# Breast Cancer Cell Lines Contain Functional Cancer Stem Cells with Metastatic Capacity and a Distinct Molecular Signature

Emmanuelle Charafe-Jauffret,[1] Christophe Ginestier,[2] Flora Iovino,[3] Julien Wicinski,[1] Nathalie Cervera,[1] Pascal Finetti,[1] Min-Hee Hur,[4] Mark E. Diebel,[2] Florence Monville,[1] Julie Dutcher,[2] Marty Brown,[2] Patrice Viens,[1] Luc Xerri,[1] François Bertucci,[1] Giorgio Stassi,[3] Gabriela Dontu,[2] Daniel Birnbaum,[1] and Max S. Wicha[2]

[1]Centre de Recherche en Cancérologie de Marseille, Laboratoire d'Oncologie Moléculaire, UMR891 Inserm/Institut Paoli-Calmettes, Université de la Méditerranée, Marseille, France; [2]Comprehensive Cancer Center, Department of Internal Medicine/Oncology, University of Michigan, Ann Arbor, Michigan; [3]Department of Surgical and Oncological Sciences, Cellular and Molecular Pathophysiology Laboratory, University of Palermo, Palermo, Italy; and [4]Cheil General Hospital, Sungkyunkwan University, Seoul, Republic of Korea

## Abstract

**Tumors may be initiated and maintained by a cellular subcomponent that displays stem cell properties. We have used the expression of aldehyde dehydrogenase as assessed by the ALDEFLUOR assay to isolate and characterize cancer stem cell (CSC) populations in 33 cell lines derived from normal and malignant mammary tissue. Twenty-three of the 33 cell lines contained an ALDEFLUOR-positive population that displayed stem cell properties *in vitro* and in NOD/SCID xenografts. Gene expression profiling identified a 413-gene CSC profile that included genes known to play a role in stem cell function, as well as genes such as *CXCR1/IL-8RA* not previously known to play such a role. Recombinant interleukin-8 (IL-8) increased mammosphere formation and the ALDEFLUOR-positive population in breast cancer cell lines. Finally, we show that ALDEFLUOR-positive cells are responsible for mediating metastasis. These studies confirm the hierarchical organization of immortalized cell lines, establish techniques that can facilitate the characterization of regulatory pathways of CSCs, and identify potential stem cell markers and therapeutic targets.** [Cancer Res 2009;69(4):1302–13]

## Introduction

The evolution of a normal cell into a fully transformed one requires the deregulation of multiple cellular processes (1, 2). According to classic models of carcinogenesis, these events can occur in any cell. In contrast, the cancer stem cell (CSC) hypothesis holds that the preferential targets of oncogenic transformation are tissue stem or early progenitor cells that have acquired self-renewal potential (3–6). These tumor-initiating cells or CSCs, in turn, are characterized by their ability to undergo self-renewal, a process that drives tumorigenesis and differentiation, which contributes to tumor cellular heterogeneity. Recent evidence supporting the CSC hypothesis has been generated using xenografts of primary human tumors. These studies have suggested that tumors are composed of a cellular hierarchy driven by the CSC component. In addition, recent data suggest that immortalized cell lines derived from both murine and human tissues may also contain a cellular population displaying stem cell properties. Most of these studies have been based on *in vitro* properties, including clonogenic potential, sphere formation, and multilineage differentiation potential (7–10). More limited studies using functional transplantation of immortalized cell lines in xenografts have also suggested the existence of such a hierarchy. These studies have generally used Hoechst dye exclusion to identify the so-called side population (SP; refs. 7, 9, 11). In addition, cell surface markers defined using primary tumor xenografts, such as CD44 and CD133, have also been used to identify similar populations in established cell lines (7, 8). However, the limitations of these techniques have precluded their application across a wide variety of cell lines representing the molecular heterogeneity of tumors such as breast cancer. In addition, a crucial question remains as to whether the stem cell components of cell lines represent a valid model for CSC biology.

To provide more definitive evidence for the existence of CSC populations within breast cancer cell lines, we have studied the expression of the stem cell marker aldehyde dehydrogenase (ALDH) in a series of 33 cell lines derived from human breast cancers and nontransformed breast cells. ALDH is a detoxifying enzyme responsible for the oxidation of intracellular aldehydes and is thought to play a role in stem cell differentiation through metabolism of retinal to retinoic acid (12). ALDH activity, as assessed by the fluorescent ALDEFLUOR assay, has been successfully used to isolate CSCs in multiple myeloma and acute myeloid leukemia, as well as from brain tumors (13, 14). We recently showed that ALDH activity can be used to isolate a subpopulation of cells that display stem cell properties from normal human breast tissue and breast carcinomas (15). We now show that the majority of breast cancer cell lines contain an ALDEFLUOR-positive population with a distinct molecular profile that displays CSC properties. These studies have important implications for the interpretation of data using cell lines and suggest that these lines may be useful for elucidating CSC regulatory pathways.

## Materials and Methods

**Cell culture.** Breast cell lines (BCL) were obtained from American Type Culture Collection[5] or from collections developed in the laboratories of Drs.

**Note:** Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

E. Charafe-Jauffret and C. Ginestier contributed equally to this work.

**Requests for reprints:** Max S. Wicha, Department of Internal Medicine/Hematology-Oncology, University of Michigan Comprehensive Cancer Center, 1500 East Medical Center Drive, Ann Arbor, MI 48109-0015. Phone: 734-647-9923, Fax: 734-647-9480; E-mail: mwicha@med.umich.edu.



doi:10.1158/0008-5472.CAN-08-2741

[5] http://www.lgcpromochem-atcc.com/common/catalog/cellBiology/cellBiologyIndex.cfm

S. Ethier[6] (SUM44, SUM52, SUM149, SUM159, SUM185, SUM190, SUM225, SUM229), V.J. Möbus (BrCa-MZ-01), and V. Catros (S68). All BCLs tested were derived from carcinomas, except for MCF10A, which is derived from fibrocystic disease, and the HMEC-derived 184A1, which was derived from normal mammary tissue. The cell lines were grown using the recommended culture conditions (Supplementary Table S1). All experiments were done with subconfluent cells in the exponential phase of growth.

**ALDEFLUOR assay and separation of the ALDH-positive population by fluorescence-activated cell sorting.** ALDH activity was assessed in 33 BCLs representing the main molecular subtypes of human breast cancer. The ALDEFLUOR kit (StemCell Technologies) was used to isolate the population with high ALDH enzymatic activity using a FACStarPLUS (Becton Dickinson), as previously described (15). Briefly, cells were incubated in ALDEFLUOR assay buffer containing ALDH substrate (1 μmol/L per $1 \times 10^6$ cells). In each experiment, a sample of cells was stained under identical conditions with 50 mmol/L of diethylaminobenzaldehyde, a specific ALDH inhibitor, as negative control. The sorting gates were established using propidium iodide–stained cells for viability. Before RNA profiling or NOD/SCID mice injection, the purity of sorted populations was checked using double sorting of 10,000 ALDEFLUOR-positive and ALDEFLUOR-negative cells in BrCa-MZ-01 and SUM159 cell lines. For both cell lines, sorted ALDEFLUOR-positive populations contained >98% of ALDEFLUOR-positive cells and no ALDEFLUOR-positive cells were detected in the ALDEFLUOR-negative population.

**Tumorigenicity in NOD/SCID mice.** Tumorigenicity of ALDELFUOR-positive, ALDELFUOR-negative, and unseparated SUM159, MDA-MB-453, and BrCa-MZ-01 cells was assessed in NOD/SCID mice. Fat pads were prepared as described (15).

**Anchorage-independent culture.** ALDEFLUOR-positive, ALDELFUOR-negative, and unseparated cells from 184A1, SUM149, and SUM159 were plated as single cells in ultralow attachment plates (Corning) at low density (5,000 viable cells/mL). Cells were grown in serum-free mammary epithelial basal medium (Cambrex Bio Science) for 3 to 7 d, as described (16). The capacity of cells to form spheres was quantified after treatment with different doses of interleukin-8 (IL-8; GenWay Biotech) added to the medium.

**RNA extraction.** Total RNA was extracted from frozen ALDEFLUOR-positive and ALDEFLUOR-negative cells using DNA/RNA All Prep Maxi kit, according to the manufacturer's instructions (Qiagen, Sample & Assay Technologies). Eight BCLs were used for transcriptional analysis: 184A1, BrCa-MZ-01, HCC1954, MDA-MB-231, MDA-MB-453, SK-BR-7, SUM149, and SUM159. RNA integrity was controlled by denaturing formaldehyde agarose gel electrophoresis and microanalysis (Agilent Bioanalyzer).

**Gene expression profiling with DNA microarrays.** Gene expression analyses used Affymetrix U133 Plus 2.0 human oligonucleotide microarrays containing over 47,000 transcripts and variants, including 38,500 well-characterized human genes. Preparation of cRNA, hybridizations, washes, and detection were done as recommended by the supplier.[7] Expression data were analyzed by the robust multichip average (RMA) method in R using Bioconductor and associated packages (17), as described (18). RMA did background adjustment, quantile normalization, and summarization of 11 oligonucleotides per gene.

