Supporting Information

Nguyen-Ngoc et al. 10.1073/pnas.1212834109

SI Materials and Methods

Isolation of Primary Human Mammary Tumor Organoids. Five primary human breast tumor specimens (T01–T05) were acquired from the Cooperative Human Tissue Network (CHTN), a program funded by the National Cancer Institute. Use of these anonymous samples was granted exemption status by the University of California at Berkeley Institutional Review Board according to the Code of Federal Regulations 45 CFR 46.101[b]. CHTN tissue samples were shipped overnight on wet ice. Two primary human breast tumor samples (T06–T07) were acquired from Johns Hopkins Hospital. Use of these deidentified samples was approved as "Not Human Subjects Research" by the Johns Hopkins School of Medicine Internal Review Board (NA00052607). Tissues from both sources were rinsed upon receipt three to five times with PBSA [1× Dulbecco's PBS supplemented with 200 U/mL penicillin/200 μg/mL streptomycin (15140–155; Invitrogen) and ∼5 μg/mL Fungizone (15290–018; Invitrogen)] to reduce traces of blood. Then they were minced into small fragments using a sterile razor blade and were incubated (typically in 10 mL of solution in a 15-mL Falcon tube) in collagenase [high-glucose DMEM (D6546; Sigma), 2 mM glutamine (5.1 mL), penicillin/streptomycin as above, and 2 mg/mL collagenase I (C2139; Sigma); for some isolations 2 mg/mL trypsin (27250–018; Gibco) was included in the digestion solution], with rocking at 37 °C. Successful isolation and culture were achieved with incubation times ranging from 6 h to overnight and in digestion solutions of collagenase alone or collagenase plus trypsin. Digested tumor fragments were pelleted in a centrifuge at 100 × g for 3 min, and the supernatant was discarded.

Human mammary epithelial tissue was cultured in mammary epithelial medium, which consisted of DMEM (Sigma D6546), 2 mM glutamine (ATCC or Invitrogen), 100 U/mL penicillin/100 μg/mL streptomycin, 10 mM Hepes (H3375-250g; Sigma), 0.075% (wt/vol) BSA (A8412; Sigma), 10 ng/mL cholera toxin (C8052; Sigma), 0.47 μg/mL hydrocortisone (H690; Sigma), 5 μg/mL insulin (I0516; Sigma), and 5 ng/mL EGF (13247-051; Invitrogen).

Preparation of Collagen I Gels. Collagen I gels were generally prepared by neutralizing rat-tail collagen solution (354236, BD Biosciences) with 1.0 N NaOH (S2770, Sigma) and 10× DMEM (D2429, Sigma) according to the ratio: 1:0.032:0.1 (vol/vol). Variations in the ratio were made by adding sterile water in an amount based on the starting concentration of the batch of collagen I so that the final concentration would be 3 mg/mL. Since the pH of the collagen I stock solution varied slightly between batches, the collagen I solution was adjusted as needed with NaOH to reach pH between 7.0 and 7.5 (salmon pink color by eye). The neutralized collagen I solution was then incubated on ice for 1–2 h until the opacity and viscosity increased slightly. At that point the neutralized collagen I solution was mixed with cells and plated into the desired format, as described in *Materials and Methods*.

ECM-Switching Experiments. In some experiments, epithelial organoids were cultured in one ECM environment (Matrigel or collagen I) and then were switched to the other after several days of culture. To remove epithelial organoids from Matrigel, the gel was transferred manually to a 1.7-mL Eppendorf tube and dispersed into culture medium by repeated pipetting. To remove organoids from collagen I, the gel was transferred manually to a 1.7-mL Eppendorf tube, optionally treated with collagenase solution (prepared as in refs. 1 and 2), and dispersed into culture medium by repeated pipetting. Recovered organoids were centrifuged at 500 × g for 1–2 min. Then the supernatant was discarded, and the organoids were re-embedded in either Matrigel or collagen I.

