

In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells

Gabriela Dontu,¹ Wissam M. Abdallah, Jessica M. Foley, Kyle W. Jackson, Michael F. Clarke, Mari J. Kawamura, and Max S. Wicha

Department of Internal Medicine, Hematology–Oncology, Comprehensive Cancer Center, University of Michigan, Ann Arbor, Michigan 48109, USA

Although the existence of mammary stem cells has been suggested by serial transplantation studies in mice, their identification has been hindered by the lack of specific surface markers, and by the absence of suitable in vitro assays for testing stem cell properties: self-renewal and ability to generate differentiated progeny. We have developed an in vitro cultivation system that allows for propagation of human mammary epithelial cells (HMECs) in an undifferentiated state, based on their ability to proliferate in suspension, as nonadherent mammospheres. We demonstrate that nonadherent mammospheres are enriched in early progenitor/stem cells and able to differentiate along all three mammary epithelial lineages and to clonally generate complex functional structures in reconstituted 3D culture systems. Gene expression analysis of cells isolated from nonadherent mammospheres revealed overlapping genetic programs with other stem and progenitor cells and identified new markers that may be useful in the identification of mammary stem cells. The isolation and characterization of these stem cells should help elucidate the molecular pathways that govern normal mammary development and carcinogenesis.

[*Keywords:* Stem cells; mammary gland development; gene expression analysis; microarray analysis]

Supplemental material is available at <http://www.genesdev.org>.

Received November 25, 2003; revised version accepted March 18, 2003.

The adult mammary gland has a lobulo-alveolar structure, composed of three cell lineages: myoepithelial cells that form the basal layer of ducts and alveoli, ductal epithelial cells that line the lumen of ducts, and alveolar epithelial cells that synthesize milk proteins (Rudland et al. 1997; Hennighausen and Robinson 2001). These cells proliferate extensively and differentiate during each pregnancy and lactation and undergo apoptosis during mammary involution (Hennighausen and Robinson 2001). As is the case in other organs, the cellular repertoire of the human mammary gland may be generated by a stem cell component. Factors intrinsic to these putative stem cells as well as their interactions with stromal cells and extracellular matrix are important in the normal morphogenesis of the gland (Wiseman and Werb 2002). Moreover, experimental and clinical data have accumulated to support the hypothesis that cancer may arise from mutations in stem cell populations (Reya et al. 2001). Indeed, the similarities between cancer cells and normal stem cells, such as increased proliferation

potential and self-renewal ability, were noticed decades ago (Rudland et al. 1997; Weissman 2000). The long-lived, slowly dividing stem cells, by their very nature more exposed to damaging agents than the more differentiated cells, can accumulate mutations that are then transmitted to the fast-proliferating progeny. The switch to carcinogenesis can occur in either the stem cells or one of their progeny that acquires the ability to self-renew. The purification and characterization of normal adult stem cells specific for the mammary gland could therefore be extremely helpful for understanding normal mammary development, as well as carcinogenesis.

The existence of self-renewing, multipotent mammary stem cells has been suggested by transplantation studies in mice and rats (Daniel et al. 1971; Kim et al. 2000). Random fragments of mammary epithelium marked with mouse mammary tumor virus (MMTV) were able to regenerate a new gland upon transplantation (Kordon and Smith 1998). Clonal dominant populations were generated, as demonstrated by the unique site of viral integration. Serial transplantation of the clonally derived outgrowth recapitulated the entire functional repertoire of the gland, demonstrating the existence of self-renewing and multipotent mammary stem cells. A recent study in mice combining long term in vivo BrdU

¹Corresponding author.

E-MAIL gdontu@umich.edu; FAX (734) 647-9480.

Article and publication are at <http://www.genesdev.org/cgi/doi/10.1101/gad.1061803>.

labeling with immunosorting and transplantation, showed that the mammary SCA-1+ (stem cell antigen 1 population) is enriched in progenitor cells, able to regenerate the gland in vivo (Welm et al. 2002).

To study the functional properties of stem cells, one needs to identify and prospectively purify them, a task that has proved technically difficult because of the scarcity of stem cells in the tissue of origin, and the lack of universal morphologic traits for stem cells (Blau et al. 2001). Stem cells are defined by their ability to self-renew and to generate differentiated progeny. In adult organs, these properties are manifest by extensive self-renewal potential, which is the basis for tissue maintenance, and the ability of stem cells to generate the entire repertoire of cell types found in a tissue (Morrison et al. 1997). These functional properties can be tested in in vitro and in vivo assays. A characteristic also shared by a number of stem cells is their ability to exclude dyes like Hoechst or rhodamine as a result of increased expression of membrane transporter proteins, such as P-glycoproteins or BCRP (breast cancer resistance proteins; Zhou et al. 2000; Bunting 2002). This increased dye exclusion identifies a small cell population called the side population (SP). The SP fraction, although heterogeneous, was shown to contain the long-term repopulating hematopoietic stem cells, the primitive myoblasts, and epidermal stem cells (Bunting 2002). Similarly, it has recently been demonstrated that in the mouse mammary gland, cells contained within the SP population can regenerate the gland upon transplantation (Alvi et al. 2002; Welm et al. 2002).

Most stem-cell enrichment protocols rely on immunosorting, and use sets of antibodies to cell-surface proteins. The subpopulations defined by various markers are then clonally assayed for the functional characteristics described above (Blau et al. 2001). The efforts to purify adult stem cells from the human mammary gland have been hindered by the lack of cell-surface markers specific for undifferentiated or differentiated mammary cells, and by the lack of suitable in vitro assays for testing stem cell properties. Several studies have described bipotent human mammary epithelial progenitor cells in normal adult breast tissue based on immunosorting using ESA (epithelial specific antigen), MUC 1 (mucin 1), and $\alpha 6$ integrin, in combination with exclusion of rhodamine dye (Stingl et al. 1998, 2001). Using these markers, a 10-fold increase in bipotent progenitors was obtained compared with unsorted cells.

Other investigators have used an alternative approach in which immortalized mammary cell lines are established from human or rodent tissues. For instance, Gudjonsson and coworkers (Gudjonsson et al. 2002) described an ESA-positive MUC 1 negative cell line derived from human mammary cells that was capable of generating ductal-acinar structures in basement membrane gel. Complete functional differentiation and synthesis of milk proteins was not shown for these cells. Although cell lines are useful for elucidating molecular pathways, the process of immortalization may introduce artifacts that significantly alter cellular properties and gene expression profiles.

An alternative, less biased approach than the use of established cell lines is the isolation and propagation of normal human mammary progenitor cells from primary tissue. This approach, however, has been limited by the lack of suitable systems that allow for the propagation of these cells in an undifferentiated state. When primary cultures of mammary epithelium from rodents or humans are cultured on solid substrata, they undergo limited replication and differentiate in a process that is regulated by hormonal factors, extracellular matrix, and cell-cell interactions (Reynolds and Weiss 1996; Muschler et al. 1999; Romanov et al. 2001; Simian et al. 2001). A major advance in neural stem cell research was achieved when it was found that an undifferentiated multipotent population of neural cells can be grown in suspension as neurospheres (Reynolds and Weiss 1996; Weiss et al. 1996). Neurospheres were shown to consist of 4%–20% stem cells, the rest of the population representing progenitor cells in various stages of differentiation. This cultivation method was instrumental in a variety of experimental systems: stem cell enrichment assays (Uchida et al. 2000), comparative gene expression profiling (Geschwind et al. 2001), and in vitro models for the development of the nervous system (Caldwell et al. 2001). We have developed a similar strategy that allows for the cultivation of undifferentiated human mammary epithelial cells in suspension. Analogous to primary neural cells, human mammary epithelial cells form spherical colonies, which we termed “nonadherent mammospheres,” when cultured on nonadherent surfaces in the presence of growth factors. We demonstrate that nonadherent mammospheres are enriched in cells with functional characteristics of stem/progenitor cells. Based on this cultivation system, we established in vitro assays suitable for testing self-renewal and differentiation along all three lineages present in the mammary gland. Transcriptional profiling of mammosphere-derived cells demonstrates significant overlap with the genetic programs of other stem cells. The use of this culture system should greatly facilitate the isolation and characterization of human mammary stem cells as well as the elucidation of the pathways that govern their self-renewal and differentiation.

Results

A small population of human mammary epithelial cells can survive and proliferate in anchorage-independent conditions

Single-cell suspensions of human mammary epithelial cells (HMECs), obtained by mechanical and enzymatic dissociation from reduction mammoplasties, were grown in conditions that do not allow adherence to a substratum, as described in Materials and Methods. As is the case for primary neural cells, most primary HMECs died under these conditions, but a small number survived and generated floating spherical colonies, after 5–7 d in culture (Fig. 1A). We have termed these colonies nonadherent mammospheres by analogy to neuro-

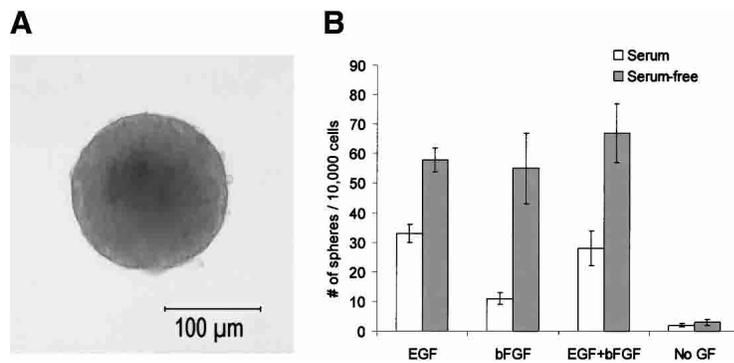


Figure 1. (A) Mammosphere morphology and optimal growth conditions. Mammosphere after 15 d of cultivation. (B) The sphere-formation efficiency is considerably higher in serum-free medium in the presence of EGF and/or bFGF than in the absence of growth factors. Data are presented as means \pm S.E.M.

spheres, and to distinguish them from the dome-shaped structures formed by HMECs on certain substrata, which have previously been termed mammospheres. For simplicity, however, we will further refer to them as mammospheres. We tested various culture conditions and determined that serum-free medium containing EGF and/or bFGF at saturating concentrations (10–20 ng/mL), supplemented with insulin, hydrocortisone, and B27, promotes optimal mammosphere formation (Fig. 1B). Mammospheres were passaged at 7–10-d intervals by mechanical and enzymatic dissociation, as described in Materials and Methods.

Cellular composition of mammospheres

To assess their cellular composition, secondary mammospheres were immunostained after 15 d of cultivation, using markers specific for the differentiated mammary cells of luminal epithelial and myoepithelial lineages. In addition, we performed immunostaining with markers identified in previous studies as candidates for mammary stem/progenitor cells. None of these markers, however, are expressed exclusively on progenitor cells or on differentiated cells. ESA, Muc 1, and cytokeratin 18 are associated with differentiated luminal epithelial cells; CD10, ASMA (α smooth muscle actin), cytokeratin 14, and integrin α 6 are associated with differentiated myoepithelial cells (Rudland et al. 1997). ESA, α 6 integrin, and CD10 were also associated with bipotent progenitors (Stingl et al. 1998, 2001; Gudjonsson et al. 2002). Cytokeratin 5 and 14 have been attributed to putative progenitor cells (Bocker et al. 2002). All the secondary mammospheres analyzed contained cells positive for α 6 integrin, cytokeratin 5, and CD10, and only some of them contained ESA-positive and cytokeratin 14-positive cells (~50% and ~30%, respectively; Fig. 2B–F). The number and distribution of the cells expressing a particular antigen appeared in a characteristic pattern. Whereas CD10, cytokeratin 5, and α 6 integrins (Fig. 2B–D) were present in a few, centrally located cells, ESA and cytokeratin 14-positive cells were randomly distributed within the mammospheres (Fig. 2E,F). We did not detect expression of Muc 1, ASMA, and cytokeratin 18 in secondary mammospheres (Fig. 2G–I). Double staining for ESA and cytokeratin 5 showed that all ESA-positive cells

also expressed cytokeratin 5, but cytokeratin 5 was more widely expressed (Fig. 2J–L).

