

# ID genes mediate tumor reinitiation during breast cancer lung metastasis

Gaorav P. Gupta\*, Jonathan Perk\*, Swarnali Acharyya\*, Paola de Candia\*, Vivek Mittal†, Katia Todorova-Manova‡, William L. Gerald§, Edi Brogi¶, Robert Benezra\*||, and Joan Massagué\*||\*\*

\*Cancer Biology and Genetics Program, †Molecular Cytology Core Facility, §Human Oncology and Pathogenesis Program, ¶Department of Pathology, and \*\*Howard Hughes Medical Institute, Memorial Sloan-Kettering Cancer Center, New York, NY 10021; and †Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724

Contributed by Joan Massagué, September 27, 2007 (sent for review September 6, 2007)

The establishment of distant metastases depends on the capacity of small numbers of cancer cells to regenerate a tumor after entering a target tissue. The mechanisms that confer this capacity remain to be defined. Here we identify a role for the transcriptional inhibitors of differentiation Id1 and Id3 as selective mediators of lung metastatic colonization in the triple negative [TN, i.e., lacking expression of estrogen receptor and progesterone receptor, and lacking Her2 (human epidermal growth factor receptor 2) amplification] subgroup of human breast cancer. Although broad expression of Id1 has recently been documented in tumors of the rare metaplastic subtype, here we report that rare Id1-expressing cells are also present in the more common TN subset of human breast tumors but not in other subtypes. We also provide evidence that Id1 expression is enriched in clinically obtained hormone receptor negative lung metastases. Functional studies demonstrate that Id1 and its closely related family member Id3 are required for tumor initiating functions, both in the context of primary tumor formation and during metastatic colonization of the lung microenvironment. *In vivo* characterization of lung metastatic progression reveals that Id1 and Id3 facilitate sustained proliferation during the early stages of metastatic colonization, subsequent to extravasation into the lung parenchyma. These results shed light on the proliferative mechanisms that initiate metastatic colonization, and they implicate Id1 and Id3 as mediators of this malignant function in the TN subgroup of breast cancers.

proliferation | stem cells | triple negative

Cancer metastasis involves multiple steps through which tumor cells ultimately colonize and overtake distant organs (1, 2). After circulating tumors cells exit the bloodstream by passing through the vascular wall into tissue parenchyma, they require a capacity to reinitiate tumor growth to establish a metastatic colony. Tumor cells that extravasate individually or in small clusters may develop into full metastatic colonies only if they have such tumor-initiating capacity. The tumor-initiating capacity that is required for the establishment of metastatic colonies is akin to the phenotypic state that is referred to as “cancer stem cell” (3). In a strict interpretation, this term may designate stem cells that have become oncogenically transformed and whose progeny build up a tumor, as it occurs in certain leukemias (4) and in intestinal carcinomas (5). In a broader sense, however, this concept also encompasses the case of tumor cells that have co-opted certain gene functions conferring upon these cells the capacity of tumor reinitiation after metastatic seeding of distant organs. This reinitiation capacity is regardless of whether such genes participate in the normal self-renewal process of authentic stem cells. Therefore, the identification of genes that confer this function is important for understanding the molecular basis for cancer metastasis.

Analysis of gene expression patterns in human tumors of different metastatic outcome has revealed gene sets, or “signatures,” whose expression in a tumor is associated with a high risk of relapse (6, 7). We recently used *in vivo* metastasis assays to

functionally filter genome-wide expression patterns of human breast cancer cells. These studies resulted in the identification of a set of candidate metastasis genes whose expression in primary breast tumors is associated with a high risk of metastasis to the lungs (8). This set of genes, the Lung Metastasis Signature (LMS), includes 18 genes that support mammary tumor growth as well as pulmonary metastasis by human breast cancer cells in mice, and their expression in estrogen-receptor-negative (ER-) tumors indicates a high risk of pulmonary relapse in breast cancer patients (8, 9). A majority of LMS genes encode cell surface and secreted products that affect the interaction of the tumor cells with the microenvironment. Some of these genes collectively support tumor angiogenesis, entry of mammary tumor cells into the circulation, and tumor cell exit from lung capillaries into the pulmonary parenchyma (10). A notable exception is the inhibitor of cell differentiation *ID1*, which is the only transcriptional regulator present in the LMS.