Antibody Staining. Organoids cultured in both Matrigel and collagen I were fixed with 4% (wt/vol) paraformaldehyde for 20 min, rinsed twice in PBS for 10 min, permeabilized with 0.5% (vol/vol) Triton X-100 in PBS for 20 min, and rinsed twice in PBS for 10 min. Samples then were embedded in Optimal Cutting Temperature compound (OCT) and frozen at −80 °C. OCT blocks were sectioned in 100-μm thicknesses by cryostat at −20 °C. Samples on slides were rinsed twice in PBS for 10 min, blocked in 10% (vol/vol) FBS in PBS for 1 h, incubated with primary antibodies overnight at 4 °C, and rinsed twice in PBS. Slides were incubated with secondary antibodies for 2–3 h and rinsed twice in PBS for 10 min. Slides were mounted with Fluoromount (F4680; Sigma) and sealed with coverslips. F-actin was stained with Alexa 647 phalloidin (1:100) (A22287; Invitrogen), and nuclei were stained with DAPI (1:1,000) (D3571; Invitrogen). Immunofluorescent staining for each antibody was done three independent times and imaged for at least 15 organoids per condition each time. Primary antibodies were mouse anti-laminin 1α (1:100) (MAB2549; R&D Systems), rabbit anti-laminin 332 (1:1,000) (gifts of Peter Marinkovich, Stanford University, Stanford, CA and Monique Aumailley, University of Cologne, Cologne, Germany), goat anti-collagen IV (1:80) (AB769; Millipore), rat anti–E-cadherin (1:250) (13-1900; Invitrogen), and FITC-conjugated mouse anti–smooth muscle α-actin (1:250) (F3777; Sigma).

Confocal Imaging. Confocal imaging was done on a Solamere Technology Group spinning-disk confocal microscope (described in ref. 2) with a 40× C-Apochromat objective lens (Zeiss Microimaging). Acquisition of both fixed and time-lapse images was done using a combination of μManager (3) and Piper (Stanford Photonics). Levels were adjusted across entire images in Adobe Photoshop to maximize clarity in the figures.

Quantification of Dissemination. Disseminated cells in each epithelial fragment were counted manually by following each entire time-lapse movie frame by frame. An epithelial cell was classified as having disseminated when it was observed gradually to leave and separate completely from its fragment over several continuous frames. Disseminating cells were characterized further as amoeboid, mesenchymal, or collective based on the morphology of the cells as they exited the epithelium, i.e., rounded, elongated, or multicellular, respectively.

Gene-Expression Analysis of Normal and Tumor Fragments in Parallel ECM Conditions. Normal fragments were obtained from normal mammary glands in FVB mice. Tumor fragments were isolated from advanced carcinomas from the MMTV-PyMT mouse model. Fragments were embedded in 3D Matrigel or collagen I and cultured in serum-free organoid medium supplemented with 2.5 nM FGF2 (described in refs. 1 and 2). In total, four different conditions were profiled (tumor vs. normal; collagen I vs. Matrigel); each condition was replicated at least three times with biologically independent replicates. Each array replicate corresponds to independent mice. These experiments were performed in two batches. In batch 1, tissue was taken from four mice, labeled "A"–"D," with each group except D having one sample from each mouse. In batch 2, epithelial tumor fragments derived from FVB

control mammary glands or from PyMT mammary tumor were allocated to different microenvironments and time-points. BWM4F RNA was hybridized twice: once in batch 1 and again in batch 2. After BWM4F was averaged, there were 13 different arrays in total.

Microarray Sample Preparation. Sample preparation, labeling, and array hybridizations were performed according to standard protocols from the University of California, San Francisco Shared Microarray Core Facilities (http://www.arrays.ucsf.edu) and Agilent Technologies (http://www.agilent.com). Total RNA quality was assessed using a PicoChip kit on an Agilent 2100 Bioanalyzer. RNA was amplified and labeled with Cy3-CTP using the Agilent low-RNA input fluorescent linear amplification kits following the manufacturer's protocol. Labeled cRNA was assessed using the NanoDrop ND-100 (NanoDrop Technologies), and equal amounts of Cy3-labeled target were hybridized to Agilent whole-mouse genome 4 × 44K Ink-jet arrays. Hybridizations were performed for 14 h, according to the manufacturer's protocol. Arrays were scanned using the Agilent microarray scanner, and raw signal intensities were extracted with Feature Extraction v9.1 software (Agilent).