Before analysis, a filtering process removed from the data set genes with low and poorly measured expression as defined by expression value inferior to 100 units in all the 16 samples, retaining 25,285 genes/ESTs. A second filter, based on the intensity of SD, was applied for unsupervised analyses to exclude genes showing low expression variation across the analyses. SD was calculated on $\log_2$-transformed data, in which lowest values were first floored to a minimal value of 100 units, i.e., the background intensity, retaining 13,550 genes/ESTs with SD superior to 0.5. An unsupervised analysis was done on 16 ALDEFLUOR-positive and ALDEFLUOR-negative cells on 13,550 genes. Before hierarchical clustering, filtered data were $\log_2$-transformed and submitted to the Cluster program (19) using data median-centered on genes, Pearson correlation as similarity metric, and centroid linkage clustering. Results were displayed using the TreeView program (19). To identify and rank genes discriminating ALDEFLUOR-positive and ALDEFLUOR-negative populations, a Mann-Whitney *U* test was applied to the 25,285 genes/ESTs and false discovery rate (FDR) was used to correct the multiple testing hypothesis (see Supplementary Table S2 for complete data set). The classification power of the discriminator signature was illustrated by classifying samples by hierarchical clustering. A leave-one-out cross-validation (LOOCV) was applied to estimate the accuracy of prediction of the identified molecular signatures and the validity of supervised analysis; each sample was excluded one by one and classified with the linear discriminant analysis (20) by using model defined on the nonexcluded samples.

**Real-time reverse transcription–PCR.** After ALDEFLUOR-positive and ALDEFLUOR-negative populations from different cell lines were sorted, total RNA was isolated using RNeasy Mini kit (QIAGEN) and used for real-time quantitative reverse transcription–PCR (qRT-PCR) assays in an ABI PRISM 7900HT sequence detection system with 384-well block module and automation accessory (Applied Biosystems). Primers and probes for the Taqman system were selected from the Applied Biosystems website.[8] The sequences of the PCR primer pairs and fluorogenic probes used are available on the Applied Biosystems website (CXCR1 assay ID: Hs_00174146_mi, FBXO21 assay ID: Hs_00372141_mi, NFYA assay ID: Hs_00953589_mi, NOTCH2 assay ID: Hs_01050719_mi, RAD51L1 assay ID: Hs00172522_mi, TBP assay ID: Hs_00427620_mi). The relative expression mRNA level of *CXCR1, FBXO21, NFYA, NOTCH2, RAD51L1* was computed with respect to the internal standard *TBP* gene to normalize for variations in the quality of RNA and the amount of input cDNA, as described previously (21).

**Invasion assay.** Assays were done in triplicate in transwell chambers with 8-μm pore polycarbonate filter inserts for 12-well plates (Corning). Filters were coated with 30 μL of ice-cold 1:6 basement membrane extract (Matrigel, BD Bioscience) in DMEM/F12 incubated 1 h at 37°C. Cells were added to the upper chamber in 200 μL of serum-free medium. For the invasion assay, 5,000 cells were seeded on the Matrigel-coated filters and the lower chamber was filled with 600 μL of medium supplemented with 10% human serum (Cambrex) or with 600 μL of serum-free medium supplemented with IL-8 (100 ng/mL). After 48 h of incubation, the cells on the underside of the filter were counted using light microscopy. Relative invasion was normalized to the unseparated corresponding cell lines under serum condition.

**Lentivirus infection.** For luciferase gene transduction, 70% confluent cells from HCC1954, MDA-MB-453, and SUM159 were incubated overnight with a 1:3 precipitated mixture of lentiviral supernatants Lenti-LUC-VSVG (Vector Core) in culture medium. On the following day, the cells were harvested by trypsin/EDTA and subcultured at a ratio of 1:6. After 1 wk of incubation, cells were sorted according to the ALDEFLUOR phenotype, and luciferase expression was verified in each sorted population (ALDEFLUOR-positive and ALDEFLUOR-negative) by adding 2 μL D-luciferin 0.0003% (Promega) in the culture medium and counting photon flux by device camera system (Xenogen; Supplementary Fig. S1).

**Intracardiac inoculation.** Six-week-old NOD/SCID mice were anesthetized with 2% isofluorane/air mixture and injected in the left ventricle of the heart with 100,000 cells in 100 μL of sterile Dulbecco's PBS lacking $Ca^{2+}$ and $Mg^{2+}$. For each of the three cell lines (HCC1954, MDA-MB-453, SUM159) and for each population (ALDEFLUOR-positive, ALDEFLUOR-negative, and unsorted), three animals were injected.

**Bioluminescence detection.** Baseline bioluminescence was assessed before inoculation and each week thereafter inoculations. Mice were anesthetized with a 2% isofluorane/air mixture and given a single i.p. dose of 150 mg/kg D-luciferin (Promega) in PBS. For photon flux counting, we used a charge-coupled device camera system (Xenogen) with a nose-cone isofluorane delivery system and heated stage for maintaining body temperature. Results were analyzed after 2 to 12 min of exposure using Living Image software provided with the Xenogen imaging system. Signal intensity was quantified as the sum of all detected photon flux counts within

[6] http://www.asterand.com/asterand/BIOREPOSITORY/hbreastcancercelllines.aspx
[7] http://www.affymetrix.com/index.affx
[8] www.appliedbiosystems.com

a uniform region of interest manually placed during data postprocessing. Normalized photon flux represents the ratio of the photon flux detected each week after inoculations and the photon flux detected before inoculation.

**Statistical analysis.** Results are presented as the mean ± SD for at least three repeated individual experiments for each group. Statistical analyses used the SPSS software (version 10.0.5). Correlations between sample groups and molecular variables were calculated with the Fisher's exact test or the one-way ANOVA for independent samples. A $P$ value of <0.05 was considered significant.

## Results

**The majority of BCLs contain an ALDEFLUOR-positive population.** We used the ALDEFLUOR assay (15) to isolate CSC from 33 BCLs representing the diverse molecular subtypes and features of breast cancer (ref. 18; Supplementary Fig. S2). As shown in Fig. 1*A*, 23 of the 33 cell lines contained an ALDEFLUOR-positive cell population that ranged from 0.2 to nearly 100%. All 16 basal/mesenchymal BCLs contained an ALDEFLUOR-positive population, whereas 7 of the 12 luminal BCLs did not contain any detectable ALDEFLUOR-positive cells ($P$ = 0.0006, Fischer's exact test; Fig. 1*B*).

**ALDEFLUOR-positive cells have tumorsphere-forming capacity.** We have previously reported that mammary epithelial stem and progenitor cells are able to survive and proliferate in anchorage-independent conditions and form floating spherical colonies that we termed mammospheres (16). Data from breast tumors, as well as cell lines, have shown that cancer stem-like cells or cancer-initiating cells can also be isolated and propagated as