The *ID* gene family (*ID1* through *ID4*) are expressed in progenitor cells in embryos and in certain adult tissues, and control cell differentiation by antagonizing the DNA binding activity of basic helix-loop-helix transcription factors (11, 12). *ID1* and *ID3* function is required for the mobilization of endothelial precursor cells from the bone marrow during pathological tumor angiogenesis, and the expression of these genes remains high in tumor neovasculature (12, 13). Broad *ID* expression has been documented in metaplastic tumors of poor clinical outcome (14). *ID* expression in more common cancer subtypes has been a matter of debate, owing to disparate observations depending on the cell line or tumor model used or the nature of the antibodies against the gene products (11, 14, 15). The identification of *ID1* as a LMS gene focused our attention on the hypothesis that *ID* gene expression confers tumor-initiating capacity to rare breast cancer cells within the more common types of breast tumors with high lung metastatic potential.

## Results

**Id1 Expression in Specific Breast Tumor Subtypes and Lung Metastases.** Immunohistochemical studies using tissue microarrays (TMA) recently revealed a substantial incidence (~25% of cases) of Id1 protein expression in the rare metaplastic breast cancer subtype. The incidence of Id1 expression was lower (2–3% of cases) in the more common subtype triple negative [TN, i.e., lacking expression of estrogen receptor and progesterone receptor, and lacking Her2 (human epidermal growth factor receptor 2) amplification] subgroup of human breast cancer.

Author contributions: G.P.G., R.B., and J.M. designed research; G.P.G., J.P., and S.A. performed research; J.P., P.d.C., V.M., K.T.-M., W.L.G., E.B., and R.B. contributed new reagents/analytic tools; G.P.G., J.P., K.T.-M., W.L.G., E.B., R.B., and J.M. analyzed data; and G.P.G., R.B., and J.M. wrote the paper.

The authors declare no conflict of interest.

¶To whom correspondence may be addressed. E-mail: benezrar@mskcc.org or massagu@mskcc.org.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0709185104/DC1](http://www.pnas.org/cgi/content/full/0709185104/DC1).

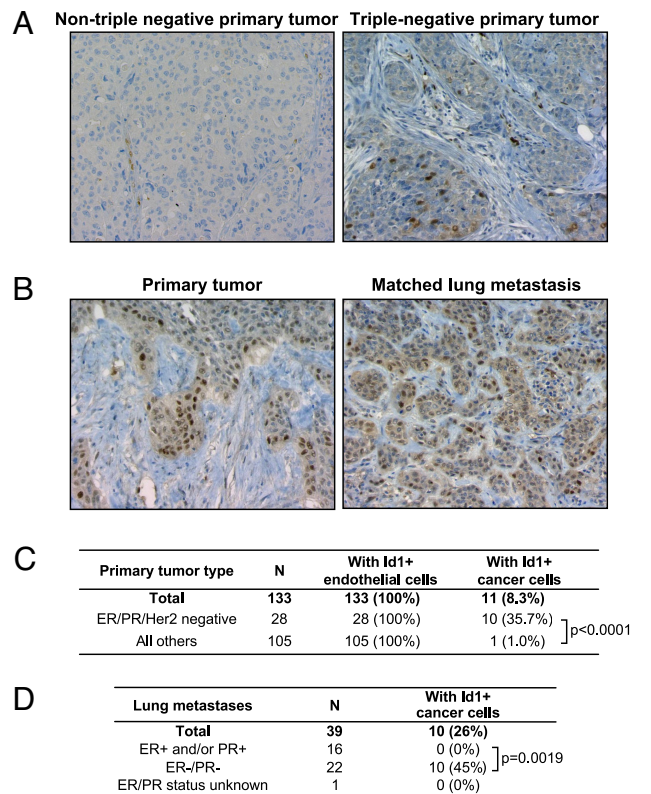
© 2007 by The National Academy of Sciences of the USA

