Cancer Cell, Volume *13*

Supplemental Data

Regulation of In Situ to Invasive

Breast Carcinoma Transition

Min Hu, Jun Yao, Danielle K. Carroll, Stanislawa Weremowicz, Haiyan Chen, Daniel Carrasco, Andrea Richardson, Shelia Violette, Tatiana Nikolskaya, Yuri Nikolsky, Erica L. Bauerlein, William C. Hahn, Rebecca S. Gelman, Craig Allred, Mina J. Bissell, Stuart Schnitt, and Kornelia Polyak

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cluster and gene ontology enrichment analysis of SAGE libraries

Differentially expressed tags (*P* < 0.05) between ITGB6+ and MUC1+ SAGE libraries from MCFDCIS xenografts were generated by Poisson analysis (available at http://genome.dfci.harvard.edu/sager/) (Cai et al., 2004). Normalized tag counts per 50,000 from these two libraries were combined with those from human breast epithelial and myoepithelial libraries (N-EPI-1, N-EPI-2, N-MYOEP-1, N-MYOEP-2, N-MYOEP-4, D-EPI-2, D-EPI-3, D-EPI-6, D-EPI-7, D-MYOEP-6, and D-MYOEP-7). Tags were further filtered to have a maximum count from all libraries above 10 per 50,000 and to have a >1.5 fold difference between ITGB6 and MUC1 libraries. Filtered data were log transformed and clustered (hierarchical, complete linkage) using the Cluster and TreeView software (Eisen et al., 1998). Color settings were adjusted to set tag count 4 per 50,000 as black. Tag counts below 4 were green and above 4 were red. Gene ontology enrichment scores for the SAGE libraries were calculated as –log(p-value) by comparing the significantly highly expressed genes in each cell type (ITGB6+ and MUC1+, or Myoep and Epi) analyzed to the background (all genes in the ITGB6+/MUC1+ libraries, or for human breast libraries, all genes with tag count >=10 per 100K in at least one library) with one-sided Fisher test.

Selection of gene list for Table 1

Genes were selected based on the following criteria: (1) statistically significant (p<0.05) difference between ITGB6+ and MUC1+ libraries; (2) statistically significant (p<0.05) difference between human MYOEP and EPI groups based on t-test or Wilcox test (for genes high in ITGB6+ and MYOEP or MUC1+ and EPI), or ratio of DMYOEP/NMYOEP ≥ 10-fold (for genes high in ITGB6+ and DMYOEP); (3) ratio of ITGB6+/MUC1+ and MYOEP/EPI are in the same direction; (4) ratio of both ITGB6+/MUC1+ and MYOEP/EPI are ≥ 2-fold; and (5) tag count ≥10 per 100k in at least one of the primary human tissue libraries.

FISH analysis

For MYC FISH LSI C-MYC (Spectrum Orange), LSI MYC Dual Color Break Apart Probe (5' Spectrum Orange, 3' Spectrum Green), CEP8 (Spectrum Aqua and Spectrum Green), and CEP10 (Spectrum Aqua) probes were purchased from Vysis, Inc. (Downers Grove, IL). Cells were treated with colcemid, harvested, and used for metaphase chromosome spreads preparations according to standard protocols. Hybridization of metaphase chromosomes was performed as previously described (Ney et al., 1993). Slides were examined using a fluorescence microscope equipped with a CytoVysion capturing system (Applied Imaging Corp., San Jose, CA).

Immuno-FISH (iFISH) using formalin fixed paraffin embedded tissue

The procedure is a modification of a protocol provided by Peters et al. (Peters et al., 2005).

I. Deparaffinization and Pretreatment

1. Bake the slides in 60^{o}C oven for 2 hrs.
2. Xylene 3x 10 min.
3. (Optional) EZ-DeWax 10 min.
4. 100% EtOH 2x 3 min.
5. 70% EtOH 1x 3 min.
6. 50% EtOH 1x 3 min.
7. H_2O 2x 3 min.
8. 1x citrate buffer (dilution of 20x in H_2O) @ 70^{o}C overnight (16 - 18 hrs), cool down to RT (30 - 45 min).
9. H_2O 2x 3 min. Use PAP pen to mark the tissue area.
10. PBST 1x 3 min.

II. Immunofluorescence (skip steps 9-17 if only perform FISH)

11. PBS +10% goat serum @ RT for 30 min in a humid chamber.
12. 1^{o} antibody (CK, SMA, CD44) diluted (1/50-1/100) in PBS+5% goat serum @ RT for 1 hr.
13. PBST 3x 3 min.
14. Biotin conjugated goat-anti-mouse 2^{o} antibody diluted (1/100 of 1 mg/ml) in PBS @ RT for 30 min.
15. PBST 3x 3 min.
16. Pacific-blue conjugated streptavidin diluted (1/100 of 1 mg/ml) in PBS @ RT for 30 min in the dark. Prepare Carnoy's, chill on ice or @ -20^{o}C.
17. PBS 3x 3 min.
18. H_2O 1x 3 min.
19. Post-fix in ice-cold Carnoy's for 10 min. Prepare Denaturation solution, heat @ 75^{o}C. Prepare ice-cold 70%, 85% and 100% EtOH.
20. H_2O 1x 3 min.
21. 70% EtOH 1x 3 min.
22. 85% EtOH 1x 3 min.
23. 100% EtOH 1x 3min.
24. Air dry for 10 min.

