High Expression of CD200 and CD200R1 Distinguishes Stem and Progenitor Cell Populations within Mammary Repopulating Units

Gat Rauner,[1,2,4,5] Tania Kudinov,[1,2,5] Shlomit Gilad,[3] Gil Hornung,[3] and Itamar Barash[1,*]

[1]Institute of Animal Science, ARO, The Volcani Center, Bet-Dagan 50250, Israel
[2]The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Jerusalem 7610001, Israel
[3]The Nancy & Stephen Grand Israel National Center for Personalized Medicine, Weizmann Institute of Science, Rehovot 7610001, Israel
[4]Present address: Baker Institute for Animal Health, Cornell University College of Veterinary Medicine, Ithaca, NY 14850, USA
[5]Co-first author
*Correspondence: itamar.barash@mail.huji.ac.il
https://doi.org/10.1016/j.stemcr.2018.05.013

SUMMARY

Aiming to unravel the top of the mammary epithelial cell hierarchy, a subset of the CD49f^{high}CD24^{med} mammary repopulating units (MRUs) was identified by flow cytometry, expressing high levels of CD200 and its receptor CD200R1. These MRU$^{CD200/CD200R1}$ repopulated a larger area of de-epithelized mammary fat pads than the rest of the MRUs, termed MRU$^{not\ CD200/CD200R1}$. MRU$^{CD200/CD200R1}$ maintained a much lower number of divergently defined, highly expressed genes and pathways that support better cell growth, development, differentiation, and progenitor activity than their MRU$^{not\ CD200/CD200R1}$ counterparts. A defined profile of hierarchically associated genes supporting a single-lineage hypothesis was confirmed by *in vitro* mammosphere analysis that assembled 114 genes with decreased expression from MRU$^{CD200/CD200R1}$ via MRU$^{not\ CD200/CD200R1}$ toward CD200$^+$CD200R1$^-$ and CD200R1$^+$CD200$^-$ cells. About 40% of these genes were shared by a previously published database of upregulated genes in mammary/breast stem cells and may represent the core genes involved in mammary stemness.

INTRODUCTION

The mammary gland is a highly regenerative organ that experiences periods of development, function, and regression dictated by embryo and newborn development. The mammary gland develops as a rudimentary ductal network headed by terminal end buds in mice, or terminal ductal lobuloalveolar units in humans and bovines, which branch upon pregnancy (Cardiff and Wellings, 1999). Hormonally stimulated lobuloalveolar structures differentiate into milk-producing alveoli during lactation and involute upon neonate weaning. These changes mainly involve the epithelial layers. Indeed, the mammary gland is an epithelial organ. Luminal epithelium lines the apical part of the ducts and encompasses the milk-producing cells upon lactation. An outer contractile basal myoepithelium forces the milk toward the nipple (Hennighausen and Robinson, 2005).

A Lin$^-$, CD24^{med}CD49f^{high}/CD24$^+$CD29^{high} cell population was identified as the entity that enables mammary gland development and regeneration (Shackleton et al., 2006; Stingl et al., 2006). This population is enriched in mammary repopulating units (MRUs) with stem-like properties that can generate an entire functional gland upon transplantation into de-epithelized mammary fat pad. Further studies reported multiple characteristics of a relatively heterogeneous entity: MRUs with relatively lower repopulation ability but rapid growth were identified during pregnancy (Asselin-Labat et al., 2010). Possibly different, actively proliferating s-SHIP-expressing stem cells with

relatively high self-renewal ability were identified at the forefront of the virgin's developing terminal end buds and in the alveolar buds during pregnancy (Bai and Rohrschneider, 2010). In contrast, slow-cycling H2b-GFP$^+$ mammary stem cells exhibiting Cd1d expression were also identified (dos Santos et al., 2013).

An additional layer of complexity at the top of the mammary cell hierarchy was added by the identification of distinct long-lived progenitors (Rios et al., 2014), which may resemble the multipotent progenitors that descend from the hematopoietic stem cells (Seita and Weissman, 2010). Further characterization of this defined population is needed to understand the initial steps of stem cell differentiation.

CD200 (OX-2), a transmembrane surface glycoprotein, is a member of the immunoglobulin superfamily. It is widely distributed across tissues, including lymphocytes, endothelial cells, keratinocytes, and neuronal cells (Ko et al., 2009). Its cognate receptor, CD200R1, is also an immunoglobulin transmembrane glycoprotein and is mainly expressed on myeloid cells. Binding of CD200 to its receptor attenuates immune activity (Minas and Liversidge, 2006). No information regarding CD200's role in the normal mammary gland is currently available. However, its expression has been detected in breast cancer (Leccia et al., 2012).

CD200 expression has been associated with stemness. In the presence of low expression levels of CD24, CD34, CD71, and CD146, high CD200 expression enabled successive enrichment of stem cells in human hair follicle bulge
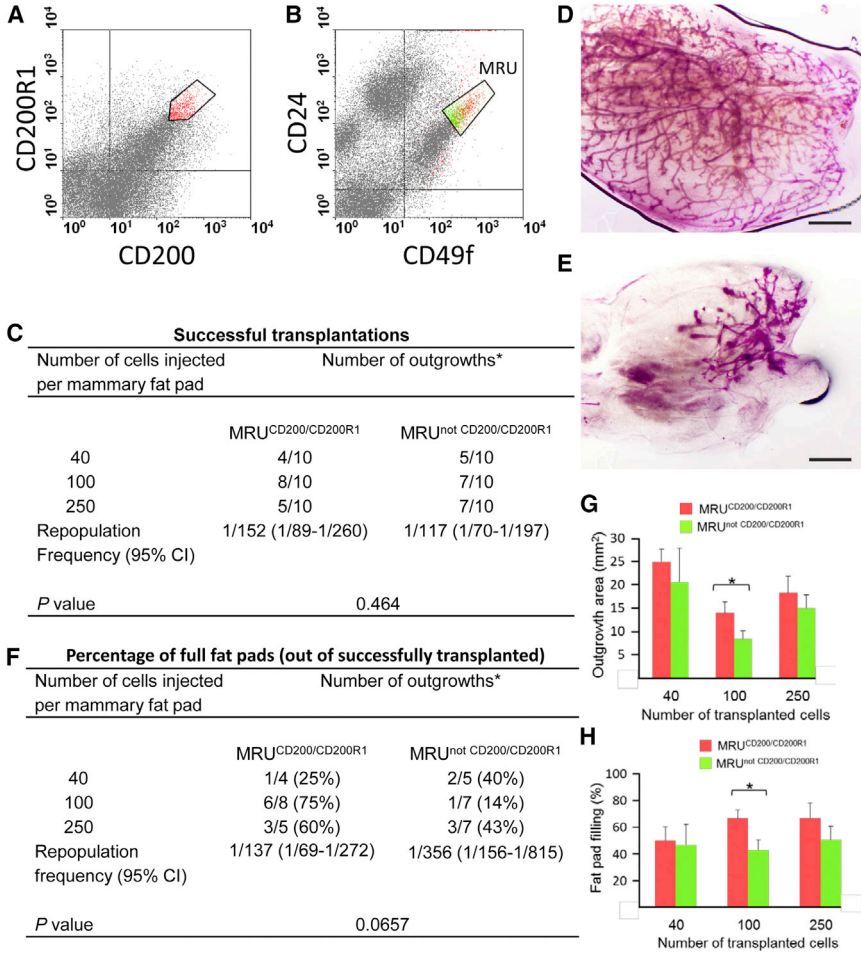
Figure 1. MRUs that Express High Levels of CD200 and CD200R1 Exhibit Better Repopulation Ability Than the Other MRUs

(A and B) Dot plots depicting the gating strategy for mouse mammary cells sorted for transplantation. Cells were analyzed simultaneously for CD200, CD200R1, CD24, and CD49f expression. (A) Identification of CD200^{high}CD200R1^{high} epithelial cell population. (B) Projection of the CD200^{high}CD200R1^{high} population on the CD49/CD24 axis identified two MRU (CD24^{med}CD49f^{high}) subpopulations: MRU$^{CD200/CD200R1}$ that expresses high levels of both CD200 and CD200R1 (red) and MRU$^{not\ CD200/CD200R1}$ that represents the rest of the MRUs (green).

(C) Limiting dilution analysis of the repopulating frequency of MRU$^{CD200/CD200R1}$ and MRU$^{not\ CD200/CD200R1}$ cells from 8-week-old virgin mice. CI, confidence interval. *Number of outgrowths per number of injected fat pads.

(D and E) Whole-mount Carmine alum staining of transplanted mouse fat pads depicting fully (D) and partially (E) reconstituted glands. Bar, 1 mm.

(F) Limiting dilution analysis of the potential for full fat pad occupancy by cells of the two MRU subpopulations. *Number of outgrowths per number of injected fat pads. For details, see (C).

(G and H) Analysis of the absolute and relative areas occupied by new epithelium after transplantation of the indicated number of cells from the two MRU subpopulations. Bars represent mean ± SEM of three replications. Asterisks mark statistically significant difference (p < 0.05).
See also Tables S1 and S5.

C

Successful transplantations

Number of cells injected per mammary fat pad	Number of outgrowths*	
	MRU$^{CD200/CD200R1}$	MRU$^{not\ CD200/CD200R1}$
40	4/10	5/10
100	8/10	7/10
250	5/10	7/10
Repopulation Frequency (95% CI)	1/152 (1/89-1/260)	1/117 (1/70-1/197)
P value	0.464	

F

Percentage of full fat pads (out of successfully transplanted)

Number of cells injected per mammary fat pad	Number of outgrowths*	
	MRU$^{CD200/CD200R1}$	MRU$^{not\ CD200/CD200R1}$
40	1/4 (25%)	2/5 (40%)
100	6/8 (75%)	1/7 (14%)
250	3/5 (60%)	3/7 (43%)
Repopulation frequency (95% CI)	1/137 (1/69-1/272)	1/356 (1/156-1/815)
P value	0.0657	

cells (Ohyama et al., 2006). Furthermore, high CD200 expression distinguished fetal- from maternal-originated placental mesenchymal stem cells (Zhu et al., 2014). In the mammary gland, however, high CD200 expression levels do not appear to mark stem cells (G.R., unpublished).

Seeking markers of mammary epithelial stem cells, a CD200^{high}CD200R1^{high} epithelial cell population was identified in the current study. Members of this population enriched the MRUs for cells with better repopulating activity compared with the rest of the MRUs. Gene-expression profiling combined with comparative dataset analysis, supported by further *in vivo* and *in vitro* studies, suggested the coexistence of two MRU subsets: a poorly metabolically active stem cell population with relatively high complement activity, and highly active differentiation-oriented progenitors.

RESULTS

An MRU Subpopulation Expressing High Levels of CD200 and CD200R1 Exhibits Enhanced Repopulation Ability

A mouse mammary cell population expressing high levels of CD200 and CD200R1 was identified by flow cytometry (Figure 1A). This population included 82.8% ± 16.6% (n = 3) epithelial cells and represented 3.3% ± 0.8% of the mammary epithelial cells. Projecting the CD200^{high}CD200R1^{high} cells on the CD24/CD49f expression axes located 49.2% ± 18.7% of the cells within CD24^{med}CD49f^{high} (MRU) boundaries (Stingl et al., 2006), representing 50.1% ± 11.9% (n = 3) of the MRUs (Figure 1B). The MRUs that expressed high CD200 and CD200R1 levels are termed here MRU$^{CD200/CD200R1}$. To examine their

repopulating potential, outgrowth development from these cells was compared with that developing from the rest of the MRUs, termed MRU$^{\text{not CD200/CD200R1}}$. As shown in Figure 1C, no difference was observed between the repopulating potential of the two subpopulations, and 40%–50% of the fat pads transplanted with 40 cells from each MRU subset were occupied by newly developed epithelium. Further analyses identified fat pads that were completely filled with outgrowths, whereas others were only partially occupied (Figures 1D and 1E, respectively). Transplantation of limiting numbers of MRU$^{\text{CD200/CD200R1}}$ and MRU$^{\text{not CD200/CD200R1}}$ into cleared mammary fat pads revealed a 2.6-fold decrease in full repopulation frequency for the MRU$^{\text{not CD200/CD200R1}}$ (versus MRU$^{\text{CD200/CD200R1}}$), which tended toward significance (p = 0.06, Figure 1F).

The absolute and relative areas of the reconstituted glands were also larger by 30% and 33%, respectively, for glands transplanted with MRU$^{\text{CD200/CD200R1}}$ compared with those transplanted with MRU$^{\text{not CD200/CD200R1}}$ (Figures 1G and 1H). Significant differences (p ≤ 0.05) between the two subpopulations were obtained for transplantation of 100 cells. Combining data from transplanting 100 and 250 cells for each of the two MRU subpopulations (Figure 1G) also resulted in significantly (p ≤ 0.05) higher occupancy rates for the MRU$^{\text{CD200/CD200R1}}$ developing cells (not shown). It is possible that transplantation of 100 cells per fat pad provided the most permissive environment, necessary for executing the outgrowth's full developmental potential.

In a serial transplantation analysis, mammary epithelial cells prepared from pooled outgrowths developed from MRU$^{\text{CD200/CD200R1}}$ occupied 11% (2/18) of the fat pads transplanted. In contrast, MRU$^{\text{not CD200/CD200R1}}$ did not give rise to new epithelium (Table S1).