terone receptor, and lacking Her2 (human epidermal growth factor receptor 2) amplification] (14). However, because of the focal pattern of Id1 expression that is frequently observed, it remained possible that evaluation of tumor samples by TMA might underestimate the proportion of Id1+ breast cancers. For this reason, we conducted Id1 immunohistochemistry with a mono-specific rabbit monoclonal antibody on whole tumor block sections from 133 patients with breast cancer of various subtypes. Among primary breast tumors that were not triple negative, only 1 of 105 exhibited Id1 expression within the tumor cell compartment (1%; Fig. 1A and B). In contrast, 10 of 28 (36%) triple negative [ER-, progesterone-receptor-negative (PR-), and Her2-] breast cancers contained cancer cells that were Id1+ (Fig. 1A and B). This proportion was even higher (6 of 7, 86%) when examining the subset of triple negative tumors that exhibited a metaplastic morphology (data not shown). Within these Id1+ samples, the frequency of Id1-expressing cancer cells was consistently higher in metaplastic TN tumors (range 10–95%) than it was in nonmetaplastic TN tumors (range 1–5%). As an internal positive control, endothelial cells in all tumor subtypes expressed abundant quantities of nuclear Id1 (Fig. 1A). These findings are consistent with previous results (14) but likely reflect a more accurate quantitation of rare Id1 positivity in human breast cancer subtypes than the frequencies that were previously attained by using TMA studies.

If Id1 augments the lung metastatic potential of breast tumor cells, one would predict that a greater proportion of lung metastases should contain Id1+ cancer cells. To address this issue, we measured Id1 protein expression in 39 lung metastasis lesions surgically extracted from breast cancer patients. Ten of 39 (26%) of these lung metastasis lesions showed Id1+ staining (Fig. 1D). In 11 cases for which primary and metastasis tissue were obtained from the same patient, 3 out of 11 metastases were Id1+, and indeed rare, focal Id1+ cells were observed in all 3 matched primaries. Hormone receptor status was available for 32 of the 39 lung metastasis samples. Intriguingly, 10 of 22 (45%) hormone-receptor-negative (ER- and PR-) lung metastases were Id1+, whereas 0 of 16 hormone-receptor-positive (ER+ and/or PR+) lung metastases expressed Id1 (Fig. 1D). Although Her2 amplification status was not available for all of these samples, 60% (6 of 10) of the confirmed TN lung metastases contained Id1-expressing cancer cells (data not shown). Interestingly, we also examined Id1 levels in 10 bone metastases, but none of these samples contained tumor cells that expressed Id1 (data not shown). These data are consistent with the hypothesis that Id1+ cells have a selectable advantage during secondary metastasis to the lungs, especially in hormone receptor negative breast cancers (see also Discussion).

**Id1 and Id3 Expression in Transgenic Mouse Models of Breast Cancer.**

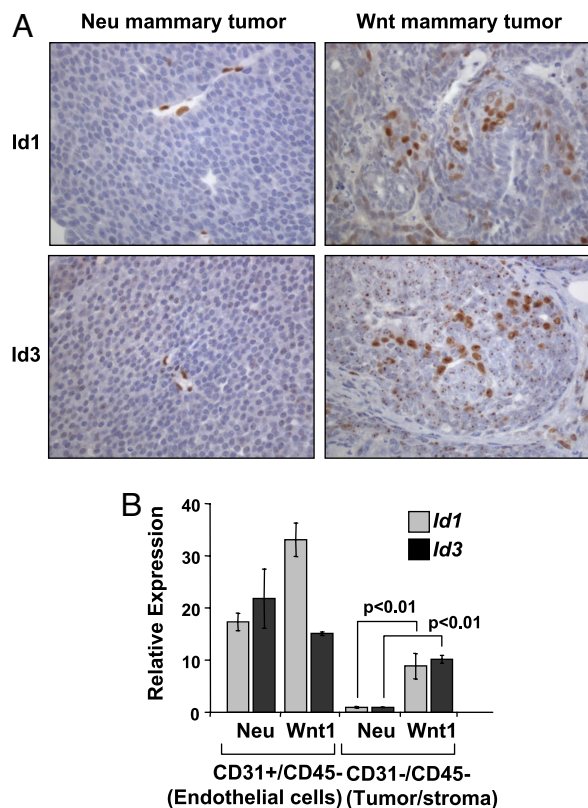
To further investigate the biology of Id proteins among different subtypes of breast cancer, we resorted to two distinct transgenic mouse models of breast cancer—MTV-NeuYD and MMTV-Wnt1. Mammary tumors induced by the Neu oncogene (16) have a uniformly luminal histology (17), require CDK4 and cyclin D1 for transformation (18, 19), and are thought to reflect human ErbB2 subtype breast cancer pathophysiology. In contrast, Wnt1-driven tumors (20) are distinguished by a prevalence of mammary progenitor-like cells (21–23), and exhibit many histopathological features in common with human basal subtype tumors, including the presence of myoepithelial cells, dense stromal and inflammatory infiltrates, and pushing margins of invasion (17, 24). Additionally, a recent cross-species comparison of global gene expression patterns has characterized Wnt1-driven mammary tumors as most similar to the basal subtype of human breast cancer (25). Although designation of the basal subtype is determined by microarray-based gene expression profiles, immunohistochemical studies of basal breast cancers have revealed that they all fall within the subgroup of TN tumors,