III. FISH

25. Label DNA probes using Nick Translation

A. Label cells of different species of origin.

- Prechill microcentrifuge tubes on ice

1 μg Human or Mouse Cot-1 DNA	1 μl
dH_2O	23.5 μl
1 mM SpectrumGreen/Red dUTP	0.5 μl
0.1 mM dTTP (1:3 dilution of 0.3 mM stock)	5 μl
0.1 mM dATP, dCTP, dGTP mix (equal volume of ea)	10 μl
10x nick translation buffer	5 μl
Nick translation enzyme	5 μl

- Incubate @ 15^{o}C for 16 hrs. Can also test conditions with 5 μl enzyme/8 hrs, 10 μl enzymes/8 hrs, and 10 μl enzyme/16 hrs. Probe should be 300 bp – 3 Kb size by gel electrophoresis.
- Stop the reaction @ 70^{o}C for 10 min. Chill on ice.
- Pool samples if doing multiple batches (n). Ethanol precipitation:
  nx 50 μl sample

nx 5 µl 3M sodium acetate
4 µl glycogen
nx 125 µl 100% EtOH

- Vortex briefly. Place on dry ice for 15 min.
- Cf @ 14K, 4°C for 30 min.
- Wash with 2x 1 ml 70% EtOH.
- Resuspend pellet in nx 15 µl dH_2O + nx 35 µl CGH hybridization buffer. Store @ -20°C in the dark.
- Mix 1:1 (v:v) human and mouse probe, denature @ 75°C for 5 min (when the slides are being dehydrated in 85% EtOH below), chill on ice, spin down briefly.

B. Hybridization with BAC probe.

- Perform the Nick Translation reaction the same as above, but use 1 µg BAC DNA per 50 µl reaction.
- Stop the reaction; add 10 µg unlabeled Cot-1 DNA per 50 µl reaction, and ethanol precipitate.
- Resuspend pellet with 1.5 µl labeled CEP probe (of different color from BAC probe), 13.5 µl dH_2O and 35 µl CGH hybridization buffer.

26. Hybridize the probe

- Rinse slides in H_2O 3x 1 min.
- Denaturation solution @ 75°C for 5 min.
- Ice cold 70% EtOH 1x 3 min.
- Ice cold 85% EtOH 1x 3 min.
- Ice cold 100% EtOH 1x 3 min.
- Dry @ 42°C for 1 min.
- Apply desired amount of denatured probe mix onto tissue section, cover w/ a piece of parafilm and seal w/ coverslips with rubber seal on edges.
- Incubate @ 37°C overnight (20 - 24 hrs) in a sealed humid chamber.

27. Wash

- Remove coverslips and parrafilm. Place slides in 0.4x SSC/0.3% NP-40 wash solution @ RT for 2 min.
- 0.4x SSC/0.3% NP-40 wash solution (pre-heat for 30 min) @ 74°C for 2 min.
- 2x SSC/0.1% NP-40 wash solution @ RT for 2 min.
- H_2O 1x 3 min.
- 2x SSC 1x 3 min.
- For Cot1 DNA probe, counterstain w/ DAPI diluted (1:10,000 of 1 mg/ml) in 2x SSC @ RT for 5 min in the dark. For BAC DNA and CEP probes, use DAPI II counterstain. (If use pacific blue, skip this step and 2x SSC washes).
- 2x SSC 1x 3 min.
- H_2O 1x 3 min.
- Place slides on their side to drain and wipe off H_2O around tissue.
- Apply anti-fade mounting solution onto section and add glass cover. Remove excess mounting media and seal with clear nail polish. Store @ 4°C in the dark.

Reagents:

EZ-DeWax: Biogenex #HK585-5K
20x Citrate buffer pH 6.0: Zymed Labs #00-5000

Human Cot-1 DNA: Invitrogen # 15279-011 (predominantly 50 to 300 bp in size and enriched for repetitive DNA sequences)
Mouse Cot-1 DNA: Invitrogen # 18440-016 (predominantly 50 to 300 bp in size and enriched for repetitive DNA sequences)
SpectrumGreen-dUTP: VYSIS # 30-803200
SpectrumRed-dUTP: VYSIS # 30-803400
Nick translation kit: VYSIS # 32-801300
CEP Hybridization Buffer: VYSIS # 32-804828
DAPI: Invitrogen # D1306.
DAPI II Counterstain: VYSIS # 32-804831 (contains antifade solution)
Antifade Solution: VYSIS # 32-804029
Mouse anti-human cytokeratin: DakoCytomation # M 3515
Mouse anti-human smooth muscle actin: DakoCytomation # M 0851
CD44 Mouse Monoclonal Antibody: NeoMarkers # MS-668-P1.
Biotin conjugated Goat anti-Mouse IgG (H+L): PIERCE # 31800
Streptavidin, Pacific Blue conjugate: Invitrogen # S-11222

10x PBS
80 g NaCl
2 g KCl
11.5 g Na_2HPO_4
2 g KH_2PO_4
dH_2O to 1 L
pH 7.4

PBST (1x PBS/0.05% Tween 20)
100 ml 10x PBS
5 ml 10% Tween-20
895 ml dH_2O

Carnoy's (3:1 EtOH:acetic acid)
36 ml 100% EtOH
12 ml glacial acetic acid
Make fresh

20x SSC (3.0 M NaCl, 0.3 M sodium citrate)
175.32 g NaCl
88.23 g sodium citrate
H_2O to 1 liter
pH 7.0 unless specified

