

Supporting Information

Haughian et al. 10.1073/pnas.1106509108

SI Materials and Methods

Primary Breast Cancers. Paraffin-embedded tissue blocks of 72 primary tumors were collected from the archives of the Department of Pathology, Helsinki University Hospital (Finland) after approval of the Ethics Committee. The clinicopathological data have been published (1). Four-micrometer sections were subjected to dual colorimetric immunohistochemistry (IHC) (Envision G/2 Doublestain; Dako) using antibodies SP2 (Neomarkers) for PR and CK5 (Leica). Tumors with $\geq 1\%$ 3, 3'-diaminobenzidine (DAB)-stained nuclei or $\geq 1\%$ Perm Red-stained cytoplasm were considered positive.

Cell Lines and Reagents. T47D cells were from Iafa Keydar (Tel Aviv University, Tel Aviv, Israel); MCF-7 breast cancer cells were from Sam Brooks (Michigan Cancer Foundation, Detroit, MI); MCF-10A, BT-474, MDA231, SKBR3, and BT-20 were from the ATCC. SUM149 and SUM159 were from Asterand. Cell lines were maintained in the media suggested by the suppliers. BJ-3Z were derived as previously reported (2, 3). Cells were tagged with ZsGreen fluor by retroviral infection as previously described (4). All parental and derived cell lines have been authenticated and deemed pure by short tandem repeat analysis. For hormone and drug inhibitor treatments, or hormone withdrawal, cells were cultured in phenol red-free MEM and 5% FBS depleted of hormones by 30-min incubation at 45 °C with dextran-coated charcoal pellets [0.25% Norit A, 0.0025% dextran in 0.01 M Tris-HCl (pH 8.0); 1 mL charcoal per 2 mL serum] followed by filtration. A complete list of reagents and antibodies can be found in Table S2.

Xenografts and Tumor-Derived Cell Lines. All animal procedures were performed according to a protocol approved by the University of Colorado Institutional Animal Care and Use Committee. Tumors were grown by injecting 10^6 ZsGreen-tagged cells with Matrigel into the mammary fat pad of ovariectomized (ovx'd) athymic *nu/nu* (Taconic) or nonobese diabetic/SCID (National Cancer Institute) mice. Hormones were delivered by s.c. silastic pellets containing cellulose (placebo), or 2 mg estradiol-17 β (E) with cellulose. Tumors were harvested at one to four in vivo passage generations. At necropsy, tumors were resected aseptically, mechanically dissociated, and enzymatically digested for 2 to 3 h at 37 °C using a solution of HBSS containing collagenase IV, hyaluronidase, and DNase I (Sigma). Dissociated tumor cells were washed extensively before both stromal and tumor cells were plated into media containing neomycin (0.1 mg/mL) to select against the stromal cells and to select for tumor cells expressing the Neo resistance marker cloned into the ZsGreen plasmid. A list of tumor-derived cell lines grown from the parental T47D line is provided in Table S1. Numbers denote cell lines derived from the same tumor. In vitro, tumor cells were divided into steroid-depleted media supplemented with 1 nM E (denoted E-#) or no E (denoted EWD-#). Some lines were supplemented with EWD plus γ -secretase inhibitor (GSI) [denoted EWD-#+DAPT, compound E (CpdE), or XIX (defined in Table S1 legend)], or E plus tamoxifen (Tam) or fulvestrant (Fulv) (Table S1). E lines grew at a rate comparable to parental cells. Proliferation of EWD cells slowed initially, entered crisis, then resumed after 30–40 d, with GSI having no overt inhibitory or stimulatory effects. EWD-8 did not undergo crisis; it was derived from a placebo-treated mouse and had adapted to EWD in vivo.