**Fig. 1.** Expression of Id1 in human breast cancer tumors and lung metastases. (A) Representative Id1 immunohistochemistry images of non-TN and TN primary breast tumors, acquired with an initial magnification of  $\times 20$ . (B) Id1 immunohistochemistry of a matched primary tumor and lung metastasis pair. Nine of these pairs were analyzed, as discussed in the text. (Original magnification:  $\times 20$ .) (C) Compiled results of Id1 immunohistochemistry performed on 133 primary breast tumors. Frequencies of Id1 positivity in endothelial cells and in cancer cells are shown. (D) Tabulated results of Id1 immunohistochemistry of 39 breast cancer lung metastases, with subset analysis based on hormone receptor expression. P values were calculated by a two-tailed Fisher's exact test.

because they lack both ER and PR expression and Her2 gene amplification (24).

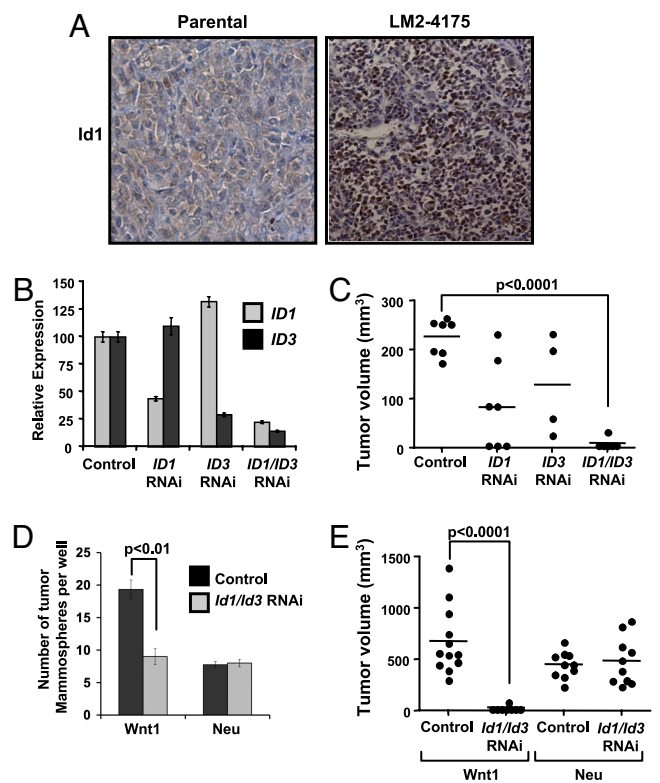
Immunohistochemical staining for Id1 was performed on mammary tumors arising in both the MMTV-Neu and MMTV-Wnt1 mouse models. In addition, we also stained these mammary tumors for Id3—the Id family member most closely related to Id1, based on genetic studies documenting functional overlap in mouse development and highly significant coexpression of Id1 and Id3 that is observed in large cohorts of human breast cancer [supporting information (SI) Table 1]. Our results indicate that, although both Neu and Wnt1 tumors exhibit Id expression in the endothelial cells lining the tumor vasculature, only Wnt1-driven tumors contain subsets of cancer cells that abundantly expressed Id1 and Id3 (Fig. 2A). These immunohistochemical findings were also confirmed by dissociation of Wnt and Neu tumors into single cell suspensions, and by using FACS to separate endothelial cells (CD31+) from nonendothelial cells (CD31-) in the tumors, after initially depleting all cells expressing hematopoietic lineage markers (by using the CD45 and Ter119 cell surface markers). Quantitative RT-PCR analysis again validated that *Id1* and *Id3* expression was restricted to the endothelial compartment of Neu tumors but was present in both endothelial and nonendothelial cells in Wnt1 mammary tumors (Fig. 2B). The enrichment of Id expression in Wnt1-induced mammary tumor cells is reminiscent of our earlier observations of selective Id1



