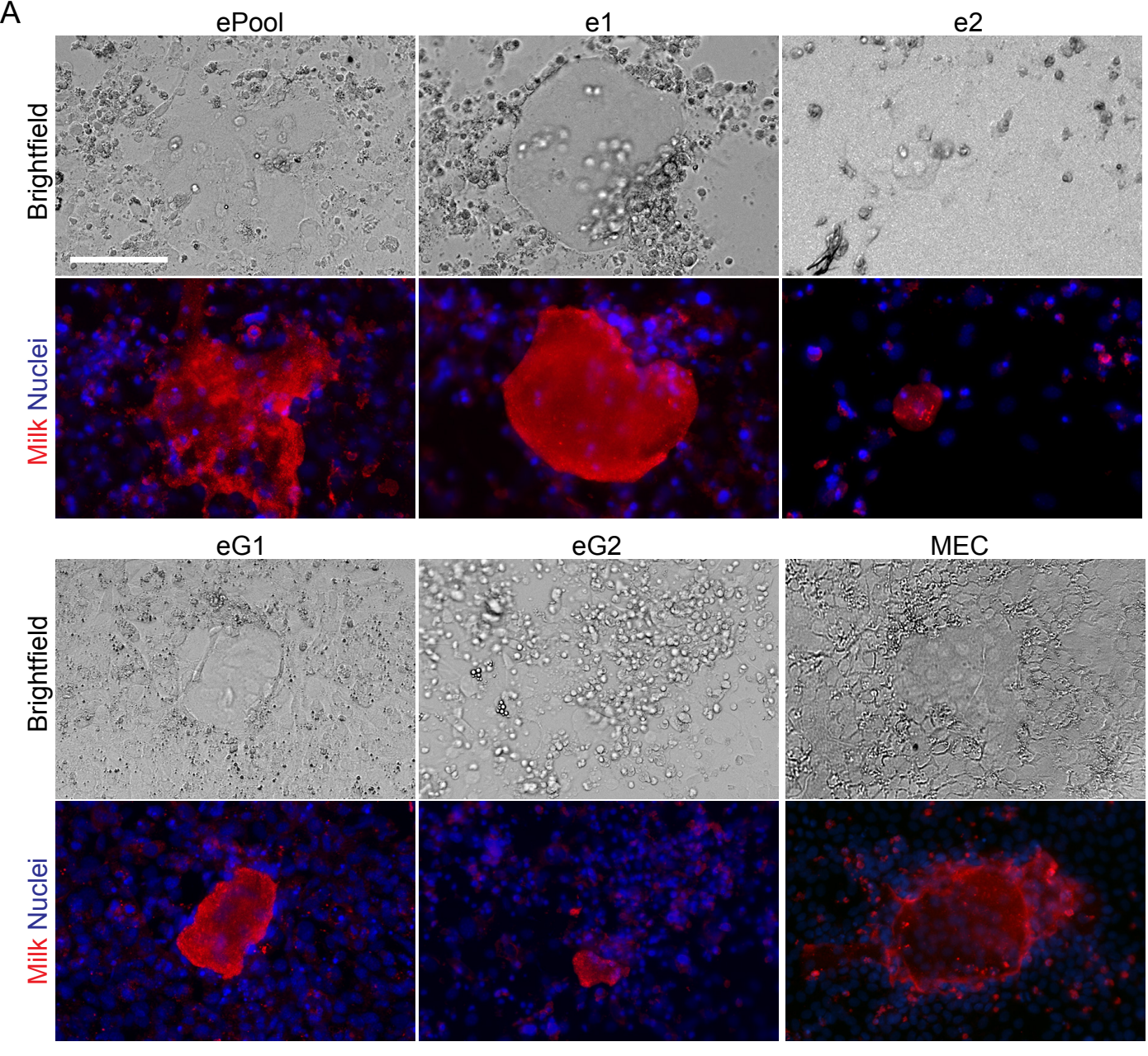

Total meshes area (pixel)
8 x 10^5
6 x 10^5
4 x 10^5
2 x 10^5
0
EBM
EGM
****
****
ns
ns
***
ns
****
ePool
e1
e2
eG1
eG2
MSC
HUVEC

