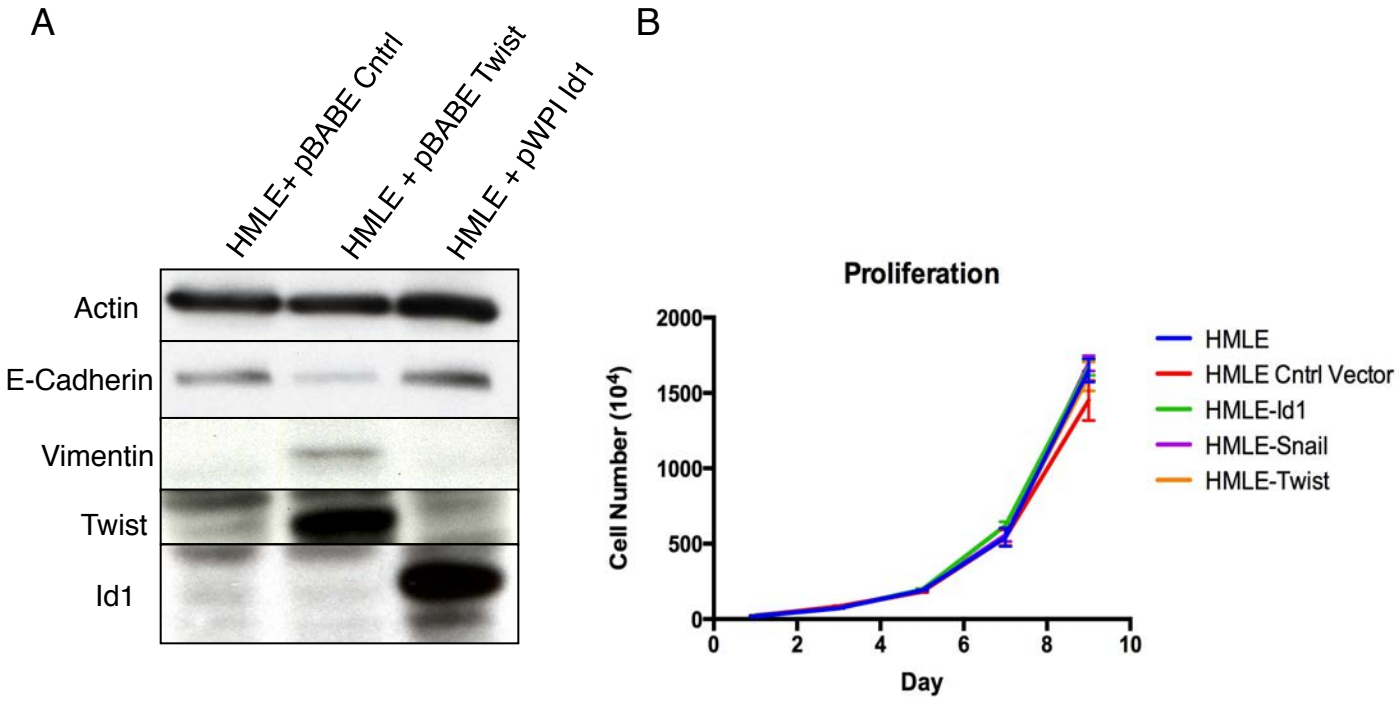
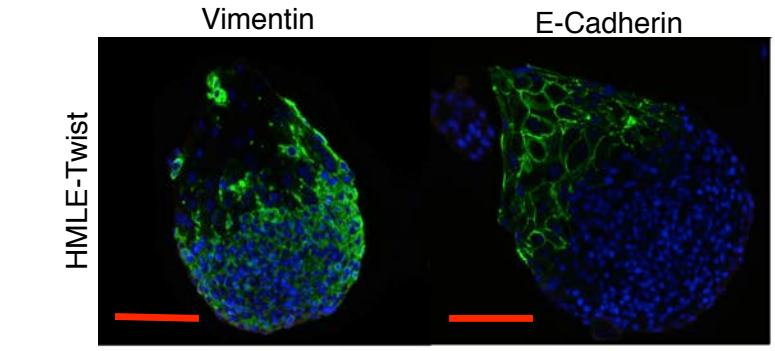
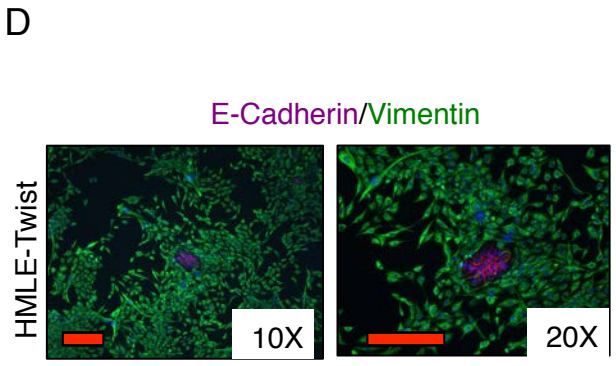
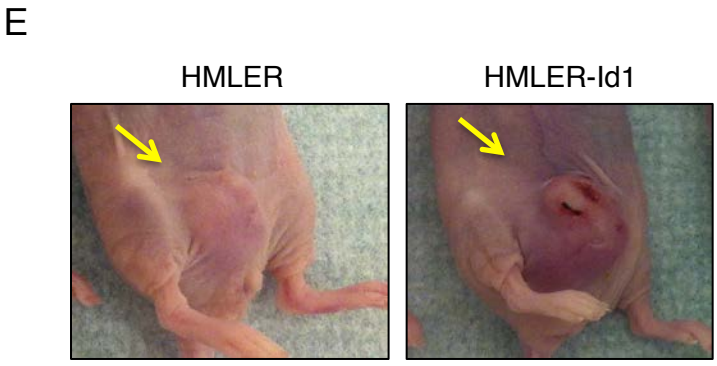