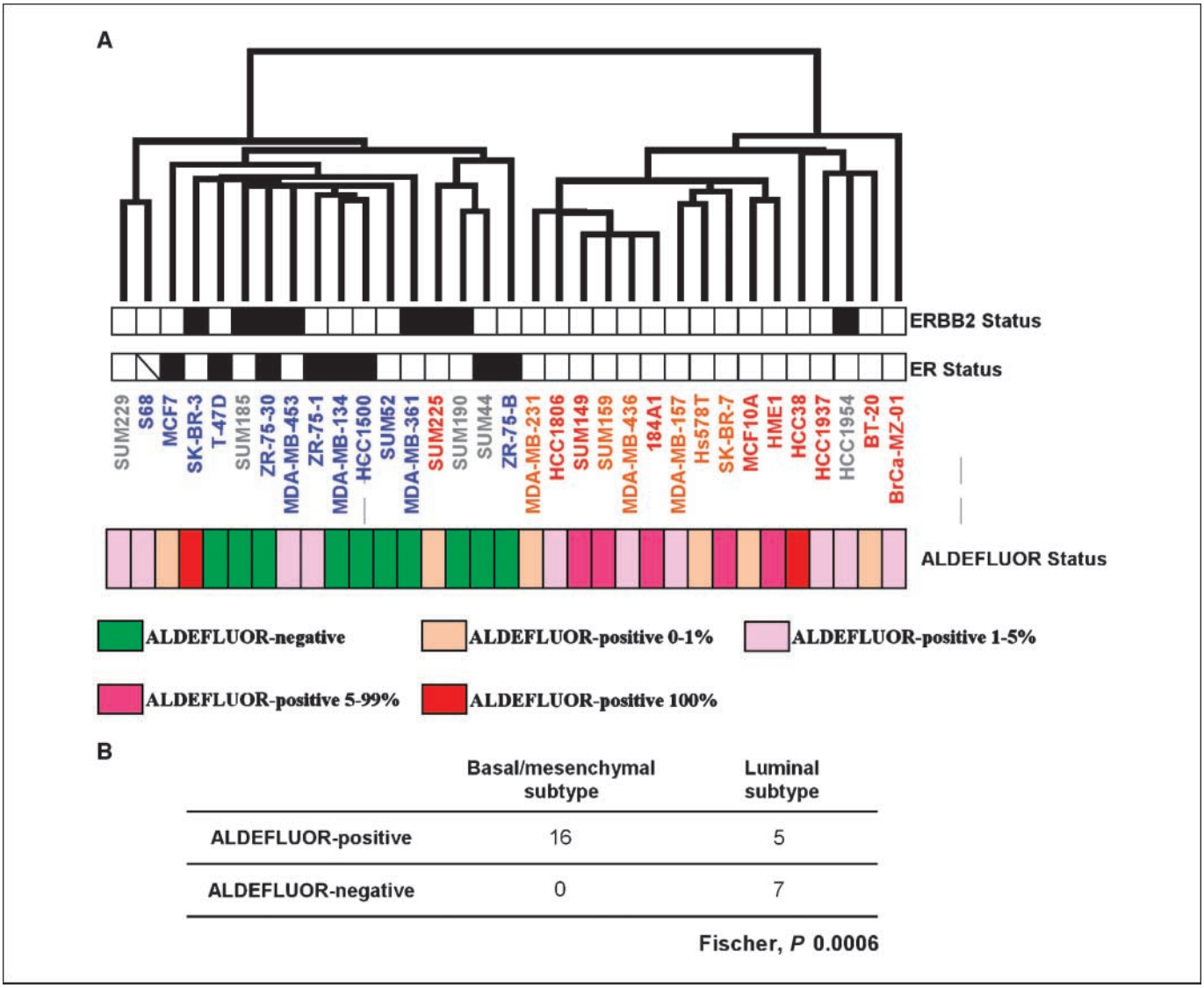


**Figure 1.** Global gene expression profiling of 33 BCLs analyzed by the ALDEFLUOR assay. Hierarchical clustering of 33 BCLs and 13,550 genes/ESTs based on mRNA expression levels. *A,* the dendrogram of samples represents overall similarities in gene expression profiles. Two large groups of samples are evidenced by clustering. Name of cell lines is colored as follows: blue for luminal ($n$ = 12), red for basal ($n$ = 10), brown for mesenchymal ($n$ = 6) cell lines according to the correlation of expression profile of each cell line with each Ross and Perou centroid (i.e., molecular subtype). Five cell lines were not attributed any subtype (*name in gray*). The ER, ERBB2, and ALDEFLUOR status of BCLs are represented according to a color ladder (for ER and ERBB2 status, negative, white; positive, black; unavailable, oblique feature; for ALDEFLUOR status, a color scale shown at the bottom of the dendrogram relates the percentage of ALDELFUOR-positive cells found in each BCL). *B,* comparison of the ALDEFLUOR status with the molecular subtypes of BCLs revealed a strong correlation between the basal/mesenchymal subtypes and the presence of ALDEFLUOR-positive cells ($P$ = 0.0006).

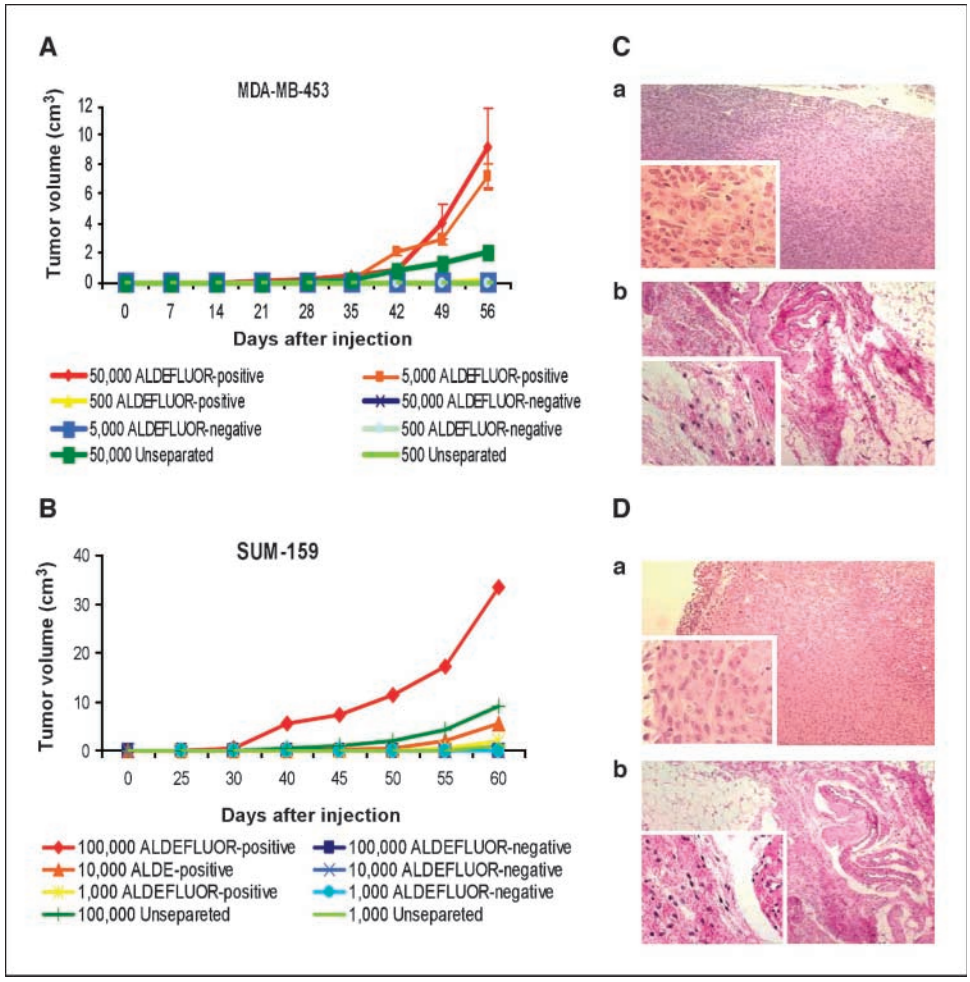


**Figure 2.** The ALDEFLUOR-positive cell populations from breast cancer cell lines (MDA-MB-453, SUM159) have CSC properties. *A* and *B,* tumor growth curves were plotted for different numbers of cells injected (for MDA-MB-453: 50,000 cells, 5,000 cells, and 500 cells; for SUM159: 100,000 cells, 10,000 cells, and 1,000 cells) and for each population (ALDEFLUOR-positive, ALDEFLUOR-negative, unseparated). Tumor growth kinetics correlated with the latency and size of tumor formation and the number of ALDEFLUOR-positive cells. *C* and *D,* H&E staining of ALDEFLUOR-positive cell injection site, revealing presence of tumor cells (*Ca,* MDA-MB-453 ALDEFLUOR-positive cell injection site; *Da,* SUM159 ALDEFLUOR-positive cell injection site). *Cb* and *Db,* the ALDEFLUOR-negative cell injection site contained only residual Matrigel, apoptotic cells, and mouse tissue (*Cb,* MDA-MB-453 ALDEFLUOR-negative cell injection site; *Db,* SUM159 ALDEFLUOR-negative cell injection site). *Points,* mean; *bars,* SD.

"tumorspheres" in similar assays (22). All mammosphere-initiating cells in the normal human mammary gland are contained within the ALDEFLUOR-positive population (15). To characterize the ALDEFLUOR-positive population from BCLs, we compared the ability of ALDEFLUOR-positive and ALDEFLUOR-negative populations from 184A1, SUM149, and SUM159 to form tumorspheres. In each cell line, the ALDEFLUOR-positive population showed increased tumorsphere-forming capacity compared with ALDEFLUOR-negative cells (Supplementary Fig. S3).

**ALDEFLUOR-positive BCL cells have CSC properties *in vivo*.** To determine the hierarchical organization of BCL, we analyzed the stem cell properties of the ALDEFLUOR-positive and ALDEFLUOR-negative populations of MDA-MB-453, SUM159, and BrCa-MZ-01 cell lines. The ALDEFLUOR-positive populations of these three BCLs constituted between 3.54 $\pm$ 1.73% and 5.49 $\pm$ 3.36% of the total cell populations (Supplementary Fig. S4). As shown in Fig. 2*A* and *B*, the size and latency of tumor formation correlated with the number of ALDEFLUOR-positive cells injected. Remarkably, 500 ALDEFLUOR-positive cells from MDA-MB-453 and 1,000 ALDEFLUOR-positive cells from SUM159 were able to form tumors. The tumor-generating capacity was maintained through serial passages demonstrating the self-renewal capacity of these cells. In contrast, ALDEFLUOR-negative cells failed to generate tumors, although limited growth was produced when 50,000 ALDEFLUOR-negative MDA-MB-453 cells were injected (Supplementary Table S3). H&E staining of the fat pad sections confirmed that tumors formed by ALDEFLUOR-positive cells contained malignant cells, whereas only residual Matrigel, apoptotic cells, and mouse tissue were seen at the sites of ALDELFUOR-negative cell injections (Fig. 2*C* and *D*). Consistent with the ALDEFLUOR-positive population having CSC characteristics, tumors generated by this population recapitulated the phenotypic heterogeneity of the initial tumor, with a similar ratio of ALDEFLUOR-positive and ALDEFLUOR-negative cells (Supplementary Fig. S4). This indicates that ALDEFLUOR-positive cells were able to self-renew, generating ALDEFLUOR-positive cells, and were able to differentiate, generating ALDEFLUOR-negative cells.

When BrCa-MZ-01 cells were separated into ALDEFLUOR-positive and ALDEFLUOR-negative components, both were capable of tumor generation. Tumors generated by the ALDEFLUOR-positive population consisted of both ALDEFLUOR-positive and ALDEFLUOR-negative cells recapitulating the phenotypic heterogeneity of the initial tumor (Supplementary Fig. S5). In contrast, tumors generated by ALDEFLUOR-negative cells gave rise to slowly growing tumors containing only ALDEFLUOR-negative cells (Supplementary Fig. S5*F* and *G*). In contrast to the ability of ALDEFLUOR-positive cells to be serially transplanted, serial passages of ALDEFLUOR-negative tumors produced decreasing tumor growth with no growth after three passages. This suggests that the ALDEFLUOR-positive component of the BrCa-MZ-01 cells