**Fig. 2.** Id1 and Id3 expression in transgenic mouse mammary tumor models. (A) Representative Id1 and Id3 immunohistochemistry images of mammary tumors from MMTV-Neu and MMTV-Wnt1 mice. (Original magnification:  $\times 20$ .) (B) Primary Neu and Wnt1 tumors were dissociated into single cell suspensions and separated by FACS by using the CD31 and CD45 cell surface markers. The CD31<sup>+</sup>/CD45<sup>-</sup> fraction is enriched for endothelial cells, and the CD31<sup>-</sup>/CD45<sup>-</sup> fraction is composed mostly of tumor cells but may also contain a minority of stromal cells. Total RNA was extracted from these samples, and *Id1* and *Id3* mRNA levels normalized to *GAPDH* were measured by qRT-PCR. Normalized expression values are shown relative to those obtained in the Neu tumor CD31<sup>-</sup>/CD45<sup>-</sup> fraction. Samples were analyzed in quadruplicate, and error bars represent 95% confidence intervals. *P* values are based on a two-tailed Student's *t* test.

expression in the TN subtype of human breast cancer and supports the hypothesis that the expression of Id1 in breast cancer is indeed nonuniform and may coincide with a more undifferentiated/hormone-receptor-negative state.

**Id1/Id3 Functional Requirement in Breast Tumor Initiation.** To test the functional importance of Id1 and Id3 in tumor initiation, we used both MDA-MB-231 human breast cancer cells and primary cells isolated from Wnt1- and Neu-driven mouse mammary tumors. The MDA-MB-231 cell line is representative of the TN subset of breast cancer and consists of cells with diverse metastatic potential that can be sorted *in vivo* for specific metastatic tropisms (8, 26). The lung metastasis gene expression signature (LMS) that includes *Id1*, and predicts a high risk of lung metastatic recurrence in breast cancer patients, was derived from this experimental system (8). Expression analysis for all four *Id* family members revealed that *Id1* and *Id3* transcript levels in the *in vivo*-selected lung-metastatic LM2-4175 subpopulation (8) were 8- and 4-fold higher, respectively, than those in the unselected MDA-MB-231 parental population, whereas *Id2* and *Id4* expression was similar in both populations (SI Fig. 6). We confirmed this trend by immunohistochemistry for Id1 in mammary tumors generated by parental MDA-MB-231 cells and LM2-4175 cells (Fig. 3A). Knockdown of



**Fig. 3.** Functional requirement for Id1/Id3 in tumor-initiating capacity. (A) Id1 immunohistochemistry of unselected MDA-MB-231 parental cells and highly lung metastatic LM2 cells grown orthotopically in the mammary fat pad of an immunocompromised mouse. (Original magnification:  $\times 20$ .) (B) qRT-PCR analysis of *Id1* and *Id3* expression in LM2 cells retrovirally transduced with short hairpin constructs targeting *Id1*, *Id3*, or both genes. Error bars indicate 95% confidence intervals of samples assayed in quadruplicate. (C) One thousand control or *Id* knockdown LM2 cells were injected orthotopically into the mammary glands of immunocompromised mice. Tumor volumes were assessed at 9 weeks after inoculation. The *P* value is based on a two-tailed *t* test. (D) Mammosphere-initiating capacity of primary Wnt1 and Neu tumor cells transduced and sorted for shId1/Id3 or shLuc expression. Cells were incubated in suspension at a concentration of 1,000 cells per ml, and mammospheres were counted after 14 days of culture. Error bars represent the standard error of the mean. The *P* value is based on a two-tailed *t* test. (E) Twenty-five thousand primary Wnt1 and 100,000 primary Neu tumor cells expressing short hairpins targeting luciferase or Id1/Id3 were injected into the mammary glands of immunocompromised mice. Tumor volumes at 3 months are represented for all cohorts. The *P* value is based on a two-tailed *t* test comparison.

