SUPPLEMENTAL FIGURE LEGENDS

Figure S1, TAZ regulates lineage identity in human MECs (Related to Figure 2). A, qRT-PCR analysis comparing expression of various lineage markers in MCF10A and MCF10F cells. Gene expression values are presented as a log-fold change (MCF10F over MCF10A, N = 3 replicates). **B,** FACS plots demonstrating cell-state transitions in MCF10F cells. MCF10F were sorted into luminal and basal subpopulations on the basis of EpCAM and CD44 expression (N = 1 experiment). **C,** The purified luminal subpopulations in **B** were assessed for expression of various luminal and basal lineage markers. **D,** Cellular morphology of purified luminal and basal MCF10F cells 12 hours after plating. **E,** Gene expression of lineage markers in MCF10F cells overexpressing constitutively active TAZ-S89A (see also Figure 2E). **F,** Heat map of nCounter gene expression analysis of MCF10A cells lacking TAZ (divided into floating or adherent fractions), showing expression of lineage-specific genes or proliferation-associated genes compared to control cells. Red shading indicates high expression, green shading indicates low expression (N = 3 biological replicates). **G,** qRT-PCR analysis of an expanded list of lineage markers in MCF10A-shTAZ cells (adherent fraction). Expression values are plotted as relative mRNA expression compared to the control cell line, which is represented by the dotted line. The colored bars represent the indicated lineages specific to each marker gene (N = 9 experiments). **H,** Western blot of MCF10A-shTAZ lysates for estrogen receptor alpha (ER), confirming up-regulation of ER seen in floating MCF10A-shTAZ cells. **I,** Morphology and **J,** qPCR analysis of MCF10A cells with stable YAP knockdown by shRNA. In all panels, error bars represent the SEM, a single asterisk indicates $p < 0.05$, and double asterisks indicate $p < 0.01$ as determined by Student's t-test.

Figure S2, YAP/TAZ activity is confined to the basal/MaSC population (Related to Figure 3). A, Relative mRNA expression of TAZ (left) and YAP (right) in sorted subpopulations of primary human MECs isolated from breast reduction tissues (see Figure 3A, N = 3 tissue samples). **B,** Quantitation of TAZ protein levels in sorted subpopulations as determined by Western blotting and densitometry. **C,** Box plots showing enrichment of the TAZ signature in additional datasets downloaded through the Gene Expression Omnibus (GEO) in which mammary subpopulations were purified by FACS and

analyzed by microarray (Related to Figure 3H). The Shehata dataset identified three subpopulations of luminal progenitor cells, which were pooled together for statistical analysis. **D,** The datasets used to generate the TAZ signature, and the datasets used to test for enrichment in mammary subpopulations (in **C** as well as in Figure 3H) are indicated in the tables. The markers used to identify and sort specific subpopulations in each study, as well as the method of lineage depletion, is shown in the table. **E,** Gene set enrichment analysis (GSEA) of three TAZ target gene sets downloaded from the Molecular Signatures Database (Broad Institute), showing enrichment of both sets in the basal/MaSC population (Lim et al. 2010).

Figure S3, Analysis of *Wwtr1*^{lacZ} mammary glands during puberty (Related to Figure 4). A, Bar graph and **B,** table showing the proportions of wild-type, *Wwtr1*^{lacZ} heterozygous, and homozygous weanlings born to heterozygous crosses. **C-D,** Representative carmine-stained whole-mount images and H&E-stained sections of 5 week old **(C)** and 8 week old (**D)** *Wwtr1*^{lacZ} mice, demonstrating similar gland architecture across all genotypes. The white dotted lines in **D** mark the border of the invading terminal end buds into the fat pad. **E,** Quantification of ductal invasion in *Wwtr1*^{lacZ} mice at various developmental stages based on whole-mount staining (quantified as the maximum percent of fat pad length invaded by epithelium). **F,** Representative low-power (upper panels) and high-power (lower panels) images of histologic sections from 16-week old mice, stained with H&E. **G,** Flow cytometry analysis of *Wwtr1*^{lacZ} MECs at 8 weeks of age. **H,** Quantification of luminal and basal subpopulations in 8-week old epithelia as depicted by the gating strategy in **G.** In all panels, error bars represent the SEM and statistical significance values were calculated by Student's t-test. For five-week old mice, N = 4 mice per group, for eight week old mice, N = 1-5 mice per group.