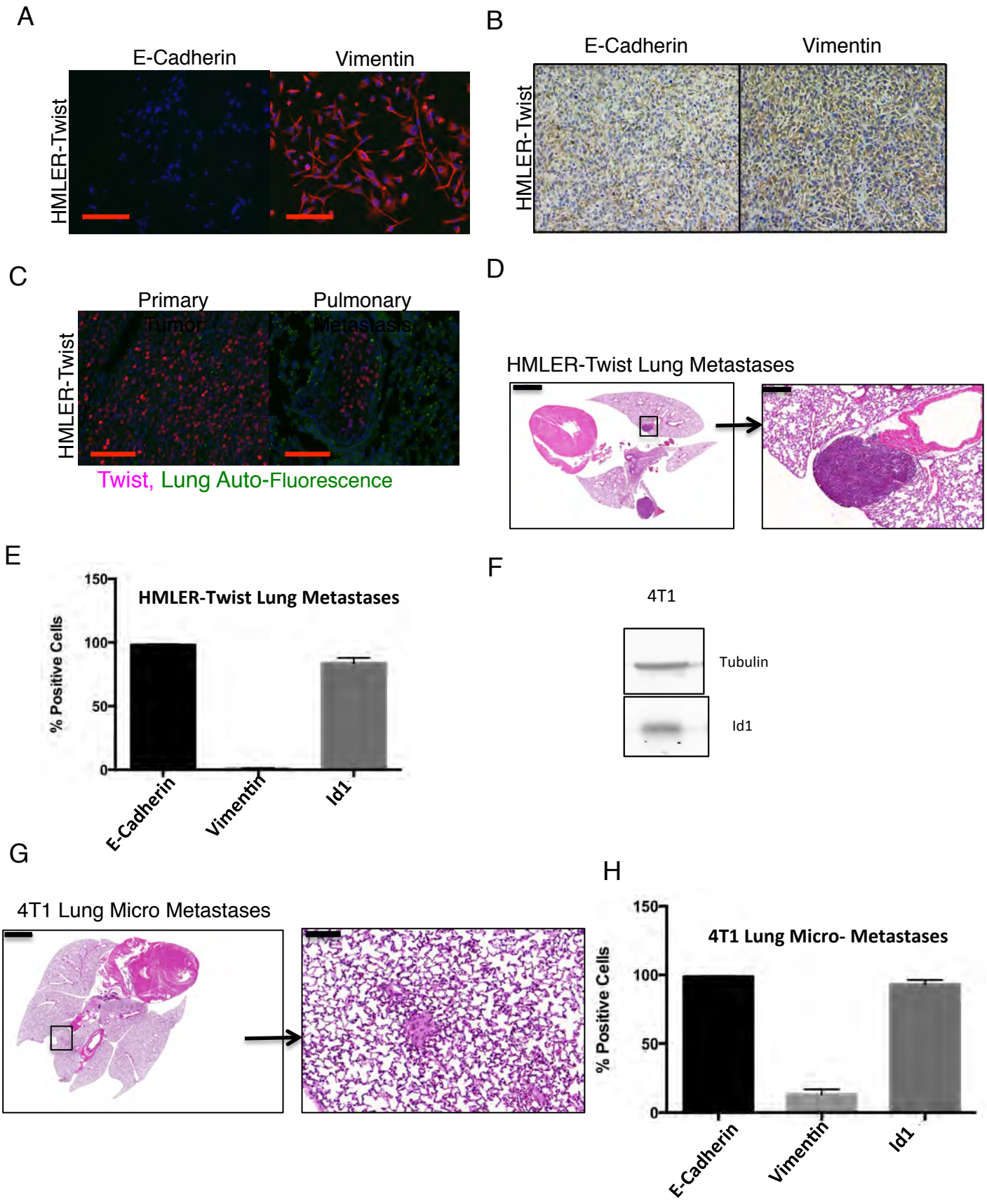




Supplemental Figure 1

Supplemental Figure 1.

(A) Protein levels of actin, E-Cadherin, Vimentin, Twist and Id1 in HMLE, HMLE-Twist, and HMLE-Id1 cells were determined by immunoblotting.
(B) Proliferation rates for HMLE expressing vector control, Id1, Snail, and Twist.
(C) IF image of mammospheres generated from HMLE-Twist cells stained for the mesenchymal marker Vimentin and the epithelial marker E-Cadherin. Scale bar = 200μm.
(D) Immunofluorescence of HMLE-Twist cells co-stained for E-cadherin and Vimentin. Scale bar = 200μm.
(E) Zoomed-in images of tumors formed by HMLER (left) and HMLER-Id1 (right) cells.