**Table 1.** Genes up-regulated and down-regulated in ALDEFLUOR-positive populations

| | Category | Symbol | Description | Cytoband | Probe set ID | Function |
|---|---|---|---|---|---|---|
| Up-regulated | Genes previously described to have a role in stem cell biology | *TPRXL* | Tetrapeptide repeat homeobox-like | chr3p25.1 | 239061_at | Early embryonic development |
| | | *NOTCH2* | Notch homologue 2 (*Drosophila*) | chr1p13-p11 | 202443_x_at | Self-renewal program |
| | | *RBM15* | RNA binding motif protein 15 | chr1p13 | 1555760_a_at | Determination of hematopoietic cell fate |
| | | *ST3GAL3* | ST3 β-galactoside α-2,3-sialyltransferase 3 | chr1p34.1 | 1555181_a_at | Maintenance of the embryonic antigens SSEA-3 and SSEA-4 |
| | | *NFYA* | Nuclear transcription factor Y, α | chr6p21.3 | 204107_at | Self-renewal program |
| | | *PCNX* | Pecanex homologue (*Drosophila*) | chr14q24.2 | 213173_at | Determination of neural cell fate of early developing embryo |
| | Signaling | *FBXO21** | F-box protein 21 | chr12q24.22 | 212231_at | Ubiquitination |
| | | *WWOX* | WW domain containing oxidoreductase | chr16q23.3-q24.1 | 210695_s_at | Protein degradation, transcription, and RNA splicing |
| | | *CAMK2B* | Calcium/calmodulin-dependent protein kinase (CaM kinase) II β | chr22q12 | 34846_at | Calcium signaling |
| | | *PNPLA2* | Patatin-like phospholipase domain containing 2 | chr11p15.5 | 39854_r_at | Triglyceride hydrolysis |
| | | *CLIC5* | Chloride intracellular channel 5 | chr6p12.1-21.1 | 213317_at | Chloride ion transport |
| | | *UGCGL1* | UDP-glucose ceramide glucosyltransferase-like 1 | chr2q14.3 | 222569_at | Protein glucosylation |
| | | *FBXL18* | F-box and leucine-rich repeat protein 18 | chr7p22.2 | 220896_at | Ubiquitination |
| | | *ADRBK1* | Adrenergic, β, receptor kinase 1 | chr11q13 | 38447_at | Phosphorylation of G protein-coupled receptors |
| | | *SLC38A2* | Solute carrier family 38, member 2 | chr12q | 1559924_at | Neutral amino acid transporter |
| | Membrane protein | *IL-8RA** | Interleukin-8 receptor, α | chr2q35 | 207094_at | Inflammatory response |
| | | *TAS2R14* | Taste receptor, type 2, member 14 | chr12p13 | 241997_at | Bitter perception |
| | | *CD300LB* | CD300 molecule-like family member b | chr17q25.1 | 1554173_at | Immune response |
| | | *GIPC3* | GIPC PDZ domain containing family, member 3 | chr19p13.3 | 236730_at | |
| | DNA repair | *RAD51L1* | RAD51-like 1 (*S. cerevisiae*) | chr14q23-q24.2 | 1570166_a_at | Homologous recombination repair |
| | Chromatin remodeling | *ARID1B* | AT rich interactive domain 1B (SWI1-like) | chr6q25.1 | 225181_at | Chromatin remodeling (SWI/SNF complex) |
| | Cytoskeleton | *EPPK1* | Epiplakin 1 | chr8q24.3 | 208156_x_at | Maintenance of the keratin intermediate filaments |
| | Extracellular matrix | *COL11A2* | Collagen, type XI, α2 | chr6p21.3 | 216993_s_at | Skeletal morphogenesis |
| | | *KLK3* | Kallikrein 3, (prostate-specific antigen) | chr19q13.41 | 231629_x_at | Protease |
| | RNA interference | *EIF2C2* | Eukaryotic translation initiation factor 2C, 2 | chr8q24 | 213310_at | Short-interfering RNA-mediated gene silencing |
| | Unknown | *ZFP41* | Zinc finger protein 41 homologue (mouse) | chr8q24.3 | 227898_s_at | Unknown |
| | | *FAM49B* | Family with sequence similarity 49, member B | chr8q24.21 | 243182_at | Unknown |
| | | *PSORS1C2* | Psoriasis susceptibility 1 candidate 2 | chr6p21.3 | 220635_at | Unknown |
| Down-regulated | Protein synthesis | *MRPL42** | Mitochondrial ribosomal protein L42 | chr12q22 | 217919_s_at | Protein synthesis within the mitochondrion |
| | | *MRPL54** | Mitochondrial ribosomal protein L54 | chr19p13.3 | 225797_at | Protein synthesis within the mitochondrion |

(*Continued on the following page*)

**Table 1.** Genes up-regulated and down-regulated in ALDEFLUOR-positive populations (Cont'd)

| Category | Symbol | Description | Cytoband | Probe set ID | Function |
|---|---|---|---|---|---|
| | *MRPL47** | Mitochondrial ribosomal protein L47 | chr3q26.33 | 223480_s_at | Protein synthesis within the mitochondrion |
| | *MRPS23** | Mitochondrial ribosomal protein S23 | chr17q22-q23 | 223156_at | Protein synthesis within the mitochondrion |
| | *EIF3S9** | eukaryotic translation initiation factor 3, subunit 9 η, 116 kDa | chr7p22.2 | 236274_at | Initiation of protein synthesis (EIF3 multiprotein complex) |
| Signaling | *ALG5** | Asparagine-linked glycosylation 5 homologue | chr13q13.3 | 218203_at | Protein glucosylation |
| | *DNAJC19** | DnaJ (Hsp40) homologue, subfamily C, member 19 | chr3q26.33 | 225359_at | Importation of mithochondrial protein |
| | *HBLD2** | HESB-like domain containing 2 | chr9q21.33 | 221425_s_at | Iron-sulfur cluster biogenesis |
| | *GART** | Phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase | chr21q22.1 | 230097_at | *De novo* purine biosynthesis |
| | *NUP37** | Nucleoporin 37 kDa | chr12q23.2 | 218622_at | Intracellular protein transport across nuclear membrane |
| | *RNF7** | Ring finger protein 7 | chr3q22-q24 | 224439_x_at | Subunit of SKP1-cullin/CDC53-F box protein ubiquitin ligases |
| | *DC2** | DC2 protein | chr4q25 | 223001_at | Protein glycosylation |
| | *USP15** | Ubiquitin-specific peptidase 15 | chr12q14 | 210681_s_at | Protein degradation |
| | *COMMD6** | COMM domain containing 6 | - | 225312_at | Inhibition nuclear factor-κB signaling |
| | *UBL5** | Ubiquitin-like 5 | chr19p13.3 | 218011_at | Ubiquitination |
| Apoptosis | *MRPL41** | Mitochondrial ribosomal protein L41 | chr9q34.3 | 225425_s_at | Stabilization of p53 protein, cell cycle arrest [p21 (WAF1/CIP1) and p27 (Kip1) dependent] |
| | *PDCD10** | Programmed cell death 10 | chr3q26.1 | 210907_s_at | Initiation of apoptosis |
| | *PDCD5** | Programmed cell death 5 | chr19q12- | 227751_at | Initiation of apoptosis |
| Differentiation program | *NACA** | Nascent polypeptide-associated complex α polypeptide | chr12q23-q24.1 | 222018_at | Erythroid differentiation |
| Cell cycle | *FAM82B** | Family with sequence similarity 82, member B | chr8q21.3 | 218549_s_at | Regulation of microtubule dynamic |
| RNA splicing | *CCNL1** | Cyclin L1 | chr3q25.32 | 1555411_a_at | Pre-mRNA processing |
| | *PRPF39** | PRP39 pre-mRNA processing factor 39 homologue (*S. cerevisiae*) | chr14q21.3 | 220553_s_at | Pre-mRNA processing |
| | *LSM3** | LSM3 homologue, U6 small nRNA associated (*S. cerevisiae*) | chr3p25.1 | 202209_at | Pre-mRNA processing |
| | *SFRS7** | Splicing factor, arginine/serine-rich 7, 35 kDa | chr2p22.1 | 213649_at | Regulation of RNA splicing |
| | *PRPF4B** | PRP4 pre-mRNA processing factor 4 homologue B (yeast) | chr6p25.2 | 202127_at | Pre-mRNA processing |
| Oxidative phosphorylation | *ATP5S** | ATP synthase, H+ transporting, mitochondrial F0 complex, subunit s (factor B) | chr14q22.1 | 206992_s_at | Subunit of mitochondrial ATP synthase |
| | *NDUFA2** | NADH dehydrogenase (ubiquinone) 1α subcomplex, 2, | chr5q31 | 209224_s_at | Components of the complex I multisubunit enzyme |
| | *ATP5J2** | ATP synthase, H+ transporting, mitochondrial F0 complex, subunit F2 | chr7q22.1 | 202961_s_at | Subunit of mitochondrial ATP synthase |

(*Continued on the following page*)

**Table 1.** Genes up-regulated and down-regulated in ALDEFLUOR-positive populations (Cont'd)

| Category | Symbol | Description | Cytoband | Probe set ID | Function |
|---|---|---|---|---|---|
| | *IMMP1L** | IMP1 inner mitochondrial membrane peptidase-like | chr11p13 | 230556_at | Proteolysis |
| Unknown | *ASTE1** | Asteroid homologue 1 (Drosophila) | chr3q22.1 | 221135_s_at | Unknown |
| | *MGC61571** | Hypothetical protein MGC61571 | chr3p24.1 | 228283_at | Unknown |
| | *WDR53** | WD repeat domain 53 | chr3q29 | 227814_at | Unknown |
| | *DKFZP686A10121** | Hypothetical protein | chr7q21.13 | 234311_s_at | Unknown |
| | *CHCHD8** | Coiled-coil-helix-coiled-coil-helix domain containing 8 | chr11q13.4 | 220647_s_at | Unknown |
| | *FLJ32745** | Hypothetical protein FLJ32745 | chr2q13 | 235644_at | Unknown |
| | *CHURC1** | Churchill domain containing 1 | chr14q23.3 | 233268_s_at | Unknown |
| | *XTP3TPA** | XTP3-transactivated protein A | chr16p11.2 | 218069_at | Unknown |
| | *FLJ37953** | Hypothetical protein FLJ37953 | chr2q33.1 | 235181_at | Unknown |
| | *SNORD50A** | Small nucleolar RNA, C/D box 50A | chr6q14.3 | 244669_at | Unknown |
| | *LOC644053** | Hypothetical protein LOC644053 | chr1q41 | 235466_s_at | Unknown |
| | *TMEM141** | Transmembrane protein 141 | chr9q34.3 | 225568_at | Unknown |
| | *C8orf59** | Chromosome 8 open reading frame 59 | chr8q21.2 | 226165_at | Unknown |

NOTE: The 28 up-regulated genes in ALDEFLUOR-positive population (*top*) and the 42 down-regulated genes in ALDEFLUOR-positive population (*bottom*), common between the 413-gene and the 49-gene signatures (Supplementary Table S2). They are grouped together according to a common biological function and ranked according to their discriminating score (DS). DS was calculated for each gene as DS = (M1 − M2) / (S1 + S2), wherein M1 and S1 respectively represent mean and SD of expression levels of the gene in subgroup 1 and M2 and S2 in subgroup 2. Because of multiple hypotheses testing, confidence levels were estimated by 100 iterative random permutations of samples as previously described with a false positive rate of 1/1.000 (50). Underlined genes are cited in the text (Results and Discussion).
*Follows the name of genes found in the 49-gene signature.

contain cells with stem cell properties, whereas the ALDEFLUOR-negative cells contain progenitor cells able to undergo limited growth but not self-renewal.

