



Published in final edited form as:

Oncogene. 2014 January 23; 33(4): 440–448. doi:10.1038/onc.2012.609.

***Ink4a/Arf*^{-/-} and *HRAS(G12V)* transform mouse mammary cells into triple-negative breast cancer containing tumorigenic CD49f⁻ quiescent cells**

Kazuharu Kai^{1,2,3}, Takayuki Iwamoto², Takashi Kobayashi³, Yoshimi Arima³, Yayoi Takamoto³, Nobuyuki Ohnishi³, Chandra Bartholomeusz^{1,2}, Rie Horii⁴, Futoshi Akiyama⁴, Gabriel N Hortobagyi², Lajos Pusztai^{1,2}, Hideyuki Saya^{3,*}, and Naoto T Ueno^{1,2,*}

¹Breast Cancer Translational Research Laboratory, The University of Texas MD Anderson Cancer Center, Houston, TX, 77030

²Departments of Breast Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, 77030

³Division of Gene Regulation, Institute for Advanced Medical Research, School of Medicine, Keio University, Tokyo Japan 160-8582

⁴Division of Pathology, The Cancer Institute of the Japanese Foundation for Cancer Research, Tokyo Japan 135-8550

Abstract

Intratumoral heterogeneity within individual breast tumors is a well-known phenomenon that may contribute to drug resistance. This heterogeneity is dependent on several factors, such as types of oncogenic drivers and tumor precursor cells. The purpose of our study was to engineer a mouse mammary tumor model with intratumoral heterogeneity by using defined genetic perturbations. To achieve this, we used mice with knockout (*-/-*) of *Ink4a/Arf*, a tumor suppressor locus; these mice are known to be susceptible to non-mammary tumors such as fibrosarcoma. To induce mammary tumors, we retrovirally introduced an oncogene, *HRAS(G12V)*, into *Ink4a/Arf*^{-/-} mammary cells *in vitro*, and those cells were inoculated into syngeneic mice mammary fat pads. We observed 100% tumorigenesis. The tumors formed were negative for estrogen receptor, progesterone receptor, and HER2. Further, they had pathological features similar to those of human triple-negative breast cancer (e.g. pushing borders, central necrosis). The tumors were found to be heterogeneous and included two subpopulations: CD49f⁻ quiescent cells and CD49f⁺ cells. Contrary to our expectation, CD49f⁻ quiescent cells had high tumor-initiating potential and CD49f⁺ cells had relatively low tumor-initiating potential. Gene expression analysis revealed that CD49f⁻ quiescent cells overexpressed epithelial-to-mesenchymal transition-driving genes, reminiscent of tumor-initiating cells and claudin-low breast cancer. Our animal model with intratumoral heterogeneity, derived from defined genetic perturbations, allows us to test novel

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

*To whom correspondence should be addressed: Naoto T. Ueno, M.D., Ph.D., F.A.C.P., nueno@mdanderson.org; Tel: (713) 792-8754; Fax: (713) 794-4385 Or Hideyuki Saya, M.D., Ph.D., hsaya@a5.keio.jp; Tel: +81(3) 5363-3981; Fax: +81(3) 5363-3982.

Conflict of interest: The authors have nothing to disclose.

molecular targeted drugs in a setting that mimics the intratumoral heterogeneity of human triple-negative breast cancer.

Keywords

Ink4a/Arf; HRAS(G12V); Triple negative breast cancer; Intratumoral heterogeneity; Tumor-initiating cell

Introduction

Tumors are heterogeneous in various ways, including their pathology, transcriptional and immunophenotypic profiles, and genetic makeup. Recent advancement in array technologies allows us to understand intertumoral heterogeneity (heterogeneity between tumors) through genome-wide analyses—e.g., oligonucleotide microarray analysis and array complementary genomic hybridization (CGH) (1, 2).

Similar distinctions are seen in intratumoral heterogeneity (heterogeneity within a tumor), a leading area of interest in cancer research during the last decade (3). One facet through which to understand intratumoral heterogeneity is cellular heterogeneity, alternatively referred to as cancer stem cell (CSC) hierarchy (4): the concept that cells have differing capacities to initiate tumors. Accumulating evidence in the CSC research field shows that intratumoral heterogeneity is consistent for specific tumor types and can be defined with tumor type-specific CSC markers (e.g. human breast cancer CSC, CD4⁺ CD24^{-low} (5, 6); human glioblastoma CSC, CD13⁺ (7)). However, there are also some exceptions to observed CSC hierarchy. For example, CD133⁻ glioblastoma cells can initiate tumors under certain conditions (8). Furthermore, in a study of genetically engineered mouse mammary tumor models, the CSC fractions of each model had distinct cell surface marker profiles (9–11). Given that in each genetically engineered mouse, mammary tumorigenesis occurs through genetic perturbation (i.e. trans- or knockout genes), intratumoral heterogeneity is also assumed to be affected by those perturbations. Clinically, this intratumoral heterogeneity could result in a lack of tumor response to many novel targeted therapies (6, 12). A mouse mammary tumor model with intratumoral heterogeneity induced by defined genetic changes would enable us to understand what is contributing to resistance to certain targeted therapies. The purpose of the present study was to determine whether intratumoral heterogeneity could be observed in our mouse mammary tumor model.

Ink4a/Arf is a tumor suppressive locus and transcribes p16^{Ink4a} and p19^{Arf} in response to physiological stresses such as oncogenic stress and oxidative stress, then consequently elicits apoptosis or senescence in non-cancer cells (13). Previous studies on genetic engineering mouse models showed that knockout of *Ink4a/Arf* didn't increase an incidence of mammary tumors (14). Meanwhile, *Ink4a/Arf*^{-/-} did accelerate tumor relapses in a drug-inducible mammary tumor model, which had similar pathological features to triple-negative breast cancer (TNBC), a subtype of human breast cancers (15). Therefore, we assumed that *Ink4a/Arf*^{-/-} is a suitable genetic background to induce mammary tumors with an additional oncogene, which allows us to mimic human breast cancer and to observe intratumoral heterogeneity in those models. In order to achieve this, we first treated *Ink4a/Arf*^{-/-} mice

mammary cells with oncogenic *HRAS(G12V)* *in vitro* (16). We picked this oncogene because the RAS pathway is activated by transmembrane receptor tyrosine kinase activation in many breast cancers although mutated *HRAS* is not commonly observed in breast cancer (17, 18).

The *HRAS(G12V)*-transformed cells formed tumors in 100% of the mice when injected into syngeneic mice. These tumors were negative for estrogen receptor, progesterone receptor, and HER2 and similar to human TNBC. Further testing revealed that our new animal model does observe intratumoral cellular heterogeneity that is defined by the expression level of integrin $\alpha 6$ (CD49f), a putative marker of human and mouse mammary CSCs (11, 19).