*Id1* and *Id3* expression, both individually and in combination (Fig. 3B), resulted in partial and complete inhibition of mammary tumor-initiating capacity, respectively (Fig. 3C). Similarly, combined knockdown of mouse Id1/Id3 in MMTV-Wnt1 primary tumor cells resulted in significant impairment of *in vitro* tumor mammosphere generation as well as *in vivo* tumor initiation (Fig. 3D and E), whereas the same knockdown construct introduced into MMTV-Neu primary cells showed neither of these defects. Thus, the selective expression of Id1 and Id3 in TN breast cancer cells coincides with a functional requirement of Id expression for sustained tumor-initiating capacity.

**Id1/Id3 Functional Requirement for Lung Colonization.** As previously described, the *in vivo*-selected LM2-4175 subpopulation of the parental MDA-MB-231 cell line exhibits aggressive lung metastatic behavior when injected either intravenously or orthotopically into immunocompromised mice, and this capacity depends on expression of genes included in the clinically validated LMS (8). Immunohistochemistry demonstrates that these emergent



promoting invasion through extracellular matrix barriers by using similar *in vitro* assays (SI Fig. 8). Collectively, our *in vitro* and *in vivo* data point toward a role for Id proteins during metastatic progression that is distinct from invasion through either extracellular matrix or endothelial barriers.

Because Id proteins have been implicated as facilitators of progenitor cell proliferation during organismal development (11), we next considered a possible role for Id1 and Id3 as regulators of disseminated cancer cell proliferation within the lung microenvironment. Toward this end, we costained relatively early stage lung metastases from mice injected with LM2–4175 control cells (day 21) for Id1 and PCNA. Consistent with a role for Id1 in proliferation, quantitative analysis of multiple metastatic foci identified a highly significant coincidence of Id1 and PCNA within tumor cell nuclei (Fig. 5C). Interestingly, although the majority of Id1-expressing cells were actively proliferating, there were also many PCNA+ cells that lacked Id1 expression (Fig. 5C).

To determine if this level of Id1 and Id3 expression is necessary for the early reinitiation of proliferation during lung metastatic colonization, we costained lungs of mice 5 and 10 days after injection with either control LM2–4175 or *ID1/ID3* knockdown LM2–4175 cells for the proliferative marker Ki-67 and a tumor cell specific antigen. Staining of control cell-injected lungs revealed multiple small foci of tumor cells that had efficiently reentered the cell cycle (Fig. 5D, upper immunofluorescence images). In contrast, lungs injected with *ID1/ID3* knockdown LM2–4175 cells predominantly contained solitary tumor cells that infrequently expressed Ki-67 (Fig. 5D, lower immunofluorescence images). Quantification of multiple lung sections revealed statistically significant reductions in proliferating tumor cells between *ID1/ID3* knockdown and control LM2–4175 metastatic cells at both day 5 and day 10 after inoculation (Fig. 5D). Extrapolation based on bioluminescence data suggests that the proliferation-defective *ID1/ID3* knockdown cells will never form metastatic colonies (refer to Fig. 4B). Collectively, these data argue that Id1 and Id3 are necessary for sustained proliferation during the early stages of metastatic colonization, subsequent to tumor cell extravasation into the lung parenchyma (Fig. 5E).

## Discussion

Metastatic colonization entails distinct biological functions that collectively enable disseminated cancer cells to thrive in a secondary site (1, 2). Thus, the role of Id1 and Id3 as providers of sustained proliferative competence during lung metastatic colonization by breast cancer cells must occur in the context of a full complement of distinct prometastatic functions. These additional functions may be contributed in part by other members of the functionally derived and clinically validated LMS gene set to which Id1 belongs. Indeed, the enforced overexpression of Id1 alone is not sufficient to render parental MDA-MB-231 cells efficiently metastatic to the lungs (8). In separate studies we have identified four other genes in the LMS set that are not regulated by Id proteins, and that cooperatively mediate tumor cell extravasation into the lung parenchyma (10). Thus, the proliferative capacity provided by Id genes is one in a set of rate-limiting functions that collectively underlie the biology of breast cancer metastasis to the lungs, which can be discerned through the mechanistic analysis of distinct genetic subsets of our previously characterized LMS (29).