**Gene expression profiling of ALDELFUOR-positive and ALDELFUOR-negative cell populations.** To determine whether ALDEFLUOR-positive cells isolated from different BCLs expressed a common set of CSC genes, we analyzed ALDEFLUOR-positive and ALDEFLUOR-negative cell populations isolated from eight BCLs (184A1, BrCa-MZ-01, HCC1954, MDA-MB-231, MDA-MB-453, SK-BR-7, SUM49, and SUM159) by using Affymetrix whole-genome oligonucleotide microarrays. Unsupervised hierarchical clustering, applied to the 16 samples and the 13,550 filtered genes/ESTs, did not separate ALDEFLUOR-positive and ALDEFLUOR-negative populations (not shown). Instead, ALDEFLUOR-positive and ALDEFLUOR-negative populations clustered with the parental cell line. This suggests that the differences in mRNA transcripts between clonal cell lines supersede differences between ALDEFLUOR-positive and ALDEFLUOR-negative cells. This further suggests that only a limited number of genes are differentially expressed between putative CSCs and their progeny.

To determine which genes discriminated ALDEFLUOR-positive and ALDEFLUOR-negative populations, the Mann-Whitney *U* test was applied to all genes but those with low and poorly measured expression, i.e., 25,285 probe sets. This test identified and ranked, after FDR correction, 413 genes/ESTs that discriminated the ALDEFLUOR-positive and ALDEFLUOR-negative cell populations. The 28 overexpressed genes corresponding to unique genes and the most frequently underexpressed genes are shown in Table 1 (*top* and *bottom*). The classification power of this discriminating signature was illustrated by classifying the 16 ALDEFLUOR-positive and ALDEFLUOR-negative samples with the 413 differentially expressed genes/ESTs. Hierarchical clustering ranked 15 of the 16 samples (Supplementary Fig. S6).

A number of genes known to play a role in stem cell biology were up-regulated in the ALDEFLUOR-positive populations (Table 1, *top*), including *NFYA, NOTCH2, PCNX, RBM15, ST3GAL3,* and *TPRXL*. Other genes encode proteins that have putative or uncharacterized role in stem cell function, such as ARID1B, RAD51L1, and the chemokine receptor CXCR1/IL-8RA (23). Genes underexpressed in the ALDEFLUOR-positive population are involved in cell differentiation, apoptosis, RNA splicing, and mitochondrial metabolism.

To increase the stringency of analysis, we raised the threshold of the Mann-Whitney analysis to the 0.5 risk and obtained a list of 49 genes/ESTs that discriminated ALDELFLUOR-positive and ALDEFLUOR-negative populations (Table 1, genes with asterisk). With this list, all of the ALDEFLUOR-positive cells, except from SK-BR-7, clustered together (Supplementary Fig. S7*A*). Among these 49 genes/ESTS, 45 corresponded to identified unique genes; only 3 of these 45 were overexpressed in the ALDEFLUOR-positive group whereas 42 were underexpressed. Characterized overexpressed genes code for an F-box protein FBXO21 and CXCR1/IL-8RA. Underexpressed genes include those coding for mitochondrial ribosomal proteins (MRPL41, MRPL42, MRPL47, MRPL54, MRPS23, IMMP1L), and differentiation (NACA) and pre-mRNA splicing factors (LSM3, pre-mRNA processing factor PRPF39, and PRPF4B).

LOOCV at 0.5% risk estimated the accuracy of prediction of the identifier molecular signature, and 88% of the samples were predicted in the right class with this CSC signature confirming the supervised analysis (Supplementary Fig. S7*B* and *C*).

**qRT-PCR assessment confirmed a significant increase of CXCR1 and FBXO21 in ALDEFLUOR-positive cells.** We performed qRT-PCR analysis of five discriminator genes overexpressed in ALDEFLUOR-positive populations (*CXCR1/IL-8RA, FBXO21, NFYA, NOTCH2*, and *RAD51L1*). Three cell lines used in the profiling analysis (BrCa-MZ-01, MDA-MB-453, SUM159) and two additional luminal cell lines (MCF7, S68) were sorted by ALDEFLUOR assay, and ALDEFLUOR-positive and ALDEFLUOR-negative populations were processed separately for qRT-PCR analysis. The qRT-PCR expression level of *CXCR1* and *FBXO21* are presented in Fig. 3*A* and *B*. Analyses of *NFYA, NOTCH2*, and *RAD51L1* are presented in Supplementary Fig. S8. Gene expression levels measured by qRT-PCR confirmed the results obtained using DNA microarrays with an increase of *CXCR1* and *FBXO21* mRNA level in the ALDEFLUOR-positive population compared with the ALDEFLUOR-negative population ($P < 0.05$).

**IL-8 promotes CSC self-renewal.** Our profiling studies suggested that the IL-8 receptor CXCR1/IL-8RA was consistently expressed in the ALDEFLUOR-positive cell population. To confirm this association, we measured the protein expression of CXCR1/IL-8RA by flow cytometry in ALDEFLUOR-positive and ALDEFLUOR-negative populations. The ALDEFLUOR-positive and ALDEFLUOR-negative populations from four different cell lines were isolated by fluorescence-activated cell sorting, fixed, and stained with a CXCR1 monoclonal antibody labeled with phycoerythrin. ALDEFLUOR-positive cells were highly enriched in CXCR1-positive cells compared with the ALDEFLUOR-negative populations (Supplementary Fig. S9).

To determine whether IL-8 signaling is important in stem cell function, we treated four BCLs with human recombinant IL-8 and determined its effect on the CSC population, as measured by the formation of tumorspheres and by ALDH enzymatic activity. As shown in Fig. 4*A*, addition of IL-8 increased the formation of primary and secondary tumorspheres in a dose-dependent manner. Furthermore, IL-8 increased the ALDEFLUOR-positive population in a dose-dependent manner in each of the four BCLs analyzed (Fig. 4*B*). This illustrates the power of the CSC signature to identify pathways that may play a role in the stem cell function.

**The IL8/CXCR1 axis is involved in CSC invasion.** The IL8/CXCR1 axis has been reported to play a role in CSC invasion (24, 25). We first used a Matrigel invasion assay, using serum as attractant, to examine the ability of ALDEFLUOR-positive and ALDEFLUOR-negative cell populations from three different cell lines (HCC1954, MDA-MB-453, SUM159) to invade. As shown in Fig. 5*A*, ALDEFLUOR-positive cells showed 6-fold to 20-fold higher invasion through Matrigel than the ALDEFLUOR-negative population ($P < 0.01$). When used as a chemoattractant, IL-8 (100 ng/mL) increased invasion of the ALDEFLUOR-positive cells ($P < 0.05$; Fig. 5*A*). In contrast to its effects on ALDEFLUOR-positive cells, IL-8 did not have any effect on the invasive capacity of ALDELFLUOR-negative cells. These results indicate that CSCs exhibited invasive behavior and, furthermore, that IL-8 facilitates this process.