The H&E staining of mammospheres showed that in addition to cells, spheres contained extracellular material (Fig. 2M). Therefore, we immunostained mammosphere sections for the extracellular matrix (ECM) components fibronectin, laminin, and collagen IV, to determine if growth in suspension is associated with secretion and deposition of ECM by the cells. None of the spheres stained positively for collagen IV and fibronectin, and ~20% stained positive for laminin (Fig. 2N–P). In contrast to the absence or low level of expression of ECM components found in the adult gland, we detected abundant expression of embryonic ECM components, including tenascin and decorin, in mammospheres (Fig. 2Q,R; Jones and Jones 2000; Kresse and Schonherr 2001).

Mammospheres are highly enriched in progenitor cells capable of differentiation along multiple lineages

To test if the mammosphere-derived cells have stem or progenitor cell properties, characterization of individual cells within the spheres was performed, following the strategy outlined in Figure 3. We developed two assays to determine the multilineage differential potential of mammosphere-derived cells (Fig. 3A), and one assay to test self-renewal (Fig. 3B). One assay for differentiation potential is based on assessing the number of multilineage colonies generated by cultivating cells at clonogenic densities in differentiating conditions, on a collagen substratum [see Material and Methods; Fig. 3A(i)]. When plated on collagen in the presence of serum, HMECs terminally differentiate, arrest, and eventually undergo senescence after a limited number of divisions (Stingl et al. 2001; Romanov et al. 2001). To test the potential to differentiate along the alveolar epithelial cell lineage, cells proliferating on a collagen substratum, in differentiating conditions, were covered by an overlaying layer of reconstituted basement membrane gel (Matrigel) in medium supplemented with prolactin [Fig. 3A(ii)]. It has been shown in several studies that the genes encoding milk proteins are expressed at very low levels in HMECs grown in 2D collagen cultures (Lee et al. 1985; Parry et al. 1985; Muschler et al. 1999). Cultivation in floating collagen gels or in Matrigel, however, allowed differen-

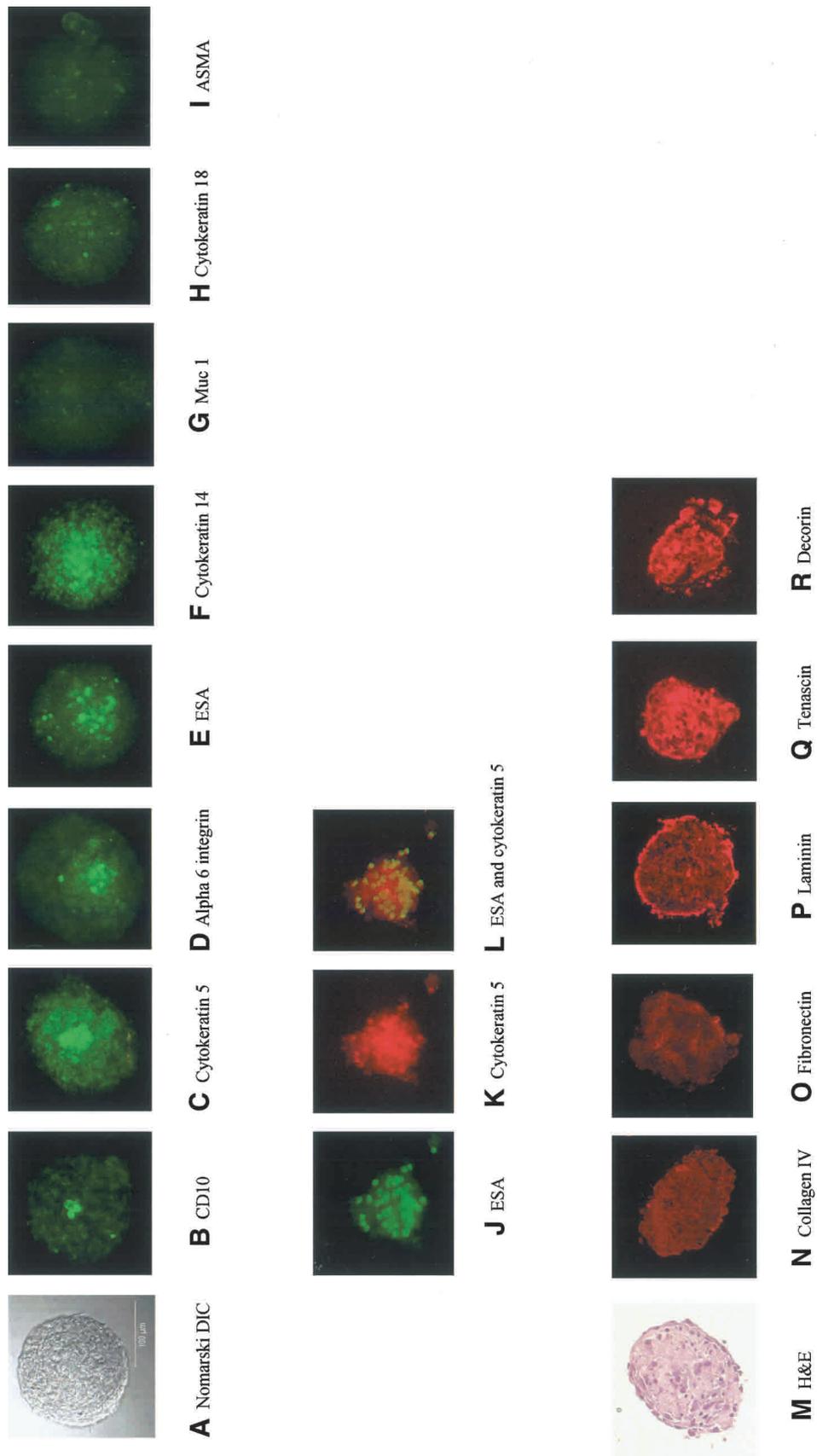


Figure 2. Cellular and ECM composition of secondary mammospheres. (A) Nomarski DIC (differential interference contrast) image of an unstained mammosphere. Fifteen-day-old secondary mammospheres immunostained in suspension (B-I), after centrifugation on slides (J-L), or as paraffin-embedded sections (N-R). (M) H&E-stained section through a paraffin-embedded sphere. The primary antibodies used are indicated *under* each picture. FITC-labeled or Texas-red-labeled secondary antibodies were used. Laminin immunostaining was positive in ~20% of mammospheres. (P) A representative positive-staining sphere is shown.

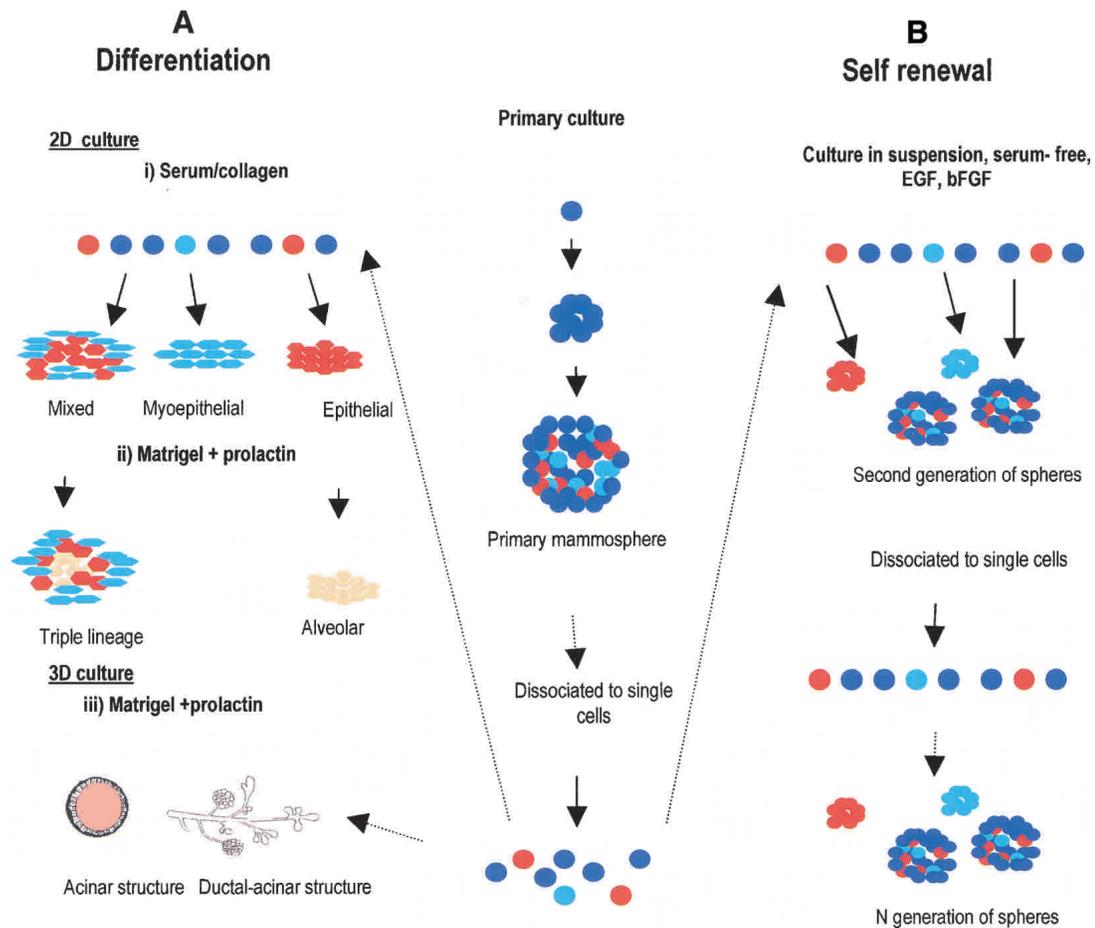


Figure 3. Experimental design for assessing the differentiation and self-renewal potential of cells grown as mammospheres. (A) Differentiation is tested by plating cells at clonogenic densities. (i) Differentiation into ductal and myoepithelial cells is assessed on cells cultivated on collagen in the presence of serum. (ii) Differentiation into ductal, myoepithelial, and alveolar cells is tested in the same conditions as in *i* with the addition of prolactin and Matrigel as an overlayer. (iii) The ability to generate complex ductal-alveolar structures is tested in 3D Matrigel culture. (B) Self-renewal is tested by evaluating the ability of mammosphere-derived cells to form new spheres, containing multipotent cells.

tiation and full functionality of alveolar cells, as tested by the production of milk proteins (Lee et al. 1985; Weaver and Bissell 1999; Zantek et al. 2001). Because 3D cultures are not suitable for immunostaining and scoring large numbers of clonally derived colonies, we devised a modified cultivation system, in which cells are allowed to grow for 5–7 d on collagen, after which they are covered with Matrigel and further cultivated for 7 d in the presence of prolactin. We assessed lineage-specific differentiation by immunostaining, using the markers specific for ductal and myoepithelial cells, as described above. For the detection of functional alveolar cells, antibodies against β -casein were used. β -Casein production is associated with the terminal phase of alveolar differentiation.