3D Culture, Immunocytochemistry (ICC), IHC, and Western Blotting. IHC of 3D Matrigel colonies was reported previously (5). Briefly, cells were plated into eight-well chambered slides (BD) precoated with 60 μ L growth-factor reduced Matrigel and incubated overnight before treatments. For assessment of cell proliferation by BrdU incorporation, the 3D colonies were treated with 0.25 mg/mL BrdU for 3 h before harvest. After treatments, the Matrigel and colonies were embedded in Histogel (Richard-Allan Scientific), formalin fixed overnight, and then switched into 70% vol/vol ethanol. The histogel and Matrigel blocks containing the fixed colonies were then paraffin embedded and cut on a microtome (Leica) into 4- μ m sections. Sections were stained by standard IHC procedures. Antibody incubations used Tris-buffered saline (pH 7.4) containing 1% Triton-X. Nuclei were stained with 1 μ g/mL DAPI dissolved in methanol, and cells were mounted in Fluoromount (Sigma) for fluorescence or Permount (Fisher) for colorimetric stains. Dual colorimetric staining with DAB and Perm Red used Envision G/2 Doublestain Kit (Dako). Images were taken on a Nikon Eclipse E600 microscope equipped with a CoolSnap fx camera (Photometrics) operated by ImagePro software (Media Cybernetics). For Western blotting, cells were solubilized in RIPA buffer [50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, and 0.1% wt/vol SDS] amended with protease and phosphatase inhibitors (20 mM β -glycerophosphate, 0.1 mM sodium orthovanadate, 10 mM sodium fluoride, 1 mM PMSF, and protease inhibitor mixture; Roche). Protein concentrations were determined by Bradford assay (Bio-rad). Fifty micrograms of total protein were resolved on 10% SDS/PAGE gels, transferred, and probed. A list of antibodies is provided in Table S2.

Gene Expression Profiling and Analysis. All microarray procedures including total cellular RNA isolation, amplification, and Cy3/Cy5 labeling of 1.5 μ g of experimental and reference RNA were done as previously described (6, 7). Microarray hybridizations were performed using Agilent human oligonucleotide 4x44K custom designed gene chips. These data have been made available for download from the Gene Expression Omnibus (GEO) database, accession no. GSE31870. The gene probes were initially filtered by requiring the lowest normalized intensity values in both sample and reference to be >10 . Probes mapping to the same gene were averaged, and of 33,257 probes, a total of 16,937 genes were examined. The expression data were analyzed using a two-class unpaired significance analysis of microarrays, and all 1,298 genes comprising the “luminobasal” signature referenced in the article were considered differentially expressed with a false discovery rate of 0. Clustering analyses were performed using Cluster v2.12 (8). Colored squares in figures containing cluster analyses indicate the relative mean transcript abundance (in log₁₀ scale), with higher expression in red and lower expression in green. For comparisons with the 52 breast cancer cell line dataset generated by Neve et al. (9), Affymetrix Human Exon 1.0 ST Arrays were performed in triplicate from RNA isolated from the E-3 and EWD-8 cell lines. A total of 12,108 common genes were identified, and the raw data were normalized using robust multiarray average. An array platform/cell line correction value was generated by normalizing the gene expression values from the E-3 line in this study with the T47D line used in the Neve dataset. This correction value was then applied to the EWD-8 arrays before further analyses.

Two separate human primary tumor microarray datasets were used in these studies. The 516 tumor dataset was generated by combining three publicly available datasets available in the GEO repository: GSE2034 (10), GSE12276 (11), and GSE2603 (12). These tumor data were combined using distance weighted discrimination (13) and then subtyped using the PAM50 and claudin-low predictors (14, 15) with 159 luminal A, 112 luminal B, 79 Her2, 101 basal-like, and 65 claudin-low subtype tumors defined in the dataset. The NKI265 tumor dataset (http://microarray-pubs.stanford.edu/wound_NKI) contains 84 luminal A, 69 luminal B, 49 Her2, 42 basal-like, and 21 claudin-low tumors as previously described (15). Cluster (v2.12) and Java Treeview (v1.1.3) were used to assess how the 1,298 gene luminobasal signature clustered tumors within the 516 and NKI265 datasets.

Karyotype Analysis. Karyotype verification of clonal origin from a single cell line was performed as previously reported (16), with minor modifications. Briefly, cells were arrested with colcemid (0.05 $\mu\text{g}/\mu\text{L}$; 2.5 h), harvested in hypotonic KCl, and air dried to precleaned slides. Giemsa/Trypsin/Leishman banding was performed 4 d later by standard methods. Images were acquired and analyzed using the CytoVision (version 3.1) system (Applied Imaging). At least five metaphases per line were analyzed, and karyotyping designation followed the International System for Human Cytogenetic Nomenclature 2009 (17). Of 16 structural abnormalities previously reported in the T47D cell line (18), the parental T47D, E-3, EWD-3, and EWD-8 lines shared a total of 13 structural abnormalities, consistent with their clonal derivation from the T47D breast cancer cell line. This analysis failed to identify gain or loss of broad chromosomal abnormalities that would distinguish luminal and luminobasal cells.