Differential Gene Expression Allocates a Larger Number of Highly Expressed Genes and Activated Pathways to the MRU$^{\text{not CD200/CD200R1}}$ Subpopulation

The observed difference in fat pad occupancy upon transplantation of the two MRU subpopulations reflected diversity within the MRUs. Therefore, the two MRU-composing populations, as well as the non-MRU CD200$^+$ population (MRU$^-$CD200R1$^-$) and CD200R1$^+$ population (MRU$^-$CD200$^-$), were sorted. Sets of 1,000 cells of each population were collected from five individual mice, lysed, and subjected to RNA sequencing (RNA-seq). On average, ~21 million reads per sample were obtained. Of these, ~95% were mapped to the mouse genome and ~57% were found on known genes. Gene-expression analysis clearly clustered the samples into four cell populations, independent of the mouse effect (Figure 2A). To further distinguish the two MRU subpopulations, the expression level of each gene in MRU$^{\text{CD200/CD200R1}}$ cells was related to its expression in the MRU$^{\text{not CD200/CD200R1}}$ cells in individual glands, and log$_2$

values were calculated. Values of >1 indicated higher expression in MRU$^{\text{CD200/CD200R1}}$ and <−1 marked higher expression in their counterpart. As shown in Figure 2B, a larger number of highly expressed genes was detected in the MRU$^{\text{not CD200/CD200R1}}$. The difference between this population and the MRU$^{\text{CD200/CD200R1}}$ population ranged from 2.3-fold with significant differences of p ≤ 0.05 up to 10.3-fold for adjusted p ≤ 0.01. Within a defined list of annotated genes with 1.5-fold difference between the two populations and adjusted p < 0.07, the six genes with the highest relative expression in the MRU$^{\text{CD200/CD200R1}}$ were *CD5l*, *Siglec1*, *Kcnn1*, *Fcna*, *C6,* and *Ccl8* (Figure 2C). These genes were expressed at 5.5- to 8.6-fold higher levels than in the MRU$^{\text{not CD200/CD200R1}}$ cells, and were all associated with immune system regulation (Asano et al., 2015; Begenisich et al., 2004; Endo et al., 2012; Maggiani et al., 2011; Martinez-Picado et al., 2016; Mayilyan, 2012). Higher differences, 64- to 194-fold their counterparts' expression, characterized genes that were highly expressed in the MRU$^{\text{not CD200/CD200R1}}$; none of these genes had a direct association with immune regulation. This list included *Igf2*, coding a protein hormone known to regulate cell proliferation, growth, migration, differentiation, and survival (Bergman et al., 2013), but whose effect on mammary gland development has not been elucidated.

Pathway activity supported by defined genes was determined using Ingenuity Pathway Analysis (IPA) software. The parameters defined for selected differences in gene expression between the two MRU populations, >1.5-fold difference, adjusted p < 0.05, provided a list of 1,278 genes. Importantly, only a few pathways were highly activated in the MRU$^{\text{CD200/CD200R1}}$ compared with MRU$^{\text{not CD200/CD200R1}}$ cells (Figures 2D and S1). Among these, the highest Z-scores were associated with genes regulating the complement and coagulation systems. Many more pathways were highly activated in the MRU$^{\text{not CD200/CD200R1}}$ population. Higher Z-scores were calculated for nitric oxide signaling, interleukin (IL)-8 signaling, and endothelial nitric oxide synthase (eNOS) (Figure 2E). Pathway activities that characterized MRU$^{\text{CD200/CD200R1}}$ cells involved inflammatory response, organismal injury, cell death, and cancer. In contrast, major activity in the MRU$^{\text{not CD200/CD200R1}}$ cells involved cell movement, tissue development, and cell development, growth, and differentiation (Figure S2).

Table 1 assembles the main molecular regulators and activators of the IPA-defined pathways. It shows that the molecular basis of the highly activated pathways in the MRU$^{\text{not CD200/CD200R1}}$ population consisted of an induced core of genes encoding multipotent kinases, GTPases, and G-protein subunits that support major cellular activities: proliferation, differentiation, and migration. Fewer encoded proteins: SRC, protein kinase C (PKC), and SHIP, maintained
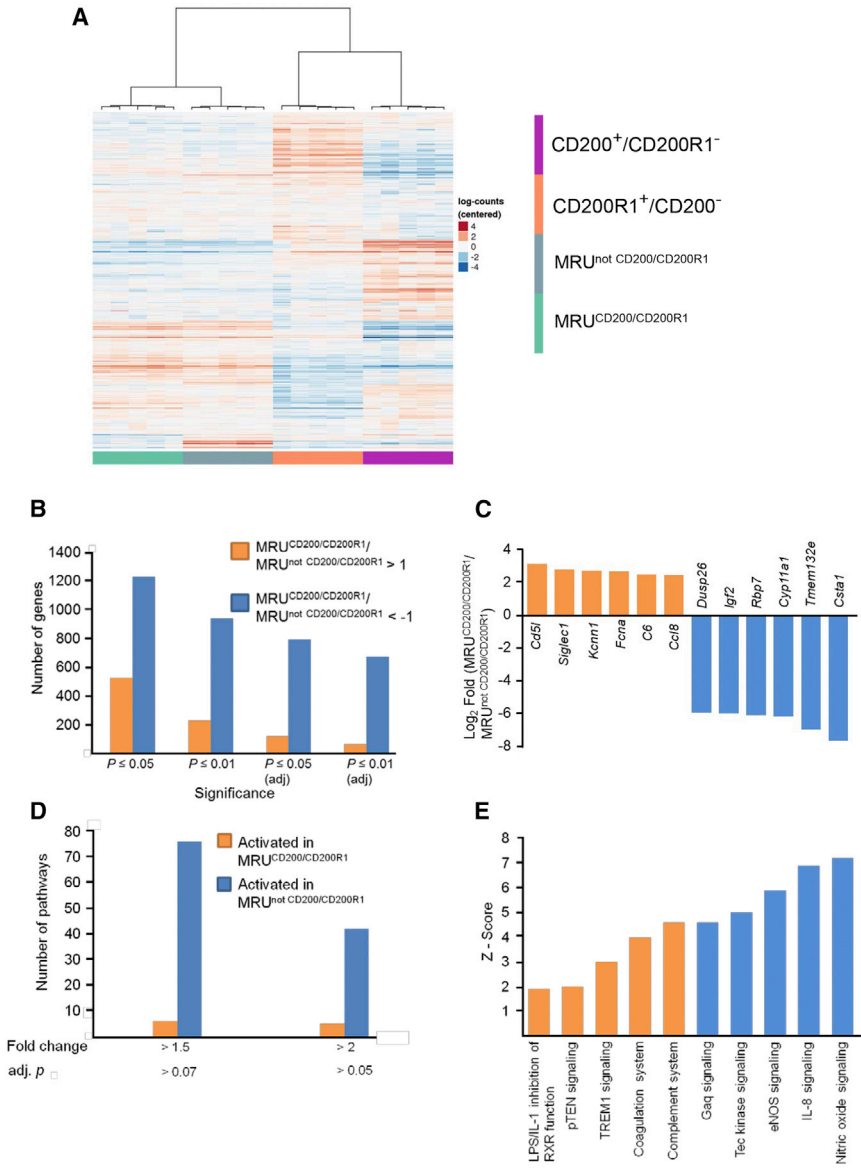
Figure 2. Unsupervised Hierarchical Clustering of Distinct Epithelial Cell Populations in This Study

The MRU$^{not\ CD200/CD200R1}$ subpopulation maintains a higher number of highly expressed genes and metabolic pathways compared with its MRU$^{CD200/CD200R1}$ counterpart.

(A) Heatmap demonstrates clear separation of the MRU$^{CD200/CD200R1}$, MRU$^{not\ CD200/CD200R1}$, CD200$^+$CD200R1$^-$ and CD200$^-$CD200R1$^+$ populations from mammary glands of five individual mice.

(B) Number of genes that are highly expressed in each of the compared subpopulations according to the defined criteria.

(C) The most highly expressed genes in each of the two MRU subpopulations.

(D) Number of pathways, determined by IPA, that are highly activated in each of the compared populations according to the defined criteria.

(E) Highly activated pathways in each of the compared populations.

See also Figures S1 and S2.

two possible hierarchical interactions with the MRU$^{not\ CD200/CD200R1}$ population. The latter may represent progenitors of the MRU$^{CD200/CD200R1}$ population (Figure 3A) or an independent subgroup of stem cells (Figure 3B). The much higher metabolic activity and, in particular, expression levels of pro-differentiation molecules such as p38 mitogen-activated protein kinase (MAPK) and ERK1/2 in the MRU$^{not\ CD200/CD200R1}$ population

the opposite activity in these pathways and may serve as negative regulators. A much more modest repertoire of activated pathways characterized the MRU$^{CD200/CD200R1}$ population. High activity of C proteins supported higher activity of the complement system. A chaperone of coagulation factor VII, von Willebrand factor (vWF), IL-6, and MAGI negatively regulate the coagulation system, and TREM1 and pTEN signaling, respectively, which were also upregulated.

Expression of CD200 and CD200R1 Marks Distinct MRU Subpopulations with Stem and Progenitor Characteristics

The higher occupancy rate of the mammary fat pad by cells of the MRU$^{CD200/CD200R1}$ population (Figure 1) suggested

supported a single hierarchy, initiated by stem cells via immediate progenitors toward differentiated CD200- and CD200R1-expressing cells.

To confirm this hypothesis, a gene subset was arrayed, with an expression profile that fit the pattern of a single lineage consisting of decreased expression from the MRU$^{CD200/CD200R1}$ cells via the MRU$^{not\ CD200/CD200R1}$ cells toward mammary epithelial cells expressing either CD200 or CD200R1 (Figure 3C). Initial gene ontology analysis of the 114 identified genes (Table S2) showed a strong association between 19 of them and epithelium. The contribution of other tissues—epidermis, muscle, and mesenchyme—also produced significant false discovery rate values (0.001–0.0.4) (data not shown). Next, the 114

Table 1. Key Encoded Proteins that Are Highly Involved in Mediating Pathway Activity in MRU[CD200/CD200R1] and MRU[not CD200/CD200R1] Populations

Key Proteins	Pathways
Key proteins highly expressed in MRU[CD200/CD200R1] population[a]	highly active pathways in MRU[CD200/CD200R1] population[a]
Complement (C1q, C3, C3a, C3b, C4b, C6)	complement pathway
Key proteins highly expressed in MRU[not CD200/CD200R1] population[b]	highly active pathways in MRU[CD200/CD200R1] population[a]
vWF	coagulation system
IL6	TREM1 signaling
MAGI	pTEN signaling
Key proteins highly expressed in MRU[not CD200/CD200R1] population[b]	highly active pathways in MRU[not CD200/CD200R1] population[b]
PI3 kinase	eNOS signaling; nitric oxide signaling; IL-8 signaling; Tec signaling; Gαq signaling; inhibition of angiogenesis by TSP1; endothelin-1 singling; anti-proliferative role of somatostatin receptor 2; thrombin signaling; HMGB1 signaling; p70s6k signaling; integrin signaling; hepatocyte growth factor (HGF) signaling
ERK 1/2	nitric oxide signaling; IL-8 signaling; antiproliferative role of somatostatin receptor 2; vascular endothelial growth factor (VEGF) family of ligand receptor interactions; HMGB1 signaling; prolactin signaling; agrin interactions; Stat3 signaling; nuclear factor (NF)-κB signaling; thrombin signaling; acute-phase response signaling; CAMP-mediated signaling; chemokine signaling; Rho signaling; role of NFAT in cardiac hypertrophy; renin-angiotensin signaling; HGF signaling; mouse embryonic stem cell pluripotency
p38 MAPK	inhibition of angiogenesis by TSP1; thrombin signaling; acute-phase response; type 1 diabetes; HMGB1 signaling; iNOS signaling; chemokine signaling; IL-1 signaling; VEGF signaling; TGFβ signaling; mouse embryonic stem cell pluripotency
PKC	nitric oxide signaling; IL-8 signaling; eNOS signaling; Inhibition of angiogenesis by TSP1; p70s6k signaling
Gα	Tec kinase signaling; endothelin 1 signaling; signaling by Rho family GTPases; IL-1 signaling
Gβ	Tec kinase signaling; IL-8 signaling; Gαq signaling; cardiac β-adrenergic signaling; antiproliferative role of somatostatin receptor 2; thrombin signaling; role of NA in cardiac hypertrophy; signaling by Rho family GTPases; α-adrenergic signaling; IL-1 signaling
GY	Tec kinase signaling; IL-8 signaling; Gαq signaling; cardiac β-adrenergic signaling; antiproliferative role of somatostatin receptor 2; thrombin signaling; role of NA in cardiac hypertrophy; signaling by Rho family GTPases; α-adrenergic signaling; IL-1 signaling
RAS	phospholipase signaling; thrombin signaling; acute-phase response; NF-κB signaling; integrin signaling; antiproliferative role of somatostatin receptor 2; thrombin signaling; glioma invasiveness phospholipase signaling; thrombin signaling; acute-phase response; NF-κB signaling; integrin signaling; HMGB1 signaling; agrin interactions; P70S6K signaling; α-adrenergic signaling; NRF-mediated stress response; prolactin signaling; p21 activated kinase (PAK) signaling

(*Continued on next page*)

Table 1. Continued

Key Proteins	Pathways
IκB	Gαq signaling; acute-phase response signaling; type 1 diabetes; iNOS signaling
JNK	IL-8 signaling; Tec signaling; inhibition of angiogenesis by TSP1; HMGB1 signaling; type 1 diabetes; Rho signaling; IL-1 signaling; PAK signaling; HGF signaling
FAK	glioma invasiveness; IL-8 signaling; Tec signaling; agrin interactions; Rho signaling; chemokine signaling; integrin signaling
RHO	glioma invasiveness; HMGB1 signaling; IL-8 signaling; Gαq signaling
VEGF and receptors 1/2/3/C	nitric oxide signaling; IL-8 signaling; VEGF signaling
CAVEOLIN-1	nitric oxide signaling; eNOS signaling; integrin signaling
eNOS	eNOS signaling; nitric oxide signaling; VEGF signaling
P21CIP1	HGF pathway
Key proteins highly expressed in MRU $^{CD200/CD200R1}$ population[a]	highly active pathways in MRU$^{not\ CD200/CD200R1}$ population[b]
SRC	IL-8 pathway; antiproliferative role of somatostatin receptor 2; thrombin pathway; agrin interactions; Stat3 pathway; integrin pathway; cAMP-mediated signaling; endothelin 1 signaling; phospholipase C signaling; role of NA in cardiac hypertrophy; p70s6k signaling; integrin signaling; VEGF signaling; Gα12-13 signaling
BTK	Gαq signaling; phospholipase C signaling; NF-κB/PI3K signaling; p70s6k signaling; Gα12-13 signaling
PKC	VEGF signaling
SHIP	NF-κB/PI3K signaling

See also Figures S1 and S2.
Divergent pathway activity in the MRU subpopulations was determined by IPA.
[a]Corresponds to MRU separation presented in Figure 1: MRU$^{CD200/CD200R1}$ expresses high levels of both CD200 and CD200R1.
[b]Corresponds to MRU separation presented in Figure 1: MRU$^{not\ CD200/CD200R1}$ represents the rest of the MRUs.

genes were analyzed by STRING (Jensen et al., 2009) to elucidate the interactions among them. Two complexes and a connecting branch were identified (Figure 4). Complex A included nine extracellular matrix-encoded proteins, mainly fibril-forming, basement-forming, FACIT and transmembrane collagens (Gelse et al., 2003). Complex A interacted with complex B via LAMA1 and its heterodimeric receptor ITGA6 that binds ACTN1. Complex B contained nine genes involved in muscle (mainly smooth muscle) architecture and function (Chereau et al., 2008; Lazar and Garcia, 1999; Rensen et al., 2000). Genes coding for proteins composing the connecting branch represented mesenchyme (*Ctgf*) and basal epithelium (*Cdh3*, *Trp63* and *Krt5*). The branch also included FCNA, which is highly expressed in the MRU$^{CD200/CD200R1}$ cells and binds proteins of the complement system (Endo et al., 2012), as well as the regulatory protein WNT10A. The latter is induced in early embryogenesis and enhances the self-renewing phenotype in cancer cells (Long et al., 2015). Also present were the

WNT receptor FZD7 that mediates WNT/β-catenin signaling for pluripotency (Fernandez et al., 2014), and HUNK, which is involved in the formation and differentiation of adult intestinal stem cells (Luu et al., 2013).