Three types of colonies were generated in differentiating conditions by mammosphere-derived cells plated at clonogenic densities: colonies containing exclusively ductal epithelial cells (Fig. 4A), colonies containing exclusively myoepithelial cells (Fig. 4B), and mixed colo-

nies, containing both lineages (Fig. 4C). Because the plating was performed at clonogenic densities, and each colony represents the progeny of a single cell, the mixed colonies are generated by bipotent progenitors. The percentage of cells with bilineage differentiation potential was substantially higher in primary mammospheres ($68\% \pm 3\%$), compared with freshly dissociated tissue ($8\% \pm 3\%$; Fig. 4D). The percentage of cells with bilineage potential increased to $98\% \pm 1\%$ in secondary and later generation mammospheres. The commitment to luminal epithelial lineage was determined by immunostaining with ESA, Muc 1, and cytokeratin 18 antibodies, and the commitment to myoepithelial lineage was determined by immunostaining with CD10, ASMA, and cytokeratin 14 antibodies. These are previously described lineage-specific markers for the mammary epithelium (Rudland et al. 1997). Consistent results were obtained with each of the luminal epithelial- and myoepithelial-specific markers.

The existence of human mammary progenitors ca-

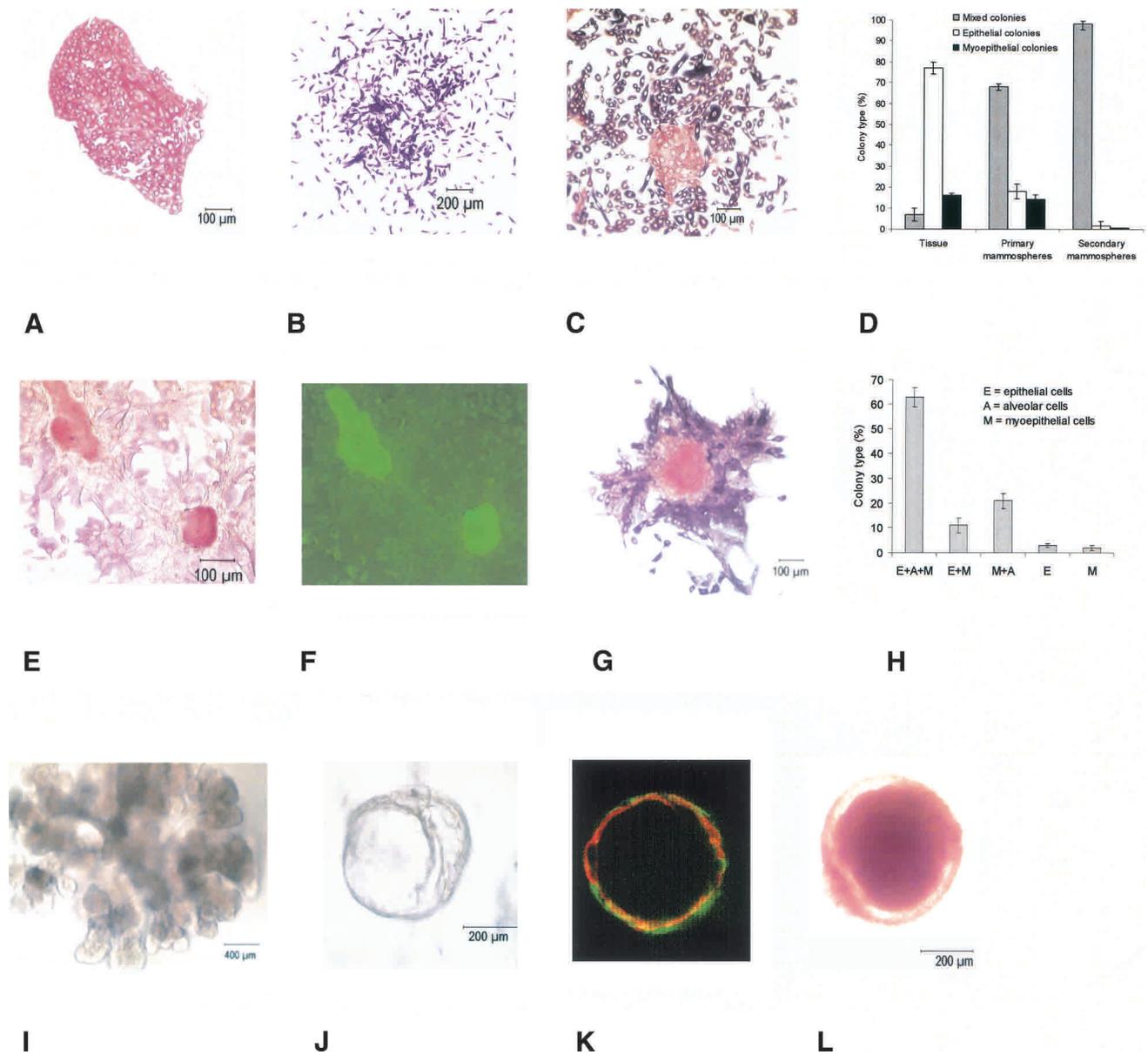


Figure 4. Mammospheres contain multipotent cells, capable of differentiating along all three lineages in the mammary gland. (A) Monolineage ductal epithelial colony immunostained for ESA. (B) Monolineage myoepithelial colony immunostained for CD10. (C) Bilineage colony (ESA, red; CD10, purple). The percentage of cells with bilineage differentiation potential is considerably higher in primary mammospheres compared with dissociated tissue. (D) Secondary mammospheres contain virtually only bipotent progenitors. (E) Triple-lineage colony (ESA, red; CD10, purple). (F) The same colony immunostained for β -casein (FITC, green) to identify the third lineage, alveolar cells. (G) Triple-lineage colony (ESA, brown; CD10, purple; β -casein, red). (H) Distribution of colony types generated by mammospheres derived cells in conditions that allow differentiation along all three lineages. (I) Ductal-acinar structure generated by a single cell in 3D Matrigel culture. (J) Acinar structure generated by a single cell in 3D Matrigel culture. (K) Acinar structure, immunostaining for myoepithelial lineage (CD10, FITC, green) and ductal epithelial lineage (ESA, Texas red). (L) Acinar structure, immunostaining for β -casein (red).

pable of differentiation into all three lineages of the adult mammary gland has not been previously demonstrated in vitro. Using the cultivation system described above, which combines growth on collagen with exposure to basement membrane components and prolactin, we found that the majority of the bipotent cells ($66\% \pm 5\%$) in the mammospheres are able to differentiate along the

third lineage, and are therefore multipotent, able to generate all three lineages present in the adult mammary gland (Fig. 4E–G). The distribution of different types of colonies obtained by plating cells derived from secondary mammospheres at clonogenic densities is shown in Figure 4H. Mixed-lineage colonies were considerably larger (5046 ± 2871 cells/colony) than single-lineage

colonies (190 ± 61 cells/colony), reflecting their higher proliferation potential.

Mammospheres contain progenitor cells able to clonally reproduce functional complex structures in vitro, which resemble the in vivo mammary tree

As an additional measure of differentiation potential, we developed a clonal assay that assesses the ability of single cells to form functional ductal-alveolar structures in Matrigel [Fig. 3A(iii)]. This 3D cultivation system enables us to test in addition the cells' ability to recapitulate the *in vivo* spatial orientation, and regenerate the architecture of the *in vivo* mammary tree. Previous studies showed that cultivation in Matrigel provides the physiological signals necessary for normal mammary morphogenesis *in vitro* (Simian et al. 2001). Two types of multicellular structures, clonally generated by primary human mammary cells, were reported in the literature: small acinus-like structures of luminal epithelial origin and solid spherical colonies of myoepithelial origin (Petersen et al. 1992). Complex, branched structures comprising both ductal and acinar morphology were obtained only when aggregates of cells, or organoids (small fragments of mammary epithelium), were implanted in Matrigel. Recently, an immortalized cell line of human mammary cells with bilineage potential was described. This cell line also has the ability to clonally generate ductal-acinar structures in Matrigel (Gudjonsson et al. 2002). The ability to develop functional alveolar cells, however, was not demonstrated. Single cells from dissociated mammospheres were plated at clonogenic densities (1000 cells/2.5 mL of gel) in 3D Matrigel culture [Fig. 3A(iii)]. After 3 wk of cultivation in Matrigel, they generated two morphologically distinct types of colonies: branched, ductal-acinar structures (Fig. 4I) or acinus-like structures (Fig. 4J). The growth was monitored microscopically daily, to ensure that each structure was generated by a single cell and that individual structures did not come in contact. The overall plating efficiency was 85%–92% with only ~ 2 –4/1000 of the structures showing combined ductal and acinar development, and all the rest acinar development. The acinar structures were significantly larger than those previously reported in the literature (500 μm compared with 100 μm) and appear to have a more complex structure. Markers for both epithelial and myoepithelial lineages were present in each of the ductal alveolar and acinar-like structures (Fig. 4K; data not shown). When prolactin was added to these cultures, β -casein was produced and secreted in the central lumen of the acinar structure, as demonstrated by immunostaining (Fig. 4L). This was also observed for the acinar part of the branched structures (data not shown). These data suggest that although the majority of the cells within the mammospheres have the potential to differentiate along all three lineages in both 2D and 3D culture, only a small number of the cells retain the intrinsic capacity to recapitulate the entire ductal-acinar architecture of the mammary tree.

Mammospheres contain cells capable of self-renewal

To test the self-renewal capacity of cells, we used a third *in vitro* assay that relies on assessing the sphere-initiation efficiency of serially passaged cells cultured as mammospheres (Fig. 3B). A similar assay has been previously used for the characterization of neural progenitors grown as neurospheres (Reynolds and Weiss 1996). Given that, as indicated below, cultivation of HMECs in suspension as mammospheres allows cells to remain undifferentiated, sphere formation efficiency in serially passaged cells can be used as a test of self-renewal. In this assay the cells' potential to differentiate along multiple lineages was also tested at each passage. Mammospheres were dissociated mechanically and enzymatically to single cells and placed in suspension culture, at 1000 cells/mL density. In parallel, for each passage, an aliquot of mammosphere-derived cells was plated on collagen, at clonogenic density, to assess their lineage differentiation potential. We performed five passages at 10–14-d intervals, with an efficiency of ~ 4 /1000 in new mammosphere formation at each passage. Most importantly, single cells from all passages, when plated on collagen, generated 98%–100% mixed lineage colonies, which demonstrates that they retained multilineage differentiation potential when cultured as mammospheres (Fig. 5A). Immunostaining of these colonies with pairs of markers specific for the ductal epithelial and myoepithelial lineages, identified as single-positive, single-negative, double-positive, and double-negative cells (Fig. 5B). The last two categories probably correspond to intermediate stages of development previously identified *in vitro* and *in vivo* (Gudjonsson et al. 2002).

To demonstrate that the majority of mammospheres are clonally derived in our experimental conditions, rather than being the result of cell aggregation, we used three independent assays.