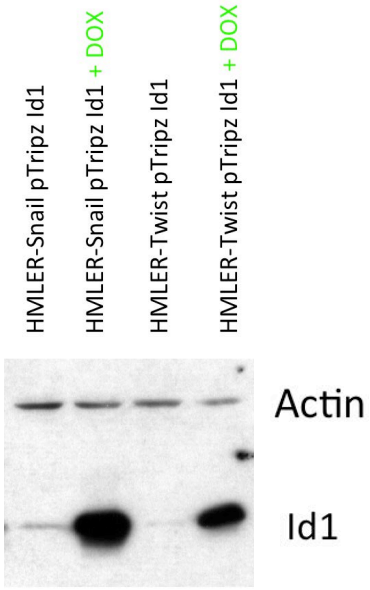


Supplemental Figure 2

Supplemental Figure 2.

(A) Immunofluorescence of HMLER-Twist cells stained for E-cadherin and Vimentin. Scale bar = 200μm.
(B) Tumors were initiated by injecting $5x10^5$ HMLER-Twist cells into the fat pads of immune-compromised mice. Immunohistochemistry reveals weak E-cadherin staining and Vimentin positive staining. Scale bar = 500μm.
(C) Tumors from (B) are positive for Twist as determined by immunofluorescence in the primary tumors and lung metastases. Scale bar = 200μm
(D) Hematoxylin-eosin staining of the whole lung from the animals that formed HMLER-Twist metastases. Scale bar = 2000 μm (left), inset scale bar = 200 μm.
(E) Percentage of E-Cadherin, Vimentin, and Id1 positive cells in HMLER-Twist lung metastases.
(F) Protein levels of tubulin and Id1 in 4T1 cells were determined by Western blotting.
(G) Hematoxylin-eosin staining of the whole lung from the animals that formed 4T1 metastases. Scale bar = 2000 μm (left), inset scale bar = 200 μm.
(H) Percentage of E-Cadherin, Vimentin, and Id1 positive cells in 4T1 lung metastases.

A

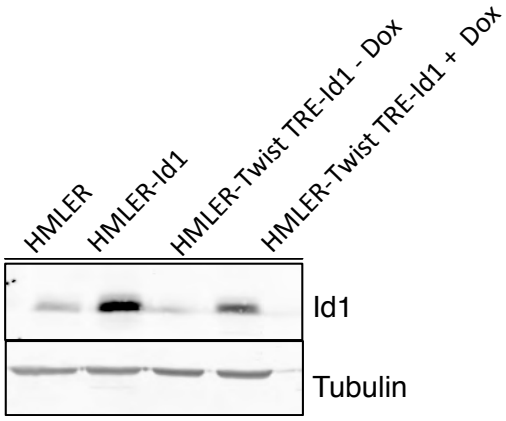


Supplemental Figure 3

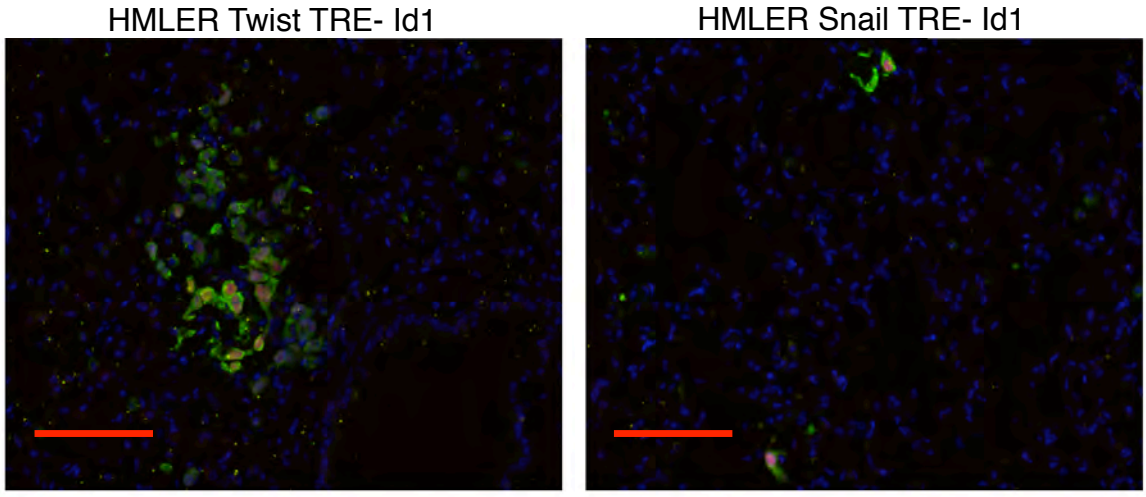
Supplemental Figure 3.

(A) Western blot showing Id1 induction with and without addition of doxycycline to the media of HMLER-Snail TRE-Id1 and HMLER-Twist TRE-Id1 cells.