Results

HRAS(G12V)* transformed *Ink4a/Arf*^{-/-} mouse mammary cells formed lethal tumors *in vivo

As previously reported, we had never seen naturally occurring mammary tumors from *Ink4a/Arf*^{-/-} mice (14). Therefore, to induce mammary tumors, we added another oncogenic event to *Ink4a/Arf*^{-/-} mouse mammary cells. Applying *in vitro* retroviral gene transfer methods, we used *Ink4a/Arf*^{-/-} mammosphere cells as target cells to be transformed (Fig. 1). The formation of cell masses called mammospheres is a well-established method for purifying primitive MECs (19, 20). We confirmed that mammosphere cells were almost all composed of lineage-negative cells (Supplementary Fig. S1A). The lineage-negative cells fully committed to luminal or basal epithelial cell lineages in a mutually exclusive manner (Supplementary Fig. S1B) and to a milk-producing cell lineage under prolactin stimulation (Supplementary Fig. S1C). Having thus confirmed that mammospheres were composed of primitive mammary epithelial cells (MECs), we then retrovirally introduced an oncogene *HRAS(G12V)* and a mock control into *Ink4a/Arf*^{-/-} or *Ink4a/Arf*^{+/+} mammosphere cells *in vitro*. We chose *HRAS(G12V)* as a candidate oncogenic driver because of the robustness for transforming epithelial cells into tumors and a potential role in human TNBC.

Next, mammosphere-derived cells, either retrovirus-infected or not-infected, were inoculated into mammary fat pads of syngeneic recipient mice (Fig. 1). Infection ratios were comparable for each combination, with a range of 10.2–12.2% (Supplementary Fig. S2). Using this method, tumors developed only from *Ink4a/Arf*^{-/-} MECs with *HRAS(G12V)* but not with mock control, neither from *Ink4a/Arf*^{+/+} MECs with *HRAS(G12V)* (Fig. 2A, B). In *Ink4a/Arf*^{-/-} and *HRAS(G12V)*-driven tumors, overexpression of HRAS was confirmed at protein and mRNA levels (Fig. 2D, Supplementary Fig. S3, S4), and overexpressed HRAS expectedly had mutation at codon12, synonymized as G12V (Fig. 2D, Supplementary Fig. S5). *Ink4a/Arf*^{-/-} and *HRAS(G12V)*-driven tumors emerged within a month in all cases and were fatal. Thus, *Ink4a/Arf*^{-/-} and *HRAS(G12V)* constitute a robust genetic combination for inducing tumors in MECs *in vivo*.

***HRAS(G12V)* and *Ink4a/Arf*^{-/-} induced mammary tumors were pathologically similar to human triple-negative breast cancer**

Induced tumors had a high mitotic index and were positive for periodic-acid schiff (PAS) staining (Fig. 2C), suggesting that they are highly proliferative and comprised of epithelial-

derived mucus-producing cells_ENREF_17. The tumors had pathological characteristics of human TNBC—pushing borders and central necrosis (Fig. 2C) (21)—and were composed of spindle-shaped cells. Furthermore, induced tumors locally invaded adjacent organs, such as the vertebrae, peritoneum, and intestinum, and spontaneously metastasized to the lung and liver (data not shown).

In immunohistochemical analyses, keratin 14 and keratin 18, which are expressed in basal and luminal MECs (22), were patchy positive (Fig. 2F), suggesting intratumoral heterogeneity in cell lineages (basal or luminal). Meanwhile, E-cadherin was negative. Vimentin and SMA, markers of mesenchymal cells, were not expressed in the tumors (Supplementary Fig. S6) (23, 24). Those three markers' status is not exactly as seen in human TNBC (25). However, we could assume that the induced tumors lost epithelial characteristics but did not fully acquire a mesenchymal phenotype yet in the process of tumorigenesis.

Extracellular signal-regulated kinases (ERKs), a downstream of ras, were highly phosphorylated; in contrast, EGFR, an upstream of ras, was not expressed at all (Fig. 2F). Finally, these tumors were negative for ER, PR, and HER2 (Fig. 2E), which allowed us to determine these tumors as triple negative tumors. Taken together, these data suggest that *Ink4a/Arf*^{-/-} and *HRAS(G12V)* transformed the mouse MECs to triple-negative mammary tumors, which shared core characteristics of human TNBC.

Comparison of molecular features of induced mouse mammary tumors and human TNBC by microarray analysis

We then analyzed the molecular features of this tumorigenesis along with *Ink4a/Arf*^{-/-} and *HRAS(G12V)*. To this end, we used *Ink4a/Arf*^{+/+} mammospheres and tumorspheres, mammospheres from *Ink4a/Arf*^{-/-} and *HRAS(G12V)* induced tumors, as phenotypic representations of non-tumor and tumor. We then compared gene expression between *Ink4a/Arf*^{+/+} mammospheres and tumorspheres (Fig. 3A). We identified 185 overexpressed genes in the tumorspheres; all of these genes had human orthologues (false discovery rates 1.00E-07). We named this gene set “*Ink4a/Arf*^{-/-} plus *HRAS(G12V)*-driven genes” and analyzed it using IPA, Ingenuity's knowledge-based pathway analysis software (26). Leading pathways that were involved were biosynthesis of steroids, cell cycle-related (ATM signaling and G1/S checkpoint regulation), and interferon response-related (activation of IRF by cytosolic pattern recognition receptors and interferon signaling) (Fig. 3B).

Given that induced tumors had some pathological findings similar to those of human TNBC, we next analyzed the correlation between the *Ink4a/Arf*^{-/-} plus *HRAS(G12V)*-driven genes and human TNBC. We performed unsupervised hierarchical clustering analysis of human breast cancer clinical data sets Transbig (27) and Wang (28) using *Ink4a/Arf*^{-/-} plus *HRAS(G12V)*-driven genes (Fig. 3C). Through this clustering, we identified two gene clusters that were highly correlated (correlation>0.4) with human TNBC samples: cell cycle genes and interferon response genes (Supplementary Table S1). To confirm that these gene sets truly reflect the intrinsic aggressiveness of human TNBC, we analyzed their prognostic significance by using another human breast cancer data set, the Mainz data set (29). Then, only the cell cycle metagene had a significant prognostic impact; the interferon response

metagene did not (Supplementary Fig. S7). Thus, we can conclude that *Ink4a/Arf*^{-/-} plus *HRAS(G12V)*-driven tumors have at least a common molecular feature with human TNBC, which can be functionally annotated as proliferation or cell-cycle related.