(1) Cells from dissociated mammospheres were plated in 96-well plates at one cell per well density, using limiting dilution, verified microscopically. These single cells generated mammospheres with an efficiency of ~ 4 /1000, identical to that found with cells plated at low density. This demonstrates that (a) a single cell is able to generate a mammosphere; (b) aggregation and/or cell-cell contact are not required for sphere formation; and (c) the majority of spheres formed at 1000 cells/mL density are probably clones, because the efficiency of sphere formation is the same as at 1 cell/well density. Because the spheres contained ~ 150 –300 cells when they were passaged, and the efficiency of sphere formation was constant, 4/1000, one can estimate that each mammosphere contains approximately one sphere-initiating cell. To verify these results, we collected individual secondary mammospheres, under microscopic control, and dissociated them into single cells. The secondary mammospheres were generated in culture of 1000 cells/mL density. The single cells obtained after dissociation were plated at 1 cell/well density, by limiting dilution, verified microscopically. Approximately 25%–50% of the cells from one mammosphere were recovered after dis-

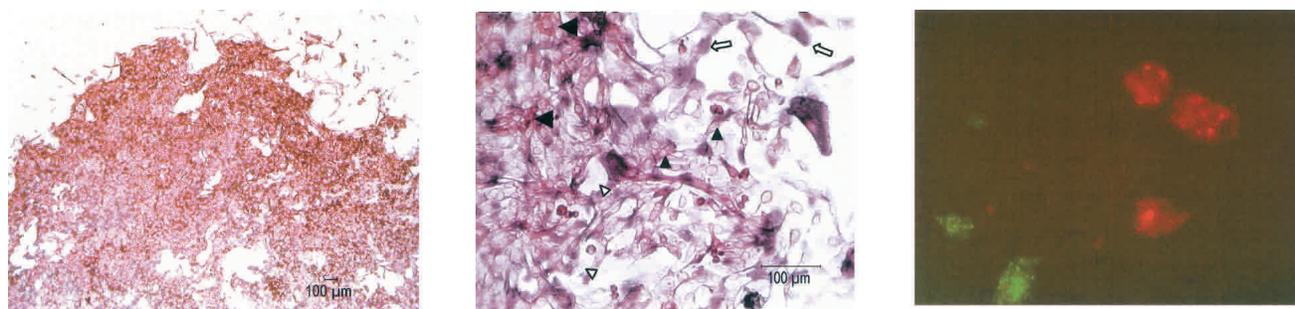
**A****B****C**

Figure 5. Secondary and later generation mammospheres contain multipotent progenitor cells. (A) Colony generated clonally by a secondary mammosphere-derived cell; immunostaining with ESA (red) and CD10 (purple). (B) Detail of the colony presented in A; ESA-positive cells (black arrows), CD10-positive cell (white arrows), cells expressing both ESA and CD10 (black arrowheads), and cells negative for both markers (white arrowheads). (C) Clonality of secondary mammospheres generated in mixed culture. HMECs grown as primary mammospheres were transfected with a retrovirus encoding for EGFP or DsRed2. Secondary mammospheres formed by the EGFP- and DsRed2 mammosphere-derived cells, mixed in equal ratio in suspension culture. The vast majority of the secondary mammospheres are monochromatic.

sociation. A total number of 54 spheres, derived from three reduction mammoplasties, were analyzed in this manner. The number of new spheres derived from one initial mammosphere varied between 1 and 4 (average 1.3 ± 0.8). Because almost all cells in these secondary mammospheres were bipotent and tripotent progenitors, one would expect the sphere-initiating cells to also be multipotent. To verify this, third-generation spheres, clonally generated, were transferred to collagen-coated dishes in differentiating conditions and allowed to attach and grow for 10 more days. The colonies thus generated were immunostained for lineage differentiation markers, as described in the previous experiments. All attached mammospheres generated mixed luminal epithelial-myoeptithelial colonies. Because these mammospheres were generated from single cells, one can conclude that the sphere-initiating cells are, indeed, multipotent.

(2) Further evidence that mammospheres were clonally derived was obtained by mixing cells obtained from mammoplasty samples from three different patients, in equal ratios (three different combinations, each containing equal number of cells from two samples). The three different samples were genotyped by PCR-based STR (short tandem repeats) amplification, using 16 different markers. Two markers that unequivocally distinguished each pair of samples were chosen and used to genotype individual mammospheres generated in the mixed cultures. Then 160 of the resulting single spheres were collected, genomic DNA was extracted, and PCR-based genotyping using two different markers was performed. The genotyping was successful in 125 cases, and showed 12 mixed and 113 nonmixed profiles, which corresponds to a calculated probability of a minimum 83% nonmixed spheres in the culture with a confidence of 99% (see Materials and Methods). One cannot exclude the possibility of collecting small numbers of dead cells or cellu-

lar debris containing DNA during the collection of individual spheres, which was then amplified by PCR. Thus, the numbers above might overstate the number of mixed spheres in the culture and therefore by this method, at least 83% of the mammospheres are genotypically homogeneous.

(3) To further demonstrate the clonal origin of mammospheres under our experimental conditions, we transfected cells with a retrovirus encoding for the red fluorescent protein (DsRed2) or enhanced green fluorescent protein (EGFP). The spheres generated by the transfected cells were then dissociated, mixed in equal ratios, and plated at 1000 cells/mL density. The vast majority of the resulting secondary spheres were monocolor expressing only DsRed2 or EGFP (Fig. 5C). Thus, by three independent assays, we demonstrated that the mammospheres are largely clonally derived, and their formation does not require cellular aggregation.

Multilineage progenitors and mammosphere initiating cells are contained in the SP population, capable of Hoechst dye exclusion

A characteristic shared by many adult stem cells, including those from hematopoietic and neural tissue, is the ability of these cells to exclude dyes such as rhodamine and Hoechst (Bunting 2002). This property identifies a small subset of cells termed the side population (SP) and is due to the expression of transporter proteins, such as BCRP and P-glycoproteins. The effect is blocked by Verapamil, a nonspecific inhibitor of membrane transport. It has been shown that the SP fraction of the hematopoietic and neural cells contains the long-term repopulating stem cells (Scharenberg et al. 2002). This has also recently been demonstrated for mouse mammary cells with SP properties, which can regenerate the gland

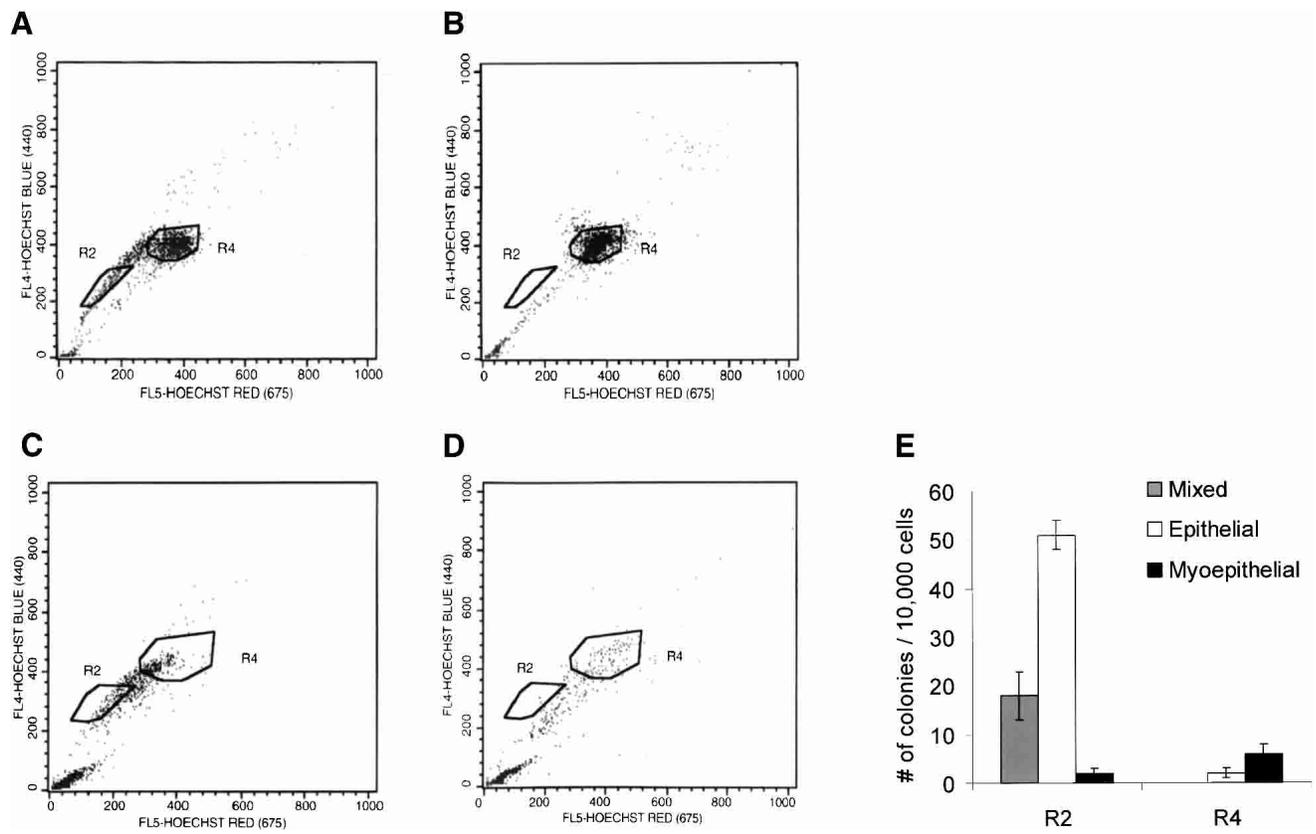


Figure 6. Hoechst 33342 staining of HMECs: uncultured cells from dissociated tissue (A,B) and mammosphere-derived cells (C,D). A low Hoechst staining population (SP) can be identified in both uncultured mammary cells (gated R2; A) and cells grown as mammospheres (gated R2; C). The majority of cells stain intensely with Hoechst (gated R4). (B,D) The exclusion of the Hoechst dye is blocked by Verapamil. Only the sorted SP cells (gated R2) generate mixed lineage colonies on collagen (E) and mammospheres in suspension culture (Table 2). Data are presented as means \pm S.E.M.

upon transplantation (Welm et al. 2002). Using flow cytometry, we identified an SP fraction in uncultured mammary cells, characterized by the ability to exclude Hoechst dye. This effect was blocked by Verapamil. The SP fraction represents \sim 1% of the total population of cells (Fig. 6A,B; Table 1a,b). In the cells cultured as mammospheres, the SP fraction represents 27% of the total population of cells (Fig. 6C,D; Table 1c,d). We compared the Hoechst low and high staining fractions (the SP and non-SP HMEC fractions) for their ability to clonally generate bilineage colonies when cultured under differentiating conditions. The results indicated that all the bipotent progenitors were present in the SP fraction (Fig. 6E). To determine if the ability of cells to form mammospheres correlates with enrichment in progenitor cells, the SP and non-SP staining population from uncultured cells were separated by FACS and plated in suspension culture. Only cells contained in the SP fraction were capable of mammosphere formation (Table 2). These results suggest that cells capable of multilineage differentiation and of mammosphere formation are present exclusively in the SP fraction. Furthermore, cultivation of cells as mammospheres enriches the progenitor-containing SP fraction by \sim 30-fold.