Denaturation solution (70% deionized formamide, 2x SSC)
35 ml formamide
5 ml 20x SSC pH 5.3
10 ml H_2O
Make fresh

0.4x SSC/ 0.3% NP-40 wash solution
20 ml 20x SSC
3 ml NP-40

H_2O to 1 liter
pH 7.0-7.5

2x SSC/ 0.1% NP-40 wash solution
100 ml 20x SSC
1 ml NP-40
H_2O to 1 liter
pH 7.0-7.5

Anti-fade mounting solution (or use the commercial one)
0.7 g DABCO in 2.4 ml H_2O
600 μl 1M Tris, pH 8.0
Vortex to dissolve
Add 27 ml glycerol, mix by inversion
Store in the dark @ 4^oC

Immunohistochemistry using formalin fixed paraffin embedded tissue

Tissue Preparation and Pretreatment

1. Deparaffinize and Rehydrate:
   Heat slides in 60-65^oC oven for 2 hrs (to soften wax and stick tissue onto slides).
   Xylene 3x 10 min; 100% EtOH 2x 4 min; 85% EtOH 1x 4 min; 70% EtOH 1x 4 min; 50% EtOH 1x 4 min; H_2O 2x 3 min.
2. Target Retrieval:
   - Heat induced epitope unmasking: pre-heat the following solutions, immerse room temperature tissue sections in and incubate for additional 20 min in steamer, then cool slides in running water to RT.

   (a) 1x antigen retrieval citra plus solution.
   (b) 1x target retrieval solution pH 9 (10 mM Tris Base, 1 mM EDTA, pH 9.0).
   (c) 1 mM EDTA pH 8.0.
   - Proteolytic digestion:

   (d) Pepsin for 5 min at 37^oC.
   (e) Proteinase K 15 min at 37^oC.
3. 3x 3min H_2O; (Optional) 3% H_2O_2 for10 min to block endogenous peroxidase activity (most commonly encountered in red blood cells, kidney, and liver tissue), 2x 3min H_2O.
4. 1x 3 min in PBS or TBS, use PAP pen to mark the tissue area. Dip slides in PBST or TBST (so later applied solution will easily spread over the section; could reuse).

Immunohistochemistry

1. Block with 10% normal serum (heat inactivated and filtered) from the same host species as the secondary antibodies in PBST or TBST at RT for 30 min to reduce background from non-specific, conserved-sequence, and/or Fc-receptor bindings.
2. (Optional) Tap off blocking buffer, wash 1x 3 min in PBST or TBST to remove any free biotin from the serum; block endogenous biotin with 2-3 drops of Avidin solution for 15 min, wash 1x 3 min in PBS or TBS, dip slides in PBST or TBST; then block remaining biotin-binding sites on the avidin with Biotin solution for 15 min. Tap off. Dip slides in PBST or TBST (could reuse). It is recommended to block endogenous biotin, biotin receptors, or avidin binding sites present in certain tissues rich in biotin, such as liver, kidney, and GI tract.

3. Apply primary antibody (diluted in 5% normal serum/PBST or TBST). Incubate at RT 1 hr or 4^oC overnight in humid box. Most antibodies work fine with 4^oC overnight incubation. To assay background, use isotype matched control IgG from the same species on spare slides.
4. Wash 4x 3 min in PBST or TBST with agitation.
5. Incubate with biotinylated secondary antibody (1:400-1:200 in PBST or TBST) at RT in a humid box for 30 min. It is said that using TBS to dilute 2^o Ab often produces weaker staining. So use TBS only for Abs with high background staining. Do not use BSA or other serum containing reagents to dilute secondary antibodies since they may bind to BSA or serum therefore reducing antibody affinity.
6. During 2^o incubation time, make A+B mix by adding 20 µl A to 1 ml PBS, vortex, then add 20 µl B, vortex, sit 30 min at RT before applying to slides.
7. Wash slides 4x 3 min in PBST or TBST.
8. Incubate in A+B 30 min at RT.
9. Wash 3x 3 min in PBS.
10. Prepare DAB by adding one DAB tablet and one urea-peroxide tablet in 5 ml H_2O, vortex to dissolve. Then add 25 µl 8% $NiCl_2$, vortex and filter (0.45 μm).
11. Incubate with DAB substrate for 5-30 min at RT. As soon as color developed, immerse slides in H_2O.
12. Counterstain in 0.5% methyl green for 5-30 min at RT. 60^oC may produce stronger stain.
13. Wash in 100% ETOH.
14. Rinse in Xylene.
15. Mount with Cytoseal. Apply coverslips.

Buffer recipes:

10x PBS: 80 g NaCl, 2 g KCl, 11.5 g Na_2HPO_4 and 2 g KH_2PO_4 to 1 L dH_2O, pH 7.4.
PBST (1x PBS/0.05% Tween 20): 100 ml 10x PBS, 5 ml 10% Tween-20 to 895 ml dH_2O.
20x TBS: 60.57 g Trizma® base ($C_4H_{11}NO_3$) and 175.32 g NaCl to 1 L dH_2O, pH 7.6.
TBST (1x TBS/0.05% Tween 20): 50 ml 20x TBS, 5 ml 10% Tween-20 to 945 ml dH_2O.
1 mM EDTA: 0.372 g $Na_2EDTA \bullet 2H_2O$ to 1 L dH_2O, pH 8.0.
Pepsin (DakoCytomation #S3002): 2 g to 500 ml 0.01 M HCl (pH 2), aliquot and store at -20^oC. Thaw prior to use.
Proteinase K (Puregene #D-50K5): 20 μg/ml (1:1000) in TE Buffer (50mM Tris Base, 1mM EDTA, pH 8). Optimum pH 7.5-10.
0.5% (w/v) Methyl Green (Sigma # M-8884): 5 g to 1L dH_2O or 0.1 M sodium acetate pH 4.2.