***Ink4a/Arf*^{-/-} and *HRAS(G12V)*-driven tumors had intratumoral heterogeneity: highly tumorigenic CD49f⁻ quiescent cells and low-tumorigenic CD49f⁺ cells**

We next evaluated intratumoral heterogeneity by assessing expression levels of tumor-initiating cell (TIC) markers. We analyzed the expression levels of candidate TIC markers CD29, CD24, CD49f, and CD44 in induced tumors, and, as a control, in primary MECs (5, 10, 11, 19, 30, 31) (Fig. 4A, Supplementary Fig. S8). Further, we characterized cell surface markers CD61 and Sca1, markers of mouse mammary progenitor cells (9, 11, 32), in the tumors (Supplementary Fig. S8). In comparison with primary MECs, in the induced tumors there were fewer CD24⁺ cells and more CD49f⁺ and CD44⁺ cells. In contrast, CD29 was positive in both primary MECs and tumors. Meanwhile, Sca1 and CD61 were positive in almost all tumor cells. We then performed a tumor cell inoculation assay in a limiting dilution manner with regard to each fraction (positive and negative) of each marker. CD29 was excluded because of its small range of expression. Lin⁺ cell-depleted GFP⁺ (Lin⁻GFP⁺) cells were sorted according to the expression level of each marker (CD44⁻, CD44⁺, CD24⁻, CD24⁺, CD49f⁻, and CD49f⁺) (Fig. 4A, B, and C), then inoculated into mammary fat pads of C57BL/6J mice at doses of 2,000, 500, 100, and 20 cells (Table 1). Unexpectedly, the tumor incidence from CD49f⁻ cells was much higher than that from CD49f⁺ cells or the other subpopulations at 100-cell inoculation. The higher tumor incidence from CD49f⁻ cells was inconsistent with previous reports that TICs can be enriched among CD49f⁺ cells (11, 19).

To better understand this phenomenon, there are two approaches. The first is to identify another positive TIC marker, and the second is to further analyze the CD49f⁻ fraction to clarify the intratumoral heterogeneity in this model. We chose the latter approach and evaluated the cell cycle status of the CD49f⁻ fraction. The presence of quiescent TICs has been proposed for some tumors, such as chronic myeloid leukemia (33). We hypothesized that tumor incidence from CD49f⁻ cells was attributable to quiescent tumor cells. To test this hypothesis, we further fractionated the CD49f⁻ cell population into CD49f⁻ quiescent and CD49f⁻ dividing cells (Fig. 4D) by staining with Hoechst 33342 and pyronin Y; quiescent cells were identified by Hoechst-Red^{low} and pyronin Y^{low} staining, and dividing cells, by Hoechst-Red^{high} and pyronin Y^{high} staining (33). We inoculated the subfractions into recipient mice. As expected, CD49f⁻ quiescent cells had higher tumorigenic activity than did CD49f⁻ dividing cells (Table 1).

CD49f⁻ quiescent cells overexpressed epithelial-to-mesenchymal transition-related genes in a gene expression analysis

Next, to understand this intratumoral heterogeneity at the molecular level, we performed expression analysis among the two subfractions: highly tumorigenic cells (CD49f⁻ quiescent) and cells with relatively low tumorigenicity (CD49f⁺). We sorted these paired fractions from a sequential series of three tumors. The genome-wide expression was compared by paired t-test between CD49f⁻ quiescent cells and CD49f⁺ cells. Then, 106

overexpressed genes (false discovery rate $<9.34E-04$) and 93 underexpressed genes (false discovery rate $<9.88E-04$) in CD49⁻ quiescent cells were analyzed with IPA pathway analysis software to understand their biological significance (Fig. 5A, B and Supplementary Table S2). The key components of overexpressed genes in CD49⁻ quiescent cells were collagen family proteins, MMP2, integrin- α V β 3, and ITGB5, which were annotated as the cellular assembly and organization network in IPA network analysis (Fig. 5C). Interestingly, CD49⁻ quiescent cells also overexpressed Thy1.1 (Supplementary Table S2), a marker of TICs in Wnt transgenic mouse mammary tumors (11). This finding suggests that we might further purify TICs with Thy1.1. Among underexpressed genes in CD49⁻ quiescent cells, as expected, cell cycle-driving genes were significantly downregulated (Fig. 5B).

Canonical pathway analysis is another knowledge-based output style of IPA; it elucidates the bioprocesses of a gene set of interest. Through this analysis, overexpressed genes in CD49⁻ quiescent cells were shown to be highly correlated with the bioprocess of hepatic stellate cell activation (Fig. 5A) (34–36), which is known to make hepatic stellate cells trans-differentiate to myofibroblast cells and results in liver fibrosis due to accumulation of collagen. Further, most of the components in this bioprocess are closely related to epithelial-to-mesenchymal transition (EMT). CD49⁻ quiescent cells overexpressed several kinds of collagens, a hallmark of EMT. Those collagens and their upstream molecules, such as TGF- β , IGF-1, and endothelin receptor type A, were also overexpressed in CD49⁻ quiescent cells (Supplementary Fig. S9) (35).

Discussion

Through genetic manipulation of mouse mammary cells *in vitro* and subsequent inoculation into recipient mice, we successfully induced the mouse TNBC model with a unique intratumoral heterogeneity, represented by both CD49⁻ quiescent TICs and CD49⁺ non-TICs. We were able to induce mammary tumors from *Ink4a/Arf*^{-/-} mouse mammary cells with *HRAS(G12V)* that have such intratumoral heterogeneity.

Previously reported mouse mammary tumor models produced spindle cell tumors in reproducible ratios; about 50% of p53-null mouse mammary tumors are spindle cell tumors although those tumors are molecularly heterogeneous (37, 38). The resultant tumor phenotype of our *Ink4a/Arf*^{-/-} and *HRAS(G12V)* model was similar to the phenotype of the spindle cell tumor model, recently called “EMT-type tumor” model (23), which also has triple-negative receptor features. The EMT-type phenotype in our model was relatively consistent or homogeneous, suggesting that *Ink4a/Arf*^{-/-} plus *HRAS(G12V)* strongly dictated this phenotype.

The populations of cells that were most tumorigenic varied in mouse mammary tumor models with intratumoral heterogeneity. In MMTV-Wnt transgenic mice, luminal progenitor-like CD61⁺ tumor cells had relatively higher tumorigenic activity than that of CD61⁻ tumor cells (9). In contrast, in a p53-null mouse mammary tumor model, mammary stem cell-like CD29^{high}CD24^{high} tumor cells had the highest tumorigenic activity (10). Combined with our data, these reports suggest that intratumoral heterogeneity partially depends on the type of genetic perturbation that drives the tumors.

Our model's unique intratumoral heterogeneity, whereby CD49f⁻ quiescent cells had much higher tumorigenicity compared to CD49f⁺ cells, however, is inconsistent with the findings from known mouse mammary tumor models. In the previously reported models, CD49f⁺ cells had relatively higher tumor-initiating potential than did CD49f⁻ cells, a pattern reminiscent of the profile of normal mammary stem cells (11, 30). The impact of the cell of origin on this intratumoral heterogeneity in the present model is unknown, however, we may at least infer that *Ink4a/Arf*^{-/-} plus *HRAS(G12V)* is sufficient to perturb stem-cell hierarchy in the axis of CD49f expression. In fact, this was the case in BRCA1-deficient plus p53^{+/-} mouse mammary tumor model. In that model, when BRCA1 was enforced to be deleted in luminal progenitor cells, CD24⁻ cells were revealed as TICs, which is inconsistent to normal stem cell hierarchy in mouse mammary glands; mouse mammary stem cells are CD24⁺. Together with our finding, certain oncogenic factors are supposed to have an impact even for changing stem cell hierarchy in occurred tumors. Consequently, this notion has been turned on the reason why we refer CD49f⁻ quiescent cells as TICs but not cancer stem cells, and rather designate intratumoral heterogeneity than CSC hierarchy.