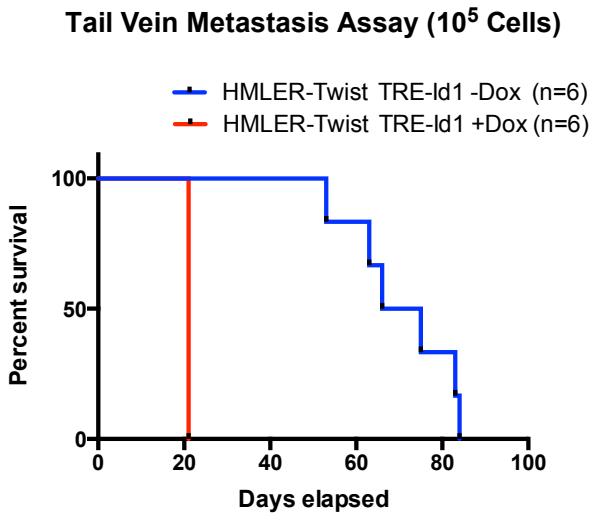
A



B



C

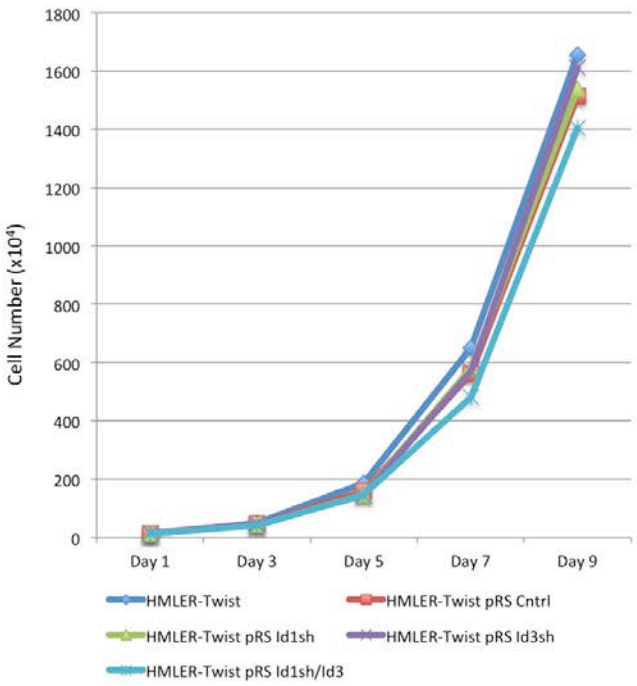


Supplemental Figure 4

Supplemental Figure 4.

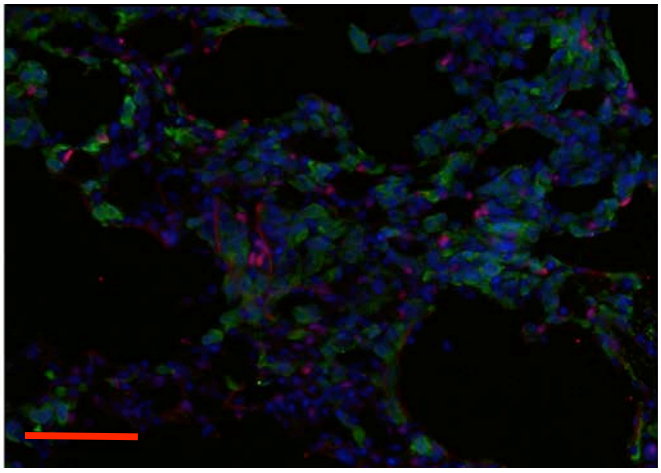
(A) Western blot showing Id1 and tubulin protein levels in HMLER-Id1 cells and HMLER-Twist TRE-Id1 cells treated with doxycycline.
(B) Immunofluorescence images of metastases from the lungs of animals injected via tail vein with 10^5 HMLER-Snail TRE-Id1 and HMLER-Twist TRE-Id1 3 weeks after they were placed on a doxycycline enriched diet. Id1 positive cells in GFP labeled tumors cells shows *in vivo* Id1 induction in both cell lines. Scale bar = 200μm.
(C) Kaplan-Meier survival curves for mice injected with 10^5 HMLE-Ras Twist TRE-Id1 via tail vein in the presence or absence of doxycycline.