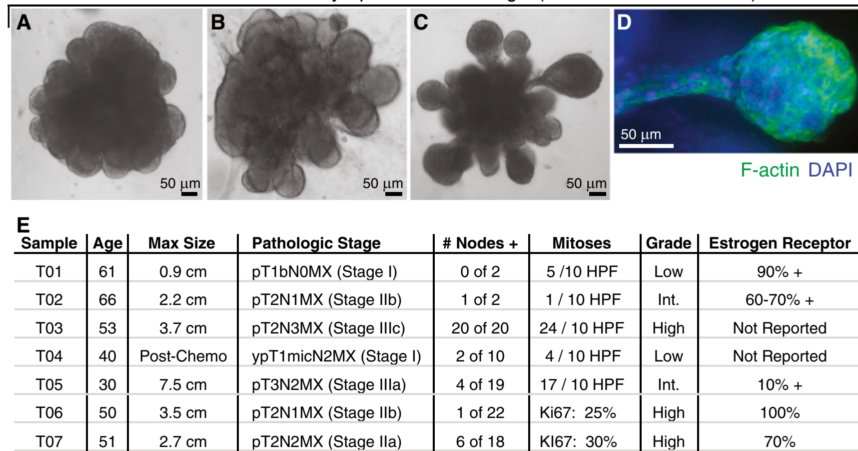
Quality Control and Normalization Analyses of Microarray Data. Analyses were conducted using R and the Bioconductor packages Agi4 × 44PreProcess, ArrayQualityMetrics, and ggplot2 (all programs available at http://www.bioconductor.org/ or http://cran.r-project.org). Plain text files generated from Agilent Feature Extraction were parsed into ExpressionSet objects. The ProcessedSignal intensities generated by Agilent Feature Extraction were used in this analysis. Array quality was assessed using box plots, hierarchical clustering, and MA plots generated by the ArrayQualityMetrics package (4). The array dataset then was normalized using quantile normalization without background subtraction. Following this normalization, as anticipated, boxplots of gMedian Intensity were all on the same scale. Batch effects were evaluated using hierarchical clustering and principle component analysis (5); arrays clearly segregate according to tumor and time, rather than by batch. Because RNA from sample BWMF4 was applied to two chips from two batches (BWMF4 and BWMF4.1), postnormalized intensities for these two samples were averaged. All other samples were biologic rather than technical replicates.

Probe to Gene Mapping. Probes were mapped to their corresponding genes based on identifiers supplied by the Agilent file GEO GPL4134-5647. RefSeq IDs, Refseq Predicted, GenBank Accession No., EmblID, Entrez Gene ID, UNIGENE_ID, Wiki_Genename, and Ensembl_transcript_ID were used to map to ENSEMBL gene IDs using Biomart, with the mouse genome sequence "ENSEMBL Genes 58, NCBIM37 Mus musculus". Altogether, probes were mapped to 19,693 genes, selecting the probe with the maximal intensity across conditions.

Clustering and Differential Gene-Expression Analyses. Hierarchical clustering and principle component analysis was done using the limma and affycoretools packages (6). Pairwise differentially expressed genes were detected using the limma package in R. Q values less than 0.05 were deemed statistically significant. A program was written in Java to generate heatmaps for publication. Positive enrichment scores such as log fold changes or modified t statistics correspond to enrichment in tumor or collagen I matrix conditions. Negative enrichment scores correspond to enrichment in normal or Matrigel conditions. Genes twofold changed or greater and with a false discovery rate (FDR) $\leq$0.05 were used as input for DAVID Gene Set Analysis (7).

Gene Family Analysis. Gene sets associated with structurally similar gene families were curated manually from Mouse Genome Informatics (http://www.informatics.jax.org/) and Interpro (http://www.ebi.ac.uk/interpro/). These gene sets include genes involved in cell–cell adhesion, cytoskeletal networks, and actin–myosin contractility. Cell-adhesion gene lists were cross-referenced further with OKCAM, an online database of cell-adhesion molecules (8). For gene family heatmaps, we constructed a linear model incorporating tissue source (normal or tumor) and microenvironment (Matrigel or collagen I) for each gene using the lmFit function in the limma package. Genes were sorted according to their enrichment with respect to normal versus tumor conditions.