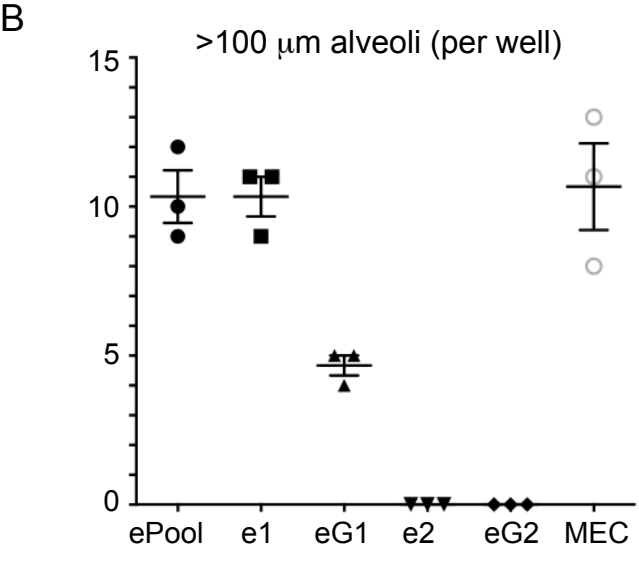

A
e1
e2
eG1
eG2

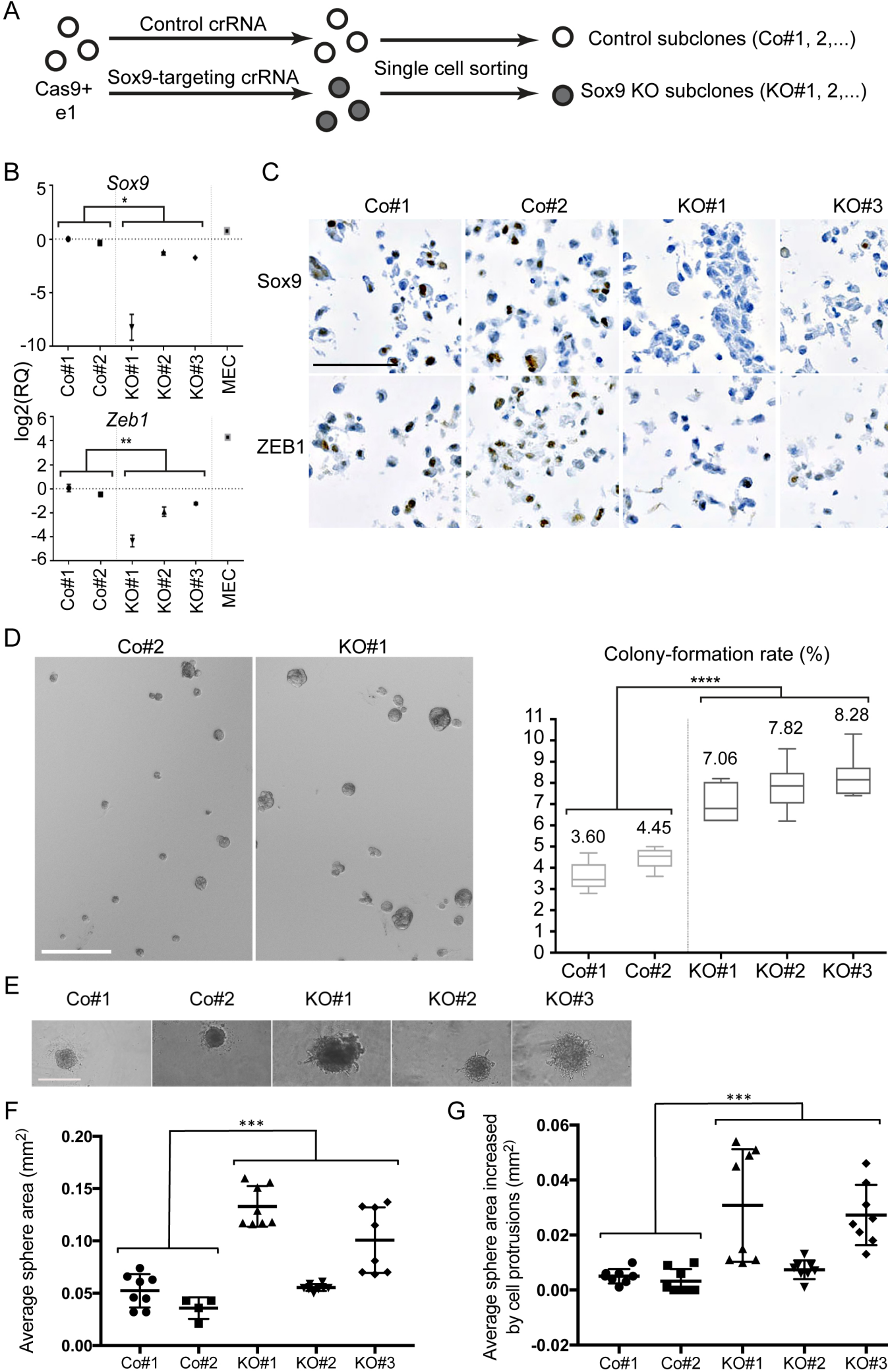

B
Colony-formation rate (%)
6
5
4
3
2
1
0
****
3.93
*
1.98
ns
0.99
**
0.69
ns
0.85
1.32
ePool
e1
e2
eG1
eG2
MSC