To test the relevance of the selected list of 114 genes to mammary stemness, it was compared with an external dataset defining genes that are upregulated in an enriched mammary stem cell gene subset in both mice and humans (Lim et al., 2010, and Table 2, row 1). An overlap of 39 genes allowed a relatively highly significant p value. Seven additional genes, *Ltbp2*, *Hal*, *Wnt10A*, *Bst1*, *Fzd7*, *Lgr4,* and *Trnp1*, which are upregulated in enriched stem cell subsets of other tissues or in embryonic stem cells, were also identified (Table S3). Most of the common genes were assembled into the aforedescribed complexes and connecting branch (Figure 4). *Cdh3*, which links the two complexes via the branch, was one of the highly expressed genes in MRU$^{CD200/CD200R1}$ cells compared with the rest of the MRU-defined cells (1.77-fold). Interestingly, two of the
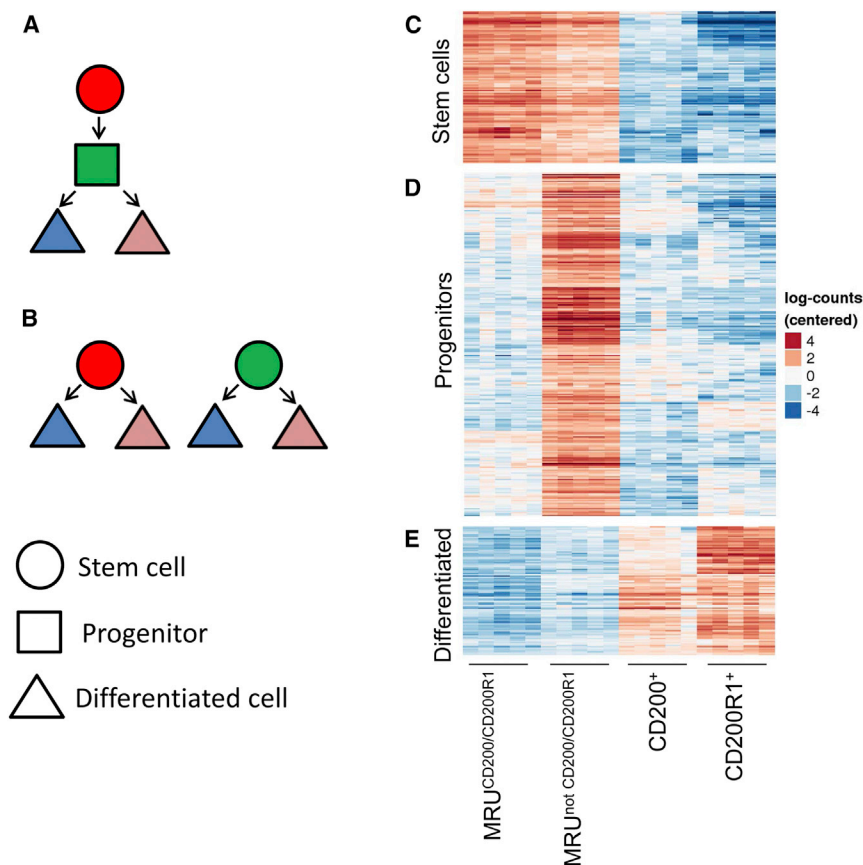
Figure 3. Co-occupancy of the MRU Fraction by MRU$^{CD200/CD200R1}$ and MRU$^{not\ CD200/CD200R1}$ May Involve Hierarchal Association or the Presence of Two Stem Cell Populations

(A and B) Demonstration of the two options.

(C) Genes with expression patterns that fit with stem cells being in a single hierarchy: expression of genes in MRU$^{CD200/CD200R1}$ is greater than in MRU$^{not\ CD200/CD200R1}$ at p < 0.05; expression of genes in MRU$^{not\ CD200/CD200R1}$ is 4-fold greater than in CD200$^+$CD200R1$^-$ and CD200R1$^+$CD200$^-$.

(D) Gene-expression pattern indicating progenitors: expression of genes in MRU$^{not\ CD200/CD200R1}$ is 2-fold greater than in MRU$^{CD200/CD200R1}$, CD200$^+$CD200R1$^-$ and CD200R1$^+$CD200$^-$.

(E) Gene-expression pattern indicating differentiated cells: expression of genes in CD200$^+$CD200R1$^-$ and CD200R1$^+$CD200$^-$ is at least 2-fold greater than in MRU$^{not\ CD200/CD200R1}$ with adjusted p ≤ 0.05; expression of genes in MRU$^{not\ CD200/CD200R1}$ is greater that in MRU$^{CD200/CD200R1}$ at adjusted p ≤ 0.75.

See also Table S2.

non-mammary-associated genes, *Lgr4* and *Lgr6* that mark stemness in the skin and intestine, respectively (Li et al., 2015; Lough et al., 2013), were not connected to the complexes. *Fzd7* and *Hall*, which were also not associated with the complexes, were the most highly expressed genes in MRU$^{CD200/CD200R1}$ versus MRU$^{not\ CD200/CD200R1}$ cells (1.97-fold and 4.89-fold, respectively).

An additional comparison with a smaller dataset of 26 mammary stem cell transcriptional regulators (Table 2, row 2) identified six genes in common with the putative stem cell list of 114 genes, thus supporting their association with stemness.

The significant number of common genes shared by the external stem cell datasets and the list of genes with decreased expression from MRU$^{CD200/CD200R1}$ via MRU$^{not\ CD200/CD200R1}$ toward CD200- and CD200R1-expressing cells supports the concept of a single-cell hierarchy that assembles the MRU$^{CD200/CD200R1}$ and the MRU$^{not\ CD200/CD200R1}$ populations as stem cells and their progenitors. A related question was further raised: does the downstream cell hierarchy toward the CD200$^+$ and CD200R1$^+$ cells encompass the complete epithelial cell population?

To answer this, an additional set of 262 genes was identified with highest expression in MRU$^{not\ CD200/CD200R1}$, aimed to mark mammary progenitors (Figure 3D). Indeed, this set of genes had 28 orders of magnitude lower compatibility to the stem cell subset (Lim et al., 2010) compared with the putative stem cell list defined here (Figure 3C and Table S4). Interestingly, this progenitor set did not share any common genes with a signature of 58 mouse/human luminal progenitors (Lim et al., 2010 and Table 2, row 3). The lack of luminal orientation was confirmed by the absence of common genes between an additional subset of genes highly expressed in the CD200$^+$ or CD200R1$^+$ differentiated cells (Figure 3E) and a list of genes upregulated in differentiated luminal cells (Lim et al., 2010 and Table 2, row 4). As expected, this subset had 30 orders of magnitude lower compatibility with the reported stem cell subset (Lim et al., 2010) as compared with the stem cell list described here (Table S4).

The latter results indicated that the proposed cell hierarchy stemming from MRU$^{CD200/CD200R1}$ does not include a defined luminal lineage. To test for basal orientation, a comparison was performed between the stem, progenitor, or differentiated subsets defined in the current study and gene sets that favor basal versus luminal expression in the mammary gland (Table 2, rows 5–7) or breast cancer (Table 2, rows 8–10). This comparison only yielded a
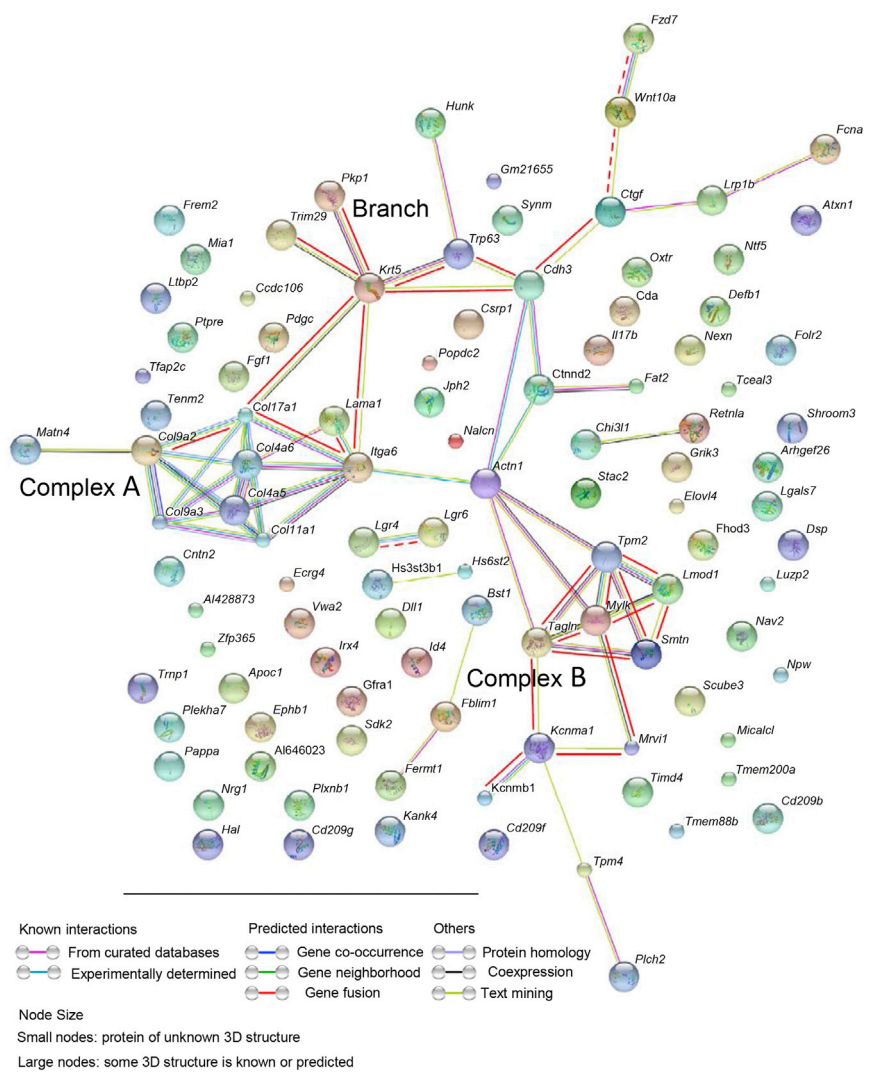
in acute myelogenous leukemia that involves induction of CD200 expression. An additional significant value (p = 0.014) was obtained by comparing the progenitor list with the signature of genes that are upregulated in double-positive lymphocytes, expressing both CD4 and CD8 coreceptors versus their double-negative counterparts (Dik et al., 2005).

In Vitro and *In Vivo* Studies of the Relationship between the MRU$^{CD200/CD200R1}$ and MRU$^{not\ CD200/CD200R1}$ Populations

CD200 and CD200R1 are expressed in epithelial cells of the basal and luminal layers of the mammary gland (Figure S3). Here we focused on the relevance of their expression levels in MRUs to unveiling the top of the mammary cell hierarchy. Sorted cells of the two MRU subpopulations were challenged for *in vitro* mammosphere generation under non-adherent conditions that indicate stem cell activity. The MRU$^{CD200/CD200R1}$ subpopulation generated a significantly higher number of mammospheres than its MRU$^{not\ CD200/CD200R1}$ counterpart (Figure 5A). The proportion of basal, CK14-stained colonies that developed after seeding diluted cells of these populations was highest for the MRU$^{CD200/CD200R1}$ cells (Figure 5B). A significant decrease of 16% in basal colony proportion was noted for the MRU$^{not\ CD200/CD200R1}$ cells, and a further 11% decline was detected for the unsorted cell population. The combined number of CK14- and CK18-stained colonies was higher by 18%–25% for the MRU$^{not\ CD200/CD200R1}$ subpopulation compared with MRU$^{CD200/CD200R1}$.

significant number of common genes for the stem cell subset defined here (Table 2, rows 5 and 8). This implies that the cell hierarchy originating from the MRU$^{CD200/CD200R1}$ cells is indeed topped by a highly enriched stem cell population with basal characteristics. However, the downstream hierarchy toward the CD200- and CD200R1-enriched populations involves a relatively narrow lineage that does not include specific basal versus luminal characteristics. This lineage may, however, include specific immune-associated characteristics of the CD200 and CD200R1 cells. Indeed, within 12 relevant immune-related databases that might shed light on the downstream characteristics of cell hierarchy (Subramanian et al., 2005), the one that provided the most significant number of common genes (p = 0.0058) related to upregulated genes in normal hematopoietic progenitors by RUNX1–RUNX1T1 translocation (Tonks et al., 2007). This translocation causes one of the most common molecular abnormalities

Table 2. Comparative Analysis of Genes that Mark Cell Status in the Hierarchy in This and Other Studies

No.	Aim	Source of Gene Set 1	Number of Genes in Gene Set 1	Source of Gene Set 2	Number of Genes in Gene Set 2	Number of Overlapping Genes	Significance (in Set of 20,000 Genes)
1	validating proposed cell hierarchy	current analysis: stem cell subset	114[a]	enriched mammary stem cell subset (Lim et al., 2010)	489	39[b]	$p = 2.75 \times 10^{-34}$
2		current analysis: stem cell subset	114	transcriptional regulators of mammary stem cells (Chakrabarti et al., 2014)	26	6	$p = 6.3 \times 10^{-9}$
3	validating involvement of luminal progenitors	current analysis: progenitor subset	262	enriched mammary luminal progenitor subset (Lim et al., 2010)	58	0	not significant
4	validating involvement of differentiated luminal cells	current analysis: differentiated cell subset	97	enriched mammary luminal differentiated cell subset (Lim et al., 2010)	116	0	not significant
5	comparing with gene expression in the breast	current analysis: stem cell subset	114	upregulated genes in basal mammary epithelial cells compared with luminal ones (Huper and Marks, 2007)	54	4	$p = 0.0004$
6		current analysis: progenitor subset	262	upregulated genes in basal mammary epithelial cells compared with luminal ones (Huper and Marks, 2007)	54	0	not significant
7		current analysis: differentiated cell subset	97	upregulated genes in basal mammary epithelial cells compared with luminal ones (Huper and Marks, 2007)	54	1	$p = 0.2$
8	comparing with gene expression of breast cancer subtypes	current analysis: stem cell subset	114	upregulated genes in basal breast cancer samples (Smid et al., 2008)	777	14	$p = 0.0001$
9		current analysis: progenitor subset	262	upregulated genes in basal breast cancer samples (Smid et al., 2008)	777	10	$p = 0.12$
10		current analysis: differentiated cell subset	97	upregulated genes in basal breast cancer samples (Smid et al., 2008)	777	4	$p = 0.19$

See also Tables S3 and S4.
[a]105 annotated.
[b]Additional stem cell-related genes were identified.