A



B

HMLER-Twist GFP Id1/3 sh



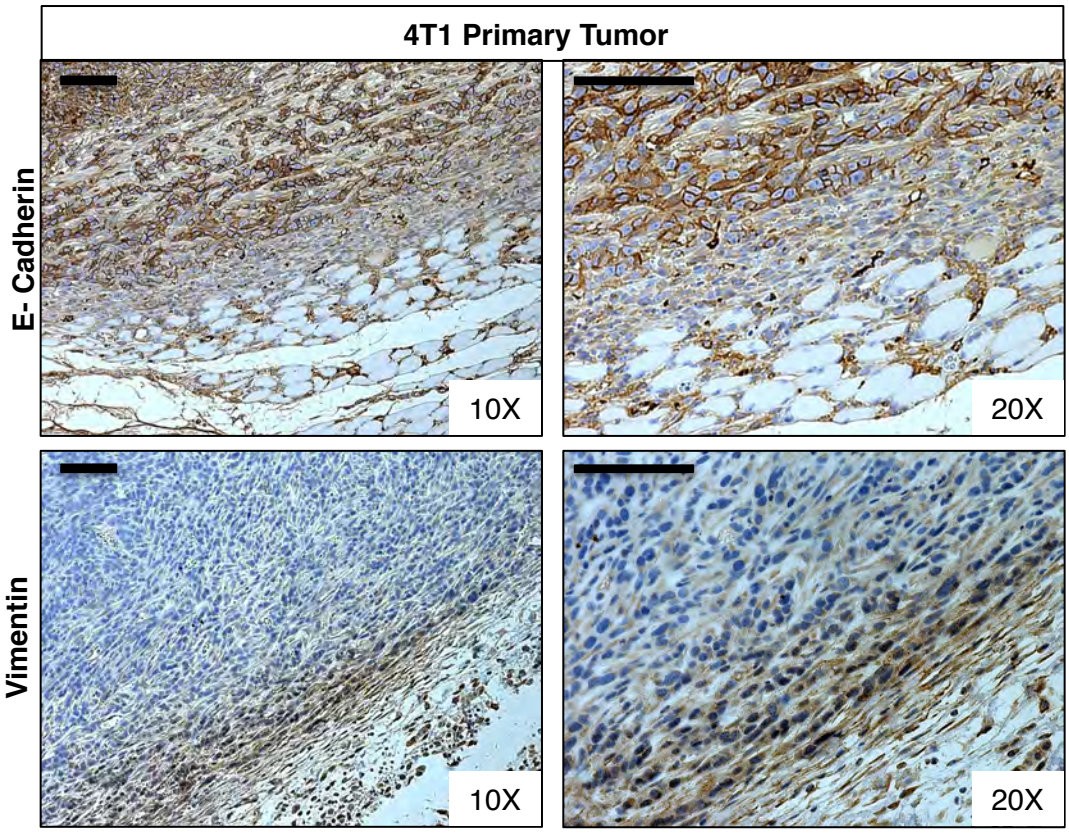
Tumor Cell, Id1, DAPI

Supplemental Figure 5

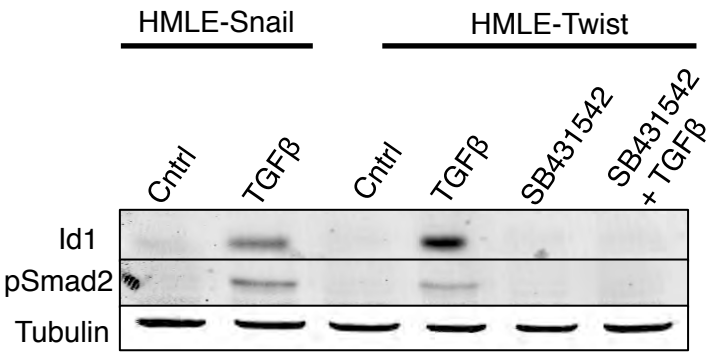
Supplemental Figure 5.

(A) Proliferation rates for HMLER-Twist transduced with retro-viruses targeting Id1 and/or Id3 or control vector.
(B) Immunofluorescent image of a lung metastasis from a mouse injected with GFP+ HMLER-Twist Id1/3sh cells shows presence of Id1 protein. Scale bar = 200μm.

A



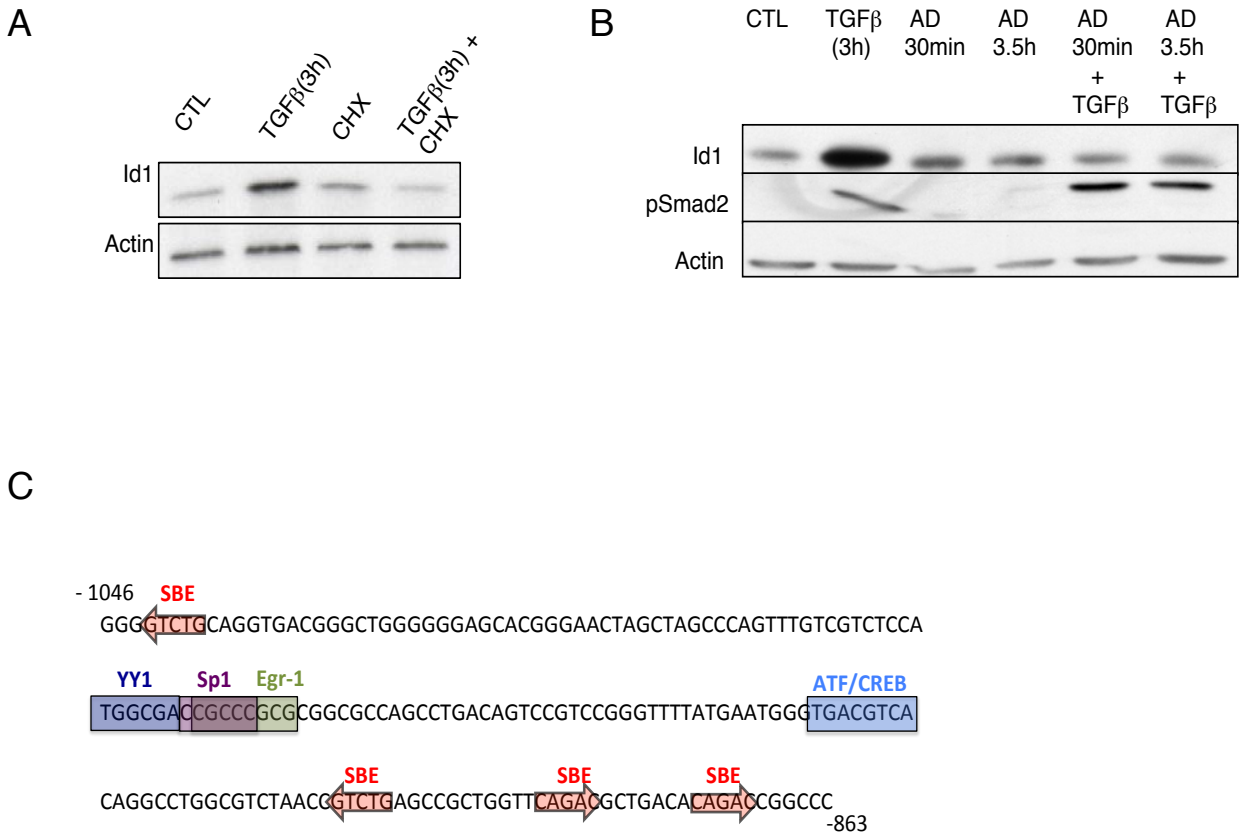
B



Supplemental Figure 6

Supplemental Figure 6.