Reagents:

Citrate buffer pH 6.0, 20x: Zymed Labs # 00-5000 (58.82 g $Na_3Citrate \bullet 2H_2O$/L, pH 6.0).
Antigen Retrieval Citra Plus, 10x: BioGenex # HK080-9K.
Target Retrieval Solution, 10x: DakoCytomation # S1699 (19.21 g Citric acid anhydrous, 7.44 g $Na_2EDTA \bullet 2H_2O$, 50 ml 10% Tween-20/L, pH 6.1).
Target Retrieval Solution, pH 9, 10x: DakoCytomation # S2367 (12.1 g Tris Base, 3.72 g $Na_2EDTA \bullet 2H_2O$/L, pH 9.0).
Target Retrieval Solution, High pH, 10x: DakoCytomation # S3307 (12.1 g Tris Base/L, pH 9.9). (very often damage tissue)
30% H_2O_2 (10x): Sigma # H1009-500ML.
Avidin/Biotin Blocking Kit: Vector Laboratories # SP-2001.
VECTASTAIN Elite ABC Kit (Standard): Vector Laboratories # PK-6100.
DAB Peroxidase Substrate: Sigma # D-4293.
Cytoseal 60 Mounting Medium: Richard-Allan Scientific # 8310-4.
PAP pen: The Binding Site, INC #AD100.1S.

Dual immunohistochemical staining

Double staining using the indicated antibody combinations was performed using EnVisionR Doublestain System (DakoCytomation) following the manufacturer's instructions.

Enrichment analysis of gene lists in functional categories

The lists of differentially expressed genes were analyzed for relative enrichment with certain categories from several functional ontologies in MetaCore, including GO and GeneGo cellular processes, metabolic pathways, diseases, or canonical pathways maps. The results were presented as histograms ranked by a -log(pValue). The pValues were calculated using the same basic formula for a hypergeometric distribution where the p-Value essentially represents the probability of particular mapping arising by chance, given the numbers of genes in the set of all genes on maps/networks/processes, genes on a particular map/network/process and genes in your list.

Scoring and prioritization of ontology categories according to the relevance of input data

In most cases, HT experimental datasets are very large and may include thousands of genes. In such cases, the issue of prioritizing pathway maps, networks and modules is important and can be based on different parameters, but follows the same procedure described below. The data set is divided into two random overlapping subsets in general where the size of the intersection between them represents a random variable with a hypergeometric distribution. We applied this for numerical scoring and prioritization of the node-centered networks. Let us consider a general data set size of *N* with *R* marked objects/events (for example, the nodes with expression data). The probability of a random sub-set size of *n* to include *r* marked events/objects is described by the distribution.

The mean of this distribution is equal to: $\mu = \sum_{r=0}^{n} r \cdot P(r,n,R,N) = \frac{n \cdot R}{N} = n \cdot q$,

Where $q = R/N$ defines the ratio of marked objects.

The dispersion of this distribution is described as:

$$\sigma^2 = \sum_{r=0}^{n} r^2 \cdot P(r,n,R,N) - \mu^2 = \frac{n \cdot R \cdot (N-n) \cdot (N-R)}{N^2 \cdot (N-1)} = n \cdot q \cdot (1-q) \cdot \left(1 - \frac{n-1}{N-1}\right).$$

It is essential that these equations are invariant in terms of exchange of *n* for *R* which means that the "subset" and "marked" are equivalent and symmetrical sets. Importantly, in the cases of $r > n$, $r > R$ or $r < R + n - N$, $P(r,n,R,N) = 0$

P-value and evaluation of statistical significance of pathway maps, GO folders and disease biomarkers folders

For a network of a certain size we can evaluate its statistical significance based on the probability of its assembly from a random set of nodes the same size as the input list. We can also evaluate the network's relevance to biological processes or any other subset of nodes. Let us consider a complete set of nodes in the network, divided into two overlapping subsets. These subsets represent the nodes linked to a certain pre-defined node list. In the general case, these subsets are different but overlapping and we assume that the intersection is large enough and non-random. We do not consider a situation where the intersection is small but non-random. The null-hypothesis states that the subsets are independent and, therefore, the size of the intersection satisfies a hypergeometric distribution. The alternative hypothesis states that there is positive correlation between the subsets. Based on these assumptions, we can calculate a p-value as the probability of intersection of two random sub-sets from the same set.

$$pVal(r,n,R,N)=\sum_{i=\max(r,R+n-N)}^{\min(n,R)}P(i,n,R,N)=\frac{R!\cdot n!\cdot(N-R)!\cdot(N-n)!}{N!}\sum_{i=\max(r,R+n-N)}^{\min(n,R)}\frac{1}{i!\cdot(R-i)!\cdot(n-i)!\cdot(N-R-n+i)!}$$

Evaluation of network topology

Degree of nodes: The number of links (interactions) connected to a node (protein) gives the node's degree. Since our network is directed, the nodes are characterized by in and out-degree, giving the number of outgoing and incoming interactions. *Average shortest path* The shortest distance between two nodes is the number of links (interactions) along the shortest path(s). The average shortest path is the average over the shortest paths for all node pairs in the network. When we calculate the shortest paths for a subset of nodes (the set of proteins for colon and breast cancer) in the global network we also consider paths crossing through nodes which are not part of the subset.