To further elucidate the relationship between the MRU subpopulations, we showed that transplantation of MRU[CD200/CD200R1] populations gives rise to outgrowths containing MRU[not CD200/CD200R1] cells (Table S5). We then asked whether outgrowths developed from the downstream MRU[not CD200/CD200R1] population encompass MRU[CD200/CD200R1] cells. Table S5 shows high CD200/CD200R1 expressors in the MRU fraction of outgrowths developed from MRU[not CD200/CD200R1]. Next we examined putative differences in differentiation potential between outgrowths developed from MRU[CD200/CD200R1] and MRU[not CD200/CD200R1]. Both populations developed
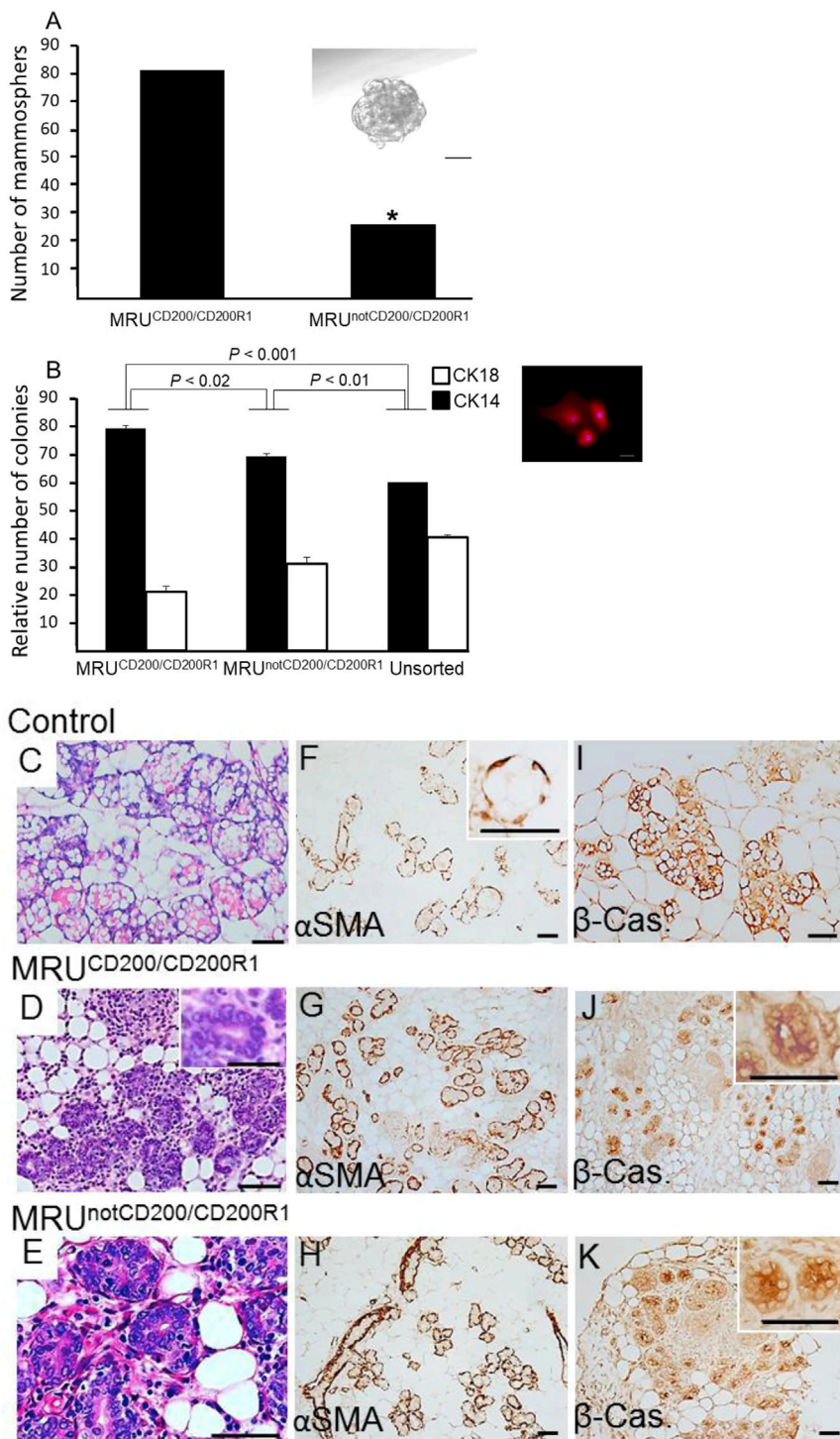
memory of their hormone-derived (virgin) state (Asselin-Labat et al., 2010), or result from lack of draining capacity via the nipple.

Of note, at this late stage of pregnancy, there was substantial infiltration of immune cells in the developing outgrowths in 42% of the glands, independent of their origin, which was not observed in the normal gland (Figures 5C–5E). It has been shown that immune cells are attracted by the chemokine receptor CCR6 (Boyle et al., 2015) and promote epithelial growth throughout the gland. Estrogen-stimulated macrophages are more potent

lobuloalveolar-like structures in late pregnancy (Figures 5C–5E). They were composed of αSMS-stained basal cells and luminal cells with β-casein-synthesizing capability (Figures 5F–5K). A much narrower lumen characterized the lobuloalveolar structures in the outgrowths relative to the intact gland. This difference may indicate a stem cell

than non-stimulated ones in promoting human outgrowth development in the de-epithelized mouse mammary gland (Fleming et al., 2012).

DISCUSSION

Characterizing the top of the epithelial cell hierarchy in the mammary gland is highly important from several perspectives: understanding developmental and regenerative processes of the gland, identifying potential tumor-initiating cells, and manipulating the capabilities of the gland toward higher milk production. In this study, we identified a heterogeneous $CD200^{high}CD200R1^{high}$ epithelial cell population. Transplantation assays demonstrated that ~50% of its constituents enriched the $CD24^{med}CD49f^{high}$ (MRU) fraction for cells with relatively high mammary repopulating activity. The higher repopulating potential of $MRU^{CD200/CD200R1}$ compared with $MRU^{not\ CD200/CD200R1}$, which encompasses the rest of the MRUs, was reflected by the larger area of the reconstituted mammary epithelium. However, it did not include changes in the number of reconstituted glands. This minor, but significant difference became highly pronounced when gene expression was profiled. RNA-seq analysis of gene expression in 1,000 pooled mammary epithelial cells clearly distinguished between the two MRU subsets, as well as additional $CD200^+CD200R1^-$ and $CD200^-CD200R1^+$ populations. With this tool, two major aspects were addressed: (1) a possible hierarchical relationship between the MRU subpopulations versus the concept of two different types of stem cells; and (2) a putative linkage between immunorelated gene expression and stemness.

The much higher number of activated genes and metabolic pathways detected in the $MRU^{not\ CD200/CD200R1}$ indicated the coexistence of two metabolically divergent populations at the top of the mammary cell hierarchy. Among the numerous pathways supporting the higher metabolic activity of $MRU^{not\ CD200/CD200R1}$, the most highly expressed ones were involved in differentiation and function of progenitors, suggesting that the $MRU^{not\ CD200/CD200R1}$ are not stem cells (Figure S4). For example, eNOS and nitric oxide signaling have been associated with osteogenic differentiation through activation of the downstream canonical WNT/β-catenin signaling pathway (Bandara et al., 2016). Higher levels of reactive oxygen species have been recently identified as a marker of progenitors when compared with stem cells in the mammary gland (Diehn et al., 2009). The strongly activated Tec kinase signaling is also highly important for T cell differentiation (Andreotti et al., 2010).

Deeper insight into the core of the mediating proteins provided further support for pro-differentiating activity

in the $MRU^{not\ CD200/CD200R1}$. For example, MAPK, which was highly activated in this population, has been associated with mammary branching morphogenesis and the expansion of cells positive for K6, a marker of hyperproliferative progenitor cells (Fata et al., 2007). Elevated MAPK levels also mark common progenitors that are downstream of the stem cell population in wild-type mice (Godde et al., 2014). PI3 kinase is also highly activated in the $MRU^{not\ CD200/CD200R1}$. Its expression in the mammary gland is regulated by the signal transducer and activator of transcription STAT5A (Schmidt et al., 2014), which marks mammary progenitor activity (Yamaji et al., 2009), and negatively regulates the expression of pTEN, which is more highly expressed in $MRU^{CD200/CD200R1}$. PKC is also known to promote epithelial progenitor cell differentiation (Rieger et al., 2016) and RAS promotes osteoprogenitor cell proliferation and bone formation (Papaioannou et al., 2016). These data converge into the IPA-defined higher metabolic activity of $MRU^{not\ CD200/CD200R1}$, which is needed for cell growth, proliferation, development, and movement. Validating the concept of a single hierarchy in which $MRU^{CD200/CD200R1}$ cells give rise to common $MRU^{not\ CD200/CD200R1}$ progenitors that differentiate into $CD200^+CD200R1^-$ and $CD200R1^+CD200^-$ cells yielded a list of 144 genes exhibiting decreasing expression patterns. The high compatibility with the previously characterized subset of stem cell markers in mouse and human glands (Lim et al., 2010), as well as with stem cell regulators (Chakrabarti et al., 2014), further supports this hypothesis. Importantly, most of the genes that were common to these two lists converged into specific collagen-associated and muscle-associated gene complexes that are probably involved in determining stem cell characteristics. *Cdh3*, a central linking gene for these complexes, was one of the most divergently expressed genes, at 1.8-fold higher levels in $MRU^{CD200/CD200R1}$ compared with $MRU^{not\ CD200/CD200R1}$. Its potency as a stem cell marker is therefore worth analyzing. Taken together, the 39 common genes, together with 7 genes that serve as stem cell markers in other tissues, may provide a core list for mammary stem cell identification.

While the use of CD200 and its receptor appeared suitable for identifying stem cells and close progenitors in the mammary epithelial cells, the downstream hierarchy defined by their expression did not include all epithelial cell populations or even defined luminal or basal ones. Rather, it was restricted to an as-yet undefined immune-related cell population.

The major differences between $MRU^{CD200/CD200R1}$ and $MRU^{not\ CD200/CD200R1}$ gene-expression profiles and resulting pathway activities led to an accumulated phenotypic outcome (Figure S4). Transplantation of $MRU^{CD200/CD200R1}$ resulted in better filling of the cleared

fat pad. Successful serial transplantation characterized only this population, indicating stem cell activity (Asselin-Labat et al., 2010). MRU$^{CD200/CD200R1}$ also maintained a 3-fold higher potential to generate mammospheres under non-adherent conditions than its MRU$^{CD200/CD200R1}$ counterpart. A more basal origin, depicted by a higher proportion of CK14-stained colonies that decreased toward the MRU$^{not\ CD200/CD200R1}$ and further on toward the unsorted cells, may be also associated with the MRU$^{CD200/CD200R1}$. Nevertheless, de-differentiation capability of the MRU$^{not\ CD200/CD200R1}$ (Clevers and Watt, 2018) or a steady state between the two subpopulations (Lanner and Rossant, 2010) at an early developmental stage cannot be excluded, due to the ability of MRU$^{not\ CD200/CD200R1}$ to re-form outgrowths containing MRU$^{CD200/CD200R1}$. Both MRU subpopulations formed outgrowths containing basal and functional luminal cells, depicting their close hierarchical state and multipotent capabilities. As such, the data presented here assemble the MRU$^{CD200/CD200R1}$ and MRU$^{not\ CD200/CD200R1}$ as putative stem cells and common progenitors, respectively (Figure S4), that together contribute to mammary homeostasis (Rios et al., 2014; Sun et al., 2014).

The second aspect of this study involves the possible link between innate immune-related gene expression and stemness. The poor metabolic activity of the MRU$^{CD200/CD200R1}$ population was accompanied by significantly higher activity of the complement system as compared with cells representing the rest of the MRUs. The complement system has long been perceived as an effector arm of innate immunity that mediates important immunoregulatory and inflammatory functions (Mastellos and Lambris, 2002). In this respect, higher activity of both CD200 and complement may be difficult to reconcile, because the former is considered to be an inhibitor of the complement system (Elward and Gasque, 2003) and the CD200–CD200R1 interaction on myeloid cells has been reported to attenuate immune activity (Minas and Liversidge, 2006). Other studies, however, suggest that the complement system may assume a non-inflammatory role in certain biological settings (Mastellos and Lambris, 2002). Indeed, complement has been implicated as a mediator of lens and limb regeneration in lower vertebrates (Kimura et al., 2003) and of mammalian liver regeneration via C3A and C5A activation (Strey et al., 2003). Others have shown a critical role for complement proteins (mainly C5a) in hematopoietic stem cell mobilization from their bone marrow niche to the blood, which was associated with the complement system's circadian activity (Ratajczak, 2015). Here we demonstrated that at least six complement proteins of the classical pathway that composes the membrane attack complex—C1 to C6—are highly expressed in the stem cell core of the mammary gland. Whether these proteins are involved in

maintaining mammary epithelial stemness or have a role in the initial steps of stem cell differentiation remains to be elucidated.