Genotype Array. Genotyping of the parental T47D, E-3, EWD-3, and EWD-8 cell lines was performed on the Illumina Infinium Genotyping platform. Genomic DNA (gDNA) was isolated using the QIAamp DNA Mini Kit (Qiagen) and quantitated using PicoGreen (Invitrogen). The gDNA was processed according to Illumina's established protocols and DNA hybridized to the HumanOmni2.5-Quad BeadChip and imaged by an iScan Reader (Illumina). Changes to SNPs between the lines were assessed using the reproducibility and heritability function in Genome Studio Software (Illumina). A total of $\approx 2.43 \times 10^6$ SNPs in the genome were evaluated. The number of variants between the parental T47D and E-3, EWD-3, and EWD-8 lines was 191, 62, and 380, respectively. For reference, differences

between the genotype of an unrelated human female compared with the parental T47D line was 12,681 variant SNPs. The limited number of variant SNPs confirms the T47D origin of the E-3, EWD-3, and EWD-8 sublines and does not suggest a high degree of genetic instability at the level of the DNA base pairs.

Antibody Labeling and Flow Cytometry. Cells were harvested using HBSS supplemented with 1 mM EDTA and blocked with 10% normal human serum (Millipore) and 0.5% BSA for 15 min at 4 °C. For intracellular markers, cells were fixed using the Cytofix/Cytoperm Fixation/Permeabilization kit (BD Biosciences) for 20 min at 4 °C. Cells were stained for 30 min and then washed with multiple wash buffer changes for at least 30 min. When necessary, primary antibodies were conjugated using the Zenon Antibody Labeling Kit (Invitrogen) per the manufacturer's instructions.

Notch Promoter Reporter and shRNA Constructs. Cellular Notch activity was assessed using the RBP-J κ -Luc Signal Reporter (SA Biosciences). Breast cancer cell lines were transfected with Lip-oD293 (SignaGen Laboratories) and cotransfected with SV-40 promoter-renilla luciferase to control for transfection efficiency. Lentiviral particles encoding nonspecific (scrambled) or Notch1 targeted shRNAs were obtained from the University of Colorado Functional Genomics core laboratory. The Notch1 shRNA construct is identified within the Mission shRNA library (Sigma) as TRCN0000003360. The scrambled control vector was the Mission Non-Target shRNA Control Vector (SHC002; Sigma). Early-passage cells from the tumor-derived E-3 line were infected by mixing lentivirus with media containing 8 $\mu\text{g}/\text{mL}$ polybrene. Stably infected cells were selected using 2 $\mu\text{g}/\text{mL}$ puromycin for 1 wk and then maintained in media containing low dose (0.2 $\mu\text{g}/\text{mL}$) puromycin. The cells were switched out of E-containing media and maintained under these EWD/low-dose puromycin conditions for ≈ 8 wk to enable expansion of the luminobasal population.

Statistical Analysis. All reported values are mean \pm SEM. Significance levels for comparisons between treatment groups, or for comparisons of gene expression values, were analyzed using an unpaired Student *t* test or ANOVA. The microarray cluster analyses were derived using Java Treeview version 1.1.3. Hierarchical clustering showing average linkage was performed using Cluster v2.12 (8). The differences were considered significant when the *P* value was <0.05 .

- Joensuu K, Heikkilä P, Andersson LC (2008) Tumor dormancy: Elevated expression of stanniocalcins in late relapsing breast cancer. *Cancer Lett* 265:76–83.
- Jacobsen BM, et al. (2006) Spontaneous fusion with, and transformation of mouse stroma by, malignant human breast cancer epithelium. *Cancer Res* 66:8274–8279.
- Sartorius CA, et al. (1994) New T47D breast cancer cell lines for the independent study of progesterone B- and A-receptors: Only antiprogesterin-occupied B-receptors are switched to transcriptional agonists by cAMP. *Cancer Res* 54:3868–3877.
- Harrell JC, et al. (2006) Estrogen receptor positive breast cancer metastasis: Altered hormonal sensitivity and tumor aggressiveness in lymphatic vessels and lymph nodes. *Cancer Res* 66:9308–9315.
- Pinto MP, et al. (2010) Vascular endothelial growth factor secreted by activated stroma enhances angiogenesis and hormone-independent growth of estrogen receptor-positive breast cancer. *Cancer Res* 70:2655–2664.
- Hu Z, et al. (2006) The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* 7:96.
- Harrell JC, et al. (2011) Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer Res Treat*, 10.1007/s10549-011-1619-7.
- Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 95:14863–14868.
- Neve RM, et al. (2006) A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* 10:515–527.
- Wang Y, et al. (2005) Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 365:671–679.
- Bos PD, et al. (2009) Genes that mediate breast cancer metastasis to the brain. *Nature* 459:1005–1009.
- Minn AJ, et al. (2005) Genes that mediate breast cancer metastasis to lung. *Nature* 436:518–524.
- Benito M, et al. (2004) Adjustment of systematic microarray data biases. *Bioinformatics* 20:105–114.
- Parker JS, et al. (2009) Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 27:1160–1167.
- Prat A, et al. (2010) Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* 12:R68.
- Varella-Garcia M, Chen L, Zheng X, Yu L, Dixon D (2006) Karyotypic characteristics of human uterine leiomyoma and myometrial cell lines following telomerase induction. *Cancer Genet Cytogenet* 170:71–75.
- Brothman AR, Persons DL, Shaffer LG (2009) Nomenclature evolution: Changes in the ISCN from the 2005 to the 2009 edition. *Cytogenet Genome Res* 127:1–4.
- Lu YJ, Morris JS, Edwards PA, Shipley J (2000) Evaluation of 24-color multifluor-fluorescence in-situ hybridization (M-FISH) karyotyping by comparison with reverse chromosome painting of the human breast cancer cell line T-47D. *Chromosome Res* 8: 127–132.