We molecularly delineated this intratumoral heterogeneity by comparing the gene expression of CD49f⁻ quiescent TICs and CD49f⁺ non-TICs. This molecular heterogeneity was represented by EMT-related genes, a pattern that was annotated as the hepatic fibrosis bioprocess in IPA analysis. The representative components of EMT-related genes include TGF-β, IGF-1, and MMP2 (39). Those are also the key components expressed in the claudin-low subtype in human breast cancers. This subtype was discovered through unsupervised hierarchical clustering analysis of human and mouse mammary tumor panels (40). It expressed lower levels of E-cadherin and tight junction proteins, including claudin 3, and is thought to be a subtype of TNBC (41). The claudin-low gene signature, which defines the claudin-low subtype, was enriched in TICs and treatment-resistant residual tumor cells; these cells are putative molecular targets (12). Interestingly, in our model, TICs shared some molecular features with those of the human claudin-low subtype, although our TIC cell surface markers were analogous to neither those of human breast cancer nor those of other mouse mammary tumor models.

In summary, we induced intratumorally heterogeneous, highly proliferative triple-negative mouse mammary tumors with the genetic combination of *Ink4a/Arf*^{-/-} and *HRAS(G12V)*. Although, as with other mouse models, this model still has the technical gap of the cell of origin being unknown, the recapitulated intratumoral heterogeneity, represented by the components of the claudin-low gene signature, allows us to identify molecular features of TICs. Further, our model may offer us the opportunity to develop drugs that can target the RAS-mediated pathway (e.g. MEK inhibitor, ADZ6244) (42) or TICs.

Materials and Methods

Animals

We bred and maintained *Ink4a/Arf*^{-/-} mice (strain B6.129-Cdkn2atm1Rdp; NCI Frederick) in our animal facility at Keio University, Tokyo, Japan (43, 44). During a last couple of years, this line had been backcrossed into C57BL/6 through at least 5 generations. C57BL/6J mice were purchased from Oriental Yeast Co., Ltd. and used as recipients. All experiments

were approved by the Animal Research Ethics Committee of Keio University School of Medicine, Tokyo, Japan. The care for all mice described above was in accordance with the institution's guidelines.

Mammary cell preparations

We dissected mammary glands from 6- to 7-week-old female mice and tumors in recipient mice. After mincing with scissors, the tissue was digested with collagenase and hyaluronidase (StemCell Technologies) in DMEM/F12 (Sigma) supplemented with 5% fetal calf serum (FCS), 5 $\mu\text{g ml}^{-1}$ insulin, and 20 ng ml^{-1} epidermal growth factor (EGF) for 3 to 4 hr at 37 °C. After vortexing and lysis of the red blood cells in NH_4Cl , we sequentially digested the resulting organoid suspension with 0.25% trypsin (2 min, 37 °C) and then with 5 mg ml^{-1} dispase and 0.1 mg ml^{-1} DNase (Sigma) (2 min). We obtained a single-cell suspension by filtration through a 40- μm cell strainer (BD Falcon). The methods for *in vitro* cellular assay are described in the Supplementary Methods.

Retroviral gene transfer

pMX-HRAS(G12V)-IRES-GFP or pMXs-IRES-GFP (mock plasmid) were transfected into Plat-E retrovirus packaging cells using Fugene 6 transfection reagent (Roche) (45). pMXs-IRES-GFP was kindly provided by T. Kitamura. Virus particle-containing supernatants were centrifuged, and virus pellets were reconstituted with floating culture medium, DMEM/F12 supplemented with 20 ng/ml EGF (PeproTech), 10 ng/ml basic fibroblast growth factor (bFGF) (PeproTech), 5 $\mu\text{g/ml}$ insulin (Sigma), 100 units/ml penicillin G, and 100 $\mu\text{g/ml}$ streptomycin (GIBCO). Single suspended cells, derived from collected primary mammospheres using 40- μm cell strainers, were placed in virus-containing medium and plated into Ultra-Low Attachment dishes (Corning). To this end, the secondary mammospheres that contains retrovirus-infected cells were formed.

Antibodies

The following antibodies against mouse antigens were purchased from eBioscience: CD24-APC, CD29-APC, CD49f-APC, CD44-APC, Sca1-biotin, CD61-biotin, CD45-biotin, CD31-biotin, and TER119-biotin. We also obtained antibodies against HRAS (Clone C-20; Santa Cruz), phospho-ERK (Cell Signaling), milk (Nordic Immunological Laboratories), keratin 14 (Covance), keratin 18 (Progen Biotechnik), E-cadherin (BD Pharmingen), smooth muscle actin (SMA; Abcam), and vimentin (Sigma). Streptavidin APC and streptavidin APC-eFluor 780 were purchased from eBioscience. Fluorochrome-conjugated secondary antibodies included anti-rabbit and anti-mouse Ig-Alexa Fluor 594 and anti-rabbit Ig-Alexa Fluor 488 (Molecular Probes).

Cell labeling, flow cytometry, and sorting

To analyze the cell-cycle status, cells were first stained with 5 $\mu\text{g/ml}$ Hoechst 33342 at 37 °C for 30 min, followed by 1 $\mu\text{g/ml}$ pyronin Y at 37 °C for 30 min. Antibody incubations were performed at 4 °C for 30 min. Cells were resuspended in 0.5 $\mu\text{g ml}^{-1}$ propidium iodide before analysis. Data analysis was performed on the single, live cell gate using Flowjo software. Cell sorting was carried out on a FACSAria II sorter (Becton Dickinson).

***In vivo* transplantation**

For tumor initiation *in vivo*, sorted cells were resuspended in phosphate-buffered saline with 0.04% trypan blue and 50% FCS, and 20- μ l volumes were injected into the right inguinal glands of 4-week-old female mice.

Immunostaining

Paraffin-embedded sections were dewaxed, washed in phosphate-buffered saline, and subjected to antigen retrieval by boiling in 10 mM citrate buffer for 15 min before blocking. After blocking, we incubated the sections sequentially with the primary antibodies, biotinylated secondary antibodies (mouse-specific and rabbit-specific), ready-to-use Vectastain avidin and biotinylated horseradish peroxidase macromolecular complex (ABC) reagent (Vector Laboratories), and 3,3'-diaminobenzidine (DAKO); finally, we counterstained with hematoxylin. For mammospheres, we spun mammospheres onto glass slides with a Cytospin 4 centrifuge (Shandon, Thermo). The cells were fixed in 4% paraformaldehyde, then, incubated with 0.2% Triton X100 in PBS for 5 minutes, and blocked with goat serum for 30 min. Following procedures commencing primary antibody reaction were performed as previously described (46).