(A) E-Cadherin and Vimentin protein expression in 4T1 primary tumors was detected by immunohistochemistry. Scale bar = 100μm.
(B) HMLE-Snail and HMLE-Twist cells were stimulated with 100pM TGFβ for 3 hours (lanes 1-4). HMLE-Twist cells were additionally treated with 2μM TβRI inhibitor SB431542 (lanes 5-6). Levels of Id1, p-Smad2 and tubulin were determined by Western blot.



Supplemental Figure 7

Supplemental Figure 7.

(A) Pleural effusion CN37 cells were treated with TGFβ in the presence or absence of 5 μg/ml protein synthesis inhibitor cycloheximide (CHX). Id1 protein levels were determined by Western Blot.
(B) CN37 cells were incubated in the absence (-) or presence (+) of 5 μg/ml actinomycin D (AD) for 30 min (prior to the TGFβ treatment) or 3.5 hours total time (30 minutes prior to the TGFβ treatment and 3 hours during the treatment). Following the 3 hour treatment with TGFβ, protein levels were evaluated by immunoblotting.
(C) A nucleotide sequence of the TGFβ-responsive region on the Id1 promoter. This region contains the following consensus sequences: four Smad binding sites (SBE-red arrows), YY1, Egr1, Sp1, and ATF/CREB binding sites.

Pulmonary lesions of triple negative breast carcinomas express epithelial marker E-Cadherin

ID No.	TMA Lung ER Pos	TMA Lung PR Pos	TMA Lung HER-2 Pos	TMA Lung 3X NEG	TMA Lung E-Cadherin Pos	TMA Lung Id1
2	0	0	0	1	1	No tumor
3	0	0	0	1	1	1
75	0	0	0	1	1	1
68	0	0	0	1	1	1
72	0	0	0	1	1	1
78	0	0	0	1	1	1
8	0	0	0	1	1	1
32	0	0	0	1	1	1
39	0	0	0	1	1	0
61	0	0	0	1	1	0
62	0	0	0	1	1	0
76	0	0	0	1	1	0
4	0	0	0	1	1	0
58	0	0	0	1	1	0
66	0	0	0	1	1	0
52	0	0	0	1	1	0
67	0	0	0	1	1	0
33	0	0	0	1	1	0
31	0	0	0	1	1	0
16	0	0	0	1	1	0
65	0	0	0	1	1	0
46	0	0	0	1	1	0
22	0	0	0	1	1	0
28	0	0	0	1	0	0
7	0	0	0	1	0	0

Supplemental Table 1

Supplemental Table 1.
Human Lung Metastatic Lesions from Triple Negative Breast Carcinomas Express E-Cadherin

Tissue microarray containing 25 cases of human pulmonary lesions from triple negative metastatic breast carcinoma. 23 cases show positive E-Cadherin expression. Id1 is expressed in 7 samples, all of which display E-Cadherin positivity.

Pulmonary lesions of triple negative breast carcinomas express epithelial marker E-Cadherin

ID No.	TMA Lung ER Pos	TMA Lung PR Pos	TMA Lung HER-2 Pos	TMA Lung 3X NEG	TMA Lung E-Cadherin Pos	TMA Lung Id1
2	0	0	0	1	1	No tumor
3	0	0	0	1	1	1
75	0	0	0	1	1	1
68	0	0	0	1	1	1
72	0	0	0	1	1	1
78	0	0	0	1	1	1
8	0	0	0	1	1	1
32	0	0	0	1	1	1
39	0	0	0	1	1	0
61	0	0	0	1	1	0
62	0	0	0	1	1	0
76	0	0	0	1	1	0
4	0	0	0	1	1	0
58	0	0	0	1	1	0
66	0	0	0	1	1	0
52	0	0	0	1	1	0
67	0	0	0	1	1	0
33	0	0	0	1	1	0
31	0	0	0	1	1	0
16	0	0	0	1	1	0
65	0	0	0	1	1	0
46	0	0	0	1	1	0
22	0	0	0	1	1	0
28	0	0	0	1	0	0
7	0	0	0	1	0	0

Supplemental Table 1

Supplemental Experimental Procedures

Animal Studies

All animal studies were performed in accordance with protocols approved by the MSKCC Institutional Animal Care and Use Committee. NOD/SCID mice between 6-8 weeks of age were used for xenograft studies. For orthotopic mammary fat pad injections, cells were re-suspended in 1:1 mixture of PBS and Matrigel (BD Biosciences). During the injections, animals were anaesthetized by aerosolized isoflurane. Mice were sacrificed when the tumors grew to more then 1cm in diameter. For lung colonization assays, $5X10^5$ cells were re-suspended in 100 μL PBS and injected into the lateral tail vein. The animals were monitored daily for any signs of duress or decreased health, at which point the animals were sacrificed.