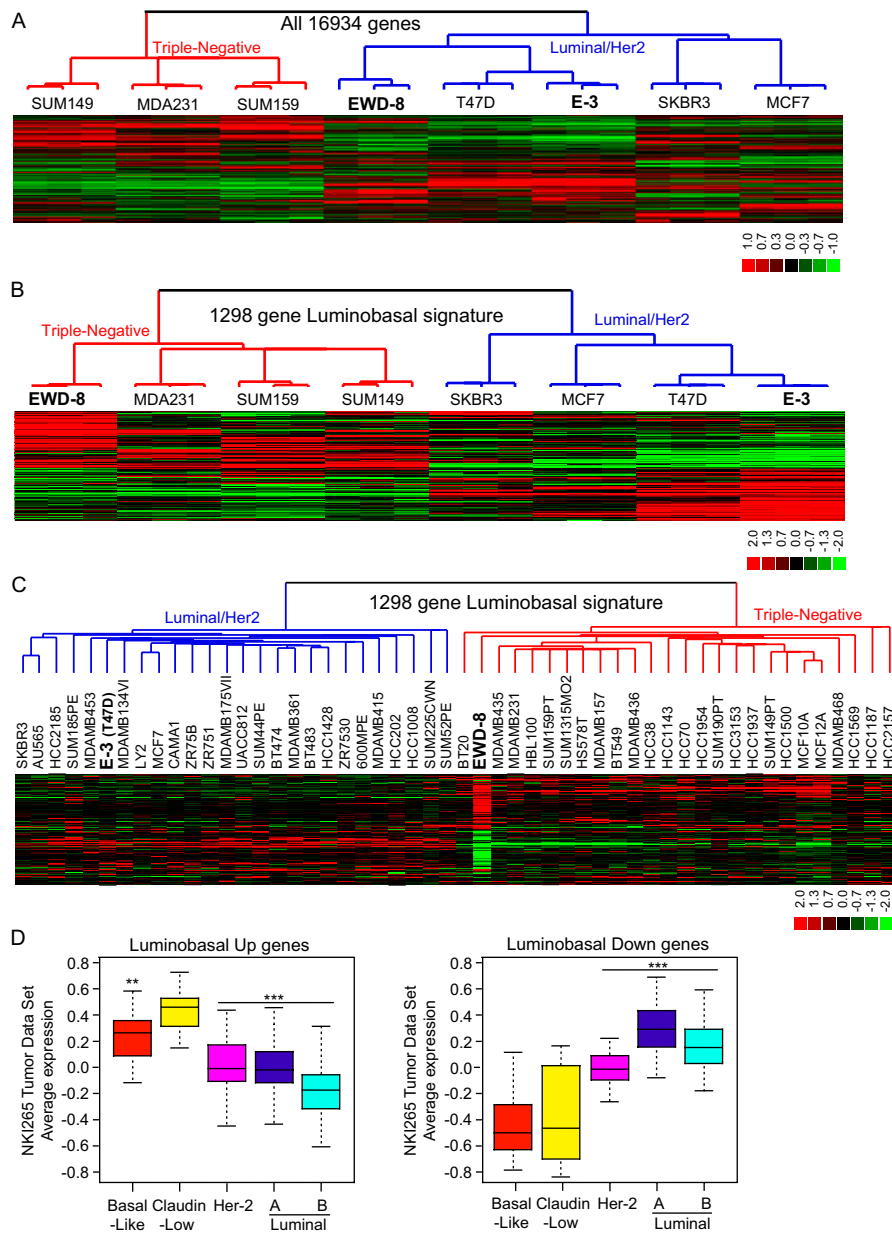


Fig. S3. Clustering on the luminobasal signature demonstrates an association with triple-negative breast cancer cell lines, whereas clustering across all genes confirms the luminal T47D origin of the luminobasal cells. (A) Clustering analysis based on all 16,934 genes present in the arrays. Gene expression was compared across the pure luminobasal EWD-8, pure luminal E-3, triple-negative (SUM149, SUM159, MDA231), Her2-positive/luminal (SKBR3), and luminal (MCF-7, T47D) cell lines. Note how the pure luminobasal EWD-8 line remains associated with the parental luminal T47D cell line. All lines were analyzed on the same microarray platform. (B) Clustering analysis of same breast cancer cell lines shown in A based on all 1,298 genes defining the luminobasal signature. Note that the pure luminobasal EWD-8 line now associates with triple-negative cell lines. (C) Clustering of the E-3 and EWD-8 lines with a large panel of breast cancer cell lines from Neve et al. (1) based on 1,298 genes in the luminobasal signature. (D) Luminobasal signature genes are associated with the triple-negative basal-like and claudin-low breast cancers in the NKI265 tumor dataset. ANOVA box plots of the average expression values for the (Left) 576 up and (Right) 526 down genes present in both the luminobasal signature and NKI265 tumors. The average was calculated across tumors within the indicated subtype. Sixty-six percent of tumors fall within the interquartile range (IQR) (colored box), where the bar indicates the median value; whiskers show the range within subtype and are 1.5*IQR. ****P < 0.001, ***P < 0.0001.**