1. Ewald AJ, Brenot A, Duong M, Chan BS, Werb Z (2008) Collective epithelial migration and cell rearrangements drive mammary branching morphogenesis. *Dev Cell* 14:570–581.
2. Ewald AJ (2010) *Practical Considerations for Long-Term Time-Lapse Imaging of Epithelial Morphogenesis in Three-Dimensional Organotypic Cultures. Imaging in Developmental Biology*, eds Wong R, Sharpe J (Cold Spring Harbor Laboratory, New York), 2nd Ed, pp 623–645.
3. Edelstein A, Amodaj N, Hoover K, Vale R, Stuurman N (2010) Computer control of microscopes using microManager. Curr Protoc Mol Biol Chapter 14:Unit14 20.
4. Reimers M (2010) Making informed choices about microarray data analysis. *PLOS Comput Biol* 6:e1000786.
5. Leek JT, et al. (2010) Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat Rev Genet* 11:733–739.
6. Smyth GK (2005) *Limma: Linear Models for Microarray Data* (Springer, New York), pp 397–420.
7. Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44–57.
8. Li CY, et al. (2009) OKCAM: An ontology-based, human-centered knowledgebase for cell adhesion molecules. *Nucleic Acids Res* 37(Database issue):D251–D260.



Sample	Age	Max Size	Pathologic Stage	# Nodes +	Mitoses	Grade	Estrogen Receptor
T01	61	0.9 cm	pT1bN0MX (Stage I)	0 of 2	5 /10 HPF	Low	90% +
T02	66	2.2 cm	pT2N1MX (Stage IIb)	1 of 2	1 / 10 HPF	Int.	60-70% +
T03	53	3.7 cm	pT2N3MX (Stage IIIc)	20 of 20	24 / 10 HPF	High	Not Reported
T04	40	Post-Chemo	ypT1micN2MX (Stage I)	2 of 10	4 / 10 HPF	Low	Not Reported
T05	30	7.5 cm	pT3N2MX (Stage IIIa)	4 of 19	17 / 10 HPF	Int.	10% +
T06	50	3.5 cm	pT2N1MX (Stage IIb)	1 of 22	Ki67: 25%	High	100%
T07	51	2.7 cm	pT2N2MX (Stage IIa)	6 of 18	KI67: 30%	High	70%

Fig. S1. Normal human mammary epithelium undergoes branching morphogenesis in Matrigel. (*A–C*) Representative bright-field images of human mammary branching morphogenesis in Matrigel. (*D*) F-actin and DAPI staining show the nonprotrusive front of a human mammary end bud in Matrigel. (*E*) Pathologic stage and characteristics of human tumor samples used in this study. Six of these samples (T01–03 and T05–T07) grew well in culture and exhibited strong ECM dependence in migration strategy and dissemination frequency. T04 was from a patient who previously had received chemotherapy; the residual tissue was largely intermediate ductal carcinoma in situ and fibroadenoma. T04 explants did not grow well in 3D culture. Human tissue was acquired from the Collaborative Human Tissue Network and the Johns Hopkins Hospital (see *SI Materials and Methods* for details).

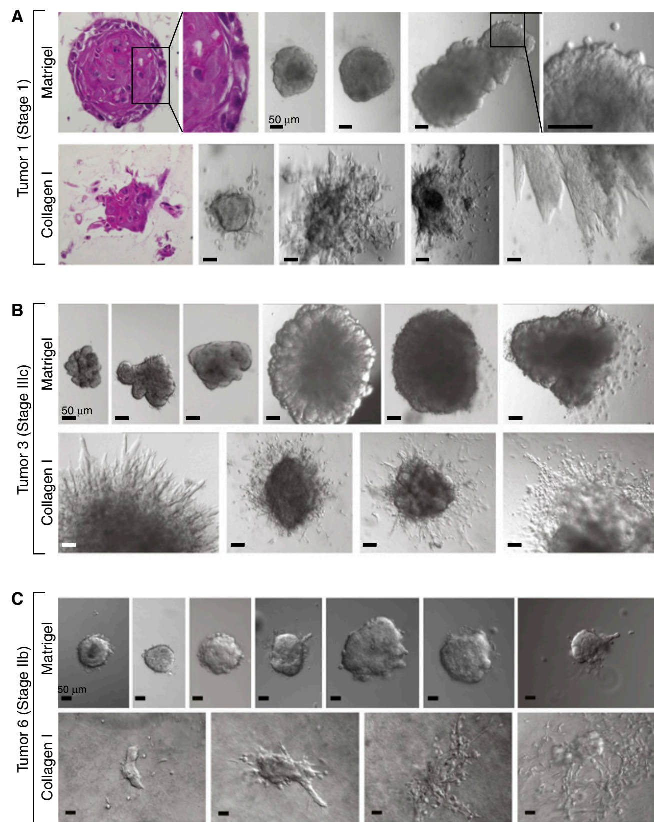


Fig. S2. Despite intra- and intertumor heterogeneity, the ECM microenvironment regulates collective migration and dissemination in human breast tumors. (*A–C*) Representative images showing the range of morphologies observed in epithelial fragments from three human tumors when cultured in Matrigel (*Upper Rows*) or collagen I (*Lower Rows*). (Scale bars, 50 µm.)