Transcriptional profile of mammosphere-derived cells

A major impediment to the isolation of human mammary stem cells is the lack of identified markers specific for undifferentiated cells. We demonstrated that at the second passage and beyond, HMECs cultivated as mammospheres contain virtually only multipotent cells. In contrast, when cells are plated on collagen, they terminally differentiate. Thus, these in vitro conditions might reflect the respective in vivo undifferentiated and differentiated cell stages in mammary development. To identify candidate markers of stem/progenitor cells, we compared the gene expression profile of undifferentiated cells grown as mammospheres to that of differentiated cells cultured on collagen, using Affymetrix microarray analysis. A similar approach, employing neurospheres, has been successfully used to establish the transcriptional profile of neural progenitors (Geschwind et al. 2001; Terskikh et al. 2001). Primary mammospheres were dissociated to single cells and divided in two aliquots that were replated in suspension culture and on collagen, in differentiating conditions, respectively. Total RNA was extracted from the second-generation spheres and from cells grown under differentiating conditions and used in

Table 1. *The SP fraction in mammosphere-derived cells and uncultivated HMECs*

Figure 6	R2 (%)	R4 (%)
a	0.82 ± 0.54	52.35 ± 2.64
b	0.06 ± 0.04	50.76 ± 11.72
c	27.46 ± 9.60	43.57 ± 10.95
d	2.45 ± 2.14	50.76 ± 11.72

Percentages correspond to the populations of cells gated R2 and R4, shown in Figure 6A–D.

Data are presented as means ± SEM.

the hybridization reaction against the HU133A Affymetrix chip. Confirmation of the expression pattern was performed for 14 genes, using real-time reverse transcriptase PCR (RT-PCR). Genes with different levels of expression in mammospheres and differentiated cells were chosen, as well as genes with similar levels of expression, for confirmation by real-time RT-PCR. Table 3 shows the changes in level of expression detected by microarray analysis in comparison with real-time PCR. As previously suggested by Yuen et al. (2002), who compared Affymetrix chips data with real-time RT-PCR data, the microarray analysis is accurate and reliable. We also selected three genes with known involvement in developmental processes, which were assigned an absent call in the microarray analysis. These transcripts for Lunatic Fringe, Manic Fringe, and Notch 4 were detected after 30 cycles of amplification in both mammosphere and differentiated cell samples in the real-time RT-PCR reaction, indicating that their level was below the detection level of the microarray analysis.

A total number of 687 probe sets detected transcripts enriched more than twofold in the mammosphere-derived cells compared with cells grown in differentiating conditions, 79 of which were assigned an absent call in the differentiated sample (Supplementary Table S4). Candidate markers for stem and progenitor cells are included in this category. Examples of genes expressed exclusively in the mammospheres are: low-density lipoprotein-related protein 1, glypican 4, frizzled homolog 2, interleukin 6, CXCR4, and fibroblast growth factor receptor 1. We also identified 81 genes expressed exclusively in mammosphere-derived cells grown in differentiating conditions, which represent negative markers for undifferentiated cells (Supplementary Table S5). Aside from these genes that are turned on or off, the ones with very high differences in expression levels could also be potential markers for different stages of development. The genes with at least twofold differential expression comprised 103 surface or membrane-associated molecules, 59 transcription factors, 46 kinases, 23 phosphatases, 20 cytoskeleton components, and the like. Examples of genes in these categories are presented in Table 4 and Supplementary Table S6, which contain genes up-regulated in mammosphere-derived cells, and Table 5 and Supplementary Table S7, which contain genes up-regulated in cells grown under differentiating conditions.

In agreement with the immunohistochemical data presented above, matrix proteins associated with early developmental processes, such as tenascin, fibulin 1, 2, 5, and decorin, had high level of expression in mammosphere cells and were absent or down-regulated in differentiated cells, whereas the components of ECM in adult tissues, such as laminin or collagen IV, had lower levels of expression or were absent in mammospheres.

Discussion

Our study reports for the first time the *in vitro* cultivation and propagation of human mammary epithelial cells in an undifferentiated state. Contrary to the dogma that epithelial cell survival is anchorage-dependent, we find that a small population of HMECs survives in suspension and generates floating spherical colonies, that we termed nonadherent mammospheres. As shown above, either the sphere-initiating cells or their undifferentiated progeny secrete and deposit ECM proteins that contribute to the composition of the mammospheres. Thus, although the sphere-initiating cells survive and proliferate in an anchorage-independent manner, the other cells in the mammosphere may depend on the presence of the ECM components for their survival. Most of the ECM components detected in mammospheres are associated with embryonic development. The ability of undifferentiated mammary cells to survive in suspension was first demonstrated by Soule and McGrath (1986), who showed that in low-calcium medium, a subpopulation of HMECs detach and survive over many passages as free-floating cells. The authors speculated that low-calcium conditions, by inhibiting terminal differentiation, select for a population of cell with stem cell properties.

Mammospheres are highly enriched in undifferentiated cells, as demonstrated by the ability of single cells, isolated from mammospheres, to generate multilineage colonies when cultured under differentiating conditions. Whereas freshly isolated mammary cells contain only 8% multilineage progenitors, these are enriched to ~68% in primary mammospheres, and virtually 100% in secondary and later-passage mammospheres. As is the case with neurospheres, mammosphere formation is facilitated by serum-free conditions and requires the presence of either EGF or bFGF. Mammospheres are able to form at clonal densities and do not require cell aggregation, as shown by limiting dilution, genotyping, and cell-mark-

Table 2. *Efficiency of mammosphere formation in the low and high Hoechst-staining populations*

Exp.	R2 (no. of spheres/10,000 cells)	R4 (no. of spheres/10,000 cells)
1	24	0
2	17	0
3	20	0
4	10	0

Breast reduction samples from four different patients were used in separate experiments.

Table 3. Comparison between changes in expression level detected by microarray analysis and real-time RT-PCR

Transcript	Change detected by microarray analysis (fold change)	Change detected by real time RT-PCR (fold change)
Up-regulated in mammospheres		
Stem cell growth factor	4.0	9.2
HGF antagonist	149.1	132.5
Stem cell growth factor β	12.9	26.5
Apolipoprotein E	26.9	16.9
Up-regulated in differentiated cells		
Parathyroid hormone like hormone	5.5	3.6
TGF α	6.2	2.4
Mammaglobin	7.4	3.2
Neuregulin 1	Absent in mammospheres, present in differentiated cells	Absent in mammospheres, present in differentiated cells
Similar expression levels (mammospheres vs. differentiated cells)		
LIF receptor	1.3	1.1
LIF	1.8	2.4
Radical fringe	1.7	4.2
Not detected by microarrays (mammospheres vs. differentiated cells)		
Lunatic fringe	Absent in both	1.15
Manic fringe	Absent in both	2.08
Notch 4	Absent in both	4.1

The fold change detected by microarrays represents the average of the two duplicate samples.

ing experiments. It has previously been demonstrated that neurospheres are generated by neural stem cells, which comprise between 4% and 20% of the mature neurospheres, the rest of the cells representing progenitor cells in various stages of differentiation (Scheffle et al. 1999). Our results are consistent with a similar model in which the mammosphere-forming cells represent mammary stem cells that undergo limited self-renewal and then give rise to mammary progenitors, still capable of multilineage differentiation. Stem cells are the only cells capable of self-renewal; all other cells differentiate while they divide and then arrest or undergo terminal differentiation after a number of cell divisions. Consequently, if mammosphere-initiating cells were early multipotent progenitors rather than true stem cells, the frequency of mammosphere formation should decrease with each passage. In addition, the size of mammospheres should decrease with each passage, as cells become more differentiated. Similarly, if stem cells are capable of generating mammospheres but are incapable of self-renewal, the number of mammospheres should decrease with serial passaging. Mammospheres contain ~150–300 cells after 7 d in culture, regardless of passage number, and the efficiency of sphere formation was constant, ~4/1000 cells through five generations. These results, indicating that each mammosphere contains on average one sphere-initiating cell, were indeed supported by the passage of single mammospheres. The frequency of sphere initiation remains constant with passaging, suggesting that the sphere-initiating cells may be stem cells. Also, the frequency of cells capable of producing complex ductal-acinar structures in 3D Matrigel cultures was the same as the frequency of sphere-initiating

cells. Whether they represent the same cell population remains to be proved, but according to the data presented, both can be used as a read-out for mammary stem cell enrichment. Our data do not rule out the possibility of in vitro dedifferentiation of cells when cultivated in suspension. Nevertheless, because mammospheres are enriched in stem/progenitor cells, they represent a valuable system for the study of normal mammary development.

In agreement with previous findings, the immunostaining of entire mammospheres suggested that antigens associated with differentiated cells of specific lineages might also be expressed in undifferentiated cells (Stingl et al. 1998, 2001; Gudjonsson et al. 2002). The pattern of expression within mammospheres is consistent with CD10, α 6 integrin, and cytokeratin 5 being present in earlier progenitors, and ESA and cytokeratin 14, on later progenitors, which are still multipotent. Muc 1, cytokeratin 18, and ASMA, are not expressed in the mammospheres, even after longer cultivation (up to 15 d in suspension culture), but they are expressed by mammosphere-derived cells transferred to collagen substrata. This suggests that these antigens are expressed only on more differentiated cells of ductal epithelial lineage (Muc 1, cytokeratin 18) and myoepithelial lineage (ASMA). These results are also consistent with previously reported data (Stingl et al. 1998, 2001; Gudjonsson et al. 2002).

A characteristic shared by many adult stem cells is increased membrane transporter activity, which can be assayed by Hoechst dye exclusion. Although this may represent only a marker of stem cells (Zhou et al. 2001), several studies suggest that the functional significance of

Table 4. *Genes up-regulated in cells grown as mammospheres*

Category	Name and abbreviation	Accession number	Function	Fold increase	Fold increase mean \pm SDEV	Expressed in other stem cells
Membrane proteins	Epithelial membrane protein 3 (EMP3)	NM_001425	Negative control of cell proliferation	6	6.1 \pm 1	
	Glypican 4 (GPC4)	NM_001448	Embryogenesis and morphogenesis	7	7.1 \pm 3.5	c
	Frizzled homolog 2 (FZD2)	NM_001466	Establishment of tissue polarity	6	6.0 \pm 2	
	Growth hormone receptor (GHR)	NM_000163	Skeletal development, growth pattern	6	5.7 \pm 5	a, b, c
	Insulin-like growth factor 2 receptor (IGF2R)	NM_000876	Embryogenesis and morphogenesis	3	3 \pm 0.3	
	Notch 3 (NOTCH3)	NM_000435	Growth patterning	3	2.7 \pm 0.5	b, c
	Fibroblast growth factor receptor 1 (FGFR1)	M60485	Embryogenesis and morphogenesis	5	4.7 \pm 2	a, b
	Platelet-derived growth factor receptor, β polypeptide (PDGF- β R)	NM_002609	Skeletal development	15	15 \pm 10.7	a, b, c
	ATP-binding cassette, subfamily A (ABC1), member 1	NM_005502	ATP-binding cassette (ABC) transporter	4	4.3 \pm 3	a, b
	Transcription factors	AE binding protein 1 (AEBPI)	NM_001129	Muscle, skeletal development	27	26.8 \pm 22
Growth arrest specific 6 (GAS6)		NM_005890	Cell cycle arrest	4	4 \pm 0.5	
p53 target zinc finger protein (WIG1)		NM_022470		3	2.9 \pm 1	a, b, c
Signaling	Small inducible-cytokine A2 (SCYA2)	S69738	Negative control of cell proliferation	61	60.6 \pm 34	
	Glycoprotein nmb (GPNMB)	NM_002510	Negative control of cell proliferation	26	25.8 \pm 8.7	
	Insulin-like growth factor binding protein 4 and 7 (IGFBP4 and 7)	NM_001552 NM_001553	Cell proliferation	9 and 6	8.5 \pm 6 and 5.6 \pm 2	c
	Stem cell growth factor (SCGF)	D86586	Positive control of cell proliferation	4	3.8 \pm 2	
	Retinotic acid receptor responder 3 (tetrazone induced) (RARRES)	NM_004585	Differentiation growth inhibition	38	38 \pm 2	
	Growth-associated protein 43 (GAP43)	NM_002045	Response to wounding, neurogenesis	24	23.8 \pm 6	
	p59fyn (FYN) oncogene	S74774	Calcium influx regulator, related to SRC, YES, FGR	4	3.9 \pm 0.2	a, b
Cytoskeleton	Wiskott-Aldrich syndrome protein interacting protein (WASPIP)	NM_003387	Cell shape and cell size control	5	5.1 \pm 3	
	Polycystic kidney disease 2 (PKD2)	NM_000297	Cation transport, organogenesis	5	4.6 \pm 1.5	a, b, c
Adhesion	Osteoblast specific factor 2 (OSF2)	D13665	Skeletal development	6	5.5 \pm 1.3	
	Tenascin C, (HXB) Nidogen and nidogen 2 (NID and NID2)	NM_002160	Cytotactin Basement membrane, cell adhesion	5 3 and 7	5 \pm 0.5 2.6 \pm 0.3 and 6.7 \pm 4.5	b, c c
Cell cycle	Cyclin D2 (CCND2)	NM_001759	Cell cycle control	2	1.9 \pm 0.3	