Average clustering coefficient. The clustering coefficient captures to what degree node's neighbors are connected. It is defined as: $C_i=\frac{2n_i}{k_i(k_i-1)}$, where n_i is the number of links among the k_i neighbors of node *i.* As $k_i(k_i-1)/2$ is the maximum number of such links, the clustering coefficient is a number between 0 and 1. The average clustering coefficient is obtained by averaging over the clustering coefficient of individual nodes. A network with high clustering coefficient is characterized by highly connected sub-graphs.

Evaluation of significance (p-value) for topological properties

We extracted the protein-protein interactions from the signaling interaction content of Metacore. We calculated the topological properties (average degree, clustering coefficient and shortest paths) of the experimentally analyzed genes and compared them with the properties of the subset of interest (such as differentially expressed genes). In order to assign statistical significance to the differences observed we generated 10,000 times sets of randomly picked genes, of the same size as the size of the subset of interest, from the experimentally analyzed genes (as these were the genes which *can* become part of the subset of interest). For example, if we had a subset of 10 genes we would calculate the average degree of these 10 genes and generate 10,000 times sets of genes (of size 10) by randomly picking genes from the experimentally analyzed set and count how many times our set of interest gives larger degree than the randomly generated sets. If our set of interest has a larger average degree than 9,500 of the random sets (and respectively smaller average degree than 500 of the random sets) we assign a p-value of 0.05 (i.e. 500/10,000), that is, our set has significantly large average degree at p=0.05 significance level.

Evaluation of significantly over (under)-connected proteins

We calculated the interactions of proteins within a set of interest (such as differentially expressed genes in cells and tissues) and compared that with the number of connection in the global protein "interactome". The goal of the analysis was to identify proteins with a significantly large number of interactions within the set of interest. We assigned statistical significance by using the cumulative hypergeometric distribution as follows: $p(k)=\sum_{i=k}^{D}P(i,D,n,N),$ where

$$P(k,D,n,N)=\frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}}.$$

N - the number of proteins (protein-based network objects) in our global “interactome” extracted from Metacore

n - number of proteins derived from the sets of genes of interest

D - the degree of a given protein in the global “interactome” database

k - the degree of a given protein within the set of interest

The p-value calculated above gives the probability of observing *k* or more interactions of a given protein (with degree D in global network) by random chance within the set of interest (of size (*n*)). The probability of observing “under-connected” connected proteins can be calculated by *1-p(k)*. The input lists of genes were converted to protein-based network objects which have been used in our analysis. The resulting network objects sets were divided by a subsets based on the molecular function (receptors, ligands, etc.).

Table S1. List of markers used. Gene symbol, description, cell type-specific expression, and references are listed.

Gene symbol	Description	Cell type-specific expression pattern	Reference
CD24	CD24 antigen	Luminal epithelial cells	Sleeman, K. E. et al. (2006). Breast Cancer Res 8, R7.
KRT18	Cytokeratin 18	Luminal epithelial cells	Moll, R. (1998). Subcell Biochem 31, 205-262.
EPCAM	Epithelial Specific Antigen	Epithelial cells	Gudjonsson, T. et al. (2002). Genes Dev 16, 693-706.
MUC1	mucin 1	Luminal epithelial cells	Stingl, J. et al. (2005). J Mammary Gland Biol Neoplasia 10, 49-59.
CDH1	E-cadherin	Luminal epithelial cells	Palacios, J. et al. (1995). Am J Pathol 146, 605-612.
CD44	CD44 antigen	Mammary epithelial progenitors	Dontu, G. et al. (2005). Stem Cell Rev 1, 207-213.
KRT19	Cytokeratin 19	Mammary epithelial progenitors	Petersen, O. W. et al. (2003). Cell Prolif 36 Suppl 1, 33-44.
ITGA6	integrin α6	Mammary epithelial progenitors	Stingl, J. et al. (2005). J Mammary Gland Biol Neoplasia 10, 49-59.
VIM	vimentin	Basal/myoepithelial cells	Azumi, N., and Battifora, H. (1987). Am J Clin Pathol 88, 286-296.
TP63	p63	Basal/myoepithelial cells	Barbareschi, M. et al. (2001). Am J Surg Pathol 25, 1054-1060.
KRT5	Cytokeratin 5	Basal/myoepithelial cells	Moll, R. et al. (1982). Cell 31, 11-24.
KRT14	Cytokeratin 14	Basal/myoepithelial cells	Takai, Y. et al. (1995). Oral Surg Oral Med Oral Pathol Oral Radiol Endod 79, 330-341.
KRT17	Cytokeratin 17	Basal/myoepithelial cells	Troyanovsky, S. M. et al. (1989). J Cell Sci 93 (Pt 3), 419-426.
ACTA2	α-smooth muscle actin	Myoepithelial cells	Savera, A. T. et al. (1997). Mod Pathol 10, 1093-1100.
MME	CD10	Basal/myoepithelial cells	Moritani, S. et al. (2002). Mod Pathol 15, 397-405.
ITGB6	integrin β6	Basal/myoepithelial cells	this current study
LAMA3, LAMB3, LAMC2	Laminin 5 (α3β3γ2)	Basement membrane protein	Zapatka, M. et al. (2007). Oncogene 26, 1417-1427.
MKI67	Ki67	Marker of proliferating cells	Vielh, P. et al. (1990). Am J Clin Pathol 94, 681-686.