Figure S4, Evolutionary conservation of the PPXY motif in SWI/SNF family members (Related to Figure 5). Clustal Omega alignment of the sequences of PPXY-containing SWI/SNF subunits comparing the homologs from various eukaryotic organisms. The regions containing the PPXY motifs are shown.

Table S1, Contents of transcription factor library (Related to Figure 1). A list of the NCBI gene symbols, ORF GenBank accessions, and Gene IDs for all transcription factors included in the TF screen.

Table S2, Results of IP/MS analysis of TAZ interacting proteins (Related to Figure 5). A list of all unique hits identified by mass spectrometry following IP of FLAG-tagged TAZ protein in 293T cells. Contains information including the number of peptides identified matching each indicated protein, and whether or not the protein contains either a PPXY motif (cognate to the WW domain of TAZ) or a PDZ domain (which can bind to the PDZ-binding motif at the C terminus of TAZ).

EXTENDED EXPERIMENTAL PROCEDURES

Isolation of Primary Human and Mouse MECs

Reduction mammoplasty tissues were obtained from the Department of Pathology at Tufts in accordance with IRB guidelines. Human MECs were isolated from reduction mammoplasty tissues as previously described (Keller et al. 2012). Briefly, the tissue was minced into ~1 mm pieces and incubated with 1.5 mg/ml collagenase and 125 units/ml hyaluronidase (Sigma), diluted in the growth medium, and incubated at 37° for 16 hours to release epithelial organoids. The organoids were separated from stromal cells by centrifugation at 200 x g, washed three times with PBS and once with RBC lysis buffer (Sigma), further digested with 0.05% trypsin/EDTA, and filtered through a 40 μm mesh filter for a single-cell suspension. Human MECs were cultured in MEGM (Lonza) for all assays, and were not passaged during culture.

Mouse MECs were isolated from the 3rd, 4th, and 5th mammary glands. After blunt dissection, glands were minced with scissors and incubated with 1.5 mg/ml collagenase and 125 units/ml hyaluronidase, diluted in Dulbecco's Modified Eagle Medium (DMEM) with 20ng/ml recombinant mouse epidermal growth factor, 10 μg/ml recombinant human insulin, 500 ng/ml hydrocortisone, and 5% calf serum, for 30-90 at 37°C. Organoids were pelleted by centrifugation at 1000 x g, washed, trypsinized and filtered as above to obtain a single-cell MEC suspension.

Cell Lines and Tissue Culture

For all experiments involving primary mammary epithelial cells, the cells were cultured in mammary epithelial basal medium (MEBM, Lonza) supplemented with the MEGM Bullet Kit additives (Lonza). MCF10A cells were also cultured in MEGM, with the addition of 100 ng/ml cholera toxin. Cells were subcultured weekly by trypsinization with 0.25% trypsin-EDTA, followed by a suspension in Dulbecco's Modified Eagle Medium (DMEM) with 10% fetal bovine serum (FBS) to inactivate trypsin. MCF10F cells were cultured in a mixture of DMEM and Ham's F12 media (DMEM/F12, 50:50) supplemented with 20 ng/ml epidermal growth factor (EGF), 10 μg/ml recombinant human insulin (Sigma), 500 ng/ml hydrocortisone (Sigma), and 100 ng/ml cholera toxin. MCF10F cells were subcultured weekly by incubation with 0.05% trypsin/EDTA.

Lentivirus Production and Generation of Stable Cell Lines

293T cells were transfected in a 10-cm dish with the appropriate lentiviral construct (3 µg DNA) and co-transfected with the vesicular stomatitis virus glycoprotein (pCMV-VSV-G, 1 µg) and lentiviral *gag-pol* genes (pCMV-Δ8.2Δvpr, 2 µg) generously provided by Inder Verma (Salk Institute, La Jolla, CA), using 20 µl FuGene 6 transfection reagent (Roche) according to the manufacturer's suggested protocol. After a 72-hour collection period, the viral supernatants were filtered through a 0.45 µm syringe filter and either used to directly infect target cells (for infection of MCF10A and MCF10F cells), or concentrated by ultracentrifugation at 27,000 RPM (to infect primary MECs). For immortalized cell lines, the cells were incubated for 4 hours in the viral supernatant with 5 protamine sulfate (Sigma), allowed to grow for 48 hours, and subsequently transferred to selection media containing either 2 µg/ml puromycin (Sigma) for 72 hours or 10 µg/ml blasticidin for one week (Life Technologies). For primary MECs, 250,000 freshly dissociated cells were centrifuged in the viral supernatant for 1 hour at 500 RPM at a multiplicity of infection (MOI) of 3 infectious particles per cell, unless otherwise indicated. Infected cells were incubated overnight with the concentrated viral supernatants in 24-well non-adherent tissue culture plates (Corning), washed, and plated for the appropriate colony-forming assays. Lentiviral vectors and their inserts, including shRNA sequences, are listed below.

Plasmid Name	Vector	Insert	Sequence Information	Source	ID Number	Refs
shTAZ #1	pLKO.1	*WWTR1* shRNA (CDS)	CCGGGCGATGAATCAGCCTCTGA ATCTCGAGATTCAGAGGCTGATTC ATCGCTTTTT	Sigma	TRCN0 000019 469	
shTAZ #2	pLKO.2	*WWTR1* shRNA (3' UTR)	CCGGTAAGCTTTATGGGTGTTAAT TCTCGAGAATTAACACCCATAAAG CTTATTTTTG	Sigma	TRCN0 000370 06	
shBRG1	pLKO.1	*SMARCA4* shRNA (CDS)	CCGGCCGAGGTCTGATAGTGAAG AACTCGAGTTCTTCACTATCAGAC CTCGGTTTTT	Sigma	TRCN0 000015 551	
shBRM #1	pLKO.1	*SMARCA2* shRNA (CDS)	CCGGGCTGAGAAACTGTCACCAA ATCTCGAGATTTGGTGACAGTTTC TCAGCTTTTT	Sigma	TRCN0 000020 329	
shBRM #2	pLKO.1	*SMARCA2* shRNA (CDS)	CCGGCGGAATCTTAGCCGATGAA ATCTCGAGATTTCATCGGCTAAGA TTCCGTTTTT	Sigma	TRCN0 000020 333	
shYAP	pLKO.1	*YAP1* shRNA (3' UTR)	CCGGCCCAGTTAAATGTTCACCAA TCTCGAGATTGGTGAACATTTAAC TGGGTTTTTG	Sigma	TRCN0 000107 265	

TAZ (WT)	pCMV5-TOPO	*WWTR1* cDNA and N-terminal 3X FLAG sequence	Matches reference sequence in RefSeq mRNA database	Addgene	24809	PMID: 18568018
TAZ (dWW)	pCMV5-TOPO	*WWTR1* cDNA and N-terminal 3X FLAG sequence	Lacks codons corresponding to AAs 111-158 of TAZ CDS	Addgene	24811	PMID: 18568018
TAZ (S89A)	pCMV5-TOPO	*WWTR1* cDNA and N-terminal 3X FLAG sequence	TCG -> GCC missense mutation in codon 89	Addgene	24815	PMID: 18568018

Transcription Factor Screen

cDNAs representing over 1000 human transcription factors were cloned into the pLenti 6.2 V5/DEST vector. Using gene annotation from EntrezGene as a guide, all proteins having a 'DNA-binding' function in the human genome were identified, and the corresponding MGC/ORF clones arrayed as a glycerol stock subset from the available human ORF collection at the Harvard Institute of Proteomics.

Entry clone DNA was prepared in 96-well format as described (Rolfs et al., 2008; Rolfs et al., 2008) and purification confirmed by 96-well DNA gel electrophoresis (E-Gel, Invitrogen/Life Technologies/Thermo). To transfer the entry clones into lentiviral vectors, 5 µl aliquots of each DNA were mixed with 20 ng of destination vector (pLenti 6.2 V5/DEST, Life Technologies), together with LR enzyme/buffer mix (Life Technologies) in a total of 10 µl, incubated for 60 minutes at 25°C, and transformed into chemically-competent *E. coli*. Transformed cells were plated individually onto LB-ampicillin agar plates, incubated over night at 37°C, and the next day individual colonies were isolated and expanded overnight in LB-ampicillin broth. Restriction digests and gel electrophoresis confirmed the presence of the appropriate inserts. All entry clones used here have been full-length sequence validated and submitted to GenBank as part of the clone collection of the ORFeome Collaboration.

The TF DNAs were pooled together and used to generate lentivirus as described above. Two million FACS-purified EpCAM+ primary human MECs were transduced with the pooled library at a multiplicity of infection of 0.3 lentiviral particles per cell, to ensure a low probability of multiple integration. For infection, the cells were divided into aliquots of 250,000 cells in each well of an ultra-low attachment 24-well tissue culture plate (Corning), centrifuged for 1 hour at 500 RPM, and incubated overnight with lentivirus. Cells were plated in a 10-cm dish for 48 hours, after which 10

μg/ml blasticidin was added to the growth medium. At this point, half of the cells were harvested immediately and constitute the “pre-screen” sample, while the other half (“screen” sample) were cultured for an additional 8 days to allow for colony formation, after which the non-adherent cells were washed off. The remaining adherent cells were harvested with trypsin-EDTA, and genomic DNA was extracted using the PureLink Genomic DNA Mini Kit (Life Technologies). TF inserts were recovered from the genomic DNA by PCR using insert-flanking primers complimentary to the pLenti 6.2/V5 vector. The resulting PCR products from the “screen” and “pre-screen” samples were ligated to bar-coded TruSeq adapters (Illumina) and sequenced together in a single lane of an Illumina Genome Analyzer IIx. The sequencing reads were aligned to the insert sequences in GenBank using Bowtie alignment software (http://bowtie-bio.sourceforge.net). For each library (“screen” and “pre-screen”), enrichment scores were calculated by determining the total number of reads aligning to each TF cDNA in the “screen” sample and dividing by the number of reads in the “pre-screen” sample. Enrichment scores greater than 2 were considered to be positive hits.

Quantitative RT-PCR

Primer sequences for all qPCR experiments are listed in the following table.

Gene Symbol	Species	Forward Primer	Reverse Primer
ACTA2	human	CAGGGCTGTTTTCCCATCCAT	GCCATGTTCTATCGGGTACTTC
ALDH1A3	human	TGTGCGGACGCTGACTTGGAC	GGCATACTCCACGCTCCGCC
ANKRD1	human	AGTAGAGGAACTGGTCACTGG	TGGGCTAGAAGTGTCTTCAGAT
BMI1	human	AGCCATTTTGATTGCTGTTTGA	CCGCTTTTAGGCATACAGATTGTA
CD14	human	ACGCCAGAACCTTGTGAGC	GCATGGATCTCCACCTCTACTG
CD24	human	CTCCTACCCACGCAGATTTATTC	AGAGTGAGACCACGAAGAGAC
CD44	human	AGATCAGTCACAGACCTGCC	GCAAACTGCAAGAATCAAAGCC
CDH1	human	GAACGCATTGCCACATACAC	GAATTCGGGCTTGTTGTCAT
CLDN4	human	GGGGCAAGTGTACCAACTG	GACACCGGCACTATCACCA
CLDN8	human	CAACCCATGCCTTAGAAACCGC	TCACGCAATTCATCCACAGTC
CNN1	human	ATGTCCTCTGCTCACTTCAACC	CCCCCTCGATCCACTCTCT
CTGF	human	AGGAGTGGGTGTGTGACGA	CCAGGCAGTTGGCTCTAATC
ELF5	human	TCTGCCTCACTCCCACAGGGTA	CCACTCCCACACATGGCGCT
EpCAM	human	AATCGTCAATGCCAGTGTACTT	TCTCATCGCAGTCAGGATCATAA
ERBB3	human	CTGATCACCGGCCTCAAT	GGAAGACATTGAGCTTCTCTGG
ESR1	human	ATTTGAAGTGGGCATGAGAACAT	CAATACCAACATCAGCCAGAAA
GATA3	human	GCGGGCTCTATCACAAAATGA	GCTCTCCTGGCTGCAGACAGC
KRT14	human	CATGAGTGTGGAAGCCGACAT	GCCTCTCAGGGCATTCATCTC
KRT15	human	TCTGCTAGGTTTGTCTCTTCAGG	CCAGGGCACGTACCTTGTC
KRT18	human	TGATGACACCAATATCACACGAC	TACCTCCACGGTCAACCCA

KRT19	human	ACCAAGTTTGAGACGGAACAG	CCCTCAGCGTACTGATTTCCT
KRT6B	human	GGCCCTCAAGGATGCTAAGAACAA	TGACGTTCATCAGCTCCTGGTACT
MME (CD10)	human	GATCAGCCTCTCGGTCCTTG	TGTTTTGGATCAGTCGAGCAG
MMP7	human	GAGTGAGCTACAGTGGGAACA	ATGTGGAGTGCCAGATGTTGC
MUC1	human	CGCCGAAAGAACTACGGGCAGCTG	CAAGTTGGCAGAAGTGGCTGCCAC
NOTCH1	human	GAGGCGTGGCAGACTATGC	CTTGTACTCCGTCAGCGTGA
OXTR	human	TGTTCGCCTCCACCTACCT	CTCGCGCAGAGAGAAGATGTG
PRLR	human	CAAAGCCCACTGTGAAATATGA	GAGCCTGAAACTCAAAGGAGAA
S100A8	human	ATGCCGTCTACAGGGATGAC	ACTGAGGACACTCGGTCTCTA
SERPINB7	human	AAATGCAGAGTTTTGCTTCAACC	GAAGAGTTTCCATATCCTGAGGC
SNAI2	human	GCATTTCTTCACTCCGAAGC	TGAATTCCATGCTCTTGCAG
TFF3	human	AGAAAAACTGTCTGGGAGCTTG	CTCATTTATGCACCGTTGTTTG
TP63 (deltaN)	human	GGAAAACAATGCCCAGACTC	GTGGAATACGTCCAGGTGGC
TP63 (TA)	human	AAGATGGTGCGACAAACAAG	AGAGAGCATCGAAGGTGGAG
TTYH1	human	CAGGTGGCTGAAAATGTGTCC	TCTGGATTGGAGCAGAAGTCA
VIM	human	GAGTCCACTGAGTACCGGAGAC	TGTAGGTGGCAATCTCAATGTC
Wwtr1/Taz	mouse	CATGGCGGAAAAAGATCCTCC	GTCGGTCACGTCATAGGACTG
WWTR1/TAZ	human	GGCTGGGAGATGACCTTCAC	CTGAGTGGGGTGGTTCTGCT
YAP1	human	TAGCCCTGCGTAGCCAGTTA	TCATGCTTAGTCCACTGTCTGT

nCounter analysis

RNA was prepared as described in Experimental Procedures. For each sample, ~100 ng of total RNA was used to measure gene expression of 110 selected breast cancer genes, including the PAM50 genes (Bernard et al., 2009), using the digital color-coded nCounter platform (Nanostring Technologies, Seattle, (Geiss et al., 2008)). Gene expression values were log base two transformed and normalized using 5 house-keeping transcripts.

Western Blotting and Immunoprecipitation

For FLAG immunoprecipitation and mass spectrometry, 5 mg of total protein lysate was extracted from 293T cells transfected with either FLAG-TAZ or an empty-vector control in lysis buffer containing 50 mM Tris (Sigma), 150 mM NaCl, 2 mM EDTA and 1% Triton-X100, for 20 minutes on wet ice with agitation, followed by 5X passage through a 20G needle. Lysates were incubated with 30 µl pre-washed FLAG-M2 agarose beads (Sigma) and incubated overnight at 4°C. The beads were precipitated by centrifugation at 10000 × g, washed five times with lysis buffer, and eluted in 5 µg/ml 3X-FLAG peptide (Sigma) with a 30 minute incubation at 4°C. 5% of the eluate was analyzed by SDS-PAGE and Western blotting to confirm the presence of FLAG-TAZ protein, while the remaining 95% of

the eluate was subjected to SDS-PAGE and stained with Coomassie blue dye. The gel sections were excised for mass spectroscopic analysis (Taplin Lab Mass Spectrometry Facility, Harvard University).

For immunoprecipitation of endogenous TAZ from MCF10A cells, nuclear extracts were prepared using the NE-PER Extraction Kit (Pierce). 2 mg of nuclear extract were incubated overnight with 3 µg mouse α-TAZ (BD), mouse α-BRG1 (Santa Cruz), or normal mouse IgG diluted to 500ul in lysis buffer (see above), followed by a 3-hour incubation with 20 µl Protein A agarose beads (RepliGen). Immune complexes were eluted off the beads by boiling in Laemmli buffer for five minutes.

For all Western blotting, protein samples were subjected to SDS-PAGE and transferred to 0.45 µm nitrocellulose (BioRad). Membranes were blocked in Tris-buffered saline (TBS, 50 mM Tris, 150 mM NaCl) with 0.1% Tween-20 detergent (Sigma), probed with the appropriate primary antibody overnight at 4°C, washed, and incubated with HRP-conjugated secondary antibodies for 1 hour at room temperature. After a subsequent washing step, membranes were developed with ECL substate (Pierce) and exposed to film for signal detection. Antibodies used for immunoprecipitation and Western blotting are listed in the following table.

Antibody	Host Species	Clone	Vendor	Catalog No.	Dilution
TAZ	Rabbit	poly	Cell Signaling	2149	1:1000
YAP/TAZ	Rabbit	D24E4	Cell Signaling	8418	1:1000
p-LATS1 (T1079)	Rabbit	D47D3	Cell Signaling	8654	1:1000
LATS1	Rabbit	C66B5	Cell Signaling	3477	1:1000
BRM/SMARCA2	Rabbit	D9E8B	Cell Signaling	11966	1:1000
BRG1/SMARCA4	Mouse	G-7	Santa Cruz	sc-17796	1:250
Beta-actin	Mouse	AC-15	Abcam	ab6276	1:20,000
Lamin A/C	Rabbit	poly	Cell Signaling	2032	1:2,000
HRP anti-Mouse IgG	Horse	-	Cell Signaling	7076	1:10,000
HRP anti-Rabbit IgG	Goat	-	Cell Signaling	7074	1:10,000

Immunostaining

For both immunofluorescence (IF) and immunohistochemical (IHC) staining of human or mouse tissues, 5-µm formalin-fixed, paraffin-embedded sections were first de-paraffinized in xylenes and dehydrated in graded ethanols, followed by an antigen retrieval step in which sections were incubated in 10mM citrate buffer (pH 6.0) at 95°C for 30 minutes. Sections were washed three times

in PBS, blocked with 2% normal horse serum and 1% bovine serum albumin (BSA) for 30 minutes, and incubated overnight at 4°C with the primary antibody (diluted in PBS + 1% BSA). The following day, sections were washed with PBS, incubated 1 hour with biotinylated or fluorescent secondary antibody, and washed again with PBS. For IHC, slides were developed using the Vectastain ABC Kit and Impact DAB substrate (Vector) according to the manufacturer's instructions, followed by counterstaining with hematoxylin. For IF, slides were incubated with 250 ng/ml DAPI for 5 minutes, washed, treated with SlowFade antifade reagent (Life Sciences) mounted with coverslips. Microscopy was performed using a Nikon Eclipse 80i microscope and SPOT image acquisition software (Micro Video Instruments). Staining for Ki67 was performed by the Histology Special Procedures Laboratory at Tufts Medical Center. A complete list of antibodies used in immunostaining is detailed below:

Immunofluorescence

Antibody	Host Species	Clone	Vendor	Catalog No.	Dilution
anti-EpCAM (mouse)	Rabbit	E144	Abcam	ab32392	1:100
anti-Smooth muscle actin	Mouse	asm-1	Vector	VP-S281	1:100
YAP/TAZ	Rabbit	D24E4	Cell Signaling	8418	1:100
KRT14	Mouse	LL002	Vector	VP-C410	1:200
Alexa 488 anti-Mouse IgG	Goat	-	Life Technologies	A-11001	1:500
Alexa 546 anti-Rabbit IgG	Goat	-	Life Technologies	A-11010	1:500

Immunohistochemistry

Antibody	Host Species	Clone	Vendor	Catalog No.	Dilution
YAP/TAZ	Rabbit	D24E4	Cell Signaling	8418	1:100
KRT14	Rabbit	-	Thermo Scientific	RB-9020-P	1:500
KRT8/KRT18	Mouse	5D3	Vector	VP-C407	1:500
Biotin anti-Mouse IgG	Horse	-	Vector	BA-2000	1:500
Biotin anti-Rabbit IgG	Goat	-	Vector	BA-1000	1:500

Flow Cytometry

Antibodies used for flow cytometry are listed below.

Antibody	Host Species	Clone	Vendor	Catalog No.	Dilution
APC-anti-CD24 (mouse)	Rat	M1/69	eBioscience	17-0242	1 ul/ 2M cells
PerCP Cy5.5-anti-CD49f (mouse)	Rat	GoH3	Biolegend	313618	2 ul/ 100K cells
PE-anti-TER119 (mouse)	Rat	-	eBioscience	12-5921	1 ul/ 100K cells
PE-anti-CD31 (mouse)	Rat	390	eBioscience	12-0311-83	2 ul/ 100K cells
PE-anti-CD45	Rat	30-F11	BD Biosciences	553081	1 ul/ 2M cells
APC-anti-EpCAM (human)	Mouse	EBA-1	BD Biosciences	347200	2 ul/ 100k cells
PE-anti-CD24 (human)	Mouse	ML5	BD Biosciences	555428	4 ul/ 100k cells

FITC-anti-CD49f	Rat	GoH3	BD Biosciences	555735	4 ul/ 100k cells
FITC-anti-EpCAM	Mouse	VU-ID9	AbD Serotec	MCA1870F	4 ul/ 100k cells
APC-anti-CD44	Mouse	G44-26	BD Biosciences	559942	4 ul/ 100k cells

Chromatin Immunoprecipitation (ChIP)

ChIP was performed using the SimpleChIP Enzymatic Chromatin IP kit (Cell Signaling Technology) according to the manufacturer's instructions. For each set of immunoprecipitations, five 15-cm tissue culture plates were seeded with cells and grown to 80% confluency to prepare chromatin. Briefly, the chromatin was cross-liked *in situ* and digested with 5 ul micrococcal nuclease for 20 minutes to yield chromatin fragments primarily consisting of mono-, di- and tri-nucleosomes. 10% of the prepared chromatin was used for each individual immunoprecipitation.

Precipitated chromatin fragments were quantified as described in 'Quantitative RT-PCR,' and the data are expressed as fold enrichment over the normal IgG control. Control primers against the *RPL30* exon 3 locus were provided in the SimpleChIP kit. The genomic primers targeting *CTGF* are as follows:

CTGF forward: 5'-GTGGTGCGAAGAGGATAGGG-3'

CTGF reverse: 5'-TGACACTCAGCATTCCTCCG-3'

Antibodies used for ChIP are given in the table below:

ChIP Antibodies

Antibody	Host Species	Clone	Vendor	Catalog No.	Dilution
TAZ	Rabbit	poly	Sigma	HPA007415	1:50
BRG1/SMARCA4	Mouse	G-7	Santa Cruz	sc-17796	1:50
BRM/SMARCA2	Rabbit	D9E8B	Cell Signaling	11966	1:100

Image Quantitation and Statistical Analyses

Quantitation of TAZ nuclear localization, ductal branching, ductal invasion luminal/basal cell staining by IF, and Ki67 staining was carried out in ImageJ software. Quantitations were carried out in the following manner:

For quantitation of the percent of cells with nuclear TAZ (Figure 3D), pixels falling into the blue or purple color spectrum were first removed from all of the images using the "Color Threshold"

function, isolating the brown signal from the IHC substrate. Next, the staining intensity of 25 nuclei from the no-antibody control sections were quantified using the "Analyze Particles" function to determine the mean background signal intensity. This procedure was repeated for at least 200 nuclei from each of the TAZ-stained tissues. Any nuclei in TAZ-stained IHC images with intensity greater than two standard deviations above the mean background level were considered to be positive for TAZ.

For quantitation of ductal branching, glands were photographed at low magnification to capture the entire gland area, and the total number of branch points was counted for each animal. "Percent ductal invasion" was computed as the length of the most distant terminal end bud or branch was divided by the total length of the gland as measured in ImageJ.

For quantitation of EpCAM and SMA staining, the number of nuclei surrounded by EpCAM or SMA signal was counted and averaged over at least five high-powered fields (60x magnification) per animal.

For Ki67 quantitation, Ki67 cells were first counted and classified as either basal or luminal based on anatomic location and morphology. The mean number of Ki67+ in at least eight low-power fields (10x magnification) was calculated for each animal. The number of positive cells was then normalized based on the percent of the gland area occupied by epithelium to account for differences in mammary gland cellularity across genotypes.

Kaplan-Meier curves were generated using the KMplot analysis tool, which also calculated logrank p-values (Gyorffy et al., 2010). All other statistical tests, including Student's t-test and the Chi-square test, were carried out using the built-in stastical functions in Microsoft Excel.

All TCGA data was obtained from the TCGA breast cancer online portal (https://tcga-data.nci.nih.gov/docs/publications/brca_2012/, (Koboldt et al., 2012). The following files were used: For microarray gene expression data: "*BRCA.exp.547.med.txt*". For reverse-phase protein array (RRPA) expression data: "*rppaData-403Samp-171Ab-Trimmed.txt*". For microarray DNA copy-number aberration data: "*brca_scna_all_thresholded.by_genes.txt*". For intrinsic subtype identification, we used the PAM50 subtype calls as provided in the TCGA portal.

Microarray data used in enrichment analyses was obtained through the GEO portal (Mootha et al., 2003; Subramanian et al., 2005, http://www.ncbi.nlm.nih.gov/geo/). Relevant GEO accession numbers for the enrichment analyses are shown in Figure S2. GSEA enrichment scores and enrichment plots were generated using the GSEA software (http://www.broadinstitute.org/gsea/index.jsp).

SUPPLEMENTARY REFERENCES

Bernard, P.S., Parker, J.S., Mullins, M., Cheung, M.C.U., Leung, S., Voduc, D., Vickery, T., Davies, S., Fauron, C., He, X.*, et al.* (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. Journal of Clinical Oncology *27,* 1160-1167.

Geiss, G.K., Bumgarner, R.E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D.L., Fell, H.P., Ferree, S., George, R.D., Grogan, T.*, et al.* (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat. Biotechnol. *26,* 317-325.

Koboldt, D.C., Fulton, R.S., McLellan, M.D., Schmidt, H., Kalicki-Veizer, J., McMichael, J.F., Fulton, L.L., Dooling, D.J., Ding, L., Mardis, E.R.*, et al.* (2012). Comprehensive molecular portraits of human breast tumours. Nature *490,* 61-70.

Mootha, V.K., Lindgren, C.M., Eriksson, K.-, Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E.*, et al.* (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. *34,* 267-273.

Rolfs, A., Hu, Y., Ebert, L., Hoffmann, D., Zuo, D., Ramachandran, N., Raphael, J., Kelley, F., McCarron, S., Jepson, D.A.*, et al.* (2008). A biomedically enriched collection of 7000 human ORF clones. PLoS ONE *3,*

Rolfs, A., Montor, W.R., Sang, S.Y., Hu, Y., Bhullar, B., Kelley, F., McCarron, S., Jepson, D.A., Shen, B., Taycher, E.*, et al.* (2008). Production and sequence validation of a complete full length ORF collection for the pathogenic bacterium Vibrio cholerae. Proc. Natl. Acad. Sci. U. S. A. *105,* 4364-4369.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America *102,* 15545-15550.