Table 4 continued on next page

Table 4. *Continued*

Category	Name and abbreviation	Accession number	Function	Fold increase	Fold increase mean \pm SDEV	Expressed in other stem cells
Metalloproteinases	Tissue inhibitor of metalloproteinase 1 and 3 (TIMP3)	NM_000362	Neurogenesis	3	3 \pm 1.4	a, c
	Disintegrin and metalloproteinase domain 12 (ADAM12), transcript variant 1, mRNA	NM_003474	Metalloproteinase, myocyte fusion	7	6.5 \pm 0	b
	Four and a half LIM domains 1	NM_001449		4	4.3 \pm 3	a, b, c
	Prostate tumor over expressed gene 1 (PTOV1)	AL046054		2	2.3 \pm 0	a
	Apolipoprotein E	NM_000041	Lipid transporter	61	60.6 \pm 34	a
	Decorin variant C (DCN)	AF138302	Histogenesis and organogenesis	8	7.9 \pm 4	c

The fold increase in expression level represents the average between the results of two independent experiments, utilizing cells derived from two different mammaplasty samples. Means \pm SDEV (standard deviation) are shown in a separate column. Genes enriched in other stem cells are indicated in the last column: (a) hematopoietic stem cells, (b) neurospheres, (c) embryonic stem cells (Ramalho-Santos et al. 2002).

increased transporter activity might extend beyond the obvious role in defense against damaging agents, to regulating the commitment of stem cells to a particular fate (Good and Kuspa 2000; Bunting 2002). This characteristic was used for identifying and sorting the most primitive subset of stem/progenitor cells from a variety of tissues including the mouse mammary gland (Storms et al. 2000; Welm et al. 2002). We identified a subpopulation of human mammary epithelial cells with SP properties representing 1% of the total population, and showed that this population contains all of the multilineage progenitor cells as well as all of the cells capable of mammosphere formation. Furthermore, the SP population is enriched 30-fold in mammospheres.

An important goal of the transcriptional profiling was to identify candidate markers for undifferentiated cells, which could be used for the prospective isolation of mammary stem cells. We identified genes expressed exclusively in the mammosphere-derived cells, grown in undifferentiating conditions, that represent the best candidates for stem/progenitor cell markers. Some of these molecules are indeed involved in stem/progenitor cell-specific functions, such as homing of hematopoietic stem cells (CXCR4), maintaining the undifferentiated state of embryonic stem cells (IL6 and gp130, which functions as a receptor for IL6 and LIF as a homodimer or heterodimer with gp190), and regulation of self-renewal (members of the Wnt/Frizzled pathway; Zandstra et al. 2000; Reya et al. 2001; Lapidot and Kollet 2002). Interestingly, abnormal expression of some of these genes is also correlated with breast cancer or/and other types of cancer, suggesting that dysregulation of stem/progenitor cell functionality may be involved in carcinogenesis (Dhingra et al. 1998; Brown 2001; Geminder et al. 2001; Muller et al. 2001; Reya et al. 2001; Taipale and Beachy

2001). We also identified genes expressed exclusively in mammosphere-derived cells grown in differentiating conditions, which could be used as negative markers for progenitor cells (Supplementary Table S5).

The differences in expression profiles identified in our studies strongly support the validity of this experimental system in reflecting the normal process of mammary development and differentiation. Molecules previously shown to be involved in normal breast development such as parathyroid-like hormone, TGF α , HGF, and members of the Notch and Wnt families were identified in our screening (Uyttendaele et al. 1998; Watson et al. 1999; Hennighausen and Robinson 2001; Kleer et al. 2002). Genes specific to mammary tissue, like mamoglobin 1 and 2, WISP3, or more universally expressed ones, like adhesion molecules, tight junction proteins, and metalloproteases, showed the expected higher level of expression in cells cultured in differentiating conditions (Watson et al. 1999; Kleer et al. 2002).

Recently, two major studies have attempted to define a common genetic program of stem cells by comparing the transcriptional profile of embryonic stem cells with that of adult neural and hematopoietic stem cells (Ivanova et al. 2002; Ramalho-Santos et al. 2002). Both studies used neurosphere-derived cells to determine genes up-regulated in neural stem cells. Because the experimental procedure they used is almost identical to that used by us, we compared our data to their findings. The mouse orthologs corresponding to the genes up-regulated in mammospheres were compared with the genes up-regulated in embryonic, neural, and hematopoietic stem cells, and in all three types of stem cells (Ramalho-Santos et al. 2002). Examples of genes with overlapping expression profiles in these four categories are presented in Table 4. A total of 335 probe sets detected transcripts

Table 5. Genes up-regulated in cells grown in differentiating conditions

Category	Name and abbreviation	Accession number	Function	Fold increase	Fold increase mean \pm SD
Membrane proteins	Met proto-oncogene (hepatocyte growth factor receptor) (MET)	NM_000245	Cell proliferation	5	5.3 \pm 0.5
	Frizzled homolog 6 (FZD6)	NM_003506	Establishment of tissue polarity	3	2.7 \pm 0.4
	Fibronectin leucine-rich transmembrane protein 3 (FLRT3)	MN_013281	Cell adhesion	5	5.4 \pm 2
Transcription factors	GATA binding protein 3 (GATA3)	NM_002051	Embryogenesis and morphogenesis	11	10.8 \pm 3
	E74-like factor 3 (ets domain transcription factor, epithelial specific) (ELF3)	AF017307	Developmental processes	2	2.4 \pm 0.6
Signaling	Pituitary tumor-transforming 1 (PTTG1)	NM_004219	Mitotic regulation	4	4.0 \pm 0
	Parathyroid hormone-like hormone (PTH LH)	NM_002820	Epidermal differentiation	6	5.6 \pm 2
	Serine/threonine kinase 17a (STK17A)	NM_004760	Induction of apoptosis	3	3.1 \pm 0.2
	Insulin receptor substrate 1 (IRS1)	NM_005544	Metabolic signaling	4	3.9 \pm 2
	Secreted frizzled-related protein 1 (SFRP1)	NM_003012	Embryogenesis and morphogenesis	4	4.3 \pm 0.4
Cytoskeleton	Epithelial V-like antigen (EVA1)	NM_005797	Embryogenesis and morphogenesis	7	7.2 \pm 3
Cell-cell and cell-matrix adhesion	Desmoglein 2 and 3 (DSG2 and 3)	NM_001943 NM_001944	Intercellular junction	4 and 2	4.3 \pm 0 and 2.3 \pm 0.8
	Desmocollin 2 and 3 (DSC2 and 3)	NM_004949 NM_001941	Intercellular junction	4 and 4	4.3 \pm 1 and 3.6 \pm 0.2
	Cyclin-dependent kinase inhibitor 3 (CDKN3)	NM_005192	Cell cycle control	3	2.7 \pm 0.1

The fold increase expression level represents the average between the results of two independent experiments, using cells derived from two different mammary samples. Means \pm SD (standard deviation) are shown in a separate column.

enriched in both neurospheres and mammospheres, 367 in embryonic stem cells and mammospheres, and 334 in hematopoietic stem cells and mammospheres (Supplementary Table S8). The global overlap between genes expressed in all three stem cells and mammospheres includes 10 up-regulated genes, among which are the ones encoding for the growth hormone receptor, the thrombin receptor, the polycystic kidney disease 2, and four and a half LIM domains 1. The authors of these studies proposed several attributes of "stemness" that we also found to be present in the mammosphere-derived cells, by microarray analysis. These include active Jak/Stat, TGF β , Notch signaling, capacity to sense growth hormone and thrombin, increased transporter activity, high resistance to stress with up-regulated DNA repair, and protein folding (Ramalho-Santos et al. 2002). Consistent with these results, we found that growth hormone receptor and thrombin receptor were up-regulated in mammospheres fivefold and eightfold, respectively. Among members of the Notch family, Notch 3 and Radical

Fringe were found to be up-regulated. Quantitative RT-PCR showed that although changes in the expression of other genes in this family occurred (Notch 1, Notch 4, the ligand Delta, and the Notch modulator Manic fringe), the level of expression in both differentiated and undifferentiated cells was below the limit of detection on the microarray (Delta is not represented on the HGU133A array). Membrane transporters were up-regulated in mammospheres compared with differentiated cells, and several metalloproteases in the ADAM family were differentially expressed. Thus, despite limitations in gene expression comparison across different species and different experimental conditions, we detected a high degree of overlap between genes expressed in mammospheres and those expressed in hematopoietic, neuronal, and embryonic stem cells.

In conclusion, the culture of human mammary epithelial cells as mammospheres should prove valuable for the isolation and characterization of mammary stem cells and early progenitors, and for the study of the pathways

involved in normal growth, differentiation, and lineage-specific commitment of human mammary epithelial cells. This system may also be amenable to the addition of stromal elements in order to study stromal–stem cell interactions involved in mammary cell commitment and differentiation. Furthermore, because mammary stem cells may be a target for cellular transformation, these systems may prove useful in studying mammary carcinogenesis. Indeed, we have recently identified a tumorigenic population of cells in human breast cancers with “stem cell” properties (Al-Hajj et al. 2003). The ability of normal neural and mammary stem cells, as well as embryonic stem cells, to proliferate in suspension suggests that stem cells may have the common property to survive and proliferate in the absence of exogenous substrata. These characteristics are also consistent with the lack of polarity of putative mammary stem cells identified by Smith and Chepko (2001), who also reported that these cells lacked membrane contacts with adjacent cells. In vivo, these characteristics could contribute to the stem cell “niche,” in which signals that dictate self-renewal and differentiation control the fate of stem cells. If this, indeed, is the case, the methodology used for mammosphere production may prove useful in isolating and characterizing stem cells from other organs.