EXPERIMENTAL PROCEDURES

Mice

FVB/N or C57BL mice were housed under a 12-hr light/dark cycle and given food and water *ad libitum*. For all surgical procedures, mice were anesthetized with isoflurane (Abbott Laboratories, Maidenhead, England) mixed with O_2 using a veterinary anesthesia machine. All animals used in this study were treated humanely. Study protocols were in compliance with the regulations of the Israeli Ministry of Health and local institution policies.

Dissociation of Mammary Tissue into Single-Cell Suspension and Flow Cytometry

Mouse mammary cells were dissociated from the fourth inguinal mammary glands of 6- to 8-week-old FVB/N virgin females. The dissociation procedure was as previously described (Rauner and Barash, 2012). Further details can be found in Supplemental Experimental Procedures.

Cell Transplantation

The endogenous mammary epithelium was surgically removed bilaterally from #4 mammary glands of 21-day-old female mice weighing 10–12 g (i.e., "clearing"). Sorted mammary epithelial cell populations were transplanted as detailed in Supplemental Experimental Procedures.

Serial Transplantation

A similar transplantation procedure was performed for serial transplantations of CD200^{high}/CD200R1^{high} MRUs. Host mice were C57BL and donor mice were C57/GFP (ubiquitin-EGFP mice [Reichenstein et al., 2016]). Further details can be found in Supplemental Experimental Procedures.

Outgrowth Analysis

For whole-mount examination, transplanted mammary fat pads were excised from sacrificed mice, fixed, and stained with Carmine Red as previously described (Rauner et al., 2013). Stained whole mounts were visualized and photographed using the binocular equipped with CellSens standard 1.4 software.

RNA Extraction and RNA-Seq Analysis

Dispersed mouse mammary cells were sorted into a 384-well plate as described in Supplemental Experimental Procedures. An aliquot of 1,000 cells in 1 μL from each population was collected into 9.5 μL RNasin lysis buffer (SMART-Seq V4 Ultra Low Input RNA Kit; Clontech, Mountain View, CA), incubated for 5 min at room temperature and kept at −80°C. For RNA-seq analysis, samples were thawed. Reverse transcription and cDNA amplification (12 cycles) of the full transcriptome were performed with the SMART-Seq V4 Ultra Low Input RNA Kit according to the manufacturer's protocol. Clean-up reactions were performed with Ampure

XP beads (Beckman Coulter, Brea, CA). The amplified cDNA products were sheared by Covaris E220X (Woburn, MA) and 4.5–12.0 ng of sheared amplified cDNA from each sample was processed as previously described (Blecher-Gonen et al., 2013). An individual barcode was ligated to each sample to allow multiplexing of 20 libraries on two sequencing lanes. Between 10 and 12 million single-end 60-base pair reads were sequenced per sample on an Illumina HiSeq 2500 v4 instrument.

Histological Analysis and Immunostaining

Immunostaining of paraformaldehyde-fixed cells in culture and of paraffin-embedded or frozen tissue sections were performed using relevant antibodies (Table S6) as previously described (Rauner and Barash, 2012) with the modifications detailed in Supplemental Experimental Procedures.

Clonal and Mammosphere Assays

These assays were essentially performed as previously described (Rauner and Barash, 2012). Further details can be found in Supplemental Experimental Procedures.

Bioinformatics Analysis

Reads were trimmed using cutadapt (Martin, 2011) and mapped to the GRCm38 genome using STAR (Dobin et al., 2013) v2.4.2a (default parameters). Counting proceeded over genes annotated in Ensembl release 82 using htseq-count (Anders et al., 2015) (mode: intersection-strict). Differential expression analysis was performed using DESeq2 (Love et al., 2014) with the betaPrior. Cook's distance cutoff and independent filtering parameters were set to False.

Statistics

Unless otherwise indicated, Student's t test was performed for statistical analyses.

ACCESSION NUMBERS

Data have been deposited in GEO under accession number GEO: GSE93961.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and six tables and can be found with this article online at https://doi.org/10.1016/j.stemcr.2018.05.013.

AUTHOR CONTRIBUTIONS

G.R., T.K., and I.B. designed the concept of the experiments. G.R. and T.K. performed the flow cytometry. S.G. performed the RNA-seq. G.H. performed the core of the bioinformatics analysis. T.K. performed data analysis. I.B. performed additional bioinformatics analysis and wrote the manuscript.

ACKNOWLEDGMENTS

REFERENCES

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Andreotti, A.H., Schwartzberg, P.L., Joseph, R.E., and Berg, L.J. (2010). T-cell signaling regulated by the Tec family kinase, Itk. Cold Spring Harb. Perspect. Biol. *2*, a002287.

Asano, K., Takahashi, N., Ushiki, M., Monya, M., Aihara, F., Kuboki, E., Moriyama, S., Iida, M., Kitamura, H., Qiu, C.H., et al. (2015). Intestinal CD169(+) macrophages initiate mucosal inflammation by secreting CCL8 that recruits inflammatory monocytes. Nat. Commun. *6*, 7802.

Asselin-Labat, M.L., Vaillant, F., Sheridan, J.M., Pal, B., Wu, D., Simpson, E.R., Yasuda, H., Smyth, G.K., Martin, T.J., Lindeman, G.J., et al. (2010). Control of mammary stem cell function by steroid hormone signalling. Nature *465*, 798–802.

Bai, L., and Rohrschneider, L.R. (2010). s-SHIP promoter expression marks activated stem cells in developing mouse mammary tissue. Genes Dev. *24*, 1882–1892.

Bandara, N., Gurusinghe, S., Lim, S.Y., Chen, H., Chen, S., Wang, D., Hilbert, B., Wang, L.X., and Strappe, P. (2016). Molecular control of nitric oxide synthesis through eNOS and caveolin-1 interaction regulates osteogenic differentiation of adipose-derived stem cells by modulation of Wnt/beta-catenin signaling. Stem Cell Res. Ther. *7*, 182.

Begenisich, T., Nakamoto, T., Ovitt, C.E., Nehrke, K., Brugnara, C., Alper, S.L., and Melvin, J.E. (2004). Physiological roles of the intermediate conductance, Ca2+-activated potassium channel Kcnn4. J. Biol. Chem. *279*, 47681–47687.

Bergman, D., Halje, M., Nordin, M., and Engstrom, W. (2013). Insulin-like growth factor 2 in development and disease: a mini-review. Gerontology *59*, 240–249.

Blecher-Gonen, R., Barnett-Itzhaki, Z., Jaitin, D., Amann-Zalcenstein, D., Lara-Astiaso, D., and Amit, I. (2013). High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. Nat. Protoc. *8*, 539–554.

Boyle, S.T., Faulkner, J.W., McColl, S.R., and Kochetkova, M. (2015). The chemokine receptor CCR6 facilitates the onset of mammary neoplasia in the MMTV-PyMT mouse model via recruitment of tumor-promoting macrophages. Mol. Cancer *14*, 115.

Cardiff, R.D., and Wellings, S.R. (1999). The comparative pathology of human and mouse mammary glands. J. Mammary Gland Biol. Neoplasia *4*, 105–122.

Chakrabarti, R., Wei, Y., Hwang, J., Hang, X., Andres Blanco, M., Choudhury, A., Tiede, B., Romano, R.A., DeCoste, C., Mercatali, L., et al. (2014). DeltaNp63 promotes stem cell activity in mammary gland development and basal-like breast cancer by

enhancing Fzd7 expression and Wnt signalling. Nat. Cell Biol. *16*, 1004–1015, 1001-1013.

Chereau, D., Boczkowska, M., Skwarek-Maruszewska, A., Fujiwara, I., Hayes, D.B., Rebowski, G., Lappalainen, P., Pollard, T.D., and Dominguez, R. (2008). Leiomodin is an actin filament nucleator in muscle cells. Science *320*, 239–243.

Clevers, H., and Watt, F.M. (2018). Defining adult stem cells by function, not by phenotype. Annu. Rev. Biochem. https://doi.org/10.1146/annurev-biochem-062917-012341.

Diehn, M., Cho, R.W., Lobo, N.A., Kalisky, T., Dorie, M.J., Kulp, A.N., Qian, D., Lam, J.S., Ailles, L.E., Wong, M., et al. (2009). Association of reactive oxygen species levels and radioresistance in cancer stem cells. Nature *458*, 780–783.

Dik, W.A., Pike-Overzet, K., Weerkamp, F., de Ridder, D., de Haas, E.F., Baert, M.R., van der Spek, P., Koster, E.E., Reinders, M.J., van Dongen, J.J., et al. (2005). New insights on human T cell development by quantitative T cell receptor gene rearrangement studies and gene expression profiling. J. Exp. Med. *201*, 1715–1723.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

dos Santos, C.O., Rebbeck, C., Rozhkova, E., Valentine, A., Samuels, A., Kadiri, L.R., Osten, P., Harris, E.Y., Uren, P.J., Smith, A.D., et al. (2013). Molecular hierarchy of mammary differentiation yields refined markers of mammary stem cells. Proc. Natl. Acad. Sci. USA *110*, 7123–7130.

Elward, K., and Gasque, P. (2003). "Eat me" and "don't eat me" signals govern the innate immune response and tissue repair in the CNS: emphasis on the critical role of the complement system. Mol. Immunol. *40*, 85–94.

Endo, Y., Takahashi, M., Iwaki, D., Ishida, Y., Nakazawa, N., Kodama, T., Matsuzaka, T., Kanno, K., Liu, Y., Tsuchiya, K., et al. (2012). Mice deficient in ficolin, a lectin complement pathway recognition molecule, are susceptible to *Streptococcus pneumoniae* infection. J. Immunol. *189*, 5860–5866.

Fata, J.E., Mori, H., Ewald, A.J., Zhang, H., Yao, E., Werb, Z., and Bissell, M.J. (2007). The MAPK(ERK-1,2) pathway integrates distinct and antagonistic signals from TGFalpha and FGF7 in morphogenesis of mouse mammary epithelium. Dev. Biol. *306*, 193–207.

Fernandez, A., Huggins, I.J., Perna, L., Brafman, D., Lu, D., Yao, S., Gaasterland, T., Carson, D.A., and Willert, K. (2014). The WNT receptor FZD7 is required for maintenance of the pluripotent state in human embryonic stem cells. Proc. Natl. Acad. Sci. USA *111*, 1409–1414.

Fleming, J.M., Miller, T.C., Kidacki, M., Ginsburg, E., Stuelten, C.H., Stewart, D.A., Troester, M.A., and Vonderhaar, B.K. (2012). Paracrine interactions between primary human macrophages and human fibroblasts enhance murine mammary gland humanization in vivo. Breast Cancer Res. *14*, R97.

Gelse, K., Poschl, E., and Aigner, T. (2003). Collagens–structure, function, and biosynthesis. Adv. Drug Deliv. Rev. *55*, 1531–1546.

Godde, N.J., Sheridan, J.M., Smith, L.K., Pearson, H.B., Britt, K.L., Galea, R.C., Yates, L.L., Visvader, J.E., and Humbert, P.O. (2014). Scribble modulates the MAPK/Fra1 pathway to disrupt luminal and ductal integrity and suppress tumour formation in the mammary gland. PLoS Genet. *10*, e1004323.

Hennighausen, L., and Robinson, G.W. (2005). Information networks in the mammary gland. Nat. Rev. Mol. Cell Biol. *6*, 715–725.

Huper, G., and Marks, J.R. (2007). Isogenic normal basal and luminal mammary epithelial cells isolated by a novel method show a differential response to ionizing radiation. Cancer Res. *67*, 2990–3001.

Jensen, L.J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., et al. (2009). STRING 8–a global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Res. *37*, D412–D416.

Kimura, Y., Madhavan, M., Call, M.K., Santiago, W., Tsonis, P.A., Lambris, J.D., and Del Rio-Tsonis, K. (2003). Expression of complement 3 and complement 5 in newt limb and lens regeneration. J. Immunol. *170*, 2331–2339.

Ko, Y.C., Chien, H.F., Jiang-Shieh, Y.F., Chang, C.Y., Pai, M.H., Huang, J.P., Chen, H.M., and Wu, C.H. (2009). Endothelial CD200 is heterogeneously distributed, regulated and involved in immune cell-endothelium interactions. J. Anat. *214*, 183–195.

Lanner, F., and Rossant, J. (2010). The role of FGF/Erk signaling in pluripotent cells. Development *137*, 3351–3360.

Lazar, V., and Garcia, J.G. (1999). A single human myosin light chain kinase gene (MLCK; MYLK). Genomics *57*, 256–267.

Leccia, F., Nardone, A., Corvigno, S., Vecchio, L.D., De Placido, S., Salvatore, F., and Veneziani, B.M. (2012). Cytometric and biochemical characterization of human breast cancer cells reveals heterogeneous myoepithelial phenotypes. Cytometry A *81*, 960–972.

Li, Z., Zhang, W., and Mulholland, M.W. (2015). LGR4 and its role in intestinal protection and energy metabolism. Front. Endocrinol. (Lausanne) *6*, 131.

Lim, E., Wu, D., Pal, B., Bouras, T., Asselin-Labat, M.L., Vaillant, F., Yagita, H., Lindeman, G.J., Smyth, G.K., and Visvader, J.E. (2010). Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. Breast Cancer Res. *12*, R21.

Long, A., Giroux, V., Whelan, K.A., Hamilton, K.E., Tetreault, M.P., Tanaka, K., Lee, J.S., Klein-Szanto, A.J., Nakagawa, H., and Rustgi, A.K. (2015). WNT10A promotes an invasive and self-renewing phenotype in esophageal squamous cell carcinoma. Carcinogenesis *36*, 598–606.