RT-PCR and mutation analysis

We extract whole RNAs from alive sample cells using RNeasy kit (Qiagen) and performed reverse transcription reactions with SuperScript III and the oligo (dT)₂₀ primer (Invitrogen). PCR was performed with NavaTaq Hot Start DNA polymerase (Novagen). Sequences of gene specific primers were as followings: *HRAS*, 5'GAGACCCTGTAGGAGGACC and 3'CATCAGGAGGGTTCAGCTTC; *GAPDH*, 5'TGAAGGTCGGTGTGAACGGATTTGGC and 3'CATGTAGGCCTAGAGGTCCACCAC. Same primer pairs were used for DNA sequencing. DNA sequencing on chain-termination method was done as previously described (47).

Microarray hybridizations

We purified total RNAs from mammospheres, tumorspheres (tumor-derived mammospheres), and sorted cell populations with Trizol and prepared them using a NucleoSpin RNA XS kit (Macherey-Nagel GmbH & Co. KG) according to the manufacturer's protocol. We quantified RNAs using a NanoDrop 1000 spectrophotometer and ascertained RNA quality with the Agilent 2100 Bioanalyzer (Agilent Technologies). We labeled 5 to 10 ng total RNAs and prepared biotinylated complementary RNAs according to the standard Affymetrix protocol (Expression Analysis Technical Manual, Affymetrix). After fragmentation, we hybridized cRNA to a GeneChipMouse 430 2.0 Genome Array for 16 hr at 45 °C. After washing, we stained the chips in an Affymetrix Fluidics Station 450 and scanned them using an Affymetrix scanner. Un-normalized summary probe profiles, with associated probe annotation, were output from Affymetrix GeneChip Operating Software, version 1.04.

Statistical analysis

Statistical significance was defined as a *P* value less than 0.05. All statistical tests and corresponding *P* values were two-sided. The data sets and the statistical methods used for the analyses of mouse and human microarray data are described in detail in the Supplementary Methods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial Support: This work was supported by grants from the Japan Society for the Promotion of Science, Japan (to K.K. and H.S.). K.K. was supported by the program “Strategic Young Researcher Overseas Visits Program for Accelerating Brain Circulation” by the Ministry of Education, Culture, Sports, Science and Technology, Japan. This research was also supported in part by the U.S. National Institutes of Health through MD Anderson’s Cancer Center Support Grant (5 P30 CA016672-36) and grant R01 CA123318-01A1 (to N.T.U.) and by the Nellie B. Connally Breast Cancer Research Fund.

We thank N. Suzuki, Y. Ito and S. Hayashi for excellent animal husbandry, I. Ishimatsu for expert assistance with histology and K. Hashimoto for expert help in the flow cytometry lab. We thank T. Suda and M. C. Hung for invaluable discussions. Thanks are also owed to M. Fujiwara for expert help with microarray analysis, T. Ishimoto for technical help with confocal microscopy and S. C. Patterson for editorial assistance.

References

- Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, Kuo WL, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell*. 2006 Dec; 10(6): 529–41. [PubMed: 17157792]
- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature*. 2000 Aug 17; 406(6797):747–52. [PubMed: 10963602]
- Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med*. 2012 Mar 8; 366(10):883–92. [PubMed: 22397650]
- Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. *Nature*. 2001 Nov 1; 414(6859):105–11. [PubMed: 11689955]
- 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 Apr 1; 100(7):3983–8. [PubMed: 12629218]
- Shipitsin M, Campbell LL, Argani P, Weremowicz S, Bloushtain-Qimron N, Yao J, et al. Molecular definition of breast tumor heterogeneity. *Cancer Cell*. 2007 Mar; 11(3):259–73. [PubMed: 17349583]
- Singh SK, Clarke ID, Terasaki M, Bonn VE, Hawkins C, Squire J, et al. Identification of a cancer stem cell in human brain tumors. *Cancer Res*. 2003 Sep 15; 63(18):5821–8. [PubMed: 14522905]
- Joo KM, Kim SY, Jin X, Song SY, Kong DS, Lee JI, et al. Clinical and biological implications of CD133-positive and CD133-negative cells in glioblastomas. *Lab Invest*. 2008 Aug; 88(8):808–15. [PubMed: 18560366]
- Vaillant F, Asselin-Labat ML, Shackleton M, Forrest NC, Lindeman GJ, Visvader JE. The mammary progenitor marker CD61/beta3 integrin identifies cancer stem cells in mouse models of mammary tumorigenesis. *Cancer Res*. 2008 Oct 1; 68(19):7711–7. [PubMed: 18829523]
- Zhang M, Behbod F, Atkinson RL, Landis MD, Kittrell F, Edwards D, et al. Identification of tumor-initiating cells in a p53-null mouse model of breast cancer. *Cancer Res*. 2008 Jun 15; 68(12):4674–82. [PubMed: 18559513]