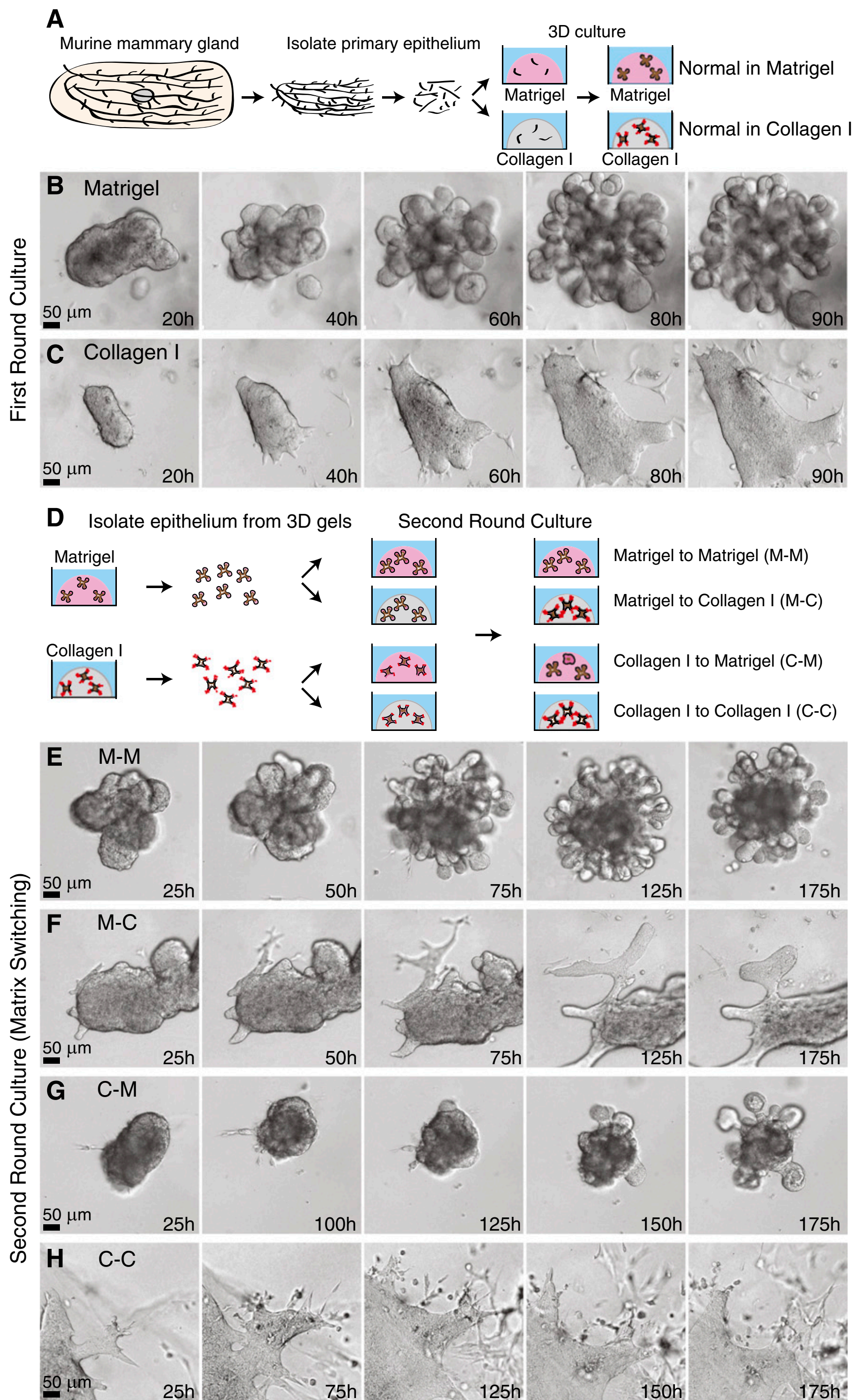
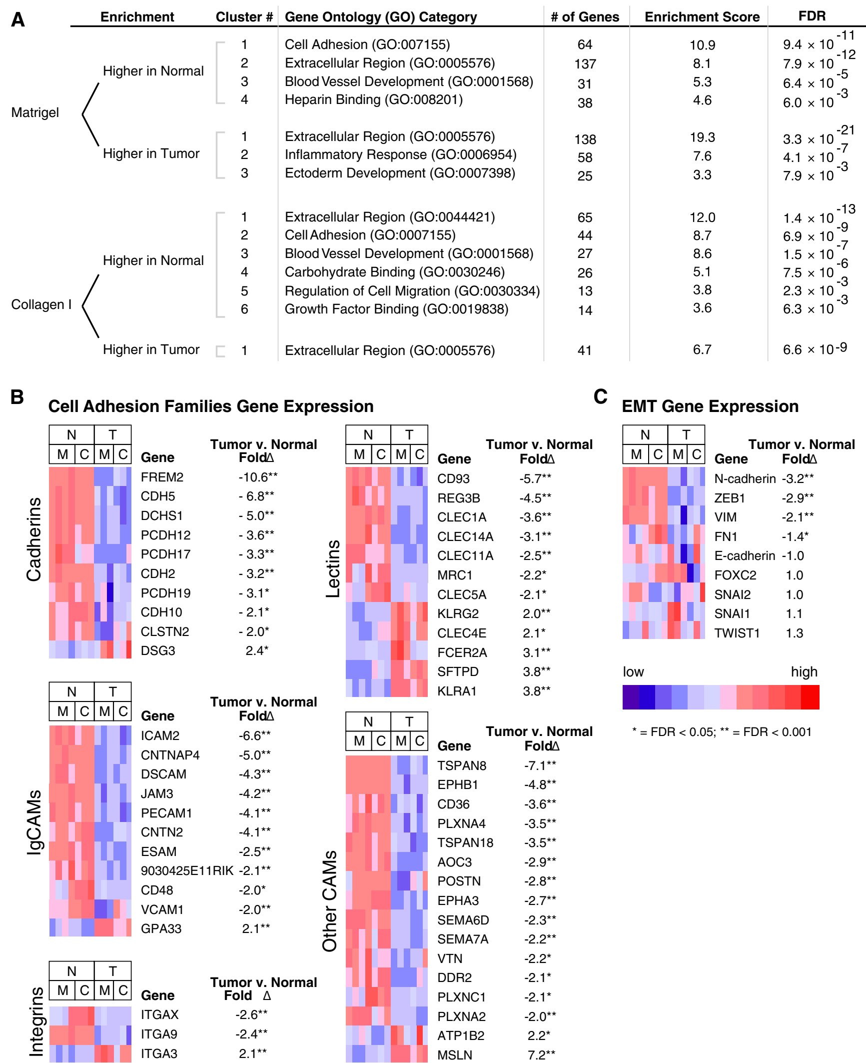