Lough, D., Dai, H., Yang, M., Reichensperger, J., Cox, L., Harrison, C., and Neumeister, M.W. (2013). Stimulation of the follicular bulge LGR5+ and LGR6+ stem cells with the gut-derived human alpha defensin 5 results in decreased bacterial presence, enhanced wound healing, and hair growth from tissues devoid of adnexal structures. Plast. Reconstr. Surg. *132*, 1159–1171.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Luu, N., Wen, L., Fu, L., Fujimoto, K., Shi, Y.B., and Sun, G. (2013). Differential regulation of two histidine ammonia-lyase genes during *Xenopus* development implicates distinct functions
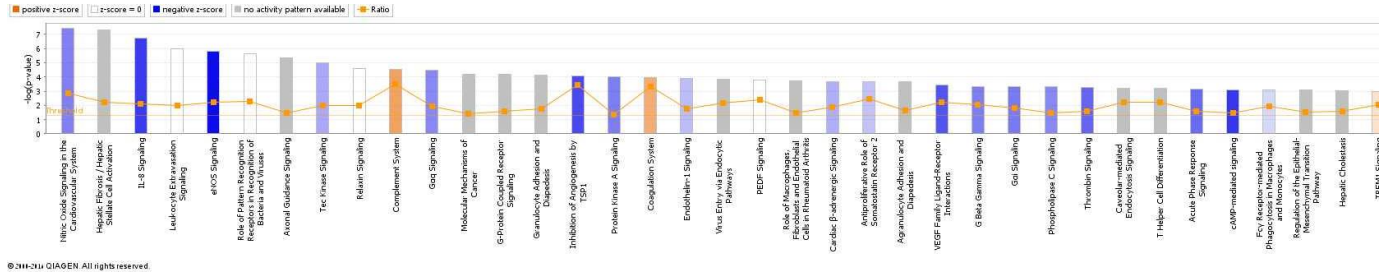
during thyroid hormone-induced formation of adult stem cells. Cell Biosci. *3*, 43.

Maggiani, F., Forsyth, R., Hogendoorn, P.C., Krenacs, T., and Athanasou, N.A. (2011). The immunophenotype of osteoclasts and macrophage polykaryons. J. Clin. Pathol. *64*, 701–705.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. *17*, 10–12.

Martinez-Picado, J., McLaren, P.J., Erkizia, I., Martin, M.P., Benet, S., Rotger, M., Dalmau, J., Ouchi, D., Wolinsky, S.M., Penugonda, S., et al. (2016). Identification of Siglec-1 null individuals infected with HIV-1. Nat. Commun. *7*, 12412.

Mastellos, D., and Lambris, J.D. (2002). Complement: more than a 'guard' against invading pathogens? Trends Immunol. *23*, 485–491.

Mayilyan, K.R. (2012). Complement genetics, deficiencies, and disease associations. Protein Cell *3*, 487–496.

Minas, K., and Liversidge, J. (2006). Is the CD200/CD200 receptor interaction more than just a myeloid cell inhibitory signal? Crit. Rev. Immunol. *26*, 213–230.

Ohyama, M., Terunuma, A., Tock, C.L., Radonovich, M.F., Pise-Masison, C.A., Hopping, S.B., Brady, J.N., Udey, M.C., and Vogel, J.C. (2006). Characterization and isolation of stem cell-enriched human hair follicle bulge cells. J. Clin. Invest. *116*, 249–260.

Papaioannou, G., Mirzamohammadi, F., and Kobayashi, T. (2016). Ras signaling regulates osteoprogenitor cell proliferation and bone formation. Cell Death Dis. *7*, e2405.

Ratajczak, M.Z. (2015). A novel view of the adult bone marrow stem cell hierarchy and stem cell trafficking. Leukemia *29*, 776–782.

Rauner, G., and Barash, I. (2012). Cell hierarchy and lineage commitment in the bovine mammary gland. PLoS One *7*, e30113.

Rauner, G., Leviav, A., Mavor, E., and Barash, I. (2013). Development of foreign mammary epithelial morphology in the stroma of immunodeficient mice. PLoS One *8*, e68637.

Reichenstein, M., Rauner, G., Kfir, S., Kisliouk, T., and Barash, I. (2016). Luminal STAT5 mediates H2AX promoter activity in distinct population of basal mammary epithelial cells. Oncotarget *7*, 41781–41797.

Rensen, S., Merkx, G., Doevendans, P., Geurts Van Kessel, A., and van Eys, G. (2000). Structure and chromosome location of Smtn, the mouse smoothelin gene. Cytogenet. Cell Genet. *89*, 225–229.

Rieger, M.E., Zhou, B., Solomon, N., Sunohara, M., Li, C., Nguyen, C., Liu, Y., Pan, J.H., Minoo, P., Crandall, E.D., et al. (2016). p300/beta-catenin interactions regulate adult progenitor cell differentiation downstream of WNT5a/protein kinase C (PKC). J. Biol. Chem. *291*, 6569–6582.

Rios, A.C., Fu, N.Y., Lindeman, G.J., and Visvader, J.E. (2014). In situ identification of bipotent stem cells in the mammary gland. Nature *506*, 322–327.

Schmidt, J.W., Wehde, B.L., Sakamoto, K., Triplett, A.A., Anderson, S.M., Tsichlis, P.N., Leone, G., and Wagner, K.U. (2014). Stat5 regulates the phosphatidylinositol 3-kinase/Akt1 pathway during mammary gland development and tumorigenesis. Mol. Cell. Biol. *34*, 1363–1377.

Seita, J., and Weissman, I.L. (2010). Hematopoietic stem cell: self-renewal versus differentiation. Wiley Interdiscip. Rev. Syst. Biol. Med. *2*, 640–653.

Shackleton, M., Vaillant, F., Simpson, K.J., Stingl, J., Smyth, G.K., Asselin-Labat, M.L., Wu, L., Lindeman, G.J., and Visvader, J.E. (2006). Generation of a functional mammary gland from a single stem cell. Nature *439*, 84–88.

Smid, M., Wang, Y., Zhang, Y., Sieuwerts, A.M., Yu, J., Klijn, J.G.M., Foekens, J.A., and Martens, J.W.M. (2008). Subtypes of breast cancer show preferential site of relapse. Cancer Res. *68*, 3108–3114.

Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., Li, H.I., and Eaves, C.J. (2006). Purification and unique properties of mammary epithelial stem cells. Nature *439*, 993–997.

Strey, C.W., Markiewski, M., Mastellos, D., Tudoran, R., Spruce, L.A., Greenbaum, L.E., and Lambris, J.D. (2003). The proinflammatory mediators C3a and C5a are essential for liver regeneration. J. Exp. Med. *198*, 913–923.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Sun, J., Ramos, A., Chapman, B., Johnnidis, J.B., Le, L., Ho, Y.J., Klein, A., Hofmann, O., and Camargo, F.D. (2014). Clonal dynamics of native haematopoiesis. Nature *514*, 322–327.

Tonks, A., Pearn, L., Musson, M., Gilkes, A., Mills, K.I., Burnett, A.K., and Darley, R.L. (2007). Transcriptional dysregulation mediated by RUNX1-RUNX1T1 in normal human progenitor cells and in acute myeloid leukaemia. Leukemia *21*, 2495–2505.

Yamaji, D., Na, R., Feuermann, Y., Pechhold, S., Chen, W., Robinson, G.W., and Hennighausen, L. (2009). Development of mammary luminal progenitor cells is controlled by the transcription factor STAT5A. Genes Dev. *23*, 2382–2387.

Zhu, Y., Yang, Y., Zhang, Y., Hao, G., Liu, T., Wang, L., Yang, T., Wang, Q., Zhang, G., Wei, J., et al. (2014). Placental mesenchymal stem cells of fetal and maternal origins demonstrate different therapeutic potentials. Stem Cell Res. Ther. *5*, 48.
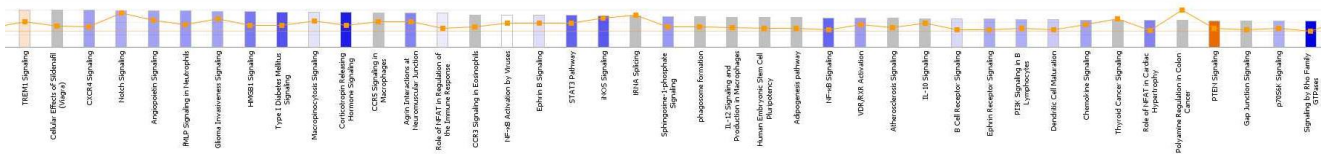
Supplemental Information

High Expression of CD200 and CD200R1 Distinguishes Stem and Progenitor Cell Populations within Mammary Repopulating Units

Gat Rauner, Tania Kudinov, Shlomit Gilad, Gil Hornung, and Itamar Barash

Cont.

Cont.

Cont.

Cont.

Induced in MRU$^{CD200/CD200R1}$
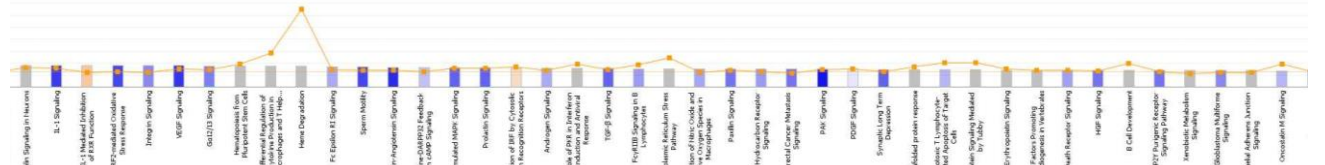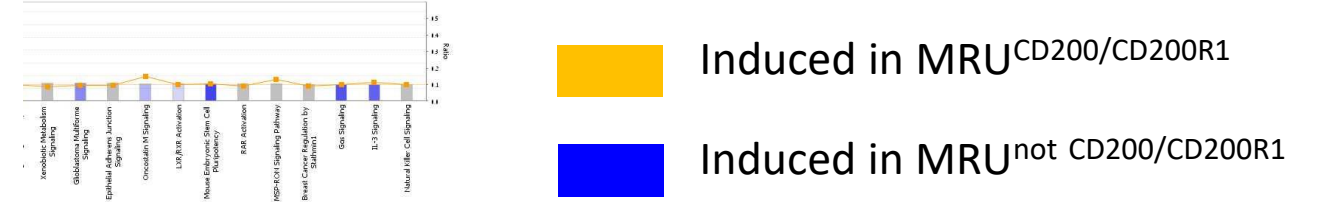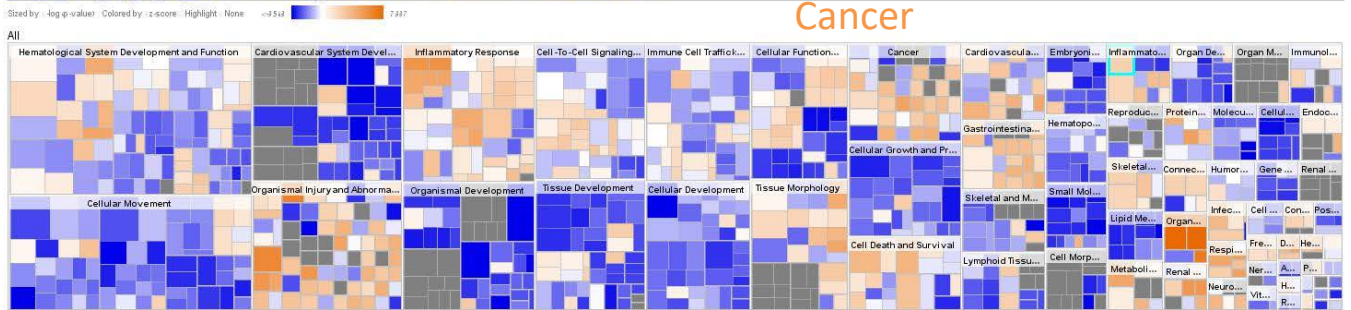
Induced in MRU$^{not\ CD200/CD200R1}$

Figure S1. Canonical Pathway Analyzed by IPA. Pathways that are Highly Induced in each of the MRU Subpopulations Relative to their Counterpart Are Marked. Related to Figure 2, and Table 1.

Inflammatory response

Cancer

both_vs_none_only_for_IPA_270116 fold 1.5 adj p 0.07- 2016-01-27 01:56 PM - Diseases & Functions

Sized by : -log (p-value) Colored by : z-score Highlight : None

All

Cellular movement

Organismal injury

Tissue devel.

Cellular devel.

Cell growth and prolif.

Cell death

© 2000-2016 QIAGEN. All rights reserved.

Highly activated in MRU$^{CD200/CD200R1}$

Highly activated in MRU$^{not\ CD200/CD200R1}$

Figure S2. IPA Analysis of Genes with Divergent Expression between MRU$^{CD200/CD200R1}$ and MRU$^{not\ CD200/CD200R1}$ that Maintain Different Diseases and Functions. Related to Figure 2 and Table 1.
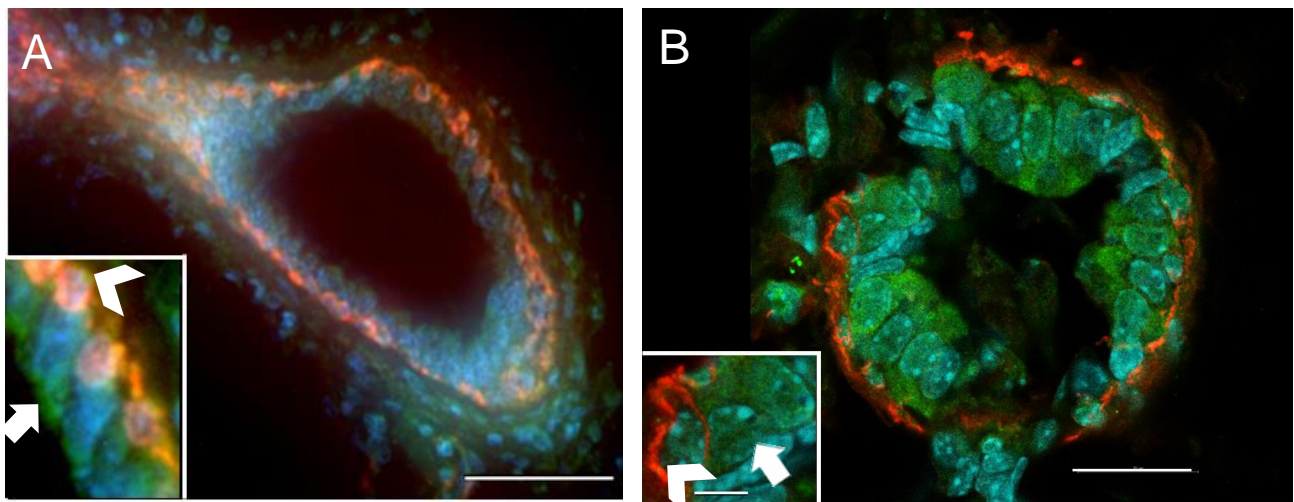
FIGURE S3. CD200 and CD200R1 are Ubiquitously Expressed in Basal and Luminal Cells of the Mammary Gland.