**ALDEFLUOR-positive cells have increased metastatic potential.** It has been proposed that CSCs play a crucial role in cancer metastasis (26, 27). The above experiments showed that

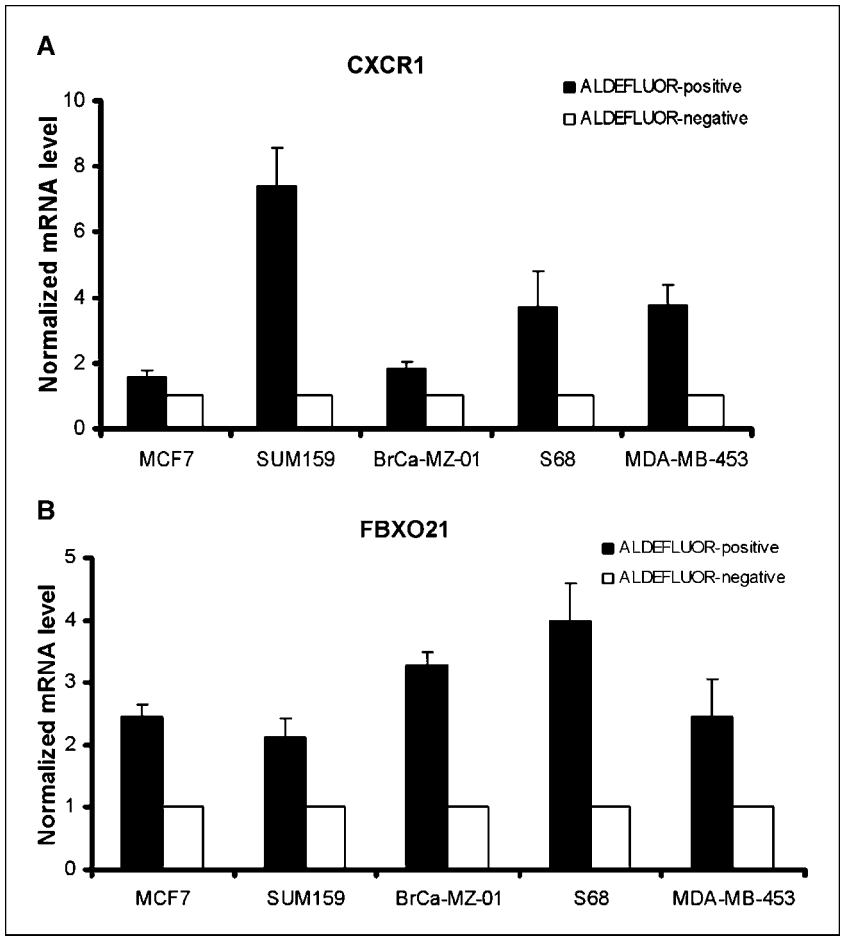


**Figure 3.** Validation of gene expression results by qRT-PCR. *A* and *B*, to confirm our gene expression results, we measured in a set of five breast cancer cell lines sorted for the ALDEFLUOR phenotype the expression of five discriminator genes overexpressed in ALDEFLUOR-positive populations (*CXCR1/IL-8RA, FBXO21, NFYA, NOTCH2,* and *RAD51L1*) by qRT-PCR. The qRT-PCR expression levels of *CXCR1* and *FBXO21* are presented here and the ones of *NFYA, NOTCH2*, and *RAD51L1* are presented in Supplementary Fig. S8. Gene expression levels measured by qRT-PCR confirm the results obtained using DNA microarrays with an increase of *CXCR1* and *FBXO21* mRNA level in the ALDEFLUOR-positive population compared with the ALDEFLUOR-negative population ($P < 0.05$).

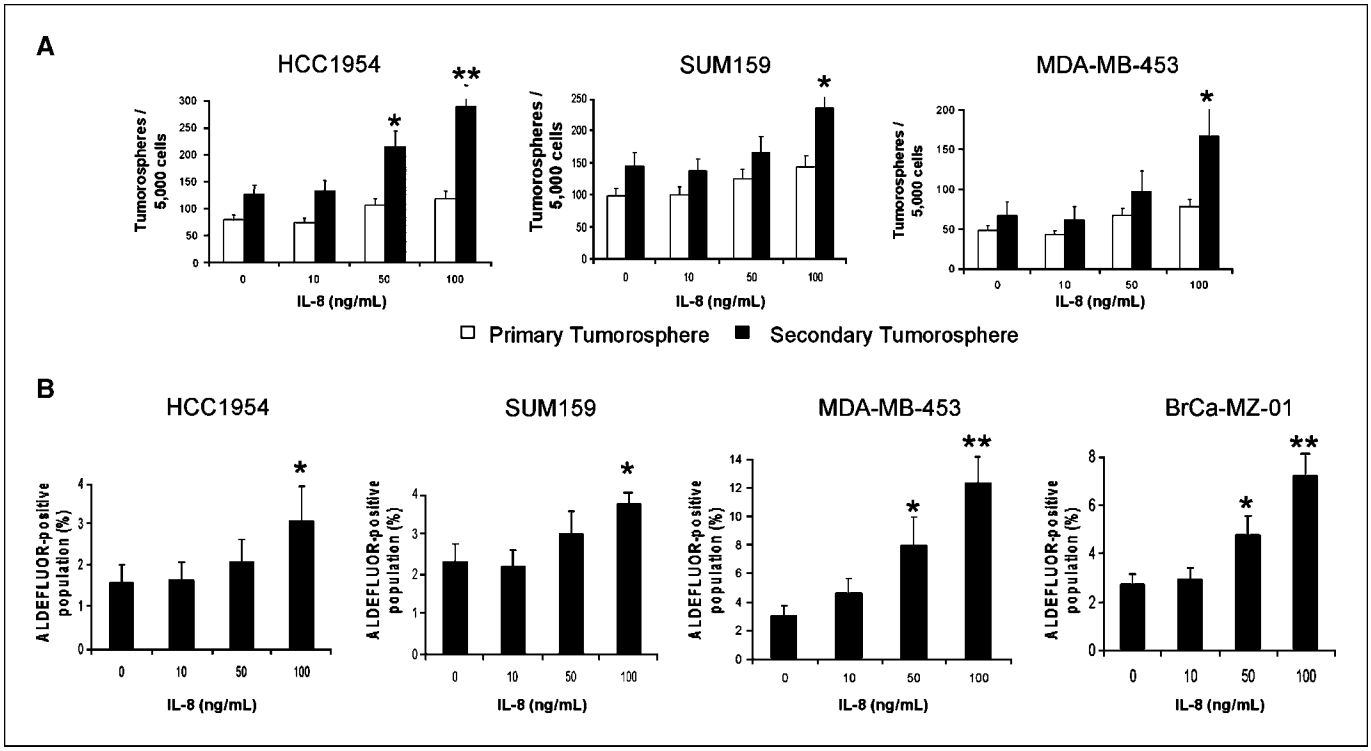


**Figure 4.** Role of the IL-8/CXCR1 axis in the regulation of breast CSCs. *A,* effect of IL-8 treatment on tumorosphere formation of three different cell lines (HCC1954, SUM159, MDA-MB-453). IL-8 treatment increased the formation of primary and secondary tumorospheres in a dose-dependent manner. *B,* effect of IL-8 treatment on the ALDEFLUOR-positive population of four different cell lines cultured in adherent conditions. IL-8 increased the ALDEFLUOR-positive population in a dose-dependent manner in each of the four cell lines analyzed. *, $P < 0.05$; **, $P < 0.01$, statistically significant differences from the control group.

ALDEFLUOR-positive cells have increased invasive capacity compared with ALDEFLUOR-negative cells. To determine the relationship between ALDEFLUOR positivity and metastatic capacity, we infected HCC1954, MDA-MB-453, and SUM159 with a luciferase lentivirus reporter system. Luciferase-infected cells were sorted using the ALDEFLUOR assay and introduced to NOD/SCID mice by intracardiac injection. A suspension of 100,000 cells from each population was injected, and metastasis was assessed by bioluminescent imaging. Mice inoculated with ALDEFLUOR-positive cells developed metastases at different sites and displayed a higher photon flux emission than mice inoculated with unseparated cells, which developed no more than one metastasis per mouse, or mice inoculated with ALDEFLUOR-negative cells, which developed only occasional metastases limited to lymph nodes (Fig. 5*B* and *C* and Supplementary Fig. S10). Histologic sections confirmed the presence of metastases at these sites (Fig. 5*C*). Thus, the metastatic capacity of BCLs is predominantly mediated by CSCs contained in the ALDEFLUOR-positive population.

## Discussion

The hypothesis that tumors are organized in a cellular hierarchy driven by CSCs has fundamental implications for cancer biology, as well as clinical implications for the early detection, prevention, and treatment of cancer. Evidence for CSCs has largely relied on primary and early passage xenograft models (28, 29). However, the success of establishing breast tumor xenograft has been low particularly for certain molecular subtypes. In contrast to primary tumors, cell lines are available in unlimited quantities and provide only carcinomatous populations for molecular analysis without normal tissue and stroma. In breast cancer, a large number of immortalized cell lines have been produced, which represent the different molecular subtypes found in primary human breast cancers (2, 18). However, a fundamental question remains as to how closely these cell lines are able to recapitulate the biology of human breast cancer.

***In vivo* evidence for stem cells in cell lines.** Recent studies have suggested that although cell lines may be clonally derived, they contain a cellular hierarchy representing different stages of cellular differentiation. Several studies have used markers, such as $CD44^+/CD24^-$ to identify CSC within breast cancer cell lines. However, their utility is limited by the observation that, frequently, a large percentage of cells within a cell line express these putative stem cell markers. For example, >90% of cells in basal breast cancer cell lines display the $CD44^+/CD24^-$ phenotype. Indeed, the $CD44^+/CD24^-$ phenotype did not isolate the tumorigenic population of these cell lines (8). An alternative approach has been to use the SP from cell lines. However, functional studies using Hoechst staining are limited by the toxicity of this agent (30). There is also evidence that the functional stem cell activity is not contained within the SP (31).

ALDH activity assessed by the ALDEFLUOR assay isolates cells with stem cell properties from various cancers (13). We show here that 23 of 33 BCLs (predominantly basal cell lines) contain an ALDEFLUOR-positive population. Lack of an ALDEFLUOR-positive population in some luminal BCLs may indicate that these luminal BCLs are derived from ALDEFLUOR-negative progenitor cells.