1. Neve RM (2006) A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* 10:515–527.

Table S2. Reagents and antibodies

Chemical/antibody (ID)	Target/application	Dose/dilution	Source
17 β -estradiol	ER	1 or 10 nM	Sigma
Medroxyprogesterone Acetate	PR	10 or 100 nM	Sigma
4-OH Tamoxifen	ER	100 nM	Sigma
ICI-182780 (Fulvestrant)	ER	100 nM	Tocris Bioscience
Bromodeoxyuridine (BrdU)	DNA	0.25 mg/mL	Sigma
DAPT (GSI IX)	Notch, GSI	2 μ M	Calbiochem/EMD Biosciences
Compound E (GSI XXI)	Notch, GSI	5 nM	Calbiochem/EMD Biosciences
GSI XIX	Notch, GSI	5 nM	Calbiochem/EMD Biosciences
CK5 (NCL-L-CK5)	IHC, IF, ICC, WB	1:100; 1:1,000	Leica Microsystems
CK5 (2290-1)	IHC, IF, ICC	1:200	Epitomics
PR (Clone 1294)	WB	1: 1,000	DAKO
PR (SP2)	IHC, ICC	1:200	Neomarkers
ER (Ab-15)	WB	1:500	Neomarkers
ER- α (SP1)	IHC, IF	1:100	Neomarkers
β -actin (AC74)	WB	1:20,000	Sigma
GATA3 (HG3-31)	IF	1:200	Santa Cruz Biotechnology
Mucin 1 (VU4H5)	IF	1:100	Santa Cruz Biotechnology
FOXA1 (ab23738)	IF	1:600	Abcam
EGFR (06847)	IF, IHC	1:400	Millipore
Notch1 (D1E11XP)	IF	1:100	Cell Signaling
SLUG (C19G7)	IF	1:50	Cell Signaling
CK8/18 (NCL-L-503)	IHC, IF	1:100	Leica Microsystems
CK14 (RB-9020)	IF	1:200	Thermo Scientific
Smooth Muscle Actin (E184)	IF	1:400	Epitomics
p63 (Ab-4)	IF	1:50	Epitomics
Her2 (SP3)	IF	1:50	Neomarkers
Vimentin (NCL-Vim-V9)	IF	1:100	Leica Microsystems
BrdU (347580)	IHC	1:50	BD Bioscience

GSI, γ -secretase inhibitor; IHC, immunohistochemistry; IF, immunofluorescence; ICC, immunocytochemistry; WB, Western blot.

Other Supporting Information Files

[Dataset S1 \(XLSX\)](#)