(A, B) Immunofluorescence analysis of FITC-labeled CD200/CD200R1 (green) and Cy3-labeled αSMA (red) expression in ductal structures of the virgin gland.

(A) CD200 is expressed at different levels in luminal and basal cells as well in the stroma. Bar = 50 μm. Inset x10 magnification. CD200-expressing basal cells are marked by arrowheads. CD200-expressing luminal cells are marked by arrow.

(B) CD200R1 expression is detected in both luminal and αSMA-expressing basal cells. Bar in main figure = 20μm. Bar in inset = 5μm. CD200R1-expressing luminal cells are marked by arrow. CD200R1-expressing basal cells are marked by arrowheads. Related to Figure 1.

A Stem cell and common progenitor
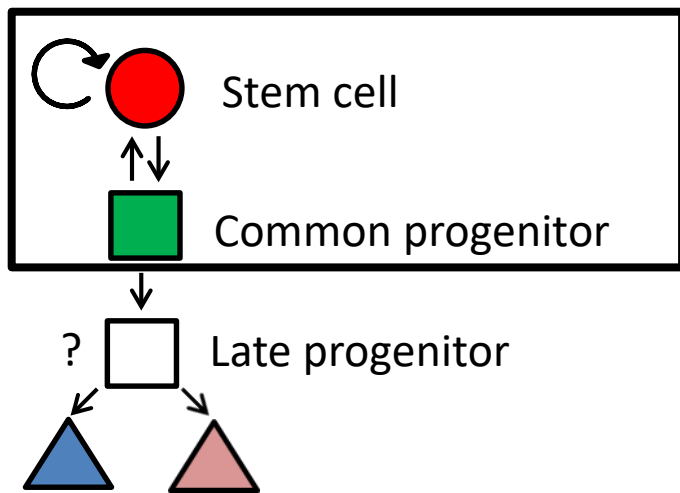
MRU$^{CD200/CD200R1}$ (stem cells, red) maintain self-renewal capability and multipotency, have higher mammosphere generation and are more basally oriented than MRU$^{not\ CD200/CD200R1}$ (common progenitors, green). They share stem cell markers. MRU$^{not\ CD200/CD200R1}$ (common progenitors, green) are multipotent but lack self-renewal capability. They are present in higher numbers after transplantation and have better differentiation capability than MRU$^{CD200/CD200R1}$. Supported by earlier studies (Rios et al., 2014).

B Quiescent and active stem cell
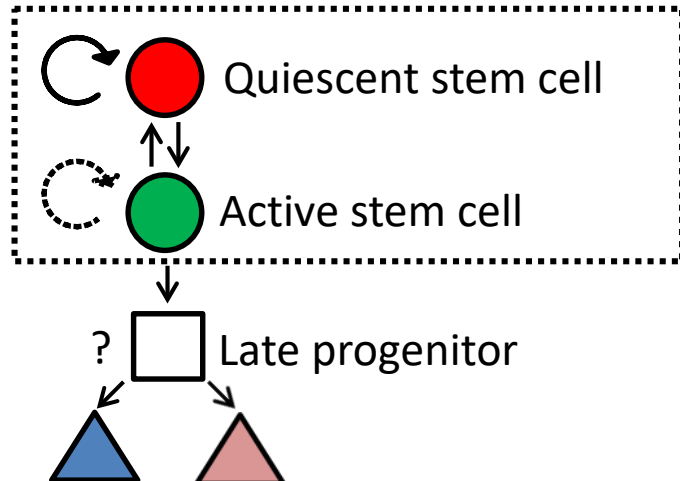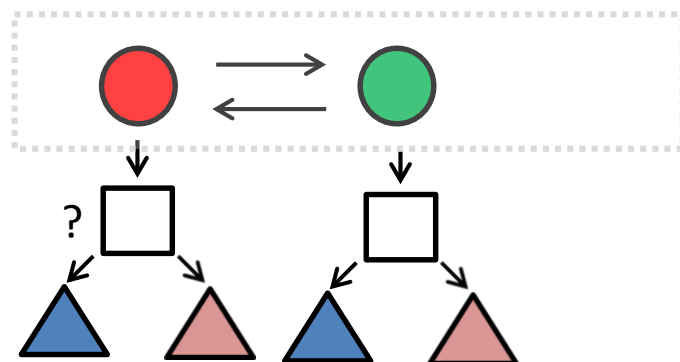

Quiescent and active stem cells have been characterized by dos Santos et al. (2013). The "active stem cells" (here, MRU$^{not\ CD200/CD200R1}$, green) do not share stem cell markers and do not maintain self-renewal capability. They differ in their basal origin from the quiescent stem cells (here, MRU$^{CD200/CD200R1}$, red).

C Two active stem cells

The two active stem cells should maintain comparable characteristics. However, self-renewal capability and stem cell markers are not shared by both active stem cells [here, MRU$^{CD200/CD200R1}$ (red) and MRU$^{not\ CD200/CD200R1}$ (green)]. There is a significant difference in the expression of genes leading to differentiation. The stem cells differ in basal origin and in mammosphere generation.

Figure S4. A Scheme of the Proposed Relationship Between MRU$^{CD200/CD200R1}$ and MRU$^{not\ CD200/CD200R1}$ in Composing the MRU Population.

A. Evidence supporting our view of a stem cell-common progenitor relationship within the MRU is in black font.

B,C. Evidence negating the alternative possibilities is in gray font. Dashed arrow indicates that this possibility cannot be formally ruled out. Related to all Figures and Tables.

Table S1. Outgrowth Development in the Host Cleared Fat Pad after Serial Transplantations[a]. Related to Figure 1.

Transplantation	Number of cells transplanted	Type of cells transplanted	Number of MRUs	Number of outgrowth-developing glands		P value
				$MRU^{CD200/CD200R1}$	$MRU^{not\ CD200/CD200R1}$	
1st	40	MRU	40	6/16 (38%)	4/14 (29%)	0.6
2nd	1000	Total mammary cell population	20 (estimated)	2/18 (11%)	0/16 (0%)	< 0.0001

[a]Initial transplantation was performed with sorted MRUs. The subsequent transplantation included dispersed mammary cells. These cells originated from small sections of outgrowth-containing tissue that were isolated under the binocular. The estimated number of MRUs in the total mammary cell populations is based on flow cytometry analyses indicating 57.0 ± 1.6% (n = 3) of epithelial cells in the total dispersed cells, of which 3.5 ± 1.0% (n = 3) are MRUs. *P* value was determined by Chi square (Pearson) analysis.

Table S2. Genes with Hierarchical-Dependent Expression in this Study[a].
Related to Figure 3 and Table 2.

Gene	Gene ID
Nalcn	ENSMUSG00000000197
Grik3	ENSMUSG00000001985
Ltbp2	ENSMUSG00000002020
Mrvi1	ENSMUSG00000005611
Fblim1	ENSMUSG00000006219
Dll1	ENSMUSG00000014773
Iist	ENSMUSG00000015143
Matn4	ENSMUSG00000016995
Stac2	ENSMUSG00000017400
Jph2	ENSMUSG00000017817
Ctgf	ENSMUSG00000019997
Hal	ENSMUSG00000020017
Kcnmb1	ENSMUSG00000020155
Smtn	ENSMUSG00000020439
Id4	ENSMUSG00000021379
Irx4	ENSMUSG00000021604
Ctnnd2	ENSMUSG00000022240
Trp63	ENSMUSG00000022510
Popdc2	ENSMUSG00000022803
Mylk	ENSMUSG00000022836
Il17b	ENSMUSG00000024578
Col17a1	ENSMUSG00000025064
Vwa2	ENSMUSG00000025082
Gfra1	ENSMUSG00000025089
1500015O10Rik	ENSMUSG00000026051
Wnt10a	ENSMUSG00000026167
Pkp1	ENSMUSG00000026413
Csrp1	ENSMUSG00000026421
Fcna	ENSMUSG00000026938
Itga6	ENSMUSG00000027111
Fermt1	ENSMUSG00000027356
Col9a3	ENSMUSG00000027570
Col11a1	ENSMUSG00000027966
Pdgfc	ENSMUSG00000028019
Pappa	ENSMUSG00000028370
Tpm2	ENSMUSG00000028464
Col9a2	ENSMUSG00000028626
Tfap2c	ENSMUSG00000028640
Cda	ENSMUSG00000028755
Plch2	ENSMUSG00000029055
Bst1	ENSMUSG00000029082

Shroom3	ENSMUSG00000029381
Synm	ENSMUSG00000030554
Micalcl	ENSMUSG00000030771
Col4a6	ENSMUSG00000031273
Col4a5	ENSMUSG00000031274
Trim29	ENSMUSG00000032013
Tagln	ENSMUSG00000032085
Elovl4	ENSMUSG00000032262
Ephb1	ENSMUSG00000032537
Folr2	ENSMUSG00000032725
Lama1	ENSMUSG00000032796
Fhod3	ENSMUSG00000034295
Ccdc106	ENSMUSG00000035228
Kank4	ENSMUSG00000035407
Fgf1	ENSMUSG00000036585
Arhgef26	ENSMUSG00000036885
Frem2	ENSMUSG00000037016
Zfp365	ENSMUSG00000037855
Trpm4	ENSMUSG00000038260
Scube3	ENSMUSG00000038677
Nexn	ENSMUSG00000039103
Apoc1	ENSMUSG00000040564
Fzd7	ENSMUSG00000041075
Sdk2	ENSMUSG00000041592
Ptpre	ENSMUSG00000041836
Lgr6	ENSMUSG00000042793
Tceal3	ENSMUSG00000044550
Defb1	ENSMUSG00000044748
Plekha7	ENSMUSG00000045659
Lrrc75b	ENSMUSG00000046807
Atxn1	ENSMUSG00000046876
Lmod1	ENSMUSG00000048096
Oxtr	ENSMUSG00000049112
Lrp1b	ENSMUSG00000049252
Tenm2	ENSMUSG00000049336
Tmem200a	ENSMUSG00000049420
Lgr4	ENSMUSG00000050199
Cd209f	ENSMUSG00000051906
Nav2	ENSMUSG00000052512
Cntn2	ENSMUSG00000053024
Hunk	ENSMUSG00000053414
Lgals7	ENSMUSG00000053522
Plxnb1	ENSMUSG00000053646
1600014C10Rik	ENSMUSG00000054676
Dsp	ENSMUSG00000054889
Fat2	ENSMUSG00000055333
Timd4	ENSMUSG00000055546

Trnp1	ENSMUSG00000056596
Cdh3	ENSMUSG00000061048
Retnla	ENSMUSG00000061100
Krt5	ENSMUSG00000061527
Hs6st2	ENSMUSG00000062184
Nrg1	ENSMUSG00000062991
Kcnma1	ENSMUSG00000063142
Luzp2	ENSMUSG00000063297
Chil1	ENSMUSG00000064246
Cd209b	ENSMUSG00000065987
Hs3st3b1	ENSMUSG00000070407
Npw	ENSMUSG00000071230
Tmem88b	ENSMUSG00000073680
Ntf5	ENSMUSG00000074121
Cd209g	ENSMUSG00000079168
Frmpd1os	ENSMUSG00000086284
Mia	ENSMUSG00000089661
Gm15867	ENSMUSG00000089812
Gm21655	ENSMUSG00000096878
Gm26615	ENSMUSG00000097088
A330048O09Rik	ENSMUSG00000097326
Gm26813	ENSMUSG00000097819
4631405K08Rik	ENSMUSG00000097850
Gm8579	ENSMUSG00000099762
Gm29157	ENSMUSG00000100635
Gm21190	ENSMUSG00000106445

[a]Gene expression decreased from MRU[CD200/CD200R1] via MRU[not CD200/CD200R1] toward CD200[+]- and CD200R1[+]-expressing cells.

Table S3. Genes with Hierarchical-Dependent Pattern of Expression that also Belong to the Previously Published Mammary/Breast Stem Cell Subset[1], or Mark Stem Cells in Other Tissues. Related to Figure 4 and Table 2.