Fig. S3. The current, local ECM microenvironment determines the collective migration pattern of murine mammary epithelium. (*A*) Schematic description of the isolation and 3D culture of normal mammary organoids. (*B* and *C*) Representative bright-field time-lapse movies of normal organoids in Matrigel (*B*) and collagen I (*C*). (*D*) Schematic description of epithelial fragment isolation and matrix switching. (*E–H*) Representative frames from bright-field time-lapse movies of normal organoids switched from Matrigel to Matrigel (M–M) (*E*), Matrigel to collagen I (M–C) (*F*), collagen I to Matrigel (C–M) (*G*), and collagen I to collagen I (C–C) (*H*).



Enrichment		Cluster #	Gene Ontology (GO) Category	# of Genes	Enrichment Score	FDR
Matrigel	Higher in Normal	1	Cell Adhesion (GO:007155)	64	10.9	9.4×10^{-11}
		2	Extracellular Region (GO:0005576)	137	8.1	7.9×10^{-12}
		3	Blood Vessel Development (GO:0001568)	31	5.3	6.4×10^{-5}
		4	Heparin Binding (GO:008201)	38	4.6	6.0×10^{-3}
	Higher in Tumor	1	Extracellular Region (GO:0005576)	138	19.3	3.3×10^{-21}
		2	Inflammatory Response (GO:0006954)	58	7.6	4.1×10^{-7}
		3	Ectoderm Development (GO:0007398)	25	3.3	7.9×10^{-3}
Collagen I	Higher in Normal	1	Extracellular Region (GO:0044421)	65	12.0	1.4×10^{-13}
		2	Cell Adhesion (GO:0007155)	44	8.7	6.9×10^{-9}
		3	Blood Vessel Development (GO:0001568)	27	8.6	1.5×10^{-7}
		4	Carbohydrate Binding (GO:0030246)	26	5.1	7.5×10^{-6}
		5	Regulation of Cell Migration (GO:0030334)	13	3.8	2.3×10^{-3}
		6	Growth Factor Binding (GO:0019838)	14	3.6	6.3×10^{-3}
	Higher in Tumor	1	Extracellular Region (GO:0005576)	41	6.7	6.6×10^{-9}