Table S2. List of antibodies used. Antigen, antibody clone name, source, catalogue number (cat#), species, application, and antigen retrieval method (AR, see above detailed Immunohistochemistry protocol) are listed. Abbreviations: FACS-Fluorescence-activated cell sorting, IB-Immunoblot, ICC-Immunocytochemistry, IF-Immunofluorescence, IHC-Immunohistochemistry, iFISH-Immunofluorescence combined with fluorescence in situ hybridization (FISH). Secondary antibodies were purchased from PIERCE (Rockford, IL) or Jackson ImmunoResearch Laboratories (West Grove, PA).

Antigen	Clone name	Source	Cat #	Species	Application	AR
β-actin	AC-74	Sigma	A2228	mouse	IB	
CD10	56C6	Ventana Medical Systems, Inc	790-2923	mouse	IHC	
CD10-RPE	SS2/36	DAKO	R0848	mouse	FACS	
CD24	ML5	BD Biosciences	555426	mouse	IHC	a
CD24-FITC	ML5	BD Biosciences	555427	mouse	FACS	
CD44	515	BD Biosciences	550988	mouse	IHC	a
CD44	156-3C11	NeoMarkers	MS-668-P1	mouse	IHC	b
CD44-PE	G44-26	BD Biosciences	555479	mouse	FACS	
CD49f-PE (integrin α6)	GoH3	BD Biosciences	555736	rat	FACS	
Cox2	CX229	Cayman Chemical	160112	mouse	IHC	b
Cytokeratin 14	LL002	Novocastra Laboratories	NCL-L-LL002	mouse	IHC, ICC, IB	a
Cytokeratin 17	E3	DAKO	M7046	mouse	IHC	b
Cytokeratin 18	C51	Novocastra Laboratories	NCL-C51	mouse	IHC	b
Cytokeratin 19	RCK108	DAKO	M0888	mouse	IHC	b
Cytokeratin 5/6	D5/16 B4	DAKO	M7237	mouse	IHC	b
E-Cadherin (CDH1)	36	BD Biosciences	610181	mouse	IHC	b
ESA	B302 (323/A3)	Biomeda	V7018	mouse	IHC	d
ESA-FITC	B29.1 (VU-ID9)	Biomeda	FM010	mouse	FACS	
Gli2	H-300	Santa Cruz Biotechnology	sc-28674	rabbit	IB	
Integrin β6	6.2A1	Dr. Shelia Violette (Biogen Idec, Inc., Cambridge, MA)		mouse	IHC, IB	d
Integrin β6	6.3G9	Dr. Shelia Violette (Biogen Idec, Inc., Cambridge, MA)		mouse	FACS, purification	
Integrin β6	ch2A1	Dr. Shelia Violette (Biogen Idec, Inc., Cambridge, MA)		human	IHC	d
Ki-67 antigen	MIB1	DAKO	M7240	mouse	IHC	a
Ki-67 antigen		Vector Laboratories	VP-K451	rabbit	IHC	a
Laminin 5	9LN5	Dr. William Brunken (Tufts University, Boston, MA)		rabbit	IHC, IB	b
MMP14		Chemicon	AB815	rabbit	IB	
Muc1	CT2	Dr. Donald Kufe (Dana Farber Cancer Institute, Boston, MA)		hamster	IB	
Muc1	DF3	Dr. Donald Kufe (Dana Farber Cancer Institute, Boston, MA)		mouse	IHC, FACS, purification	a
p63	4A4	Calbiochem	OP132	mouse	IHC, IB	a
p63	4A4	Chemicon	MAB4135	mouse	IHC, IB	c
Pan Cytokeratin	AE1/AE3	DAKO	M3515	mouse	IHC, iFISH	a
SMAD4		Cell Signaling	9515	rabbit	IB	
Smooth muscle actin	1A4	DAKO	M0851	mouse	IHC, iFISH	b
TGFβ Receptor II		Cell Signaling	2518	rabbit	IB	
Vimentin	V9	DAKO	M0725	mouse	IHC, IB	b

Table S3. List of genes differentially expressed between ITGB6+ and MUC1+ cells from MCFDCIS xenografts and from myoepithelial and epithelial cells in human breast tissue. Gene symbols, protein names, functional categories, ratios in expression, and p values of the observed differences are listed.