11. Cho RW, Wang X, Diehn M, Shedden K, Chen GY, Sherlock G, et al. Isolation and molecular characterization of cancer stem cells in MMTV-Wnt-1 murine breast tumors. *Stem Cells*. 2008 Feb; 26(2):364–71. [PubMed: 17975224]
12. Creighton CJ, Li X, Landis M, Dixon JM, Neumeister VM, Sjolund A, et al. Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proc Natl Acad Sci U S A*. 2009 Aug 18; 106(33):13820–5. [PubMed: 19666588]
13. Sherr CJ. The INK4a/ARF network in tumour suppression. *Nat Rev Mol Cell Biol*. 2001 Oct; 2(10):731–7. [PubMed: 11584300]
14. Serrano M, Lee H, Chin L, Cordon-Cardo C, Beach D, DePinho RA. Role of the INK4a locus in tumor suppression and cell mortality. *Cell*. 1996 Apr 5; 85(1):27–37. [PubMed: 8620534]
15. Debies MT, Gestl SA, Mathers JL, Mikse OR, Leonard TL, Moody SE, et al. Tumor escape in a Wnt1-dependent mouse breast cancer model is enabled by p19Arf/p53 pathway lesions but not p16 Ink4a loss. *J Clin Invest*. 2008 Jan; 118(1):51–63. [PubMed: 18060046]
16. Seeburg PH, Colby WW, Capon DJ, Goeddel DV, Levinson AD. Biological properties of human c-Ha-ras1 genes mutated at codon 12. *Nature*. 1984 Nov 1–7; 312(5989):71–5. [PubMed: 6092966]
17. von Lintig FC, Dreilinger AD, Varki NM, Wallace AM, Casteel DE, Boss GR. Ras activation in human breast cancer. *Breast Cancer Res Treat*. 2000 Jul; 62(1):51–62. [PubMed: 10989985]
18. Eckert LB, Repasky GA, Ulku AS, McFall A, Zhou H, Sartor CI, et al. Involvement of Ras activation in human breast cancer cell signaling, invasion, and anoikis. *Cancer Res*. 2004 Jul 1; 64(13):4585–92. [PubMed: 15231670]
19. Pece S, Tosoni D, Confalonieri S, Mazzarol G, Vecchi M, Ronzoni S, et al. Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. *Cell*. 2010 Jan 8; 140(1):62–73. [PubMed: 20074520]
20. Liao MJ, Zhang CC, Zhou B, Zimonjic DB, Mani SA, Kaba M, et al. Enrichment of a population of mammary gland cells that form mammospheres and have in vivo repopulating activity. *Cancer Res*. 2007 Sep 1; 67(17):8131–8. [PubMed: 17804725]
21. Rakha EA, Ellis IO. Triple-negative/basal-like breast cancer: review. *Pathology*. 2009 Jan; 41(1):40–7. [PubMed: 19089739]
22. Shackleton M, Vaillant F, Simpson KJ, Stingl J, Smyth GK, Asselin-Labat ML, et al. Generation of a functional mammary gland from a single stem cell. *Nature*. 2006 Jan 5; 439(7072):84–8. [PubMed: 16397499]
23. Cardiff RD. The pathology of EMT in mouse mammary tumorigenesis. *J Mammary Gland Biol Neoplasia*. 2010 Jun; 15(2):225–33. [PubMed: 20521088]
24. Zhang D, Lafortune TA, Krishnamurthy S, Esteva FJ, Cristofanilli M, Liu P, et al. Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor Reverses Mesenchymal to Epithelial Phenotype and Inhibits Metastasis in Inflammatory Breast Cancer. *Clin Cancer Res*. 2009 Oct 13.
25. Sarrío D, Rodríguez-Pinilla SM, Hardisson D, Cano A, Moreno-Bueno G, Palacios J. Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype. *Cancer Res*. 2008 Feb 15; 68(4):989–97. [PubMed: 18281472]
26. Iwamoto T, Bianchini G, Booser D, Qi Y, Coutant C, Shiang CY, et al. Gene pathways associated with prognosis and chemotherapy sensitivity in molecular subtypes of breast cancer. *J Natl Cancer Inst*. 2011 Feb 2; 103(3):264–72. [PubMed: 21191116]
27. Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst*. 2006 Sep 6; 98(17):1183–92. [PubMed: 16954471]
28. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet*. 2005 Feb 19–25; 365(9460):671–9. [PubMed: 15721472]
29. Schmidt M, Böhm D, von Törne C, Steiner E, Puhl A, Pilch H, et al. The humoral immune system has a key prognostic impact in node-negative breast cancer. *Cancer Res*. 2008 Jul 1; 68(13):5405–13. [PubMed: 18593943]

30. Stingl J, Eirew P, Ricketson I, Shackleton M, Vaillant F, Choi D, et al. Purification and unique properties of mammary epithelial stem cells. *Nature*. 2006 Feb 23; 439(7079):993–7. [PubMed: 16395311]
31. Lim E, Vaillant F, Wu D, Forrest NC, Pal B, Hart AH, et al. Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med*. 2009 Aug; 15(8):907–13. [PubMed: 19648928]
32. Welm BE, Tepera SB, Venezia T, Graubert TA, Rosen JM, Goodell MA. Sca-1(pos) cells in the mouse mammary gland represent an enriched progenitor cell population. *Dev Biol*. 2002 May 1; 245(1):42–56. [PubMed: 11969254]
33. Ito K, Bernardi R, Morotti A, Matsuoka S, Saglio G, Ikeda Y, et al. PML targeting eradicates quiescent leukaemia-initiating cells. *Nature*. 2008 Jun 19; 453(7198):1072–8. [PubMed: 18469801]
34. Dooley S, Delvoux B, Lahme B, Mangasser-Stephan K, Gressner AM. Modulation of transforming growth factor beta response and signaling during transdifferentiation of rat hepatic stellate cells to myofibroblasts. *Hepatology*. 2000 May; 31(5):1094–106. [PubMed: 10796885]
35. Cho JJ, Hocher B, Herbst H, Jia JD, Ruehl M, Hahn EG, et al. An oral endothelin-A receptor antagonist blocks collagen synthesis and deposition in advanced rat liver fibrosis. *Gastroenterology*. 2000 Jun; 118(6):1169–78. [PubMed: 10833492]
36. Svegliati-Baroni G, Ridolfi F, Di Sario A, Casini A, Marucci L, Gaggiotti G, et al. Insulin and insulin-like growth factor-1 stimulate proliferation and type I collagen accumulation by human hepatic stellate cells: differential effects on signal transduction pathways. *Hepatology*. 1999 Jun; 29(6):1743–51. [PubMed: 10347117]
37. Cui XS, Donehower LA. Differential gene expression in mouse mammary adenocarcinomas in the presence and absence of wild type p53. *Oncogene*. 2000 Dec 7; 19(52):5988–96. [PubMed: 11146550]
38. Herschkowitz JI, Zhao W, Zhang M, Usary J, Murrow G, Edwards D, et al. Comparative oncogenomics identifies breast tumors enriched in functional tumor-initiating cells. *Proc Natl Acad Sci U S A*. 2012 Feb 21; 109(8):2778–83. [PubMed: 21633010]
39. Taube JH, Herschkowitz JI, Komurov K, Zhou AY, Gupta S, Yang J, et al. Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc Natl Acad Sci U S A*. 2010 Aug 31; 107(35):15449–54. [PubMed: 20713713]
40. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol*. 2007; 8(5):R76. [PubMed: 17493263]
41. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res*. 2010 Sep 2.12(5):R68. [PubMed: 20813035]
42. Bartholomeusz C, Oishi T, Saso H, Akar U, Liu P, Kondo K, et al. MEK1/2 inhibitor selumetinib (AZD6244) inhibits growth of ovarian clear cell carcinoma in a PEA-15-dependent manner in a mouse xenograft model. *Mol Cancer Ther*. 2012 Feb; 11(2):360–9. [PubMed: 22144664]
43. Shimizu T, Ishikawa T, Sugihara E, Kuninaka S, Miyamoto T, Mabuchi Y, et al. c-MYC overexpression with loss of Ink4a/Arf transforms bone marrow stromal cells into osteosarcoma accompanied by loss of adipogenesis. *Oncogene*. 2010 Oct 21; 29(42):5687–99. [PubMed: 20676132]
44. Sugihara E, Shimizu T, Kojima K, Onishi N, Kai K, Ishizawa J, et al. Ink4a and Arf are crucial factors in the determination of the cell of origin and the therapeutic sensitivity of Myc-induced mouse lymphoid tumor. *Oncogene*. 2012 Jun 7; 31(23):2849–61. [PubMed: 21986948]
45. Kitamura T, Koshino Y, Shibata F, Oki T, Nakajima H, Nosaka T, et al. Retrovirus-mediated gene transfer and expression cloning: powerful tools in functional genomics. *Exp Hematol*. 2003 Nov; 31(11):1007–14. [PubMed: 14585362]
46. Kai K, Nagano O, Sugihara E, Arima Y, Sampetean O, Ishimoto T, et al. Maintenance of HCT116 colon cancer cell line conforms to a stochastic model but not a cancer stem cell model. *Cancer Sci*. 2009 Dec; 100(12):2275–82. [PubMed: 19737148]