Fig. S4. Cell–cell adhesion and extracellular genes are down-regulated in tumor epithelium. (*A*) Analysis using DAVID functional annotation clustering. Genes with fold changes ≥2 and FDR ≤0.05 were used as input into DAVID. The most highly enriched categories include genes whose protein products are involved in cell adhesion, are localized to the extracellular space, or are involved in the inflammatory response. (*B*) Expression of structurally related genes implicated in cell–cell and cell–matrix adhesion. C, collagen 1; M, Matrigel; N, normal tissue; T, tumor. (*C*) Expression of genes associated with epithelial-to-mesenchymal transition (EMT). EMT genes either are up-regulated in normal or are not significantly differentially expressed. For all heatmaps, $*P < 0.05$; $**P < 0.001$.

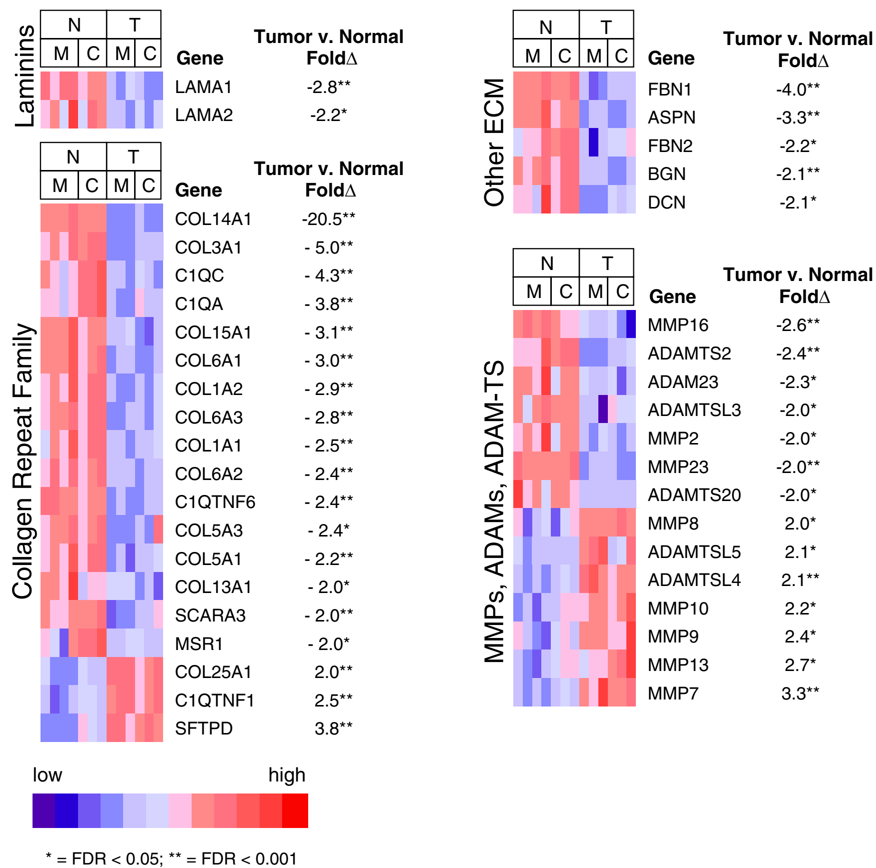


Fig. S5. ECM genes and metalloproteinases are expressed differentially by normal and tumor epithelium. The majority of genes differentially expressed in laminin, collagen, and other ECM gene sets were up-regulated in normal epithelium. Approximately equal numbers of differentially expressed metalloproteinase genes were down-regulated in normal and tumor epithelium. For all heatmaps, $^{*}P < 0.05$; $^{**}P < 0.001$.