We used *in vivo* assays in NOD/SCID mice to show the stem cell properties of the ALDEFLUOR-positive populations. Self-renewal was shown by serial passage in NOD/SCID mice, and differentiation was shown by the ability of ALDEFLUOR-positive but not

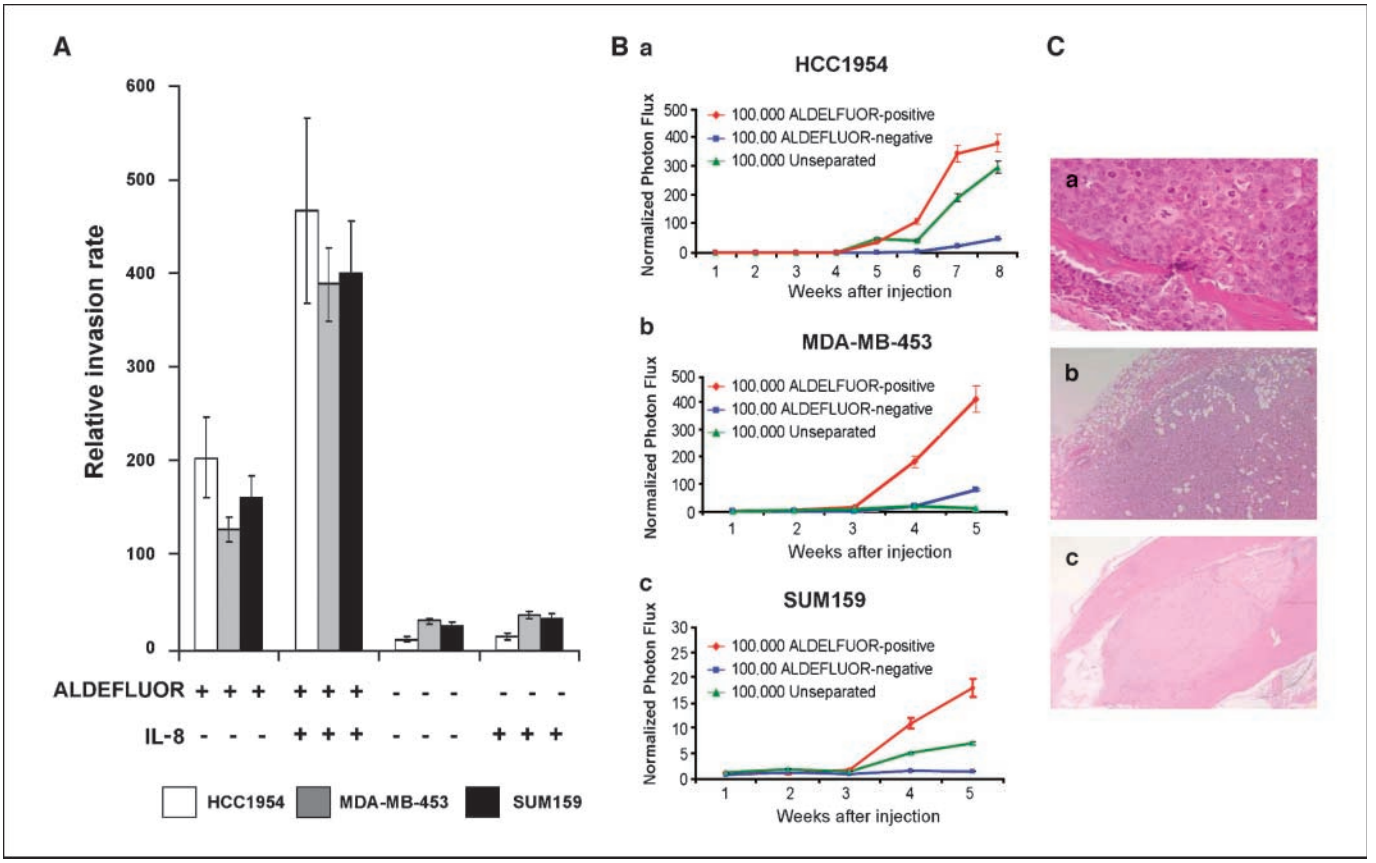


**Figure 5.** ALDEFLUOR-positive cells display increased metastatic potential. *A,* the IL-8/CXCR1 axis is involved in CSC invasion. The role of the IL-8/CXCR1 axis in invasion was assessed by a Matrigel invasion assay using serum or IL-8 as attractant for three different cell lines (HCC1954, MDA-MB-453, SUM159). ALDEFLUOR-positive cells were 6-fold to 20-fold more invasive than ALDEFLUOR-negative cells ($P < 0.01$). When using IL-8 (100 ng/mL) as attractant, we observed a significant increase of ALDEFLUOR-positive cells invading through Matrigel compared with serum as attractant ($P < 0.05$). In contrast, IL-8 had no effect on the invasive capacity of the ALDEFLUOR-negative population. *B* and *C,* the ALDEFLUOR-positive population displayed increased metastatic potential. *Ba–c,* quantification of the normalized photon flux measured at weekly intervals after inoculation of 100,000 luciferase-infected cells from each group (ALDEFLUOR-positive, ALDEFLUOR-negative, unseparated). Mice inoculated with ALDEFLUOR-positive cells developed several metastasis localized at different sites (bone, muscle, lung, soft tissue) and displayed a higher photon flux emission than mice inoculated with unseparated cells, which developed no more than one metastasis per mouse. In contrast, mice inoculated with ALDEFLUOR-negative cells developed only an occasional small metastasis, which was limited to lymph nodes. *C,* histologic confirmation, by H&E staining, of metastasis in bone (*Ca*), soft tissue (*Cb*), and muscle (*Cc*) resulting from injection of ALDEFLUOR-positive cells.

ALDEFLUOR-negative cells to regenerate the cellular heterogeneity of the initial tumor.

**A breast CSC signature.** Utilizing eight BCLs, we identified 413 genes whose expression discriminates ALDEFLUOR-positive and ALDEFLUOR-negative cells. This signature contained a number of genes known to play a role in stem cell biology. Genes overexpressed in the ALDEFLUOR-positive population include *NOTCH2* (*Notch homologue 2*), which regulates self-renewal and differentiation of mammary stem cells (16, 32); *NFYA*, known to regulate self-renewal and differentiation of stem cells (33); pecanex homologue *PCNX, RBM15/OTT*, which plays a pleiotropic role in hematopoietic stem cells (34) and affects myeloid differentiation via NOTCH signaling (35); and homeobox-like factor *TPRXL* involved in embryonic development, *ST3GAL3,* which codes for a stage-specific embryonic antigen-4 synthase, associated with fetal development and renal and gastric carcinogenesis (36). Notably, stage-specific embryonic antigen-4 protein (SSEA-4) is expressed in stem cell populations, such as $CXCR4^+/CD133^+/CD34^+/lin^-$ stem cells in human cord blood and quiescent mammary stem cells (37).

Genes underexpressed in the ALDEFLUOR-positive population are involved in cell differentiation, apoptosis, and mitochondrial oxidation. They include genes coding for nascent polypeptide-associated complex α subunit *NACA*, programmed death proteins *PDCD5* and *PDCD10*, mitochondrial ribosomal protein *MRPL41*, which induces apoptosis through P53-dependent and P53-independent manner via BCL2 and caspases, and proteins involved in mitochondrial processes, such as oxidative phosphorylation (NDUFA2, ATP5J2, IMMP1L) and protein synthesis in the mitochondrion (MRPL42, MRPL47, MRPL54, MRPS23). Down-regulation of apoptotic genes in CSCs may play a role in the resistance of these cells to radiation and chemotherapy (38). *ALDH1A1* was not identified as a differentially expressed gene in the ALDEFLUOR-positive signature. However, examination of gene expression profile of individual BCLs revealed that, although some showed differential expression of *ALDH1A1* in the ALDEFLUOR-positive population, others showed differential expression of *ALDH1A3*, a different ALDH isoform in this population. This suggests that the expression of different ALDH isoforms could contribute to the ALDEFLUOR-positive phenotype.

**From chemokines to stemokines.** The expression of CXCR1, a receptor for IL-8, is increased in a variety of cancers (39, 40). Although IL-8 expression is associated with ER-negative breast

cancer (41), this chemokine has not previously been reported to play a role in stem cell function. Its implication in the regulation of growth and metastasis is well-established in androgen-independent prostate cancer (42). Furthermore, the expression level of IL-8 is associated with tumorigenicity and metastasis through vascular endothelial growth factor production and angiogenesis (43, 44). We validated the gene expression data in three ways. First, qRT-PCR analysis confirmed a significant increase of *CXCR1* mRNA in ALDEFLUOR-positive populations from cell lines both included and not included in profiling analysis. Second, we showed, using flow cytometry, that CXCR1-containing cells were found exclusively within the ALDEFLUOR-positive population. Third, recombinant IL-8 increased mammosphere formation and the percentage of ALDEFLUOR-positive cells in BCLs. The IL-8/CXCR1 axis may, thus, regulate mammary stem cell proliferation or self-renewal. Because endothelial and stromal cells secrete IL-8, this chemokine may play a role in mediating interactions between tumor stem cells and the tumor microenvironment.

Recent studies have suggested a role for interleukins/chemokines in the regulation of CSCs (45, 46). This includes a role for IL-6 in breast CSCs and IL-4 in mediating chemoresistance of colon CSCs (46–48). These factors may be involved in the association between inflammation and cancer. This also includes a role for CCL5 (RANTES), a chemokine secreted by mesenchymal stem cells, which acts as a paracrine factor to enhance breast cancer cell motility, invasion, and metastasis (45).

**The roots of metastasis.** CSCs may be responsible for mediating tumor metastasis. A link between CSC and metastasis was first suggested with the identification of stem cell genes in an 11-gene signature generated using comparative profiles of metastatic and primary tumors in a transgenic mouse model of prostate cancer and in cancer patients (49). This signature was also a powerful predictor of disease recurrence, death after therapy, and distant metastasis in a variety of cancer types. We have shown that ALDEFLUOR-positive cells are more metastatic than ALDEFLUOR-negative cells and that IL-8, previously reported to play a role in tumor metastasis, promotes the invasion and chemotaxis of CSCs, which preferentially express the IL-8 receptor CXCR1. The ability to isolate metastatic CSC from cell lines should facilitate studies of the molecular mechanisms by which CSCs mediate tumor metastasis.