Statistical Analysis

Data are presented as mean +/- SEM. Statistical significance was determined using the Student's t test (between two groups). P values < 0.05 were considered statistically significant.

Cell Proliferation assay

The proliferation assay was carried out by plating 10^4 cells into 6-well plates (for time points 1-5) or 10 cm plates (for time points 6-9). At each 24-hour time point cells were collected and counted using a standard hemocytometer protocol with the addition of trypan blue to exclude the dead cells.

Mammosphere assay

Mammosphere assays were performed as previously described (Dontu et al., 2003). Briefly, cells were grown under non-adherent conditions using Ultra-Low Attachment (Costar) 24-well plates in MEBM (Mammary Epithelial Cell Growth Medium)(Lonza) media supplemented with hEGF (20 ng/ml), human bFGF (20 ng/ml), hydrocortisone (0.5μg/ml), insulin (5 μg/ml), GA-1000 (Gentamicin 30 μg/ml and Amphotericin 15 ng/ml), heparin (4 μg/ml) and B27 Supplement. Mammosphere cultures were grown for 10-14 days at 37°C and 5% CO_2. Total mammosphere numbers were determined by counting all spheroid bodies with a diameter greater than 50 μm found in a single well.

RNA and Protein Analysis

RNA was extracted using the RNeasy kit (Qiagen). cDNA was generated from 1μg of RNA using SuperScript III First-Strand Synthesis System (Invitrogen). Quantitative PCR was performed using SYBR Green QuantiTect Primer Assay (Qiagen) according to manufacturer's instructions in a 7900HT Fast-Real Time PCR System Instrument (Applied Biosystems). Primer pairs for the individual genes were obtained from the bioinformatically validated QuantiTect Library and are as follows: *ID1* (QT00230650), ID3 (QT01673336), *SNAI1* (QT00010010), *TWIST1* (QT00097223), *CDH1* (QT00080143), *CDH2* (QT00063196), *VIM* (QT00095795), and *GAPDH* (QT00079247). The fold changes in gene expression were calculated using the ΔΔCt method.

For immunoblotting, cells were washed with PBS and lysed in RIPA buffer (50 nM Tris-HCl, pH 7.4, 1% Nonidet P-40, 0.25% Na-deoxycholate, 150 mM NaCl, 1 mM EDTA) supplemented with protease inhibitor (Roche). Proteins were separated by SDS-PAGE, transferred to a nitrocellulose membrane and probed with primary antibodies. Proteins were visualized using chemiluminescence (ECP-Plus, Amersham Biosciences) or LI-COR Odyssey Infrared Imaging detection system. The following antibodies were utilized for the detection of proteins by immunoblotting: ATF-3 (188, Santa Cruz Biotechnology), Id1 (Biocheck), Id3 (Biocheck), pSmad2 (3108, Cell Signalling), Smad3 (51-1500, Invitrogen), Smad4 (7966, Santa Cruz Biotechnology), p300 (584, Santa Cruz Biotechnology), Actin (Sigma), and Tubulin (Sigma).

Antigenic Phenotype Analysis using Flourescent Activated Cell Sorting

To identify $CD44^{high}/CD24^{low}$ subpopulation, the cells were incubated for 30 min on ice with 20 μL of the following primary labeled antibodies: PE mouse anti-human CD24 and APC mouse anti-human CD44 (BD Pharmingen). PE mouse IgG2A κ and APC mouse IgG2A κ antibodies (BD Pharmingen) were used as controls. Cells were then washed with HBBS, collected by centrifugation, and analyzed by FACS.

Immunofluorescence

Cells were seeded at 10^4 cells per well density in a 4-well glass slides. Once the cells reached 50-90% confluency, they were washed twice with phosphate-buffered saline (PBS) and fixed with 4% paraformaldehyde (PFA) for 20 min at room temperature followed by 3 more washes with PBS. The cells were then blocked with 10% normal goat serum (NGS), 1% BSA and 0.05% Triton X-100 for 1 hour at room temperature. Primary antibodies were diluted in blocking buffer and incubated at 4°C overnight. The following primary antibody dilutions were

used: Id1 (BioCheck, 1:200), E-Cadherin (Sigma, 1:500), Vimentin (Cell Signalling, 1:500) and β-Catenin (Santa Cruz, 1:500). After several washes with PBS, the cells were incubated with fluorescently-conjugated secondary antibodies (Alexa-488 and Alexa-568, 1:1000, Invitrogen). Following two washes with PBS, cells were incubated with DAPI for 5 minutes, washed with PBS, and mounted. Images were acquired using the Zeiss Axio2 Imaging microscope.

For immunofluorescent staining of mammospheres, mammospheres were embedded in OTC and sectioned at 5 μm thickness using a cryostat.

In order to perform immunofluorescence on animal tumor tissues, mice were sacrificed and perfused with PBS and 4% PFA. The tissues were fixed in 4% PFA overnight at 4°C, and embedded in OCT for frozen section (5 μm) preparation.

Immunohistochemistry

Tissue sections of formalin-fixed and paraffin-embedded human lung metastases were obtained from the Department of Pathology in accordance with protocols approved by the MSKCC Institutional Review Board.