					ITGB6+/MUC1+		Myoepi/Epi	
#	Gene Symbo	Protein	Class	Protein name	Signal	P_value	Signal	P_value
1	APOE	APOE_HUMAN		Apolipoprotein E precursor	9.088	0	6.052	0.001
2	AXL	UFO_HUMAN		Tyrosine-protein kinase receptor UFO precursor	8.179	0.006	15.65	0
3	AZGP1	ZA2G_HUMAN		Zinc-alpha-2-glycoprotein precursor	-41.81	0	-9.025	0.01
4	CDKN1A	CDN1A_HUMA		Cyclin-dependent kinase inhibitor 1	4.09	0.01	2.465	0.002
5	CFL1	COF1_HUMAN		Cofilin-1	3.635	0.009	-6.574	0.025
6	COL4A2	CO4A2_HUMA		Collagen alpha-2(IV) chain precursor [Contains:	13.63	0.002	14.77	0
7	COL7A1	CO7A1_HUMA		Collagen alpha-1(VII) chain precursor	13.63	0.002	4.082	0.009
8	CTNNBIP1	CNBP1_HUMA		Beta-catenin-interacting protein 1	9.088	0.015	-3.107	0.018
9	DUSP6	DUS6_HUMAN		Dual specificity protein phosphatase 6	3.938	0.034	4.283	0.009
10	EFCAB4A	EFC4A_HUMA		EF-hand calcium-binding domain-containing protein 4A	-6.602	0.001	-7.269	0.005
11	FXYD3	FXYD3_HUMA		FXYD domain-containing ion transport regulator 3	-2.106	0.001	-4.718	0.002
12	GABRP	GBRP_HUMAN		Gamma-aminobutyric-acid receptor subunit pi precursor	-2.32	0	-15.24	0.018
13	GPC1	GPC1_HUMAN		Glypican-1 precursor	9.088	0.003	7.143	0.028
14	HAX1	HAX1_HUMAN		HS1-associating protein X-1	3.272	0.021	-2.9	0.005
15	JMJD3	JMJD3		jumonji domain containing 3	2.726	0.027	1.96	0.032
16	LGALS1	LEG1_HUMAN		Galectin-1	1.912	0	4.804	0.008
17	LOC124220	U773_HUMAN		Protein UNQ773/PRO1567 precursor	-45.11	0	-9.574	0.003
18	MAL2	MAL2_HUMAN		Protein MAL2	-15.41	0	-4.566	0.002
19	MRC2	MRC2_HUMAN		Macrophage mannose receptor 2 precursor	4.544	0	11.14	0.001
20	MST150	NID67_HUMAN		Putative small membrane protein NID67	9.997	0.002	13.43	0.034
21	MT2A	MT2_HUMAN		Metallothionein-2	20.3	0	6.83	0
22	MYL9	MLRN_HUMAN		Myosin regulatory light chain 2, smooth muscle isoform	9.088	0.003	4.507	0
23	NNMT	NNMT_HUMAN		Nicotinamide N-methyltransferase	8.179	0.028	10.28	0
24	S100A14	S10AE_HUMA		Protein S100-A14	1.601	0.049	-3.884	0.006
25	SCGB1D2	SG1D2_HUMA		Secretoglobin family 1D member 2 precursor	-4.035	0.003	-7.051	0.025
26	SERPINA1	A1AT_HUMAN		Alpha-1-antitrypsin precursor	4.771	0	-13.28	0.012
27	SERPINA3	AACT_HUMAN		Alpha-1-antichymotrypsin precursor	-1.695	0.013	-5.83	0.004
28	SERPINH1	SPH2_HUMAN		Collagen-binding protein 2 precursor	7.27	0.008	4.667	0.001
		SERPH_HUMA		Serpin H1 precursor				
29	SPARC	SPRC_HUMAN		SPARC precursor	42.71	0	15.11	0.001
30	SPDEF	SPDEF_HUMA		SAM pointed domain-containing Ets transcription factor	-3.081	0.05	-37.8	0.003
31	SREBF1	SRBP1_HUMA		Sterol regulatory element-binding protein 1	-2.455	0.021	-3.673	0.01
32	STAP2	STAP2_HUMA		Signal-transducing adaptor protein 2	-3.081	0.05	-5.124	0.005
33	SYNGR2	SNG2_HUMAN		Synaptogyrin-2	-3.026	0.014	-5.377	0.002
34	TAGLN	TAGL_HUMAN		Transgelin	16.36	0.001	16.8	0
35	TIMP1	TIMP1_HUMAN		Metalloproteinase inhibitor 1 precursor	6.059	0	3.006	0.037
36	TM9SF2	TM9S2_HUMA		Transmembrane 9 superfamily protein member 2	-7.702	0.042	-3.704	0.002
37	TPM2	TPM2_HUMAN		Tropomyosin beta chain	16.36	0.001	7.576	0.001
38	TRIM26	TRI26_HUMAN		Tripartite motif-containing protein 26	-7.702	0.042	-1.996	0.041
39	UCK2	UCK2_HUMAN		Uridine-cytidine kinase 2	-4.035	0.042	3.681	0.03
40	VIM	VIME_HUMAN		Vimentin	2.908	0	7.695	0

Table S4. List of top scoring pathway maps, GO processes, and biomarkers differentially represented between ITGB6+ and MUC1+ cells from MCFDCIS xenografts and from myoepithelial and epithelial cells in human breast tissue.