Gene ID	Gene name	Symbol	Log_2 fold change $\text{MRU}^{CD200/CD200R1}/$ $\text{MRU}^{not\ CD200/CD200R1}$
ENSMUSG00000061527	keratin 5	Krt5	0.54863591
ENSMUSG00000034295	formin homology 2 domain containing 3	Fhod3	0.75778299
ENSMUSG00000022510	tumor protein p63	Trp63	0.7799422
ENSMUSG00000025064	collagen, type XVII, alpha 1	Col17a1	0.61716078
ENSMUSG00000027470	myosin light chain kinase	Mylk	0.95655246
ENSMUSG00000028464	tropomyosin 2 (beta)	Tpm2	0.50264019
ENSMUSG00000032085	transgelin	Tagln	0.53413686
ENSMUSG00000048096	leiomodin 1 (smooth muscle)	Lmod1	0.58008322
ENSMUSG00000005611	murine retrovirus integration site 1 homolog	Mrvi1	0.42691702
ENSMUSG00000030554	synemin, intermediate filament protein	Synm	0.56189161
ENSMUSG00000030554	smoothelin	Smtn	0.56189161
ENSMUSG00000032013	tripartite motif containing 29	Trim29	0.63381269
ENSMUSG00000061048	cadherin 3, type 1, P-cadherin (placental)	Cdh3	0.83388473
ENSMUSG00000021604	iroquois homeobox 4	Irx4	0.70640527
ENSMUSG00000026413	plakophilin 1 (ectodermal dysplasia/skin fragility syndrome)	Pkp1	0.81727681
ENSMUSG00000021379	inhibitor of DNA binding 4, dominant negative helix-loop-helix protein	Id4	0.70828064
ENSMUSG00000026051	chromosome 2 open reading frame 40	1500015O10Rik	0.8787596
ENSMUSG00000063142	potassium large conductance calcium-activated channel, subfamily M, alpha member 1	Kcnma1	0.79965976
ENSMUSG00000062991	neuregulin 1	Nrg1	0.40250429
ENSMUSG00000019997	connective tissue growth factor	Ctgf	0.55736619
ENSMUSG00000028626	collagen, type IX, alpha 2	Col9a2	0.59256227
ENSMUSG00000032796	laminin, alpha 1	Lama1	0.80467718
	lectin, galactoside-binding, soluble, 7	Lgals7	0.52174104
ENSMUSG00000024578	interleukin 17B	Il17b	0.57630163
ENSMUSG00000089661	melanoma inhibitory activity	Mia	0.7526307
ENSMUSG00000027111	integrin, alpha 6	Itga6	0.47572854
ENSMUSG00000041836	protein tyrosine phosphatase, receptor type E	Ptpre	0.27070909
ENSMUSG00000032262	ELOVL fatty acid elongase 4	Elovl4	0.55469019
ENSMUSG00000049112	oxytocin receptor	Oxtr	0.70664276
ENSMUSG00000032537	EPH receptor B1	Ephb1	0.60780788
ENSMUSG00000017400	SH3 and cysteine rich	Stac2	0.48402756

Gene ID	Gene name	Symbol	Log₂ fold change	
	domain 2			
ENSMUSG00000035407	KN motif and ankyrin repeat domains 4	*Kank4*	0.70168944	
ENSMUSG00000014773	delta-like 1 (*Drosophila*)	*Dll1*	0.57855187	
ENSMUSG00000020155	potassium large conductance calcium-activated channel, subfamily M, beta member 1	*Kcnmb1*	0.73793513	
ENSMUSG00000042793	leucine-rich repeat containing G protein-coupled receptor 6	*Lgr6*	0.50525771	
ENSMUSG00000054676	chromosome 19 open reading frame 12	*1600014C10Rik*	0.61677566	
ENSMUSG00000029055	phospholipase C, eta 2	*Plch2*	0.71703952	
ENSMUSG00000041592	sidekick homolog 2 (chicken)	*Sdk2*	0.56928968	
ENSMUSG00000022803	Popeye domain containing 2	*Popdc2*	0.68476	

Gene ID	Gene name	Symbol	Log₂ fold change: $MRU^{CD200/CD200R1}$/ $MRU^{not\ CD200/CD200R1}$	Tissue	Reference
ENSMUSG00000002020	latent transforming growth factor beta binding protein 2	*Ltbp2*	0.76	Chondrocytes	Goessler et al., 2005
ENSMUSG00000020017	histidine ammonia-lyase	*Hal*	2.29	Xenopus laevis intestine	Luu et al., 2013
ENSMUSG00000026167	Wnt family member 10a	*Wnt10A*	0.63	Esophageal squamous cell carcinoma	Long et al., 2015
ENSMUSG00000029082	bone marrow stromal cell antigen 1	*Bst1*	0.86	Mesenchyma	Aumatuso et al., 2014
ENSMUSG00000041075	frizzled class receptor 7	*Fzd7*	0.98	Intestinal epithelial Lgr5⁺ cells	Flanagan et al., 2015
ENSMUSG00000050199	leucine-rich repeat-containing G-protein-coupled receptor 4	*Lgr4*	0.63	Embryonic intestine	Kinzel et al., 2014
ENSMUSG00000056596	TMF1-regulated nuclear protein 1	*Trnp1*	0.56	Cereberal cortex	Sthal et al., 2013

[1]Lim et al., 2010

Table S4. Validation of Proposed Cell Hierarchy[a]. Related to Figures 3 and to Table 2.

#	Source of genes in set 1	Number of genes in set 1	Source of genes in set 2	Number of genes in set 2	Number of common genes	Significance (in set of 20,000 genes)
1	Current analysis: Stem cell subset	114	Stem cell subset (Lim et al., 2010)	489	39	$P = 2.75 \times 10^{-34}$
2	Current analysis: progenitor subset	262	Stem cell subset (Lim et al., 2010	489	20	$P = 5 \times 10^{-6}$
3	Current analysis: differentiated cell subset	97	Stem cell subset (Lim et al., 2010	489	10	$P = 1 \times 10^{-4}$

[a]The gene set identifying stem cells is much more compatible with the previously published stem cell list (Lim et al., 2010) than the progenitor and differentiated cell subsets.

Table S5. Transplantation of MRU$^{CD200/CD200R1}$ population results in outgrowths containing MRU$^{not\ CD200/CD200R1}$ and vice versa. Related to Figures 1,3 and 5.

Population	MRU composition in the resulting outgrowths (%)	
	MRU$^{CD200/CD200R1}$	MRU$^{not\ CD200/CD200R1}$
Control	17.0	83.0
Transplanted MRU$^{CD200/CD200R1}$	14.2	85.8
Transplanted MRU$^{not\ CD200/CD200R1}$	28.1	71.9

MRU$^{CD200/CD200R1}$ and MRU$^{not\ CD200/CD200R1}$ populations were sorted from depressed mammary epithelial cells collected from 36 #4 mammary glands as described in Experimental procedures. From each population, 100 cells were transplanted into each one of 10 de-epithelized mammary fat pads of recipient virgin females. Eight weeks after transplantation, fat pads of each group were collected and digested for dispersed epithelial cells. These were sorted for the presence of MRU$^{CD200/CD200R1}$ and MRU$^{not\ CD200/CD200R1}$ together with cells extracted and pooled from six #4 mammary glands of 6-week-old control virgin mice.

Table S6. List of Antibodies Used in this Study.

Antigen/ Purpose	1st Antibody and dilution	Manufacturer	2nd antibody and dilution	Manufacturer
CD24 FACS	PE-conjugated, rat monoclonal, clone M1/69 (1:30)	StemCell Technologies		
CD49f FACS	FITC-conjugated, rat monoclonal, clone GoH3 (1:30)	StemCell Technologies		
CD200 FACS	APC-conjugated, rat monoclonal, clone OX-90 (1:30)	BioLegend, San Diego, CA		
CD200R1 FACS	Alexa Fluor 700-conjugated, rat monoclonal, clone OX-110 (1:30)	Novus Biologicals, Littleton, CO		
CD200 IP	FITC-conjugated rat monoclonal LS-C534810 (1:50)	LSBio, Seattle, WA		
CD200R1 IP	Rabbit polyclonal LS-C378947 (1:50)	LSBio	Alexa Flour 488 goat anti rabbit A1108 (1:500)	Life Technology Paisley, UK
αSMA IP	Mouse monoclonal SC-32251 (1:75)	Santa Cruz, Dallas, TX	Cy3 goat anti mouse 115-165-003 (1:100)	Jackson Immunoresearch Baltimore Pike, PA
αSMA IHC	Mouse monoclonal SC-32251 (1:75)	Santa Cruz	EnVision-labeled HRP polymer	Dako Cytomation, Glostrup, Denmark
β-casein IHC	Rabbit monoclonal (1:100)	Barash's Lab.	EnVision-labeled HRP polymer	Dako Cytomation,
CK18 IP	Mouse monoclonal (ab668) (1:100)	Abcam, Cambridge, UK	Cy3 goat anti mouse 115-165-003 (1:100)	Jackson Immunoresearch
CK14 IP	Mouse monoclonal clone LL002 (1:50)	Biorad, Perth, UK	Cy3 goat anti mouse 115-165-003 (1:100)	Jackson Immunoresearch

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Dissociation of Mammary Tissue into Single-Cell Suspension and Flow Cytometry

Mouse mammary cells were dissociated from the fourth inguinal mammary glands of 6- to 8-week-old FVB/N virgin females. The dissociation procedure was as previously described (Rauner and Barash, 2012), except that enzymatic digestion was shortened from 3 h to 1 h. Lin⁻ cell suspension was prepared using the EasySep mouse mammary enrichment kit (StemCell Technologies, Vancouver, Canada) according to the manufacturer's protocol, as previously described (Rauner and Barash, 2012). Epithelial cells were resuspended in HF solution containing Hank's balanced salt solution (Biological Industries, Bet-Haemek, Israel) supplemented with 0.1% (v/v) Hepes (Biological Industries) and 2% (v/v) fetal bovine serum (Biological Industries) at a concentration of 10^7 cell/ml, and incubated for 90 min on ice with fluorescent conjugated antibodies. The antibodies and their dilutions are listed in Table S6. Cell viability was tested by staining with BD Horizon Fixable Viability Stain 450 (FVS450, BD Biosciences, San Jose, CA) according to the manufacturer's protocol. Cell sorting and cell analyses were performed in a FACSAria II or III cell sorter and LSR II cell analyzer (BD Biosciences) at the Department of Biological Services of the Weizmann Institute of Science (Rehovot, Israel). Resulting data were visualized and analyzed by FACSDiva (BD Biosciences) and WinMDI 2.9 software (Scripps Research Institute, La Jolla, CA).

Soring $MRU^{CD200/CD200R1}$ and $MRU^{not\ CD200/CD200R1}$ populations:

Live $MRU^{CD200/CD200R1}$ and $MRU^{not\ CD200/CD200R1}$ cells were sorted from Lin⁻ mammary epithelial cells that were simultaneously analyzed for CD49f, CD24, CD200, CD200R1 and FVS450 (which separates live from dead cells). $CD200^{high}CD200R1^{high}$ cells which express both proteins at high levels were gated. The gated $CD200^{high}CD200R1^{high}$ population was projected onto the plot depicting expression of CD24 and CD49f, to collect only those $CD200^{high}CD200R1^{high}$ cells that were also MRU ($MRU^{CD200/CD200R1}$). The rest of the MRUs are refereed as $MRU^{not\ CD200/CD200R1}$.

Some of the $CD200^{high}CD200R1^{high}$ cells were not MRU (and were not collected), but rather demonstrate a $CD49f^{high}/CD24^{low}$ phenotype, suggesting that they may be myoepithelial cells (Stingl et al. 2006, Nature 23:439). We did not identify $CD200^{high}CD200R1^{high}$ cells expressing $CD49f^{neg/low}/CD24^{med/high}$ as would be expected from luminal cells or luminal progenitors (Stingl et al. 2006, Nature 23:439).

Cell Transplantation

The endogenous mammary epithelium was surgically removed bilaterally from #4 mammary glands of 21-day-old female mice weighing 10–12 g (i.e., "clearing"). Sorted mammary epithelial cell populations were resuspended in 1:1 (v/v) HF solution:Matrigel (BD Biosciences) and the cell suspension (20 µl) was injected into the cleared mammary fat pad using a 50-µl Hamilton syringe (Hamilton Company, Reno, NV) equipped with a 21-gauge needle. Outgrowths were allowed to develop for 8.5 weeks before the transplanted fat pad was removed.

Serial Transplantation

A similar transplantation procedure was performed for serial transplantations of $CD200^{high}/CD200R1^{high}$ MRUs. Host mice were C57BL and donor mice were C57/GFP [ubiquitin-EGFP mice (Reichenstein et al., 2016)] that express enhanced GFP under the direction of the human ubiquitin C promoter in all tissues; the expression is uniform within a cell-type lineage and remains constant throughout development. GFP-expressing primary outgrowths were visualized in whole mounts of host mammary fat pad using a binocular (Olympus SZX16, Tokyo, Japan) equipped with CellSens standard 1.4 software (Olympus). The observed outgrowths were dissected from the host fat pad and dissociated into single cells as described above. The single-cell suspension inevitably included non-GFP-expressing stromal cells from the host mouse. Cells (n = 1000) were transplanted as described above into the cleared fat pad of C57BL secondary hosts, and developed outgrowths were visualized after 7 weeks using a binocular equipped with CellSens standard 1.4 software.

Outgrowth Analysis

For whole-mount examination, transplanted mammary fat pads were excised from sacrificed mice, fixed and stained with Carmine Red as previously described (Rauner et al., 2013). Stained whole mounts were visualized and photographed using the binocular equipped with CellSens standard 1.4 software.

Histological Analysis and Immunostaining

Immunostaining of paraformaldehyde-fixed cells in culture was performed as previously described (Rauner and Barash, 2012). Tissue immunostaining was performed on paraffin-embedded (Rauner and Barash, 2012) or frozen tissue sections. For the latter, tissue biopsies were fixed in 4% paraformaldehyde for 24 h at 4ºC and then transferred to 15% and 30% sucrose solution for 6 h and 24 h, respectively. Tissues were submerged in OCT compound (Sakura Finetek, Alphen aan den Rijn, The Netherlands) and stored at -80ºC. Immunostaining was performed on 5-mm paraffin-embedded or frozen sections. Antigen retrieval for paraffin-embedded sections was performed by boiling in 0.01 M citrate buffer for 10 min. For frozen sections, it was performed by submerging the slides in 1% SDS solution pH 7.4. The reactions with primary and secondary fluorescence-labeled antibodies (Table S6) followed the protocol described in (Rauner and Barash, 2012). For immunohistochemistry, paraffin-embedded sections were treated with 3% hydrogen

peroxide for 30 min and boiled for 10 min in 0.01 M citrate. Sections were incubated with the primary antibody (Table S6), followed by incubation with EnVision-labeled HRP polymer (Dako Cytomation, Glostrup, Denmark) for 1 h at room temperature. Signal was generated with 3,39-diaminobenzidine (DAB) substrate kit (Vector Laboratories, Burlingame, CA) according to the manufacturer's protocol.

Clonal and Mammosphere Assays
For clonal assay, sorted or unsorted cells were seeded in a 24-well culture plate (Corning, Lowell, MA) at a density of 2500 cell/well and cultured for 7 days, in mammary medium (Rauner and Barash, 2012). Developing colonies were fixed in 4% paraformaldehyde supplemented with 0.03 M sucrose, permeabilized with 0.5% Triton X-100 and stained with antibodies to CK14 or CK18 as described in the Histological Analysis and Immunostaining section. Clones consisting of at least three individual adjacent cells were counted using an inverted fluorescence microscope (Eclipse Ti, Nikon Instruments, Melville, NY) and characterized as luminal or basal.

For mammosphere assay, sorted populations were individually seeded in 96-well ultra-low-attachment plates (Corning) at a density of 100 cell/well and supplemented with mammary medium (Rauner and Barash, 2012) for 8 days. Developing mammospheres were analyzed and counted as described above.