## Disclosure of Potential Conflicts of Interest

M.S. Wicha has financial holdings in and is a scientific advisor for OncoMed Pharmaceuticals. The other authors disclosed no potential conflicts of interest.

## Acknowledgments

Received 7/16/2008; revised 10/17/2008; accepted 11/11/2008; published OnlineFirst 02/03/2009.

**Grant support:** NIH grants CA66233 and CA101860, University of Michigan Cancer Center NIH support grant 5 P 30 CA46592, INSERM, Institut Paoli-Calmettes, Taubman Institute, Ministries of Health and Research grants PHRC 2006 and PHRC2007, Institut National du Cancer-Institut Lilly grant AO 2006, Institut National du Cancer grants PL2005 and ACI2007, and Ligue Nationale Contre le Cancer Label DB.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank the University of Michigan Cancer Center Flow Cytometry core.

## References

1. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100:57–70.
2. Neve RM, Chin K, Fridlyand J, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 2006;10:515–27.
3. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med 1997;3:730–7.
4. Glinsky GV. Stem cell origin of death-from-cancer phenotypes of human prostate and breast cancers. Stem Cell Rev 2007;3:79–93.
5. Jaiswal S, Traver D, Miyamoto T, Akashi K, Lagasse E, Weissman IL. Expression of BCR/ABL and BCL-2 in myeloid progenitors leads to myeloid leukemias. Proc Natl Acad Sci U S A 2003;100:10002–7.
6. Krivtsov AV, Twomey D, Feng Z, et al. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature 2006;442:818–22.
7. Christgen M, Ballmaier M, Bruchhardt H, von Wasielewski R, Kreipe H, Lehmann U. Identification of a distinct side population of cancer cells in the Cal-51 human breast carcinoma cell line. Mol Cell Biochem 2007;306:201–12.
8. Fillmore CM, Kuperwasser C. Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. Breast Cancer Res 2008;10:R25.
9. Kondo T, Setoguchi T, Taga T. Persistence of a small subpopulation of cancer stem-like cells in the C6 glioma cell line. Proc Natl Acad Sci U S A 2004;101:781–6.
10. Setoguchi T, Taga T, Kondo T. Cancer stem cells persist in many cancer cell lines. Cell Cycle 2004;3:414–5.
11. Patrawala L, Calhoun T, Schneider-Broussard R, Zhou J, Claypool K, Tang DG. Side population is enriched in tumorigenic, stem-like cancer cells, whereas ABCG$^{2+}$ and A. Cancer Res 2005;65:6207–19.
12. Chute JP, Muramoto GG, Whitesides J, et al. Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells. Proc Natl Acad Sci U S A 2006;103:11707–12.
13. Cheung AM, Wan TS, Leung JC, et al. Aldehyde dehydrogenase activity in leukemic blasts defines a subgroup of acute myeloid leukemia with adverse prognosis and superior NOD/SCID engrafting potential. Leukemia 2007;21:1423–30.
14. Corti S, Locatelli F, Papadimitriou D, et al. Identification of a primitive brain-derived neural stem cell population based on aldehyde dehydrogenase activity. Stem Cells 2006;24:975–85.
15. Ginestier C, Hur MH, Charafe-Jauffret E, et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell 2007;1:555–67.
16. Dontu G, Abdallah WM, Foley JM, et al. *In vitro* propagation and transcriptional profiling of human mammary stem/progenitor cells. Genes Dev 2003;17:1253–70.
17. Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4:249–64.
18. Charafe-Jauffret E, Ginestier C, Monville F, et al. Gene expression profiling of breast cell lines identifies potential new basal markers. Oncogene 2006;25:2273–84.
19. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95:14863–8.
20. Hua J, Balagurunathan Y, Chen Y, et al. Normalization benefits microarray-based classification. EURASIP. J Bioinform Syst Biol 2006;43056.
21. Ginestier C, Cervera N, Finetti P, et al. Prognosis and gene expression profiling of 20q13-amplified breast cancers. Clin Cancer Res 2006;12:4533–44.
22. Ponti D, Costa A, Zaffaroni N, et al. Isolation and *in vitro* propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. Cancer Res 2005;65:5506–11.
23. Ringe J, Strassburg S, Neumann K, et al. Towards *in situ* tissue repair: human mesenchymal stem cells express chemokine receptors CXCR1, CXCR2 and CCR2, and migrate upon stimulation with CXCL8 but not CCL2. J Cell Biochem 2007;101:135–46.
24. Hughes L, Malone C, Chumsri S, Burger AM, McDonnell S. Characterisation of breast cancer cell lines and establishment of a novel isogenic subclone to study migration, invasion and tumourigenicity. Clin Exp Metastasis 2008;25:549–57.
25. Itoh Y, Joh T, Tanida S, et al. IL-8 promotes cell proliferation and migration through metalloproteinase-cleavage proHB-EGF in human colon carcinoma cells. Cytokine 2005;29:275–82.
26. Gupta GP, Perk J, Acharyya S, et al. ID genes mediate tumor reinitiation during breast cancer lung metastasis. Proc Natl Acad Sci U S A 2007;104:19506–11.
27. Li F, Tiede B, Massague J, Kang Y. Beyond tumorigenesis: cancer stem cells in metastasis. Cell Res 2007;17:3–14.
28. Al Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A 2003;100:3983–8.
29. Ricci-Vitiani L, Lombardi DG, Pilozzi E, et al. Identification and expansion of human colon-cancer-initiating cells. Nature 2007;445:111–5.
30. Montanaro F, Liadaki K, Schienda J, Flint A, Gussoni E, Kunkel LM. Demystifying SP cell purification: viability, yield, and phenotype are defined by isolation parameters. Exp Cell Res 2004;298:144–54.
31. Stingl J, Eirew P, Ricketson I, et al. Purification and unique properties of mammary epithelial stem cells. Nature 2006;439:993–7.
32. Farnie G, Clarke RB. Mammary stem cells and breast cancer-role of Notch signalling. Stem Cell Rev 2007;3:169–75.
33. Zhu J, Zhang Y, Joe GJ, Pompetti R, Emerson SG. NF-Ya activates multiple hematopoietic stem cell (HSC) regulatory genes and promotes HSC self-renewal. Proc Natl Acad Sci U S A 2005;102:11728–33.

**34.** Raffel GD, Mercher T, Shigematsu H, et al. Ott1(Rbm15) has pleiotropic roles in hematopoietic development. Proc Natl Acad Sci U S A 2007;104:6001–6.

**35.** Ma X, Renda MJ, Wang L, et al. Rbm15 modulates Notch-induced transcriptional activation and affects myeloid differentiation. Mol Cell Biol 2007;27:3056–64.

**36.** Peiffer I, Eid P, Barbet R, et al. A sub-population of high proliferative potential-quiescent human mesenchymal stem cells is under the reversible control of interferon $\alpha/\beta$. Leukemia 2007;21:714–24.

**37.** Villadsen R, Fridriksdottir AJ, Ronnov-Jessen L, et al. Evidence for a stem cell hierarchy in the adult human breast. J Cell Biol 2007;177:87–101.

**38.** Hambardzumyan D, Becher OJ, Holland EC. Cancer stem cells and survival pathways. Cell Cycle 2008;7:1371–8.

**39.** Maxwell PJ, Gallagher R, Seaton A, et al. HIF-1 and NF-$\kappa$B-mediated upregulation of CXCR1 and CXCR2 expression promotes cell survival in hypoxic prostate cancer cells. Oncogene 2007;26:7333–45.

**40.** Murphy C, McGurk M, Pettigrew J, et al. Nonapical and cytoplasmic expression of interleukin-8, CXCR1, and CXCR2 correlates with cell proliferation and microvessel density in prostate cancer. Clin Cancer Res 2005;11:4117–27.

**41.** Freund A, Chauveau C, Brouillet JP, et al. IL-8 expression and its possible relationship with estrogen-receptor-negative status of breast cancer cells. Oncogene 2003;22:256–65.

**42.** Inoue K, Slaton JW, Eve BY, et al. Interleukin 8 expression regulates tumorigenicity and metastases in androgen-independent prostate cancer. Clin Cancer Res 2000;6:2104–19.

**43.** Balbay MD, Pettaway CA, Kuniyasu H, et al. Highly metastatic human prostate cancer growing within the prostate of athymic mice overexpresses vascular endothelial growth factor. Clin Cancer Res 1999;5:783–9.

**44.** Kim SJ, Uehara H, Karashima T, Mccarty M, Shih N, Fidler IJ. Expression of interleukin-8 correlates with angiogenesis, tumorigenicity, and metastasis of human prostate cancer cells implanted orthotopically in nude mice. Neoplasia 2001;3:33–42.

**45.** Karnoub AE, Dash AB, Vo AP, et al. Mesenchymal stem cells within tumour stroma promote breast cancer metastasis. Nature 2007;449:557–63.

**46.** Schafer ZT, Brugge JS. IL-6 involvement in epithelial cancers. J Clin Invest 2007;117:3660–3.

**47.** Todaro M, Alea MP, Di Stefano AB, et al. Colon cancer stem cells dictate tumor growth and resist cell death by production of interleukin-4. Cell Stem Cell 2007;1:389–402.

**48.** Sansone P, Storci G, Tavolari S, et al. IL-6 triggers malignant features in mammospheres from human ductal breast carcinoma and normal mammary gland. J Clin Invest 2007;117:3988–4002.

**49.** Glinsky GV, Berezovska O, Glinskii AB. Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer. J Clin Invest 2005;115:1503–21.

**50.** Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;286:531–7.