47. Kai K, Zhang Z, Yamashita H, Yamamoto Y, Miura Y, Iwase H. Loss of heterozygosity at the ATBF1-A locus located in the 16q22 minimal region in breast cancer. *BMC Cancer*. 2008; 8:262. [PubMed: 18796146]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

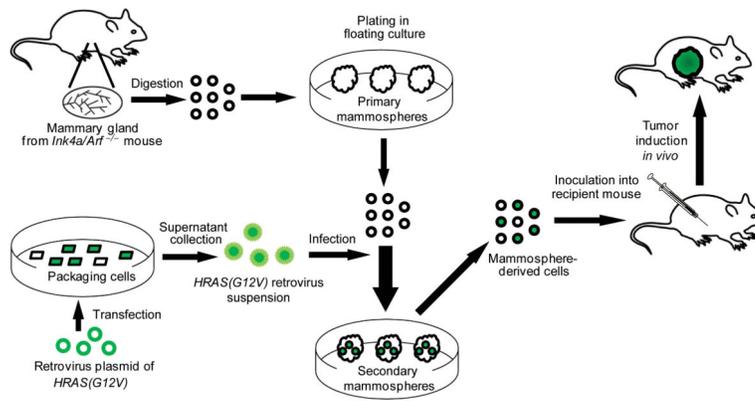


Figure 1.
Experimental strategy to induce mammary tumors from *Ink4a/Arf*^{-/-} mouse MECs.

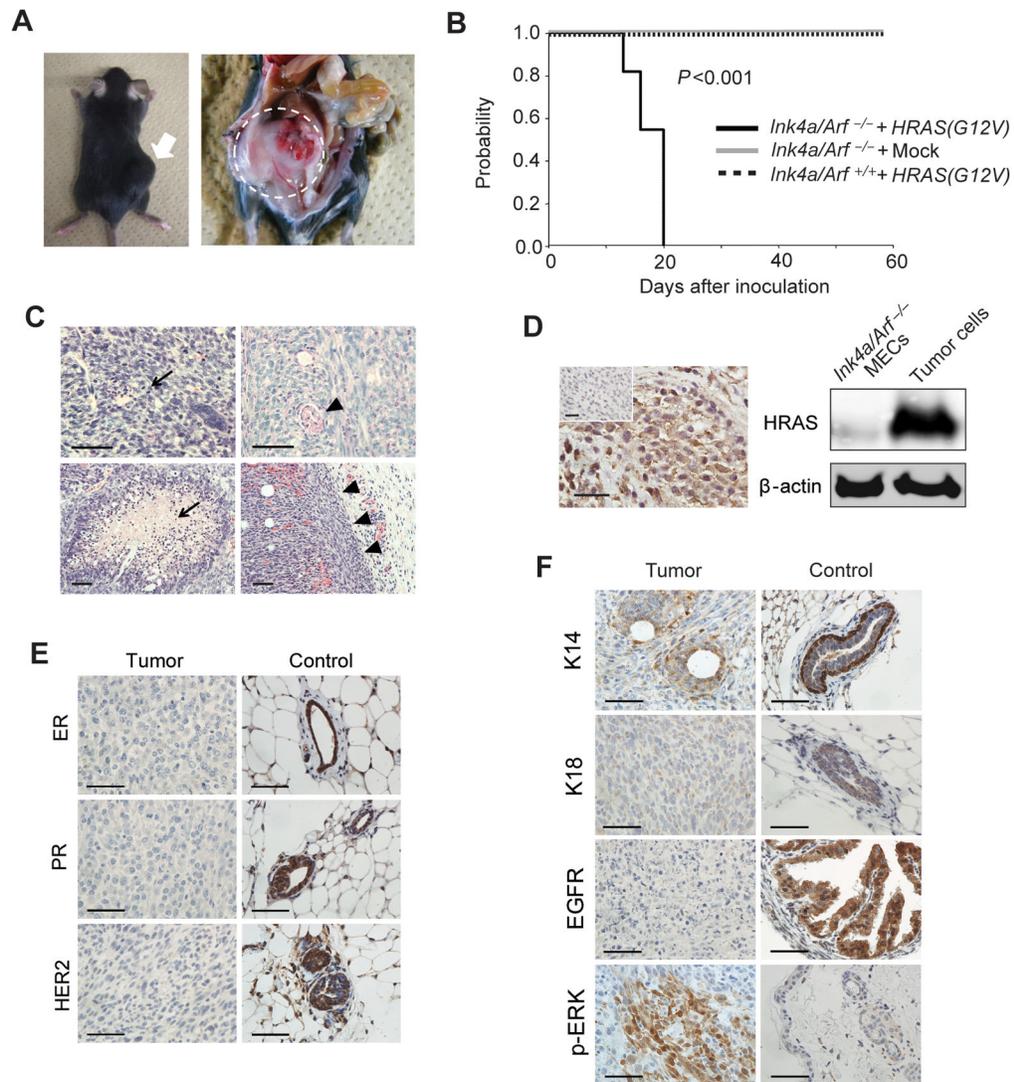


Figure 2. *HRAS(G12V)* transforms *Ink4a/Arf*^{-/-} MECs to form triple-negative mammary tumors
 A, Left, Tumor-bearing mouse that was inoculated with *HRAS(G12V)*-introduced *Ink4a/Arf*^{-/-} MECs into the right inguinal mammary fat pad. Arrow indicates tumor. Right, Intraperitoneal image of the same mouse. Tumor is delineated with a dotted line. B, Tumor-free survival after mice underwent transplantation with 1.0×10^5 mammosphere cells with each indicated profile. The *P* value was obtained by log-rank statistical analysis. C, Microscopic images of induced-tumor sections with hematoxylin and eosin staining (upper left and lower images) and periodic-acid Schiff (PAS) staining (upper right). A mitotic cell (arrow) and a PAS-positive mucus deposit (arrowhead) are visible in the upper images. Central necrosis (arrow) and pushing border (arrowheads) are visible in the lower images. D, Left, Immunohistochemical staining (IHC) of induced tumor with anti-*HRAS* antibody and with isotype rabbit IgG (Inset). Right, Western blot analysis of *HRAS* in *Ink4a/Arf*^{-/-} MECs and tumor cells. β -actin is shown as a loading control. E, F, IHC of induced tumor with antibodies (F) to keratin 14, keratin 18, EGFR, and phosphorylated ERK (p-ERK), and (E) to estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor

receptor-2 (HER2). The images in the right columns of E and F are positive controls except for p-ERK in (F). Controls for ER, PR, HER2, K14, and K18 are from mouse mammary ducts; for EGFR is from mouse endometrium. For p-ERK, the peritoneal wall in the same section is shown as a negative control. Scale bars, 50 μm .