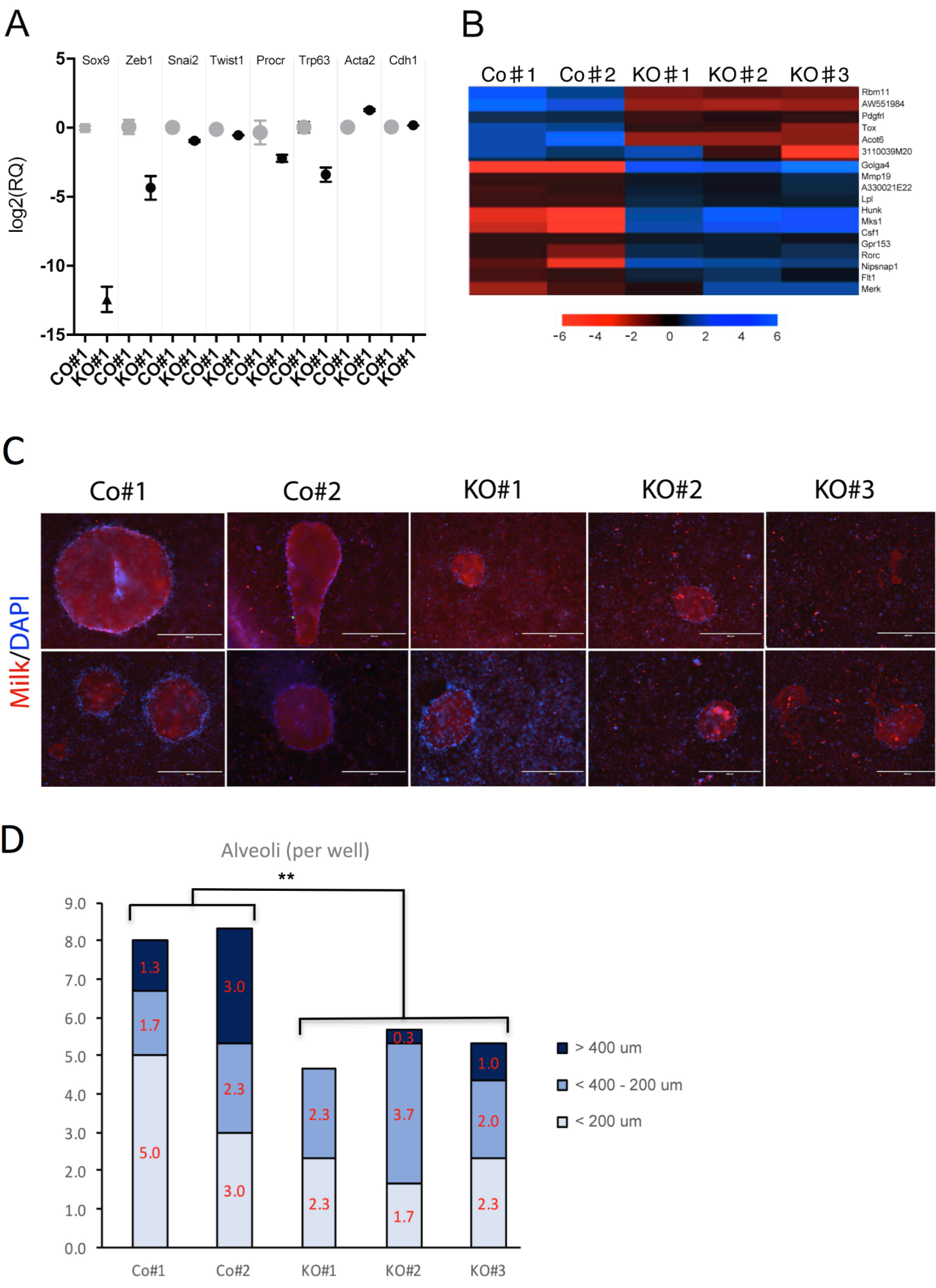

C
Form factor
0.90
0.85
0.80
0.75
0.70
0.65
0.76
****
0.80
e1
eG1


D
e1
eG1
CD44
CD44 DAPI


A
ePool
e1
e2
Brightfield
Milk Nuclei
eG1
eG2
MEC
Brightfield
Milk Nuclei


B
>100 μm alveoli (per well)
15
10
5
0
ePool
e1
eG1
e2
eG2
MEC


A
Control crRNA
Sox9-targeting crRNA
Cas9+ e1
Single cell sorting
Control subclones (Co#1, 2,...)
Sox9 KO subclones (KO#1, 2,...)
B
Sox9
*
log2(RQ)
Co#1
Co#2
KO#1
KO#2
KO#3
MEC
Zeb1
**
C
Co#1
Co#2
KO#1
KO#3
Sox9
ZEB1
D
Co#2
KO#1
Colony-formation rate (%)
****
3.60
4.45
7.06
7.82
8.28
E
Co#1
Co#2
KO#1
KO#2
KO#3
F
***
Average sphere area (mm2)
G
***
Average sphere area increased by cell protrusions (mm2)


A
Sox9
Zeb1
Snai2
Twist1
Procr
Trp63
Acta2
Cdh1
log2(RQ)
5
0
-5
-10
-15
CO#1
KO#1
B
Co#1
Co#2
KO#1
KO#2
KO#3
Rbm11
AW551984
Pdgfrl
Tox
Acot6
3110039M20
Golga4
Mmp19
A330021E22
Lpl
Hunk
Mks1
Csf1
Gpr153
Rorc
Nipsnap1
Flt1
Merk
-6
-4
-2
0
2
4
6
C
Co#1
Co#2
KO#1
KO#2
KO#3
Milk/DAPI
D
Alveoli (per well)
**
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
1.3
1.7
5.0
3.0
2.3
3.0
2.3
2.3
0.3
3.7
1.7
1.0
2.0
2.3
Co#1
Co#2
KO#1
KO#2
KO#3
> 400 um
< 400 - 200 um
< 200 um