Materials and methods

Dissociation of mammary tissue

Normal breast tissue from reduction mammoplasties was dissociated mechanically and enzymatically, as previously described (Stingl et al. 1998). Cells were sieved sequentially through a 100- μ m and a 40- μ m cell strainer (Falcon) to obtain a single cell suspension. An aliquot of the single cell suspension was loaded in the hemocytometer, and doublets, triplets, and groups of higher numbers of cells were counted. A minimum of 1000 cells or groups of cells was counted. The number of single cells was >99% in all experiments. The cell suspension was resieved in the cases in which it was lower (usually 95%–99%).

Mammosphere culture

Single cells were plated in ultralow attachment plates (Corning) at a density of 20,000 viable cells/mL in primary culture and 1000 cells/mL in passages. Cells were grown in a serum-free mammary epithelial growth medium (MEGM, BioWhittaker), supplemented with B27 (Invitrogen), 20 ng/mL EGF and 20 ng/mL bFGF (BD Biosciences), and 4 μ g/mL heparin (Sigma). Bovine pituitary extract was excluded. Mammospheres were collected by gentle centrifugation (800 rpm) after 7–10 d and dissociated enzymatically (10 min in 0.05% trypsin, 0.53 mM EDTA-4Na, Invitrogen) and mechanically, using a fire-polished Pasteur pipette. The cells obtained from dissociation were sieved through a 40- μ m sieve and analyzed microscopically for single-cellularity. Typically, the number of doublets, triplets, and groups of cells was between 30 and 150 for 10,000 single cells. If groups of cells were present at a frequency >100 for 10,000 single cells, mechanical dissociation and sieving were repeated. For the experiments regarding differentiation on collagen and in Matrigel, colonies were monitored microscopically daily to ensure that

they were derived from single cells and that they did not become confluent during the experiment.

Differentiating culture conditions

Single cell suspensions were plated on collagen-coated plates at a density of 2000 viable cells/10-cm-diameter dish. Cells were grown in Ham's F-12 medium (BioWhittaker) with 5% FBS, 5 μ g/mL insulin, 1 μ g/mL hydrocortisone, 10 μ g/mL cholera toxin (Sigma), 10 ng/mL EGF (BD Biosciences), and 1 \times Pen/Strep/Fungizone Mix (BioWhittaker). After 5 d, a layer of Matrigel (BD Biosciences) was added as well as prolactin (Sigma) at 1 μ g/mL concentration in the assays testing for alveolar differentiation. 3D cultures were performed as previously described (Weaver and Bissell 1999). The complex acinar and branched-acinar structures formed in Matrigel only if cells were allowed first to attach for 30 min on a substratum before plating. We used laminin, collagen type I, fibronectin, and tissue-culture-treated plastic, with equivalent results. Attachment to a substratum before suspension in Matrigel appears to be necessary for the activation of permissive pathways involved in morphogenesis.

Immunohistochemistry

To assess lineage composition of the colonies, cells were fixed in the plates for 20 min in methanol, at -20°C , then stained using Peroxidase Histostain-Plus and Alkaline-phosphatase Histostain-Plus kits (Zymed), according to the manufacturer's protocol. Immunostaining of the mammospheres was performed in suspension, after fixation in methanol–acetone 1:1, then following the same protocol as above. The immunostaining of ECM proteins was performed on mammospheres fixed in neutral formaldehyde; then they were embedded in paraffin and sectioned with the microtome at 3 μ m thickness. The acinar and branched-acinar structures were recovered from Matrigel by incubation for 3–4 h in Matrisperse (BD Biosciences), followed by fixation in acetone–methanol and staining in suspension. The primary antibodies, ESA and cytokeratin 18 for epithelial cells, CD10 and cytokeratin 14 (Novocastra) for myoepithelial cells, and the casein antibody (NeoMarkers) for alveolar cells were used at the dilutions indicated by the manufacturer. The staining with laminin, collagen IV antibodies (NeoMarkers), fibronectin, tenascin (Chemicon), and decorin (R&D) antibodies was performed using antibody dilution and antigen retrieval techniques indicated by the manufacturers. AEC and DAB (Zymed) were used as substrates for peroxidase and NBT/BCIP (GIBCO-BRL) for alkaline phosphatase. For fluorescent microscopy, FITC-labeled and Texas-red-labeled secondary antibodies (Zymed) were used.

Hoechst 33342 staining and FACS analysis

The Hoechst staining was performed as previously described (Storms et al. 2000). Briefly, cells from dissociated mammospheres or from dissociated tissue were sieved through a 40- μ m cell strainer (Falcon), then the cells were resuspended in medium (100,000 cells/mL) and allowed to recover at 37°C for 1 h before treatment with Hoechst. The same medium as for differentiating on collagen was used. Hoechst 33342 (Sigma) was added at a final concentration of 2.5 μ g/mL and incubated at 37°C for 1 h. In the control reactions, Verapamil (Sigma) was added at a final concentration of 50 μ g/mL. After incubation, cells were washed once with HBSS (BioWhittaker), and resuspended in HBSS supplemented with 5% FBS. Cells were stained with 1 μ g/mL propidium iodide (Sigma) for viability, then ana-

lyzed and sorted by flow cytometry using FACStar^{PLUS} (Becton Dickinson).

Analysis of mammospheres by genotyping

DNA from three different mammaplasty samples was isolated using the Wizard SV Genomic DNA Purification System (Promega), as described by the manufacturer. The samples were used for genotyping by PCR amplification of STRs using 16 markers (D1S423, D11S4151, D10S1652, D12S364, D5S419, D8S549, D13S219, D14S1007, D6S289, D9S1853, D17S849, DXS991, D18S1147, D163046, D22S423, D15S878). A pair of primers that distinguished each two samples was chosen for further genotyping of individual mammospheres (D18S1147 and D22S423, D18S1147 and D163046, D18S1147 and D15S878). Single cells obtained from the different reduction mammaplasties were mixed at a ratio of 1:1 and plated on ultralow attachment plates at a density of 1000 viable cells/mL. Cells from the three reduction mammaplasties were mixed in this manner to obtain three different combinations, each containing cells from two different samples. Cells were grown in suspension for 7 d, and the mammospheres were collected, washed, and resuspended in a diluted volume of fresh medium. Individual mammospheres were collected microscopically into individual tubes. A total of 160 mammospheres were collected from the combined cultures. DNA was isolated using the Wizard SV Genomic DNA Purification System (Promega) as described by the manufacturer. A 3700 Sequencer (Applied Biosystems) was used for the analysis of the amplified STRs. The probability of the representation $113:12$ homogeneous versus mixed spheres, $p(x)$, where x takes values between 0 and 1, was calculated using the formula

$$A = \int_0^1 C_{125}^{113} x^{113} (1-x)^{12} dx$$

and imposing the condition

$$\frac{A_1}{A} = \frac{1-P}{1}.$$

For $P = 0.99$, $t = 0.8266$ (i.e., the probability that the homogeneous spheres represent a minimum of 82.66% of the total population is 99%).

Retrovirus-mediated transfection and expression of EGFP and DsRed2 in HMECs

To transfect the cells with a fluorescent protein, we used a retroviral delivery system. Retroviral vectors pLEGFP-C1 and pDsRed2-C1 (Clontech) were transfected in the PT67 packaging cell line (Clontech) with the CalPhos Mammalian Transfection Kit (Clontech). The supernatant was harvested after 3 d starting 48 h after transfection, filtered through a 48- μ m syringe filter, and used to infect HMECs in primary culture. Single cells were plated at a density of 2000 cells/mL in standard conditions for suspension culture, adding 1 mL of retrovirus-containing supernatant for each 3 mL of medium. Polybrene was added to a final concentration of 8 μ g/mL. The medium was changed after 24 h of incubation. After 12 d of cultivation in suspension culture, mammospheres expressing EGFP and DsRed2 were collected, dissociated to single cells, mixed at a ratio of 1:1, and plated in suspension at 2000 cells/mL density. After 10 d of cultivation, mammospheres were observed and scored.

RNA isolation, amplification, and labeling for microarray analysis

Mammospheres generated after 10 d of initial plating were dissociated enzymatically and mechanically to yield single cells.

An aliquot of the cells was cultivated in suspension, and another one was plated on collagen cultured for 5 d and then overlaid with Matrigel. After 5 d, total RNA was isolated using RNeasy Mini Kit (QIAGEN). A linear amplification protocol, consisting of two rounds of double-strand cDNA synthesis and in vitro transcription, was used, according to Affymetrix protocol. The Enzo BioArray High Yield RNA Transcript Labeling Kit (Affymetrix) was used for labeling. The labeled cRNA (10 μ g total) was fragmented and used in the hybridization reaction to GeneChip probe array HG U133 A according to the manufacturer's protocol. The experiment was performed in duplicate using cells derived from two different breast reduction samples. Only the genes that were up- or down-regulated more than two-fold in both experiments were taken into consideration. The correlation factor between the expression level of individual transcripts in experiments using amplified versus nonamplified RNA was 0.93 for one patient and 0.92 for the other. The correlation coefficient for the two replicates was 0.96 for both mammosphere and differentiated cell replicates, underscoring the reproducibility of the results. The MAS 5.0 software from Affymetrix was used for statistical analysis. Out of 22,247 probe sets, 7721 (34.7%) were assigned a present call in both samples from mammospheres and 8386 (38%) in both samples from cells grown in differentiating conditions, which shows a good RNA integrity. The genes that were up- or down-regulated more than twofold in both replicates were categorized using NetAffx.com and NCBI databases. To compare our data with those presented by the studies on overlapping transcriptional profile of stem cells, we identified mouse-human pairs of orthologs by using NetAffx.com, which establishes a link between set probes on the HG U133 human array and the mouse U74AV2. The comparison between mouse and human orthologs up-regulated in hematopoietic stem cells, in the same experimental system, showed a 39% overlap, representing the stem molecular traits conserved between the two species (Ivanova et al. 2002).

Real-time RT-PCR

Real-time PCR was performed on a SmartCycler (Cepheid), using the QuantiTect Sybr Green RT-PCR kit (QIAGEN) with pairs of primers specific for each transcript (sequences available upon request) and following the manufacturer's protocol. The ratio between expression levels in the two samples was calculated by relative quantification, using β -actin as a reference transcript for normalization.