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

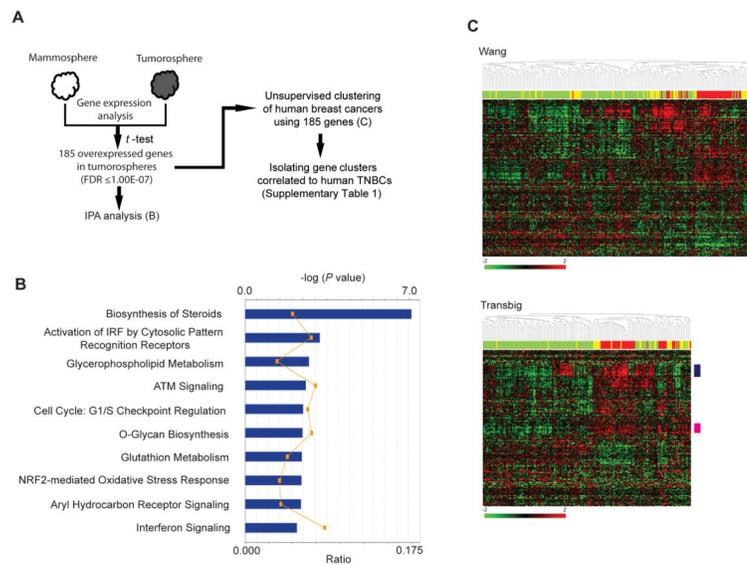


Figure 3. The molecular features of induced tumors partially overlap those of human triple-negative breast cancer

A, Strategic schema to analyze molecular similarities between induced tumors and human TNBC. B, Canonical pathway analysis of 185 overexpressed genes in tumorspheres (induced tumor-derived mammospheres) compared with mammospheres by IPA (Ingenuity pathway analysis). Data for the 10 most statistically significant pathways are presented. The upper *x*-axis corresponds to data for the bars; these data are logarithms of *P* values that were calculated by Fisher exact test. The lower *x*-axis corresponds to data in the line graph; these data represent the ratio of the number of molecules in a given pathway to the total number of molecules that make up that pathway. C, Heat maps of breast cancer in the Wang (upper) and Transbig (lower) breast cancer microarray data sets using 185 tumorsphere-overexpressed genes. The three subtypes are colored accordingly: triple-negative, red; HER2+, yellow; ER+, light green. The bars on the side denote gene clusters (correlation > 0.4) upregulated in TNBC (dark blue, interferon response genes; pink, cell cycle-related genes).

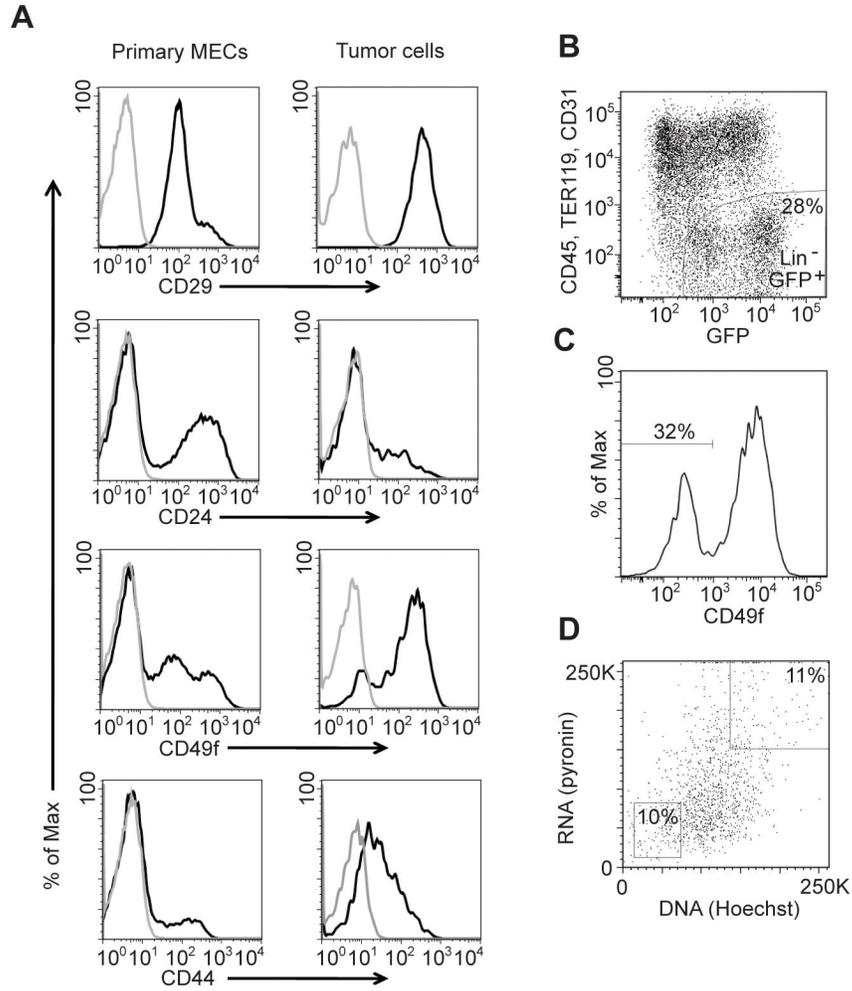


Figure 4. Tumor-initiating cell markers and gating strategy to purify tumor-initiating cells
 A, Differences in expression of CD29, CD24, CD49f, and CD44 between primary MECs and induced tumors. Populations depleted of dead cells and Lin⁺ cells, and GFP⁺ gated in the case of tumors, were subjected to a histogram cell count. Gray lines show isotype labeling. B, C, and D, Gating strategy to purify CD49f⁻ quiescent and dividing cells. B, Gating to purify Lin⁻ and GFP⁺ cells. C, Histogram of CD49f expression in Lin⁻GFP⁺ cells. D, Further gating of CD49f⁻ cells into CD49f⁻ quiescent cells (left lower box) and CD49f⁻ dividing cells (right upper box).

Table 1

Limiting dilution analysis of mouse mammary cancer-initiating cells

Cell profile	Number of cells injected			
	2,000	500	100	20
CD44 ⁺	1/4	1/7	0/7	
CD44 ⁻	5/5	2/7	0/7	
CD24 ⁺	4/5	4/7	0/7	
CD24 ⁻	5/5	5/7	2/7	
CD49f ⁺	8/8	4/7	0/8	
CD49f ⁻	4/4	4/5	5/8	
CD49f ⁻ quiescent			4/7	3/13
CD49f ⁻ dividing			0/6	0/13

The indicated cell populations were injected at the dosages listed. Denominators in the table represent the number of injections, and numerators represent the number of resultant tumors from the injected tumor cells. Cell cycle status was determined from the staining patterns of Hoechst-Red and pyronin Y as follows: quiescent, Hoechst-Red^{low}pyronin Y^{low}; dividing, Hoechst-Red^{high}pyronin Y^{high}.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript