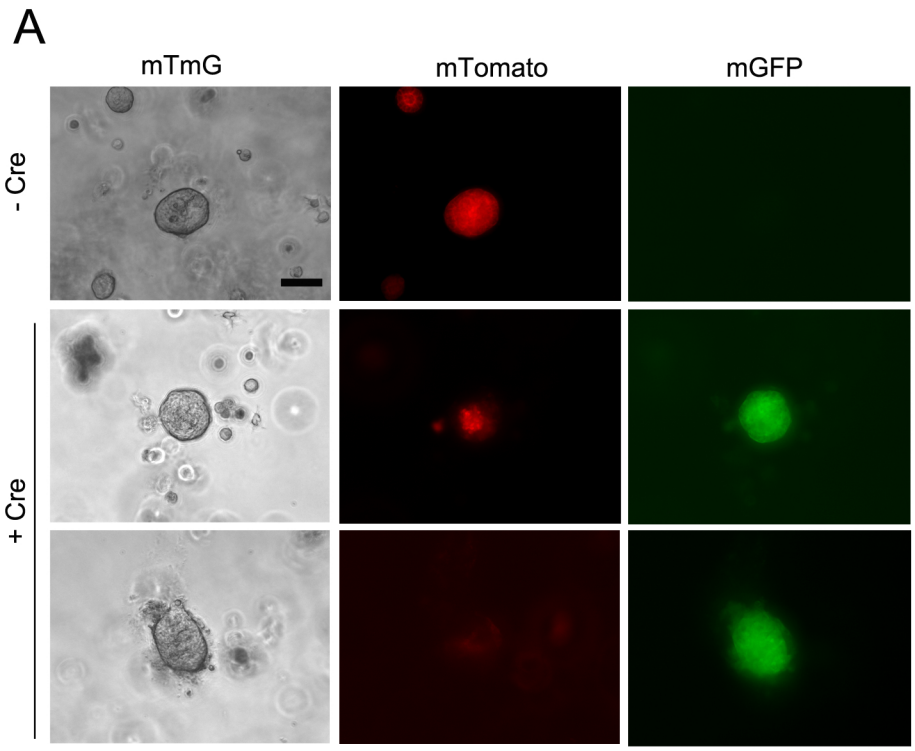
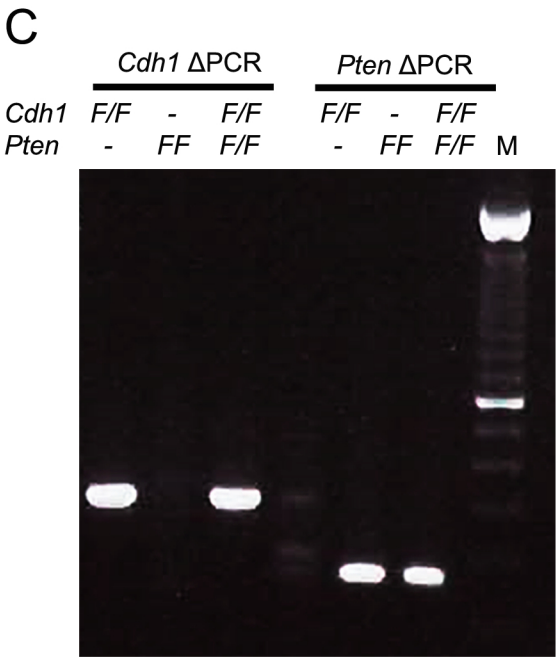
Cell Reports, Volume *16*

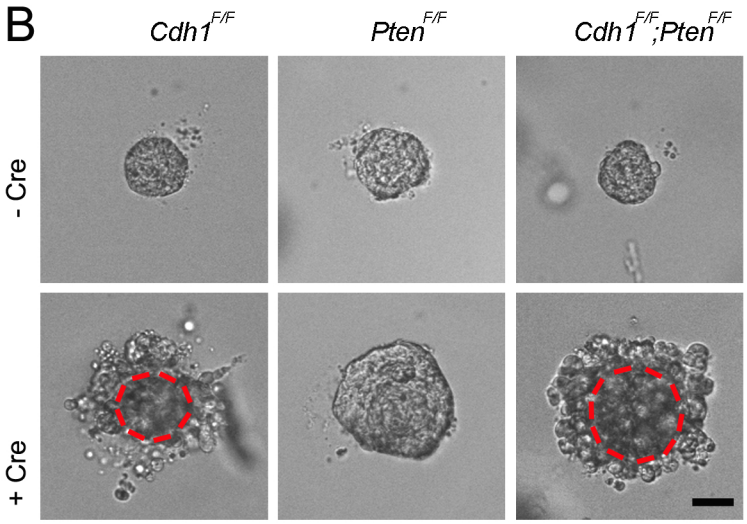
Supplemental Information

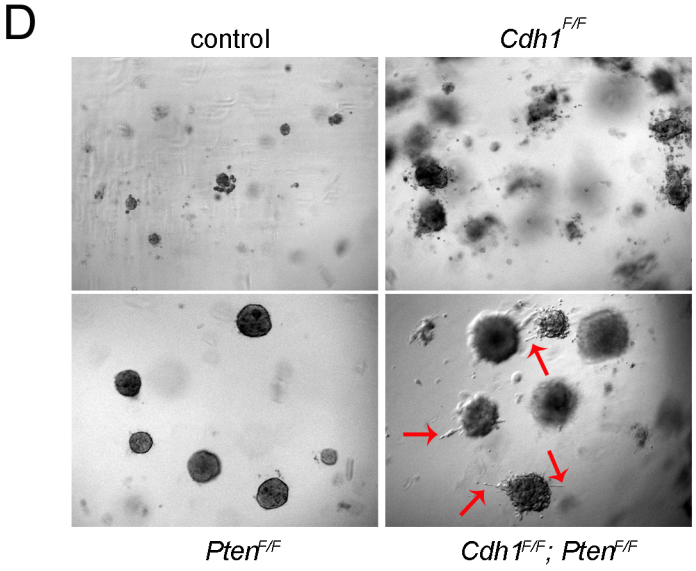
PTEN Loss in E-Cadherin-Deficient Mouse Mammary Epithelial Cells Rescues Apoptosis and Results in Development of Classical Invasive Lobular Carcinoma

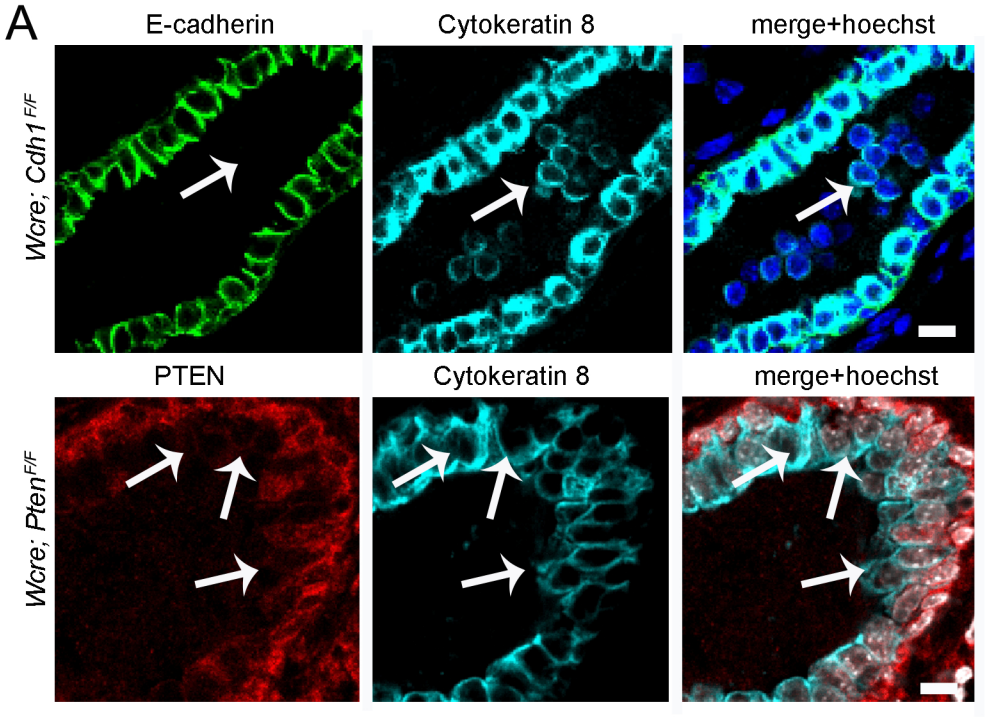
Mirjam C. Boelens, Micha Nethe, Sjoerd Klarenbeek, Julian R. de Ruiter, Eva Schut, Nicola Bonzanni, Amber L. Zeeman, Ellen Wientjens, Eline van der Burg, Lodewyk Wessels, Renée van Amerongen, and Jos Jonkers

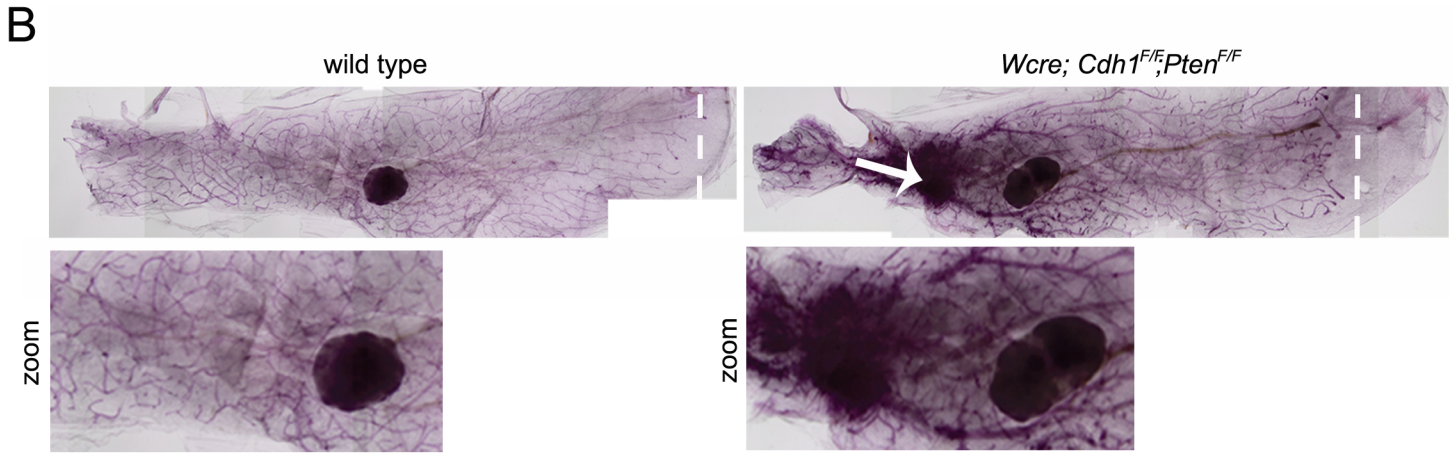

A
mTmG
mTomato
mGFP
- Cre
+ Cre

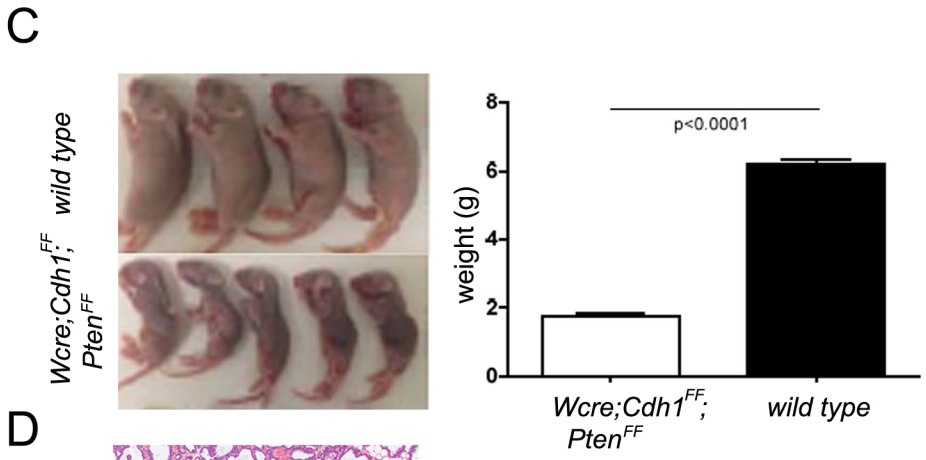

C
Cdh1 ΔPCR
Pten ΔPCR
Cdh1 F/F - F/F F/F - F/F
Pten - FF F/F - FF F/F M

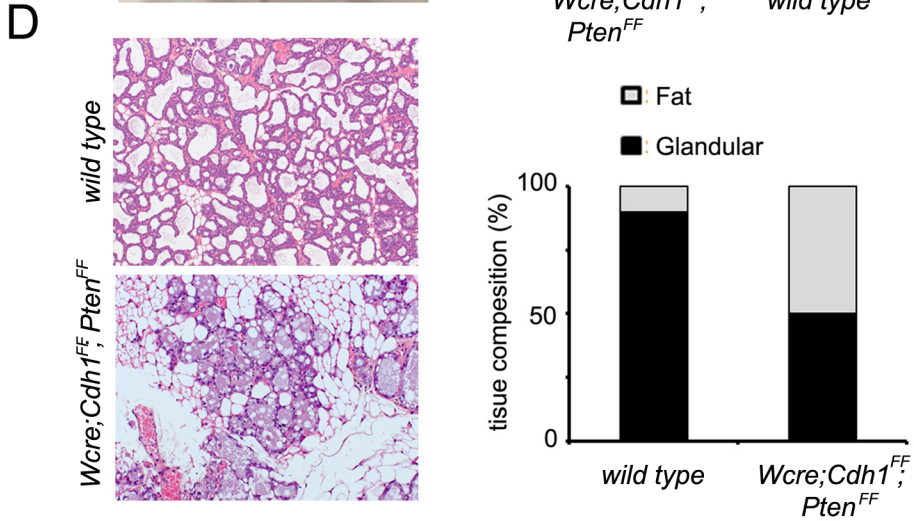

B
Cdh1F/F
PtenF/F
Cdh1F/F;PtenF/F
- Cre
+ Cre


D
control
Cdh1F/F
PtenF/F
Cdh1F/F; PtenF/F

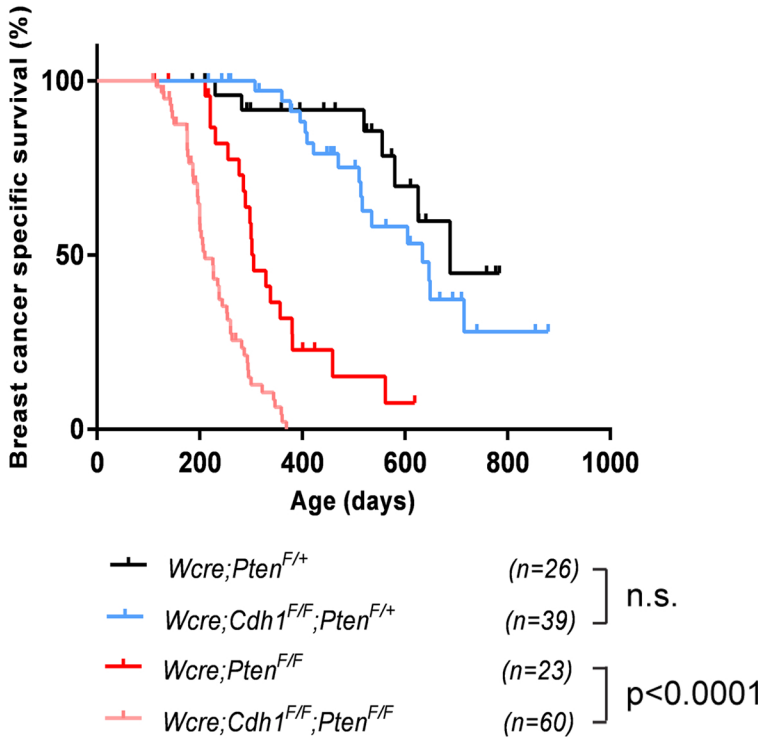

A
E-cadherin
Cytokeratin 8
merge+hoechst
Wcre; Cdh1F/F
PTEN
Cytokeratin 8
merge+hoechst
Wcre; PtenF/F


B
wild type
Wcre; Cdh1F/F;PtenF/F
zoom
zoom

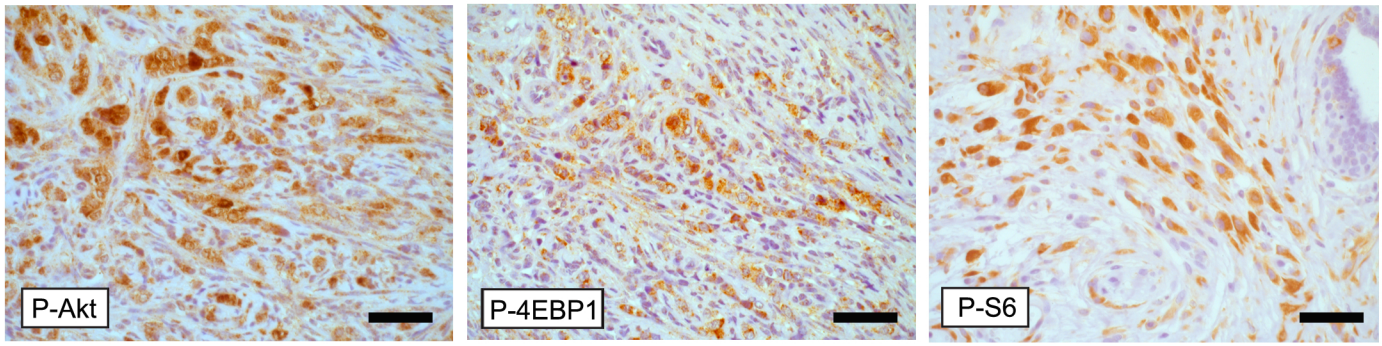

C
wild type
Wcre;Cdh1FF; PtenFF
p<0.0001
weight (g)
8
6
4
2
0
Wcre;Cdh1FF; PtenFF
wild type


D
wild type
Wcre;Cdh1FF; PtenFF
Fat
Glandular
tisue compesition (%)
100
50
0
wild type
Wcre;Cdh1FF; PtenFF

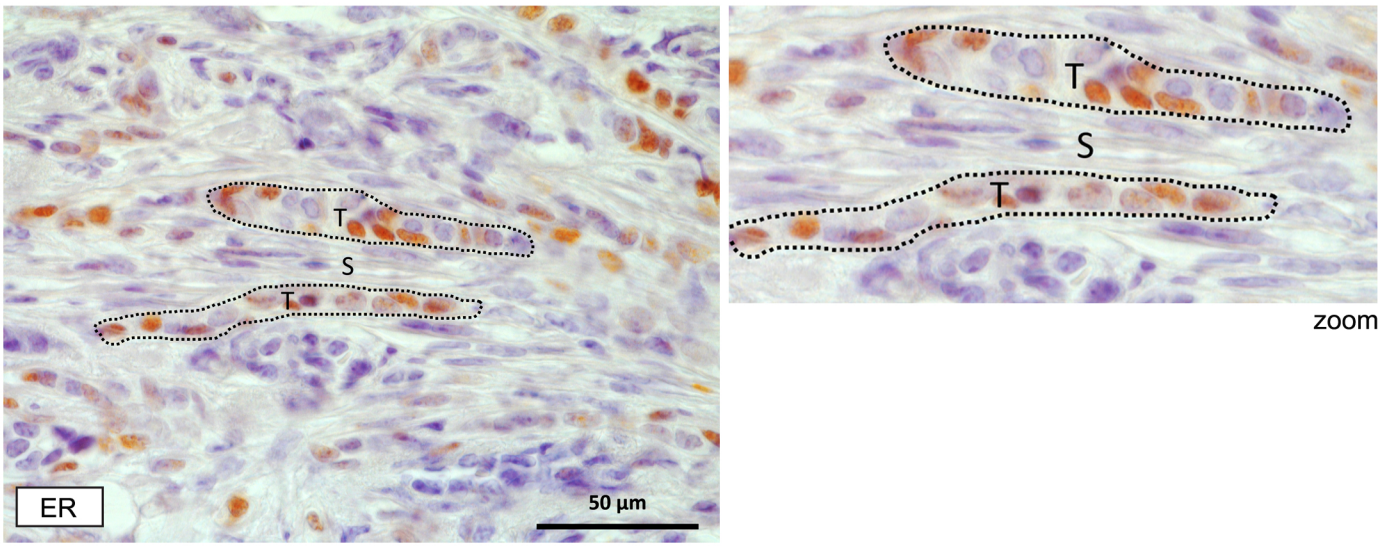
A


Breast cancer specific survival (%)
100
50
0
0
200
400
600
800
1000
Age (days)
Wcre;Pten F/+ (n=26)
Wcre;Cdh1 F/F ;Pten F/+ (n=39)
n.s.
Wcre;Pten F/F (n=23)
Wcre;Cdh1 F/F ;Pten F/F (n=60)
p<0.0001

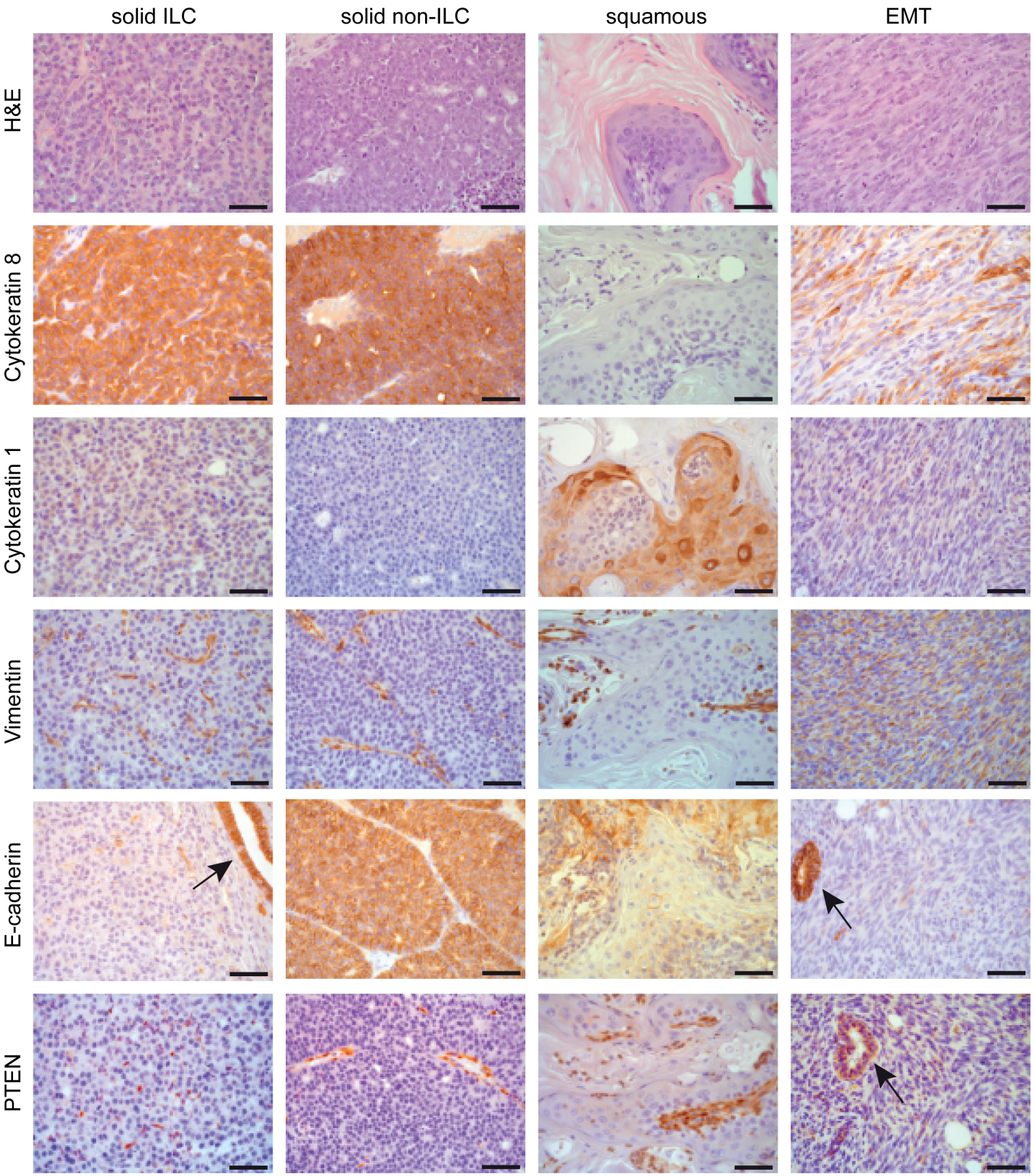
B


P-Akt
P-4EBP1
P-S6

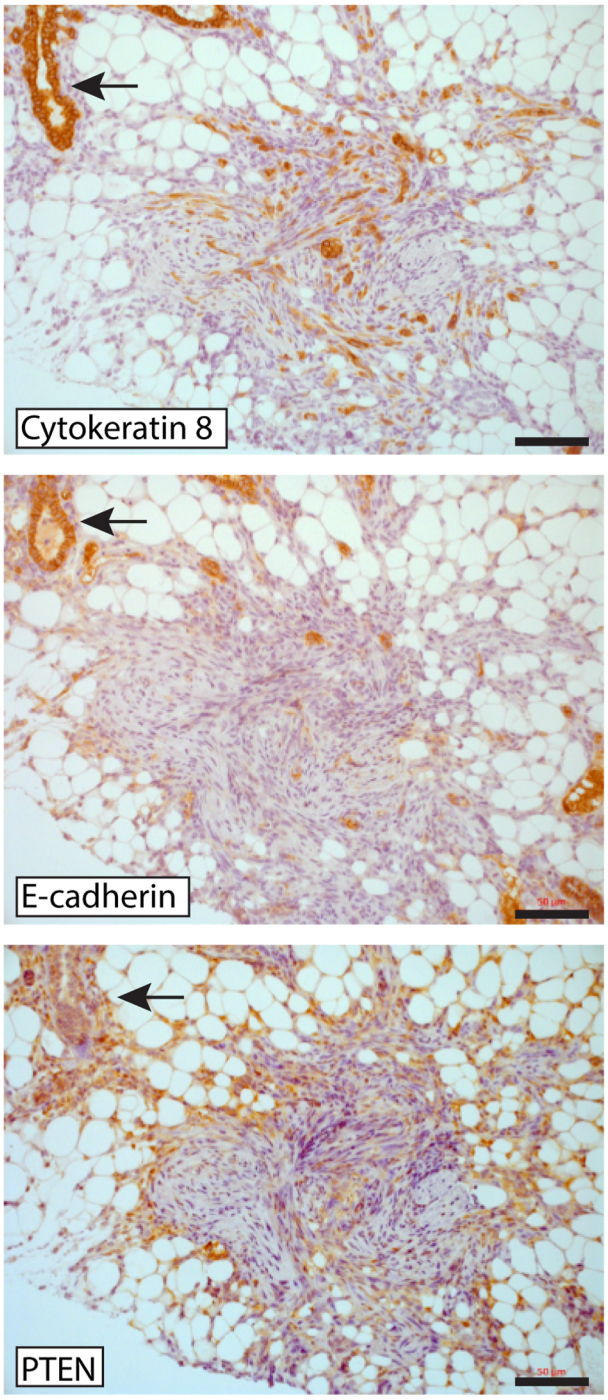
C


T
S
T
ER
50 µm
T
S
T
zoom

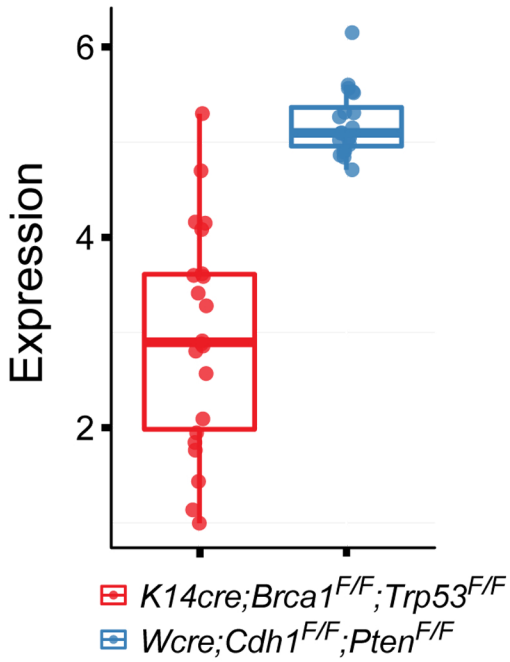
D


solid ILC
solid non-ILC
squamous
EMT
H&E
Cytokeratin 8
Cytokeratin 1
Vimentin
E-cadherin
PTEN

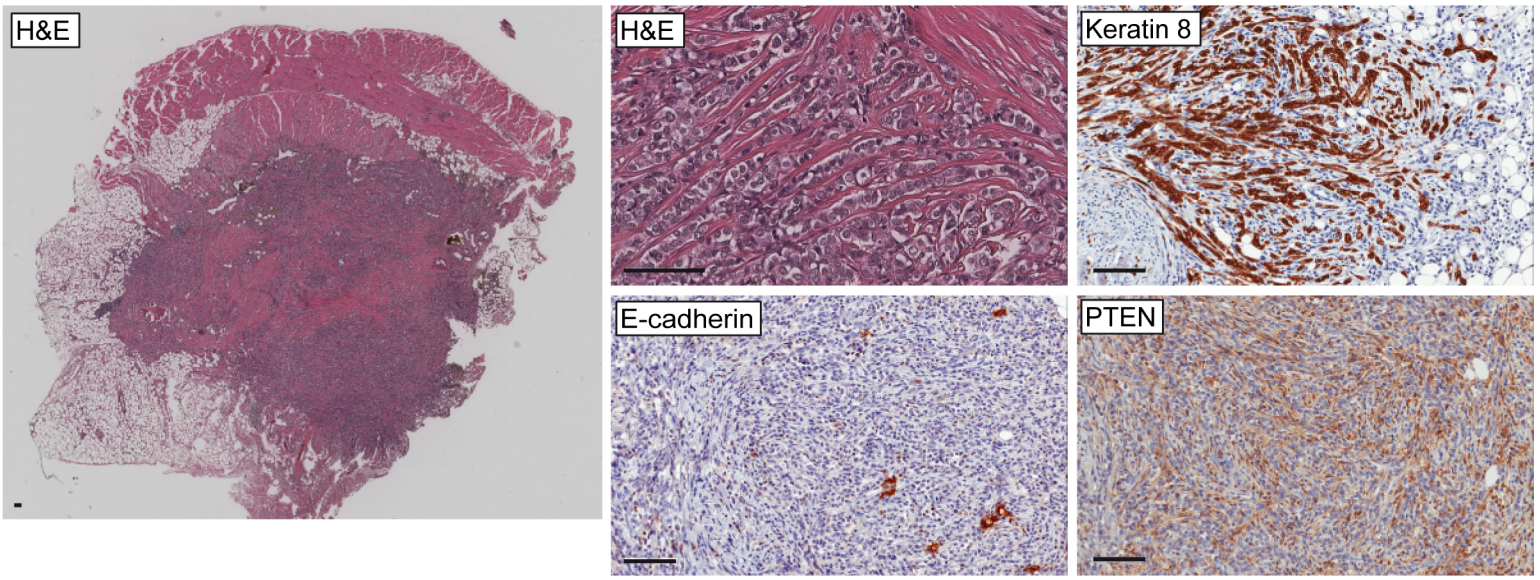
E


Cytokeratin 8
E-cadherin
PTEN

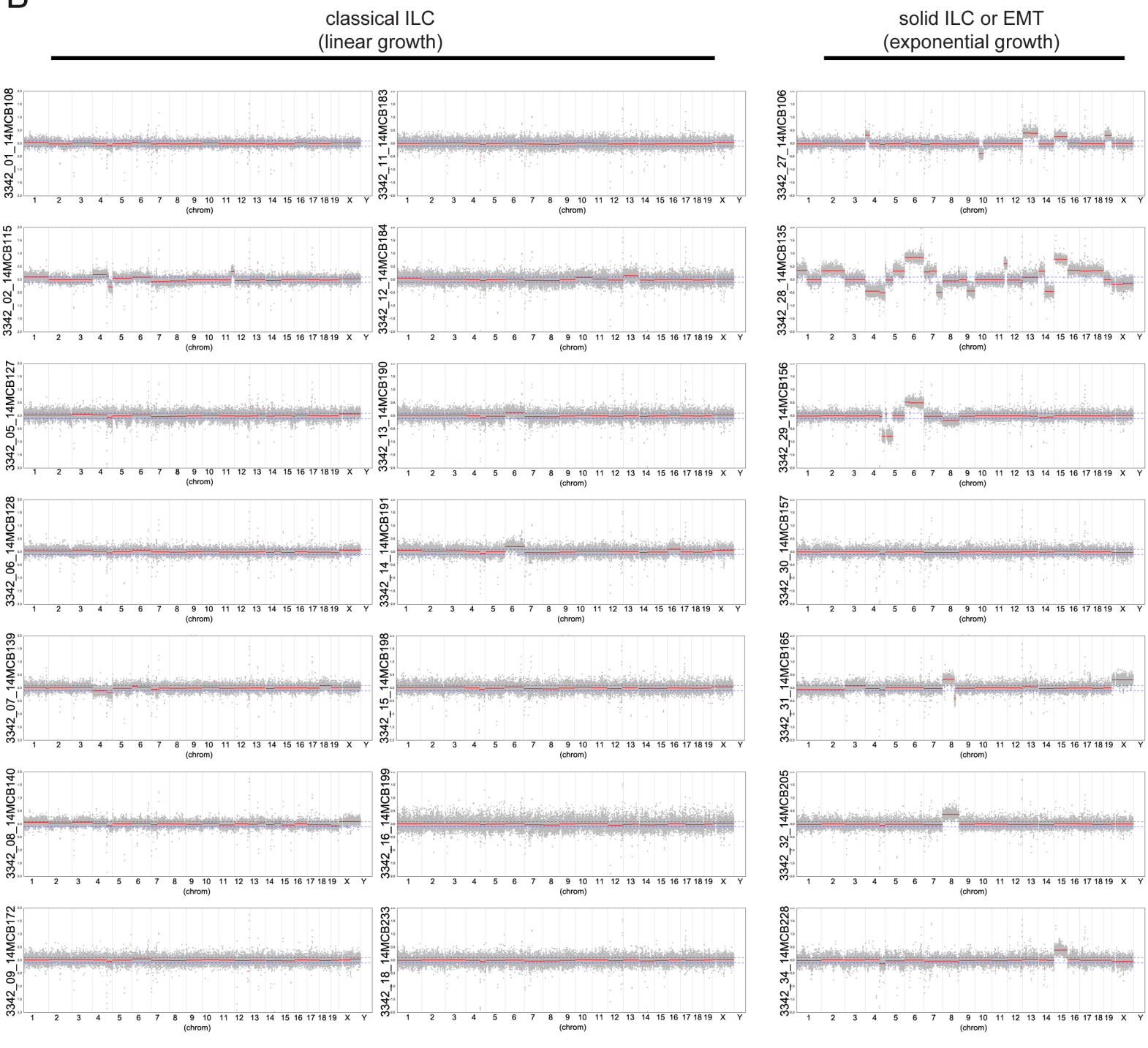
F


Expression
6
4
2
K14cre;Brca1F/F;Trp53F/F
Wcre;Cdh1F/F;PtenF/F

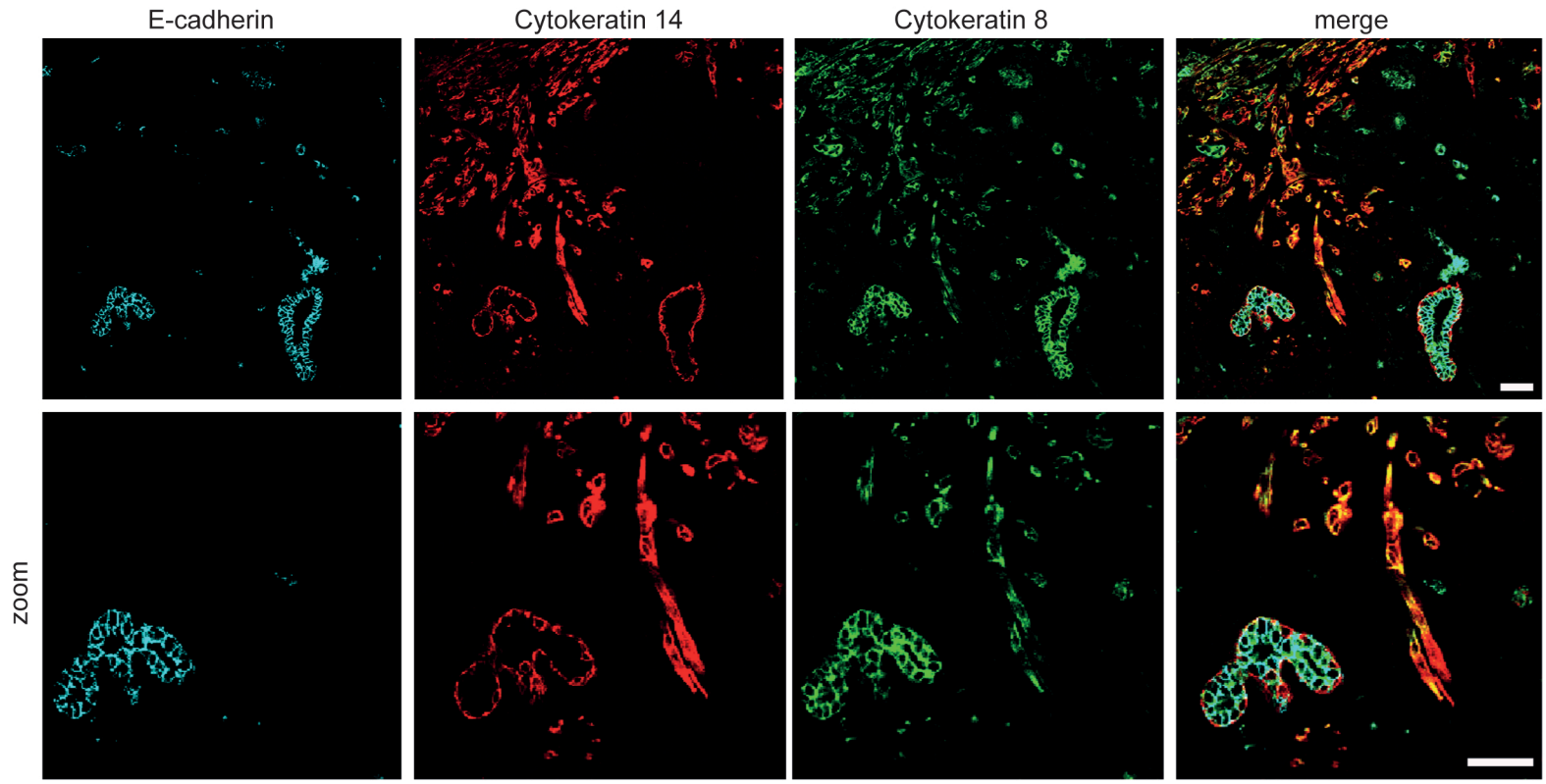
A


H&E
H&E
Keratin 8
E-cadherin
PTEN

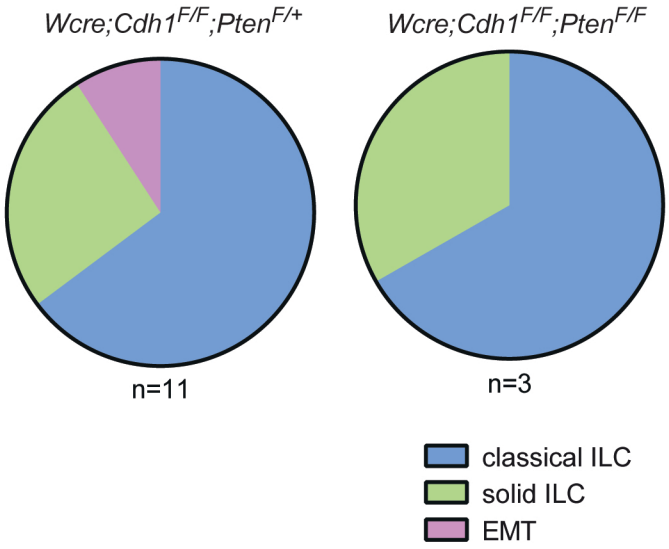
B


classical ILC
(linear growth)
solid ILC or EMT
(exponential growth)
3342_01_14MCB108
3342_11_14MCB183
3342_27_14MCB106
3342_02_14MCB115
3342_12_14MCB184
3342_28_14MCB135
3342_05_14MCB127
3342_13_14MCB190
3342_29_14MCB156
3342_06_14MCB128
3342_14_14MCB191
3342_30_14MCB157
3342_07_14MCB139
3342_15_14MCB198
3342_31_14MCB165
3342_08_14MCB140
3342_16_14MCB199
3342_32_14MCB205
3342_09_14MCB172
3342_18_14MCB233
3342_34_14MCB228
(chrom)

A



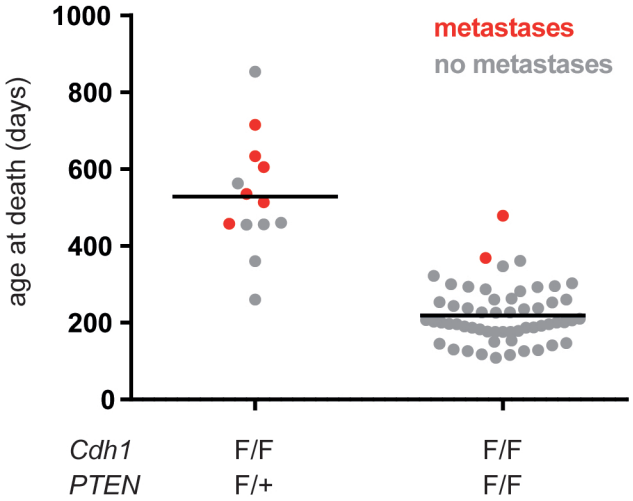
B



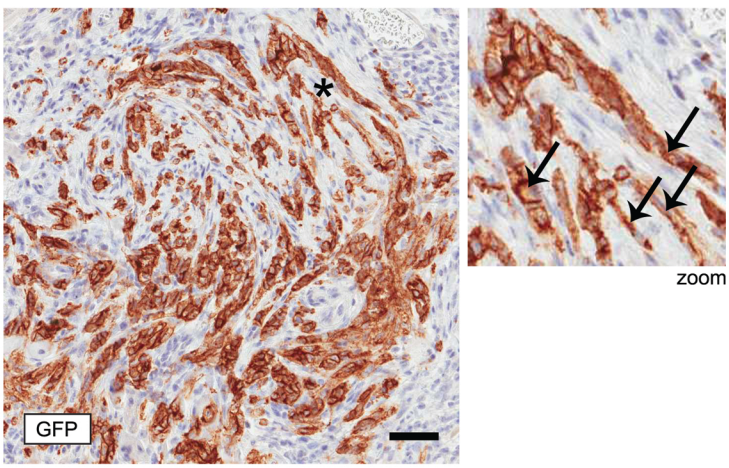
C

	median latency	median survival	Δ days
Wcre;Cdh1$^{F/F}$;Pten$^{F/+}$	109	210	101
Wcre;Cdh1$^{F/F}$;Pten$^{F/F}$	405	634	229

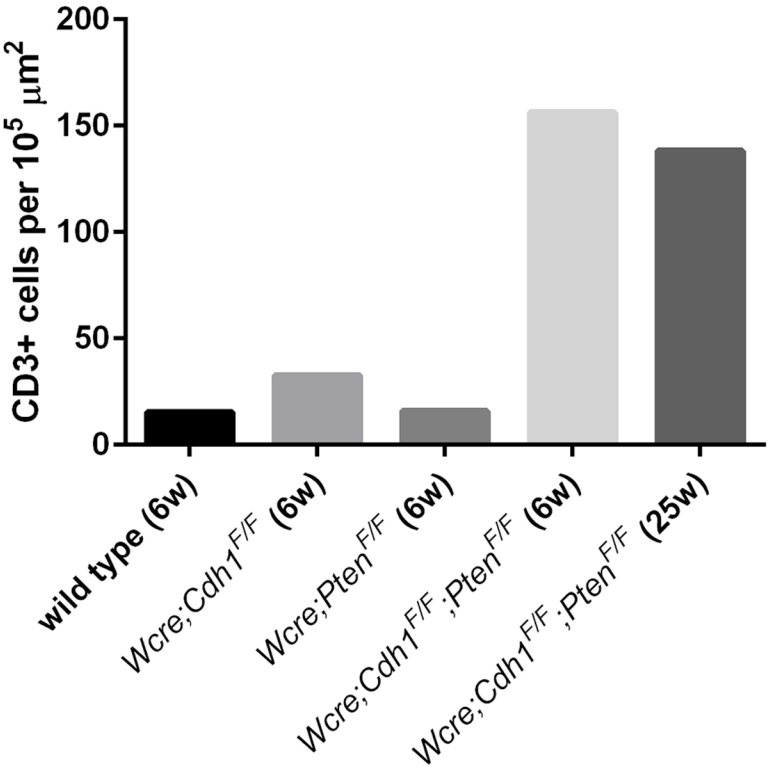
D



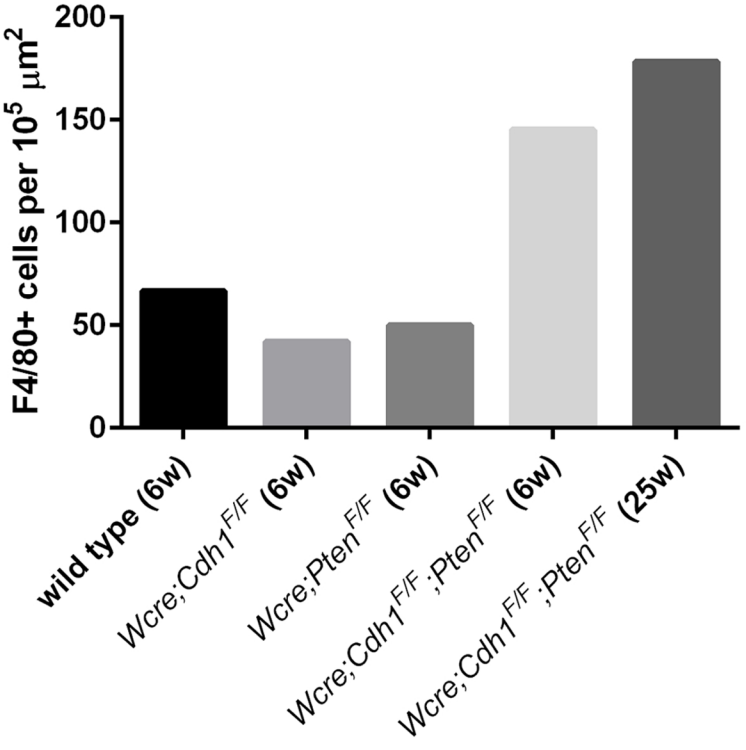
E



A


CD3+ cells per 10^5 µm^2
200
150
100
50
0
wild type (6w)
Wcre;Cdh1F/F (6w)
Wcre;PtenF/F (6w)
Wcre;Cdh1F/F;PtenF/F (6w)
Wcre;Cdh1F/F;PtenF/F (25w)

B


F4/80+ cells per 10^5 µm^2
200
150
100
50
0
wild type (6w)
Wcre;Cdh1F/F (6w)
Wcre;PtenF/F (6w)
Wcre;Cdh1F/F;PtenF/F (6w)
Wcre;Cdh1F/F;PtenF/F (25w)

Boelens_Suppl table 1, Identified oncogenic mutations by exome sequencing, Related to Figure 4

#	Tumor phenotype	Gene symbol	Cds position	Nucleotide substitution	Protein position	Amino acids	Consequence	Allele freq
1	EMT	Ccnd3	850	Ccc/Tcc	284	P/S	missense	0.333
2	Solid ILC	Kras	38	gGc/gAc	13	G/D	missense	0.588
3	Solid ILC	Asxl1	1979	ggA/gg	639	G/X	frameshift	0.436
4	EMT	Myh9	232-244	CCCAAGTTCTCCAag/ag	78-82	PFKXK/X	frameshift	0.269
		Ptpn11	188	tAt/tGt	63	Y/C	missense	0.448
5	EMT	Trp53	808	Cgt/Tgt	270	R/C	missense	0.496
6	EMT	Pwwp2a	2111	cGc/cAc	704	R/H	missense	0.297
		Ptpn11	1532	aCa/aAa	511	T/K	missense	0.302
7	EMT	Ptpn11	1393	Gct/Act	465	A/T	missense	0.222

Boelens_Suppl table 2. Antibodies used in this study

Antibody	Manufacturer	Application	Species	Antigen retrieval	Dilution
Akt	Cell Signaling #9272	WB	mouse	NA	1:1000
Cleaved Caspase 3	Cell Signaling #9661	IF	mouse	TRIS/EDTA pH9.0	1:100
CD3	Thermo Scientific RM-9107	IHC	mouse	TRIS/EDTA pH9.0	1:600
CD3	Thermo Scientific RM-9107	IHC	human	Cell Conditioning 1	1:100
CD68	DAKO M0814	IHC	human	Cell Conditioning 1	1:2000
Cytokeratin 1	Covance PRB-165P	IHC	mouse	Citrate pH6.0	1:200
Cytokeratin 8	DSHB Univ of Iowa TROMA-I	IHC	mouse	Citrate pH6.0	1:800
Cytokeratin 8	DSHB Univ of Iowa TROMA-I	IF	mouse	TRIS/EDTA pH9.0	1:200
Cytokeratin 14	Covance PRB-155P	IF	mouse	TRIS/EDTA pH9.0	1:200
E-cadherin	Cell Signaling #3195	IHC	mouse	TRIS/EDTA pH9.0	1:200
E-cadherin	BD Biosciences #610181	IF	mouse	TRIS/EDTA pH9.0	1:100
E-cadherin	BD Biosciences #610181	WB	mouse	NA	1:1000
ERα	Santa Cruz sc-542	IHC	mouse	TRIS/EDTA pH9.0	1:2000
F4/80	Serotec MCA497	IHC	mouse	Proteinase K	1:400
HSP90	Santa Cruz sc-9747	WB	mouse	NA	1:2000
Ki-67	Cell Signaling #9449	IF	mouse	Citrate pH6.0	1:100
PDGFRβ	Cell Signaling #3169	IHC	mouse	TRIS/EDTA pH9.0	1:50
PDGFRβ	Abcam ab32570	IHC	human	Cell Conditioning 1	1:100
phospho-Akt	Cell Signaling #4060	IHC	mouse	Citrate pH6.0	1:8000
phospho-Akt	Cell Signaling #4060	WB	mouse	NA	1:2000
phospho-S6	Cell Signaling #2211	IHC	mouse	Citrate pH6.0	1:400
phospho-S6	Cell Signaling #2211	WB	mouse	NA	1:1000
PTEN	Cell Signaling #9559	IHC	mouse	Citrate pH6.0	1:100
PTEN	Cell Signaling #9188	IF	mouse	TRIS/EDTA pH9.0	1:100
PTEN	Cell Signaling #9188	WB	mouse	NA	1:1000
S6	Cell Signaling #2217	WB	mouse	NA	1:1000
Vimentin	Cell Signaling #5741	IHC	mouse	TRIS/EDTA pH9.0	1:200

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Isolation of primary mammary epithelium and organoid generation

In brief, mammary glands were harvested from 8- to 12-week-old FVB mice, minced with a pair of scissors, and shaken for 20 min at 37°C in collagenase solution: DMEM (10565–018; Gibco) with 2 mg/ml Collagenase type IV (17104-019; Gibco), 2 mg/ml trypsin (27250–018; Gibco), 5% FBS (F0926; Sigma-Aldrich), 5 µg/ml insulin (I9278; Sigma-Aldrich), and 50 µg/ml gentamicin (15750; Gibco). Suspensions were centrifuged at 1500 rpm to remove a floating layer of adipocytes, and pellets were treated with 2 U/µl DNase (D4263; Sigma-Aldrich) to detach organoids from stromal cells. Enzymes and single cells were removed by four spins of 1 min at 1500rpm by which pellets were obtained that mainly contained small mammary epithelial fragments (organoids). Organoids were embedded in Matrigel (354230; BD) and plated as 100-µl suspensions in Millicell Ezslide 8-well Glass wells (PEZGS0816; Millipore). Gels were allowed to polymerize for 30 min at 37°C before supplementing organoid medium: DMEM with 1% insulin-transferrin-selenium (51500–056; Gibco) and 1% penicillin-streptomycin (P4333; Sigma-Aldrich). Images of organoids were obtained by using the Axio Observer Z1 ZEISS microscope. To obtain organoid growth, organoid size was measured weekly using the ZEISS ZEN Software. Organoid whole-mount imaging was performed by removing culture medium to wash and fix the organoids in 3.7% Paraformaldehyde at RT for 30'. Organoids were permeabilized with 1% Triton X-100, 10% glycerol in PBS for 30 min at RT. Organoids were incubated with primary antibody against; Rabbit Cleaved Caspase-3 (1:100, Cell Signaling #9661) and mouse ki67 (1:100, Cell Signaling #9449) in PBS containing 0.1% triton and 10% glycerol o/n at 4°C. Organoids were subsequently washed 3 times with PBS and stained with secondary antibodies anti-Rabbit-AlexaFluor 568 (1:500, Invitrogen #A11011) and anti-mouse-AlexaFluor 488 (1:500, Molecular Probes #A21141) and Hoechst o/n at 4°C. Organoids were subsequently washed and mounted using Vectashield (Vector Laboratories H-1000). Images were acquired using a Leica TCS SP5 Confocal and images were analyzed using LAS AF Version 2.6.3 software.

Adenoviral delivery of Cre recombinase

Before embedding in Matrigel, Adcre was added to isolated mammary organoids to drive Cre-recominase activity during subsequent organoid development in matrigel (1×10^8 Transducing Units (TU); Gene Transfer Vector Core, University of Iowa). Seven days later, DNA was isolated using direct PCR lysis buffer (301-C; Viagen) to analyze Cre-driven recombination efficiency, and protein was isolated using NP40 lysis buffer (20 mM Tris pH 7.4, 100 mM NaCl, 1% NP40, 10% glycerol, 10 mM EDTA complemented with protease inhibitors (Roche) to analyze inactivation of E-cadherin and/or PTEN.

Immunoblotting

Tissue was homogenized using a microhomogenizer and protein lysates were prepared using a RIPA lysis buffer (20mM Tris pH 8.0, 300mM NaCl, 10mM EDTA, 20% glycerol, 2% NP40) supplemented with protease and phosphatase inhibitors (Roche) and quantified using the BCA Protein Assay Kit (Pierce). Equal amounts of protein were separated by 4-12% SDS-PAGE (Invitrogen), transferred onto a nitrocellulose membrane (Biorad), blocked with 5% BSA in PBS-Tween (0.05%), and probed for primary antibodies as listed in Table S1 and secondary antibodies HRP goat-anti-rabbit (DAKO). Protein was visualized using ECL (Pierce ECL-Western blotting 32106).

Histopathology

Mouse tissues were formalin-fixed in 10% neutral buffered formalin for 48 hours, embedded in paraffin, sectioned and stained with hematoxylin and eosin (H&E) or Masson's trichrome. IHC was performed as previously described (Henneman et al., 2015). Primary antibody details and antigen retrieval methods are described in Supplemental information Table S1. Secondary antibodies that were used are HRP anti-rabbit Envision (DAKO K4011), Biotin anti-rabbit (DAKO E0432), and Biotin anti-rat (Santa Cruz sc-2041). All slides were digitally processed using the Aperio ScanScope (Aperio, Vista, CA, USA) and captured using ImageScope software version 12.0.0 (Aperio).

Immunofluorescence

Formalin-fixed and paraffin-embedded sections were processed as described (Pasic *et al.*, 2011) and incubated overnight at 4°C with primary antibodies. Antibody details and antigen retrieval methods are described in Supplemental Table S1. Secondary antibodies anti-Rat-AlexaFluor 647 (1:1000, Invitrogen #A21247), anti-Rabbit-

AlexaFluor 568 (1:1000, Invitrogen #A11011) and anti-mouse-AlexaFluor 488 (1:1000, Molecular Probes #A21141) were incubated overnight at 4°C. Sections were subsequently stained with Hoechst (1:1000, Thermo Scientific #62249) for 5 min and mounted using Vectashield (Vector Laboratories H-1000). Images were acquired using a Leica TCS SP5 Confocal and analyzed using LAS AF Version 2.6.3 software.

Whole-mount staining

Inguinal mammary glands were dissected, fixed for 4 hours in methacarn (60% methanol, 30% chloroform, 10% acetic acid), and subsequently stained overnight in 0.2% carmine alum. After dehydration in 70%, 90% and 100% ethanol series, the glands were cleared in xylene and photographed using an Olympus SZX12 microscope.

RNA sequencing

Total tumor tissues were homogenized in TRIzol reagent (15596-018, Ambion life technologies) using a polytron (DI 18 Disperser, IKA) according to the manufactures protocol. Briefly, 0.2x volumes of chloroform (Chloroform stab./Amylene, Biosolve) was added to the homogenate and the tube(s) were shaken vigorously. The tube(s) were incubated for 2-3 minutes at room temperature and centrifuged (Hettich, rotanta 46 RS) for 1 hour (4120 RCF, 4 °C) . Approximately 70% of the upper aqueous phase was transferred to a clean 15 mL tube and 0.5x volume of isopropanol (33539, Sigma-Aldrich,) was added. The tube(s) were incubated overnight at -20 °C and centrifuged for 30 minutes (4120 RCF, 4 °C). The supernatant was removed and the pellet was washed twice with 80% ethanol (32221-2.5L, Sigma-Aldrich). The total RNA pellet was air-dried and dissolved in an appropriate volume of nuclease free water (AM9937, Ambion life technologies) and quantified using Nanodrop UV-VIS Spectrophotometer (Thermo Fisher Scientific Inc.). Total RNA was purified using the MinElute Cleanup Kit (74204, Qiagen) according to the manufactures instructions. Quality and quantity of the total RNA was assessed by the 2100 Bioanalyzer using a Nano chip (Agilent, Santa Clara, CA). Total RNA samples having RIN>8 were subjected to library generation.

Illumina TruSeq mRNA libraries were generated using the TruSeq RNA Library Preparation Kit v2 sample preparation kit (Illumina Inc., San Diego, Cat.No RS-122-2001/2) according to the manufacturer's instruction (Part # 15026495 Rev. B) Briefly, polyadenylated RNA from intact total RNA was purified using oligo-dT beads. Following purification the RNA was fragmented, random primed and reverse transcribed using SuperScript II Reverse Transcriptase (Invitrogen, part # 18064-014). Second strand synthesis was accomplished by using Polymerase I and RNaseH. The generated cDNA fragments were 3' end adenylated and ligated to Illumina Paired-end sequencing adapters and subsequently amplified by 15 cycles of PCR. The libraries were analyzed on a 2100 Bioanalyzer using a 7500 chip (Agilent, Santa Clara, CA), diluted and pooled equimolar into a 10nM sequencing pool containing 8-10 samples each. The libraries were sequenced with 50 base single reads on a HiSeq2000 using V3 chemistry (Illumina Inc., San Diego). The obtained data will available for future analysis at www.xxx.org.

Generating PAM50 signature of mCLC

In order to investigate molecular similarities between mCLC from *Wcre*;*Cdh1*$^{F/F}$;*Pten*$^{F/F}$ mice and human breast cancer samples, the TCGA BRCA dataset was integrated with RNAseq data of mouse mammary tumor samples from *Wcre*;*Cdh1*$^{F/F}$;*Pten*$^{F/F}$ and *K14cre;Brca1*$^{F/F}$*;Trp53*$^{F/F}$ mice. Gene expression profiles for each sample were collected from RNAseq data normalized by TMM (Robinson and Oshlack, 2010). Batch effects were corrected using ComBat (Johnson et al., 2007). Unsupervised hierarchical clustering was performed using genes from the PAM50 signature (Parker et al., 2009).

Whole-exome sequencing

Due to high stromal contamination in mCLC, the tumor cell population was enriched through FACS-sorting based on expression of epithelial marker EpCAM. Tumor tissues of *Wcre;Cdh1*$^{F/F}$*;Pten*$^{F/F}$ mice were processed as described above. Single cells were stained with directly conjugated EpCAM-FITC (eBioscience 11-5791) for 30 min at 4 °C in the dark in PBS/1% BSA. DAPI (1:20) was added to exclude dead cells. EpCAM-positive cells were collected using a BD FACSAria IIu flow cytometer. Genomic DNA was isolated from this cell population using QIAamp DNA micro kit (Qiagen) according to manufacturer's instructions. For exponentially growing non-CLC tumor tissue, genomic DNA was isolated using proteinase K lysis and phenol-chloroform organic extraction. For whole-exome sequencing, up to 2 ug of genomic DNA samples was sheared to a length between the 200-300 bp

using the covaris S2 system. The sheared DNA samples were cleaned using 1.8x of the Agencourt Ampure XP beads (part.no A63881). The DNA library preparation was performed with the Kapa HTP DNA library preparation kit (KK8234), according to protocol. Input library prep sample of FACS-sorted CLC tumor samples was 5-150 ng dsDNA; Input library prep sample of non-CLCs was 1000 ng dsDNA. The enrichment assay was performed using the SureSelect XT2 Mouse All Exon kit (part no. 5190-4681). One adjustment to the hybridization protocol was made, we used a 1:1 diluted Mouse All Exon capture library.
The obtained data will available for future analysis at www.xxx.org

Extraction of mutations from whole-exome sequencing data

Sequence reads from the exome capture were mapped to the mouse reference genome (mm10), using BWA (Li and Durbin, 2009). The mapped sequence reads were further processed according to GATK best practices using Picard (version 1.128) and GATK (version 3.4) (Van der Auwera et al., 2013) to perform duplicate marking, indel realignment and base quality score recalibration. SNPs and indels were called from these processed alignments using HaplotypeCaller (version 3.4.0). Variants were filtered to remove calls from sequencing artefacts using GATKs recommended practice for hard-filtering of variant calls. To remove variants that are part of the mouse background, we filtered variants that were known FVB variants (Wong et al., 2012) and variants that occurred with a variant allele frequency greater than 0.1 in at least one sample in a panel of four normal (spleen) samples. The remaining variants were then annotated using VEP and filtered to select only deleterious variants (missense variants, frame shift variants, stop gained/lost variants and UTR variants). To prioritize cancer-related variants, these variants were annotated for their presence in the Cancer Gene Census (Futreal et al., 2004).

CNV analysis

We used the CopywriteR tool (Kuilman et al., 2015) to obtain copy number profiles from the exome alignments, which effectively extracts off-target reads from the exome capture and uses these to estimate copy number profiles. The log2 copy number estimates from CopywriteR were smoothed and segmented using the circular binary segmentation (CBS) algorithm as implemented in the DNACopy R package (Venkatraman and Olshen, 2007). In this segmentation, segments that were smaller than 5 probes or that deviated less than 3 standard deviations were pruned to reduce the number of spurious segments. For each sample, we then centered the (log2) copy number estimates of these segments by subtracting the segment copy number value in each sample. Using these centered values, we then determined the number of chromosomal aberrations in each sample as the number of segments with copy number values above 0.1 or below -0.1. To remove systematic artefacts from the copy number calling and/or segmentation, we did not count any segments that were present in more than 15 samples. Similarly, to further reduce the noise in the counts, we removed segments that spanned less than 10 probes, as these were considered more likely to represent noise than actual focal amplifications/deletions.

SUPPLEMENTAL FIGURE LEGENDS

Supplemental Figure S1. Analysis of Cre recombinase activity in AdCre transduced mammary organoids

(A) Fluorescent analysis of untransduced and AdCre transduced mTmG organoids to asses organoid transduction efficiency as displayed by conversion of RFP into GFP-driven fluorescence at 7 days of organoid culture. Bar=50µm. (B) Representative images of untransduced or AdCre transduced *Cdh1*$^{F/F}$, *Pten*$^{F/F}$ and *Cdh1*$^{F/F}$;*Pten*$^{F/F}$ mammary organoids at 7 days of organoid culture. Bar=50µm. (C) PCR analysis of recombined floxed *Cdh1* and *Pten* alleles (ΔPCR) in AdCre transduced *Cdh1*$^{F/F}$, *Pten*$^{F/F}$ and *Cdh1*$^{F/F}$;*Pten*$^{F/F}$ mammary organoids. (D) Representative overview of untransduced and AdCre transduced *Cdh1*$^{F/F}$, *Pten*$^{F/F}$ and *Cdh1*$^{F/F}$;*Pten*$^{F/F}$ mammary organoids at day 15 of organoid culture, one day after renewal of culture medium.

Supplemental Figure S2. Analysis of somatic inactivation of E-cadherin or PTEN in mouse mammary epithelium

(A) Immunofluorescence analyses of CK8, E-cadherin and PTEN in mammary epithelial tissue sections of 6-week-old *Wcre;Cdh1*$^{F/F}$ and *Wcre;Pten*$^{F/F}$ mice. Arrows highlight accumulation of E-cadherin deficient CK8-positive luminal cells in the duct of the mammary epithelium of *Wcre;Cdh1*$^{F/F}$ mice in contrast to PTEN-inactivated mammary luminal cells in *Wcre;Pten*$^{F/F}$ mice. Bar=25µm. (B) Comparison of mammary epithelial outgrowth by Carmine-alum whole-mount staining of mammary glands from 8-week-old wildtype (n=9) and *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/F}$ (n=9) mice. The dashed lines mark the leading edge and the arrow highlights lesion formation. (C) Size and weight of 6-day-old pups nursed by wildtype and *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/F}$ dams, showing impaired growth of pups nursed by *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/F}$ dams. (D) Comparison of mammary gland development during pregnancy by H&E staining of wildtype and *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/F}$ pregnant mice as quantified by measuring glandular and fat tissue.

Supplemental Figure S3. Somatic inactivation of E-cadherin and PTEN induces mammary tumors that closely resemble human CLC

(A) Mammary tumor specific survival plot of *Wcre;Pten*$^{F/+}$ (n=26), *Wcre;Cdh1*$^{F/F}$ (n=39), *Pten*$^{F/+}$ (n=23), *Wcre;Pten*$^{F/F}$ (n=21) and *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/F}$ (n=60) mice, revealing decreased survival of *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/F}$ compared to *Wcre;Pten*$^{F/F}$ mice (p<0001). Latencies are depicted in days of age. (B) Representative images of IHC staining for p-AKT, p-4ebp1 and p-S6 in tumor cells of mCLC in *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/F}$ mice, confirming activation of PI3K signaling. Bar=50µm. (C) Representative images of IHC staining for ERα of mammary gland sections with mCLC in *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/F}$ mice to exemplify the differences in cell morphology between neoplastic cells (encircled by a dashed line) and surrounding stromal cells that were used to score the ER+ tumor cells. Bar=50µm. (D) Representative examples of mammary tumors classified as solid ILC, solid carcinoma, squamous metaplastic carcinoma and EMT as determined by H&E and IHC staining of CK8, CK1, Vimentin, E-cadherin and PTEN. Arrows highlight E-cadherin or PTEN proficient cells. Bar=25µm. (E) Representative IHC staining for CK8, E-cadherin and PTEN in mammary tumors from 12-week-old *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/+}$ mice. The arrow highlights CK8-marked luminal cells expressing E-cadherin and PTEN in the mammary duct, in contrast to CK8-marked tumor cells in the adjacent lesion, which lost E-cadherin and PTEN protein expression. Bar=100µm. (F) *Esr1* mRNA expression levels extracted from RNA sequencing data of mammary tumors from *K14cre;Brca1*$^{F/F}$;*Trp53*$^{F/F}$ (n=22) and *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/F}$ (n=20) mice and compared to FACS-sorted Epcam-positive tumor cells from *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/F}$ mammary tumors (n=18).

Supplemental Figure S4. Characterization of mammary tumors from mice transplanted with precancerous *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/F}$ mammary gland fragments

(A) Representative images of IHC staining of CK8, E-cadherin and PTEN in tumors from mice transplanted with precancerous *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/F}$ mammary gland fragments. Bar=100µm. (B) DNA copy-number profiles of linear-growing mCLCs (n=14) and exponential-growing solid ILCs and EMT tumors (n=7) from mice transplanted with precancerous *Wcre;Cdh1*$^{F/F}$;*Pten*$^{F/F}$ mammary gland fragments.

Supplemental Figure S5. Metastatic behavior of *Wcre;Cdh1$^{F/F}$;Pten$^{F/F}$* mammary tumors

(A) IF analyses of mammary tumors by immunostaining of CK8, CK14 and E-cadherin revealing dual positive CK8/CK14 tumor cells at the invasive front of mammary tumors obtained from 12-week-old *Wcre;Cdh1$^{F/F}$;Pten$^{F/F}$* mice. Bar=50µm (B) Pie charts reflecting the primary tumor phenotype of metastasis-harboring *Wcre;Cdh1$^{F/F}$;Pten$^{F/+}$* (n=11) and *Wcre;Cdh1$^{F/F}$;Pten$^{F/F}$* (n=3) mice. (C) Analyses of median tumor latency and median survival of *Wcre;Cdh1$^{F/F}$;Pten$^{F/+}$* and *Wcre;Cdh1$^{F/F}$;Pten$^{F/F}$* mice to compare average days of time between tumor detection and sacrifice of the mice. (D) Age at death and metastatic status plotted for *Wcre;Cdh1$^{F/F}$;Pten$^{F/+}$* and *Wcre;Cdh1$^{F/F}$;Pten$^{F/F}$* mice to examine the correlation between life span and metastatic incidence. (E) IHC staining for GFP in tumor cells of mCLCs from *Wcre;Cdh1$^{F/F}$;Pten$^{F/F}$;mTmG* mice. Bar = 25µm.

Supplemental Figure S6. Quantification of stromal components in mouse CLC

(A) Quantification of IHC staining of CD3-marked immune cells in mammary gland sections of wildtype, *Wcre;Cdh1$^{F/F}$*, *Wcre;Pten$^{F/F}$* and *Wcre;Cdh1$^{F/F}$;Pten$^{F/F}$* mice, displayed as number of cells per 10^5 µm^2 analyzed mammary gland area. (B) Quantification of IHC staining of F4/80-marked macrophages in mammary gland sections of wildtype, *Wcre;Cdh1$^{F/F}$*, *Wcre;Pten$^{F/F}$* and *Wcre;Cdh1$^{F/F}$;Pten$^{F/F}$* mice displayed as number of cells per 10^5 µm^2 analyzed mammary gland area.

SUPPLEMENTAL REFERENCES

Johnson, W.E., Li, C., Rabinovic, A., 2007. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8, 118–127.

Futreal, P.A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Rahman, N., Stratton, M.R., 2004. A census of human cancer genes. Nat. Rev. Cancer 4, 177–183.

Kuilman, T., Velds, A., Kemper, K., Ranzani, M., Bombardelli, L., Hoogstraat, M., Nevedomskaya, E., Xu, G., de Ruiter, J., Lolkema, M.P., Ylstra, B., Jonkers, J., Rottenberg, S., Wessels, L.F., Adams, D.J., Peeper, D.S., Krijgsman, O., 2015. CopywriteR: DNA copy number detection from off-target sequence data. Genome Biol. 16, 49.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.

Parker, J.S., Mullins, M., Cheang, M.C.U., Leung, S., Voduc, D., Vickery, T., Davies, S., Fauron, C., He, X., Hu, Z., Quackenbush, J.F., Stijleman, I.J., Palazzo, J., Marron, J.S., Nobel, A.B., Mardis, E., Nielsen, T.O., Ellis, M.J., Perou, C.M., Bernard, P.S., 2009. Supervised risk predictor of breast cancer based on intrinsic subtypes. J. Clin. Oncol. 27, 1160–1167.

Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K.V., Altshuler, D., Gabriel, S., DePristo, M.A., 2013. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr. Protoc. Bioinformatics 11, 1–33.

Venkatraman, E.S., Olshen, A.B., 2007. A faster circular binary segmentation algorithm for the analysis of array CGH data. Bioinformatics 23, 657–663.