Acknowledgments

Thanks are due to Dr. Thomas Giordano for tissue procurement, Dr. Robert Lyons and the DNA sequencing core for genotyping analysis, the University of Michigan Cancer Center flow cytometry core, and Dr. Sean Morrison for technical advice and critical review of the paper. This work was supported by NIH grant CA075136 and in part by the University of Michigan's Cancer Center NIH Support Grant (5 P 30 CA46592).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. 2003. Prospective identification of tumori-

- genic breast cancer cells. *Proc. Natl. Acad. Sci.* **100**: 3983–3988.
- Alvi, A., Clayton, H., Joshi, C., Enver, T., Ashworth, A., Vivanco, M.M., Dale, T., and Smalley, M. 2002. Functional and molecular characterisation of mammary side population cells. *Breast Cancer Res.* **5**: E1.
- Blau, H.M., Brazelton, T.R., and Weimann, J.M. 2001. The evolving concept of a stem cell: Entity or function? *Cell* **105**: 829–841.
- Bocker, W., Moll, R., Poremba, C., Holland, R., Van Diest, P., Dervan, P., Burger, H., Wai, D., Ina, D.R., Brandt, B., et al. 2002. Common adult stem cells in the human breast give rise to glandular and myoepithelial cell lineages: A new cell biological concept. *Lab. Invest.* **8**: 737–746.
- Brown, A.M.C. 2001. Wnt signaling in breast cancer: Have we come full circle? *Breast Cancer Res.* **3**: 351–355.
- Bunting, K.D. 2002. ABC transporters as phenotypic markers and functional regulators of stem cells. *Stem Cells* **20**: 11–20.
- Daniel, C., Young, L., Medina, D., and DeOme, K. 1971. The influence of mammogenic hormones on serially transplanted mouse mammary gland. *Exp. Gerontology* **6**: 95–101.
- Dhingra, K., Sahin, A., Emami, K., Hortobagyi, G., and Estrov, Z. 1998. Expression of leukemia inhibitory factor and its receptor in breast cancer: A potential autocrine and paracrine growth regulatory mechanism. *Breast Cancer Res. Treat.* **48**: 165–174.
- Geminder, H., Sagi-Assif, O., Goldberg, L., Meshel, T., Rechavi, G., Witz, I., and Ben-Baruch, A. 2001. A possible role for CXCR4 and its ligand, the CXC chemokine stromal cell-derived factor-1, in the development of bone marrow metastases in neuroblastoma. *J. Immunology* **167**: 4747–4757.
- Geschwind, D.H., Ou, J., Easterday, M.C., Dougherty, J.D., Jackson, R.L., Chen, Z., Antoine, H., Terskikh, A., Weissman, I.L., Nelson, S.F., et al. 2001. A genetic analysis of neural progenitor differentiation. *Neuron* **29**: 325–339.
- Good, J.R. and Kuspa, A. 2000. Evidence that a cell-type-specific efflux pump regulates cell differentiation in *Dictyostelium*. *Dev. Biol.* **220**: 53–61.
- Gudjonsson, T., Villadsen, R., Nielsen, H.L., Rønnov-Jessen, L., Bissell, M.J., and Petersen, O.W. 2002. Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties. *Genes & Dev.* **16**: 693–706.
- Hennighausen, L. and Robinson, G.W. 2001. Signaling pathways in mammary gland development. *Dev. Cell* **1**: 467–475.
- Ivanova, N.B., Dimos, J.T., Schaniel, C., Hackney, J.A., Moore, K.A., and Lemischka, I.R. 2002. A stem cell molecular signature. *Science* **298**: 601–604.
- Jones, F. and Jones, P. 2000. The tenascin family of ECM glycoproteins: structure, function, and regulation during embryonic development and tissue remodeling. *Dev. Dynamics* **218**: 235–259.
- Kim, N.D., Oberley, T.D., Yasukawa-Barnes, J., and Clifton, K.H. 2000. Stem cell characteristics of transplanted rat mammary clonogens. *Exp. Cell Res.* **260**: 146–159.
- Kleer, C.G., Zhang, Y., Pan, Q., van Golen, K.L., Wu, Z.-F., Livant, D., and Merjaver, S.D. 2002. WISP3 is a novel tumor suppressor gene of inflammatory breast cancer. *Oncogene* **21**: 3172–3180.
- Kordon, E.C. and Smith, G.H. 1998. An entire functional mammary gland may comprise the progeny from a single cell. *Development* **125**: 1921–1930.
- Kresse, H. and Schonherr, E. 2001. Proteoglycans of the extracellular matrix and growth control. *J. Cell. Phys.* **189**: 266–274.
- Lapidot, T. and Kollet, O.T. 2002. The essential roles of the chemokine SDF-1 and its receptor CXCR4 in human stem cell homing and repopulation of transplanted immune-deficient NOD/SCID and NOD/SCID/B2m(null) mice. *Leukemia* **16**: 1992–2003.
- Lee, E., Lee, W., Kaetzel, C., Parry, G., and Bissell, M. 1985. Interaction of mouse mammary epithelial cells with collagen substrata: Regulation of casein gene expression and secretion. *Proc. Natl. Acad. Sci.* **82**: 1419–1423.
- Morrison, S.J., Shah, N.M., and Anderson, D.J. 1997. Regulatory mechanisms in stem cell biology. *Cell* **88**: 287–298.
- Muller, A., Homey, B., Soto, H., Ge, N., Catron, D., Buchanan, M., McClanahan, T., Murphy, E., Yuan, W., Wagner, S., et al. 2001. Involvement of chemokine receptors in breast cancer metastasis. *Nature* **410**: 50–56.
- Muschler, J., Lochter, A., Roskelley, C.D., Yurchenco, P., and Bissell, M.J. 1999. Division of labor among the $\alpha 6 \beta 4$ integrin $\beta 1$ integrins, and an E3 laminin receptor to signal morphogenesis and β -casein expression in mammary epithelial cells. *Mol. Biol.* **10**: 2817–2828.
- Parry, G., Lee, E., Farson, D., Koval, M., and Bissell, M. 1985. Collagenous substrata regulate the nature and distribution of glycosaminoglycans produced by differentiated cultures of mouse mammary epithelial cells. *Exp. Cell Res.* **156**: 487–499.
- Petersen, O., Rønnov-Jessen, L., Howlett, A., and Bissell, M. 1992. Interaction with basement membrane serves to rapidly distinguish growth and differentiation pattern of normal and malignant human breast epithelial cells. *Proc. Natl. Acad. Sci.* **89**: 9064–9068.
- Ramallo-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R.C., and Melton, D.A. 2002. “Stemness”: Transcriptional profiling of embryonic and adult stem cells. *Science* **298**: 597–600.
- Reya, T., Morrison, S.J., Clarke, M.F., and Weissman, I.L. 2001. Stem cells, cancer, and cancer stem cells. *Nature* **414**: 105–111.
- Reynolds, B.A. and Weiss, S. 1996. Clonal and population analyses demonstrate that an EGF-responsive mammalian embryonic CNS precursor is a stem cell. *Dev. Biol.* **175**: 1–13.
- Romanov, S.R., Kozakiewicz, B.K., Holst, C.R., Stampfer, M.R., Haupt, L.M., and Tlsty, T.D. 2001. Normal human mammary epithelial cells spontaneously escape senescence and acquire genomic changes. *Nature* **409**: 633–637.
- Rudland, P.S., Barraclough, R., Fernig, D.G., and Smith, J.A. 1997. Mammary stem cells in normal development and cancer. In *Stem Cells* (ed. C.S. Potten), pp. 147–232. Academic Press, San Diego, CA.
- Scharenberg, C.W., Harkey, M.A., and Torok-Storb, B. 2002. The *ABC G2* transporter is an efficient Hoechst 33342 efflux pump and is preferentially expressed by immature human hematopoietic progenitors. *Blood* **99**: 507–512.
- Scheffle, B., Horn, M., Blumcke, I., Laywell, E., Coomes, D., Kukekov, V., and Steindler, D. 1999. Marrow-mindedness: A perspective on neurogenesis. *Trends Neurosci.* **22**: 348–357.
- Simian, M., Hirai, Y., Navre, M., Werb, Z., Lochter, A., and Bissell, M.J. 2001. The interplay of matrix metalloproteinases, morphogens and growth factors is necessary for branching of mammary epithelial cells. *Development* **128**: 3117–3131.
- Smith, G.H. and Chepko, G. 2001. Mammary epithelial stem cells. *Microscopy Res. Tech.* **52**: 190–203.
- Soule, H. and McGrath, C. 1986. A simplified method for passage and long-term growth of human mammary epithelial cells. *In Vitro Cell. Dev. Biol.* **22**: 6–12.
- Stingl, J., Eaves, C.J., Kuusk, U., and Emerman, J.T. 1998. Phenotypic and functional characterization in vitro of a multi-

- potent epithelial cell present in the normal adult human breast. *Differentiation* **63**: 201–213.
- Stingl, J., Eaves, C.J., Zandieh, I., and Emerman, J.T. 2001. Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. *Breast Cancer Res. Treat.* **67**: 93–109.
- Storms, R.W., Goodell, M.A., Fisher, A., Mulligan, R.C., and Smith, C. 2000. Hoechst dye efflux reveals a novel CD7⁺CD34⁻ lymphoid progenitor in human umbilical cord blood. *Blood* **96**: 2125–2133.
- Taipale, J. and Beachy, P.A. 2001. The hedgehog and Wnt signalling pathways in cancer. *Nature* **411**: 349–354.
- Terskikh, A.V., Easterday, M.C., Li, L., Hood, L., Kornblum, H.I., Geschwind, D.H., and Weissman, I.L. 2001. From hematopoiesis to neurogenesis: Evidence of overlapping genetic programs. *Proc. Natl. Acad. Sci.* **98**: 7934–7939.
- Uyttendaele, H., Soriano, J.V., Montesano, R., and Kitajewski, J. 1998. Notch4 and Wnt-1 proteins function to regulate branching morphogenesis of mammary epithelial cells in an opposing fashion. *Dev. Biol.* **196**: 204–217.
- Watson, M.A., Dintzis, S., Darrow, C.M., Voss, L.E., DiPersio, J., Jensen, R., and Fleming, T.P. 1999. Mammaglobin expression in primary, metastatic, and occult breast cancer. *Cancer Res.* **59**: 3028–3031.
- Weaver, V.M. and Bissell, M.J. 1999. Functional culture models to study mechanisms governing apoptosis in normal and malignant mammary epithelial cells. *J. Mammary Gland Biol. Neoplasia* **4**: 193–201.
- Weiss, S., Reynolds, B.A., Vescovi, A.L., Morshead, C.M., Craig, C.G., and van der Kooy, D. 1996. Is there a neural stem cell in the mammalian forebrain? *Trends Neurosci.* **19**: 387–393.
- Weissman, I.L. 2000. Stem cells: Units of development, units of regeneration, and units in evolution. *Cell* **100**: 157–168.
- Welm, B.E., Tepera, S.B., Venezia, T., Graubert, T.A., Rosen, J.M., and Goodell, M.A. 2002. Sca-1^(pos) cells in the mouse mammary gland represent an enriched progenitor cell population. *Dev. Biol.* **245**: 42–56.
- Wiseman, B.S. and Werb, S. 2002. Stromal effects on mammary gland development and breast cancer. *Science* **296**: 1046–1049.
- Yuen, T., Wurmbach, E., Pfeffer, R., Ebersole, B., and Sealfon, S. 2002. Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. *Nucleic Acids Res.* **30**: e48.
- Zandstra, P.W., Le, H.-V., Daley, G.Q., Griffith, L.G., and Lauffenburger, D.A. 2000. Leukemia inhibitory factor (LIF) concentration modulates embryonic stem cell self-renewal and differentiation independently of proliferation. *Biotech. Bioeng.* **69**: 607–661.
- Zantek, N.D., Walker-Daniels, J., Stewart, J., Hansen, R.K., Robinson, D., Miao, H., Wang, B., Kung, H.-J., Bissell, M.J., and Kinch, M.S. 2001. MCF-10A-NeoST: A new cell system for studying cell–ECM and cell–cell interactions in breast cancer. *Clin. Cancer Res.* **7**: 3640–3648.
- Zhou, F.C., Kelley, M.R., Chiang, Y.H., and Young, P. 2000. Three to four-year-old nonpassaged EGF-responsive neural progenitor cells: Proliferation, apoptosis, and DNA repair. *Exp. Neurology* **164**: 200–208.
- Zhou, S., Schuetz, J.D., Bunting, K.D., Colapietro, A.-M., Sam-path, J., Morris, J.J., Lagutina, I., Grosveld, G.C., Osawa, M., Nakauchi, H., et al. 2001. The ABC transporter *Bcrp1/ABCG2* is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. *Nat. Med.* **7**: 1028–1034.