Animal tumors were dissected and placed into a 4% PFA in PBS and fixed overnight. Tissue was then dehydrated in 70% ethanol, embedded in paraffin and sectioned into 5 μm or 10 μm sections. Tissue sections were deparaffinized, rehydrated, and stained with hematoxylin and eosin (H&E), as well as with the following primary antibodies: Id1 (BioCheck, 1:200), E-Cadherin (Sigma, 1:500), Vimentin (Cell Signalling, 1:500). Detection was performed using the Vectastain Elite ABC kit (Vector), diaminobenzidine was used as the chromogen, and Harris Modified Hematoxylin was used for nuclear counterstaining.

Immunohistochemistry of mammospheres was carried out using rabbit monoclonal anti p-Smad2/3 antibody (Cell Signalling, 9410).

Viral Production and Cell infections

To generate HMLE cells over-expressing Id1 protein, ID1 gene was cloned into the lentiviral delivery plasmid pWP1 (Addgene, #12254). Lentiviral particles were produced via Calcium Phosphate transfection method in 293T cells. 48 hours after transfection, virus was collected, filtered, and added to the target cells in the presence of 8 μg/ml polybrene. Infected cells were then selected for GFP positive cells using FACS. HMLE-Snail and HMLE-Twist cells were generated in a similar fashion, where pBabe-Puro-mTwist (Addgene #1783) and pBabe-Puro-Snail (Addgene # 23347) retroviral plasmids were used for the production of viral

particles and infected cells were selected using puromycin. For ATF-3 overexpression, the ATF3 cDNA was cloned into pBabe-puro, and retroviral particles were used to infect CN37 pleural effusion cells.

For the generation of the transformed HMLE cell lines, HMLE cells were transduced with the pBABE-HcRed-HRasV12 retrovirus. Viral particles were produced as described above. Cell selection was performed using FACS. For the generation of inducible Id1 HMLER cell lines, HMLE cells were transduces with pTRIPZ-Id1 lentiviral particles to produce HMLER-Id1 cell line. HMLER-Id1 cells were then transduced with pBabe-Puro-Twist and pBabe-Puro-Snail plasmids in order to generate HMLER-Twist-Id1 and HMLER-Snail-Id1 cells, respectively. Cells were selected for GFP positive cells using FACS and cultured for 2 weeks until they underwent a morphological change evident of EMT induction.

Luciferase Assay

CN37 cells were transiently transfected using Lipofectamine (Invitrogen) with various ID1 promoter reporter constructs. *Renilla* luciferase (pRL-TK*)* plasmid was used as internal control.

Gene Knockdown

Id1 and Id3 knockdown in HMLER-Twist cells was achieved with retroviral vectors that were validated previously (Gupta et al., 2007) (pRS-Puro-Id1sh, pRS-Hygro-Id3sh, and pRS-Puro-Control). Small interfering RNA (siRNA) transfections in CN37 cells were carried out using INTERFERin (Polyplus-transfection) according to the manufacturer's instructions. The siRNA oligos targeting Smad2 (M-003561) and Smad3 (M-020067) were obtained from Dharmacon, Thermo Scientific. Non-targeting siRNA (D001206) (Dharmacon) was used as a control. For the generation of CBP/p300 dKO cells, *CBPflox/flox*;*p300flox/flox* MEFs were treated with adenovirus expressing Cre recombinase or GFP control (Kasper et al., 2010). A pool of 3 target-specific lentiviral vector plasmids designed to knock down Smad4 was used for Smad4 knockdown experiments (Santa Cruz, 29484-SH), where scrambled shRNA sequence was used as a control.

Cell Culture

MCF-7, MDA-MB-157, MDA-MB-231, MDA-MB-361, MDA-MB-435, BT-549 breast cancer cell lines were grown in DMEM (Dulbecco's Modified Eagle Medium) media supplemented with 10% FBS (fetal bovine serum), 1% Penicillin-Streptomycin and 2mM L-Glutamine. HMLE cells and their derivatives were cultured in MEBM (Mammary Epithelial Cell Growth Medium)(Lonza), supplemented with BPE (Bovine Pituitary Extract) (70 μg/ml), hEGF (human

Epidermal Growth Factor) (5ng/ml), hydrocortisone (0.5 μg/ml), insulin (5 μg/ml), GA-1000 (Gentamicin 30 μg/ml and Amphotericin 15 ng/ml), hTransferrin (5 μg/ml), and isoproterenol (6 μM). CBP/p300 dKO MEFs were cultured in DMEM media supplemented with 10% FBS (fetal bovine serum), 1% Penicillin-Streptomycin and 2mM L-Glutamine.

References

Dontu, G., Abdallah, W. M., Foley, J. M., Jackson, K. W., Clarke, M. F., Kawamura, M. J., and Wicha, M. S. (2003). In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. Genes & development *17*, 1253-1270.

Gupta, G. P., Perk, J., Acharyya, S., de Candia, P., Mittal, V., Todorova-Manova, K., Gerald, W. L., Brogi, E., Benezra, R., and Massague, J. (2007). ID genes mediate tumor reinitiation during breast cancer lung metastasis. Proc Natl Acad Sci U S A *104*, 19506-19511.

Kasper, L. H., Lerach, S., Wang, J., Wu, S., Jeevan, T., and Brindle, P. K. (2010). CBP/p300 double null cells reveal effect of coactivator level and diversity on CREB transactivation. The EMBO journal *29*, 3660-3672.