A number of studies have pointed to a role for the Id proteins in the metastatic progression of breast cancer. Overexpression of Id1 has been shown to enhance the invasiveness of breast cancer cells in *in vitro* assays, and facilitates sustained proliferation of transfected cells that are cultured in serum-starved conditions (11). Conversely, inhibition of Id1 expression in breast cancer cells has a significant impact on the ability of breast cancer cells to metastasize to the lungs in xenograft models (8, 28). Although it has been suggested that invasiveness through extracellular

matrix material *in vitro* is impaired in tumor cells expressing full-length antisense Id1 (27, 28), we do not discern any such defect when using short hairpin RNAi targeting Id1 expression in aggressively lung metastatic MDA-MB-231 cells. Furthermore, our *in situ* analyses of extravasation and early proliferation indicate that Id1 and Id3 are required *in vivo* for the sustained proliferative activity of metastatic tumor cells during the reinitiation of secondary tumors in the lung. The present insights into the biology of Id proteins draw mechanistic parallels between the sustained proliferation of tissue progenitor cells in development and the proliferative competency required for tumor reinitiation during metastatic colonization.

The role for Id proteins that we describe here applies to breast cancer cells that express Id. Although early studies using commercially available polyclonal antibodies described broad overexpression of Id1 protein in a majority of human primary breast tumors, recent data from studies using a new highly specific rabbit monoclonal antibody indicate that the expression of Id1 is not nearly as widespread (14). Our current analysis of whole tumor block sections has strengthened these initial findings, but it has also revealed a reduced sensitivity in detecting focal Id1 expression when examining sample cores in the context of tissue microarrays. Collectively, these studies have established that Id1 expression in cancer cells is rarely seen in tumors characterized by hormone receptor expression or Her2 gene amplification, but it is observed in 36% of primary breast cancers that are negative for these biomarkers (i.e., TN), with the greatest prevalence occurring within the subset of TN tumors that exhibit a metastatic morphology (86%). Interestingly, all of the breast tumors examined (regardless of subtype) exhibited Id1 positivity within endothelial cells lining the tumor vasculature, consistent with a role for this protein in tumor angiogenesis (11–13). This selective pattern of Id1 expression was also observed in mouse mammary tumors induced by the Neu versus Wnt1 oncogenes, with Wnt1-driven tumors containing an abundance of tumor cells that express Id1. Wnt1-driven mammary hyperplasias and tumors have previously been shown to contain mammary progenitor-like cells (21–23). It remains possible that Id1 expression in human breast tumors may also coincide with an increased frequency of intratumoral mammary stem-like cells. Consistent with this notion, recent gene expression studies of human breast tumors identified coexpression of Id family genes and stem-like genes in a subset of breast tumor-initiating cells (30). However, rigorous investigation of a potential causal link between Id expression and progenitor-like properties in breast cancer should be the subject of future studies.

We are providing a demonstration that Id1 protein is expressed in human breast cancer lung metastases. Consistent with a causal role in lung colonization, we observe an increased prevalence of Id1 positivity among lung metastases compared with primary tumors. Because population-based studies show that TN tumors represent  $\approx 16\%$  of all breast tumor cases (31), the frequency of Id1+ staining that we observed in our primary tumor cohort predicts that  $\approx 6\%$  (36% of 16%) of all human breast tumors will show focal Id1 staining. Given that 26% of the lung metastasis lesions that we analyzed were Id1+, this represents an enrichment of Id1 at the lung metastatic site relative to the primary tumor. This noteworthy clinical correlation implies that the proliferative functions mediated by Id genes may be a limiting requirement in the development of lung metastases in patients with hormone-receptor-negative breast cancer. We also note that none of the ER+ breast cancers that developed lung metastases exhibited Id1 protein expression, suggesting that either reinitiation of proliferation at secondary sites is not a rate-limiting step for ER+ breast cancer cells or that these cells adopt alternative mechanisms for tumor reinitiation. Thus, mechanisms of metastasis may be distinct for different subtypes