Myoepithelial/epithelial cells		**ITGB6+/MUC1+ cells**	
Map	**p-Value**	**Map**	**p-Value**
Oxidative phosphorylation	1.43E-06	* Cell adhesion_ECM remodeling	2.92E-04
* Cell adhesion_ECM remodeling	4.40E-05	Cytoskeleton remodeling_Regulation of actin cytoskeleton by Rho GTPases	3.22E-04
* Cell adhesion_Integrin-mediated cell adhesion	1.48E-04	* Cytoskeleton remodeling_Slit-Robo signaling	2.19E-03
* Cytoskeleton remodeling_Keratin filaments	6.32E-04	* Cytoskeleton remodeling_Cytoskeleton remodeling	4.02E-03
* Cell adhesion_Role of tetraspanins in the integrin-mediated cell adhesion	8.51E-04	* Niacin-HDL metabolism	5.37E-03
* Cell adhesion_Endothelial cell contacts by non-junctional mechanisms	3.28E-03	* Development_MAG-dependent inhibition of neurite outgrowth	7.45E-03
* Cytoskeleton remodeling_Cytoskeleton remodeling	3.89E-03	Transcription_Role of Akt in hypoxia induced HIF1 activation	1.00E-02
* Cell adhesion_Chemokines and adhesion	4.04E-03	* Cell adhesion_Chemokines and adhesion	1.25E-02
* Cytoskeleton remodeling_Fibronectin-binding integrins in cell motility	9.51E-03	* Cell adhesion_Role of tetraspanins in the integrin-mediated cell adhesion	2.36E-02
* Niacin-HDL metabolism	9.67E-03	Oxidative stress_Role ASK1 under oxidative stress	2.53E-02
* Cytoskeleton remodeling_Slit-Robo signaling	1.21E-02	* Regulation of lipid metabolism_Regulation of lipid metabolism by niacin and isoprenaline	2.66E-02
* Cytoskeleton remodeling_Alpha-1A adrenergic receptor-dependent inhibition of PI3K	1.58E-02	* Cytoskeleton remodeling_TGF, WNT and cytoskeletal remodeling	2.87E-02
* Cytoskeleton remodeling_Neurofilaments	1.82E-02	* Glycolysis and gluconeogenesis p. 2 / Human version	3.20E-02
Cytoskeleton remodeling_Regulation of actin cytoskeleton by Rho GTPases	2.92E-02	* Immune response _CCR3 signaling in eosinophils	3.49E-02
* Cytoskeleton remodeling_TGF, WNT and cytoskeletal remodeling	2.97E-02	* Cell adhesion_Integrin-mediated cell adhesion	3.72E-02
Process	**p-Value**	**Process**	**p-Value**
cell adhesion	3.22E-10	regulation of actin polymerization and/or depolymerization	3.65E-09
biological adhesion	3.22E-10	regulation of actin filament length	7.50E-09
cellular component assembly	1.38E-08	glycolipid transport	3.03E-08
actin cytoskeleton organization and biogenesis	8.21E-08	negative regulation of biological process	1.55E-07
negative regulation of cellular process	1.24E-07	actin polymerization and/or depolymerization	2.08E-07
negative regulation of biological process	1.58E-07	negative regulation of protein metabolic process	1.58E-06
cell development	4.62E-07	negative regulation of cellular process	2.16E-06
actin filament-based process	4.64E-07	multicellular organismal development	5.76E-06
negative regulation of cell adhesion	5.59E-07	lipid transport	9.07E-06
myoblast differentiation	6.19E-07	regulation of cellular component organization and biogenesis	9.22E-06
positive regulation of cell migration	9.46E-07	regulation of actin filament polymerization	1.29E-05
regulation of cell adhesion	1.10E-06	response to extracellular stimulus	2.58E-05
cell differentiation	1.33E-06	regulation of multicellular organismal process	3.09E-05
cellular developmental process	1.33E-06	steroid catabolic process	3.48E-05
negative regulation of cell proliferation	1.53E-06	regulation of apoptosis	4.21E-05
muscle fiber development	4.12E-06	sequestering of actin monomers	4.25E-05
Disease / disease group	**p-Value**	**Disease / disease group**	**p-Value**
Urogenital neoplasms	2.75E-09	Carcinoma	9.70E-13
Breast neoplasms	6.65E-09	Neoplasms, Glandular and Epithelial	2.40E-11
Breast diseases	6.90E-09	Digestive System Neoplasms	6.60E-11
CarcinomaUrological and male genital diseases	9.70E-09	Endocrine System Diseases	2.30E-10
Neoplasms, Grandular and Epithelial	9.90E-09	Adenocarcinoma	2.00E-09
Urogenital diseases	1.40E-08	Urogenital Diseases	2.20E-09
Prostatic neoplasms	1.60E-08	Flaviviridae Infections	3.60E-09
Genital neoplasms, female	2.70E-08	Endocrine Gland Neoplasms	4.50E-09
Neoplasms by site	3.20E-08	Urogental Neoplasms	6.50E-09
Genital neoplasms, male	4.90E-08	Genital Neoplasms, Female	6.60E-09
Wounds and injuries	5.20E-08	Breast Neoplasms	7.90E-09
Endocrine Gland Neoplasms	6.90E-08	Breast Diseases	1.20E-08
Neoplasms	2.60E-07	Neoplasms by Histologic Type	1.40E-08

Figure S1. Comparison of subcutaneous and mammary fat pad injections. A-B: Histology of the tumors (H&E) and the expression of SMA and p63 analyzed at the indicated time points after injection to the mammary fat pad (**A**) and subcutaneous site (**B**). Although overall progression to invasion appears to be similar at the two sites, much higher fraction of the epithelial cells are positive for p63 in xenograft developed in the mammary fat pad indicating site-dependent differences in the differentiation of the cells. **C:** Histological and immunohistochemical analyses of the indicated xenografts derived from cell lines injected alone (-) or co-injected with normal myoepithelial cells (+HME). No DCIS histology and myoepithelial cell layer was detected in any of these tumors. SMA (smooth muscle actin) positive cells are mouse myofibroblasts and endothelial cells. Scale bars correspond to 100 μm.

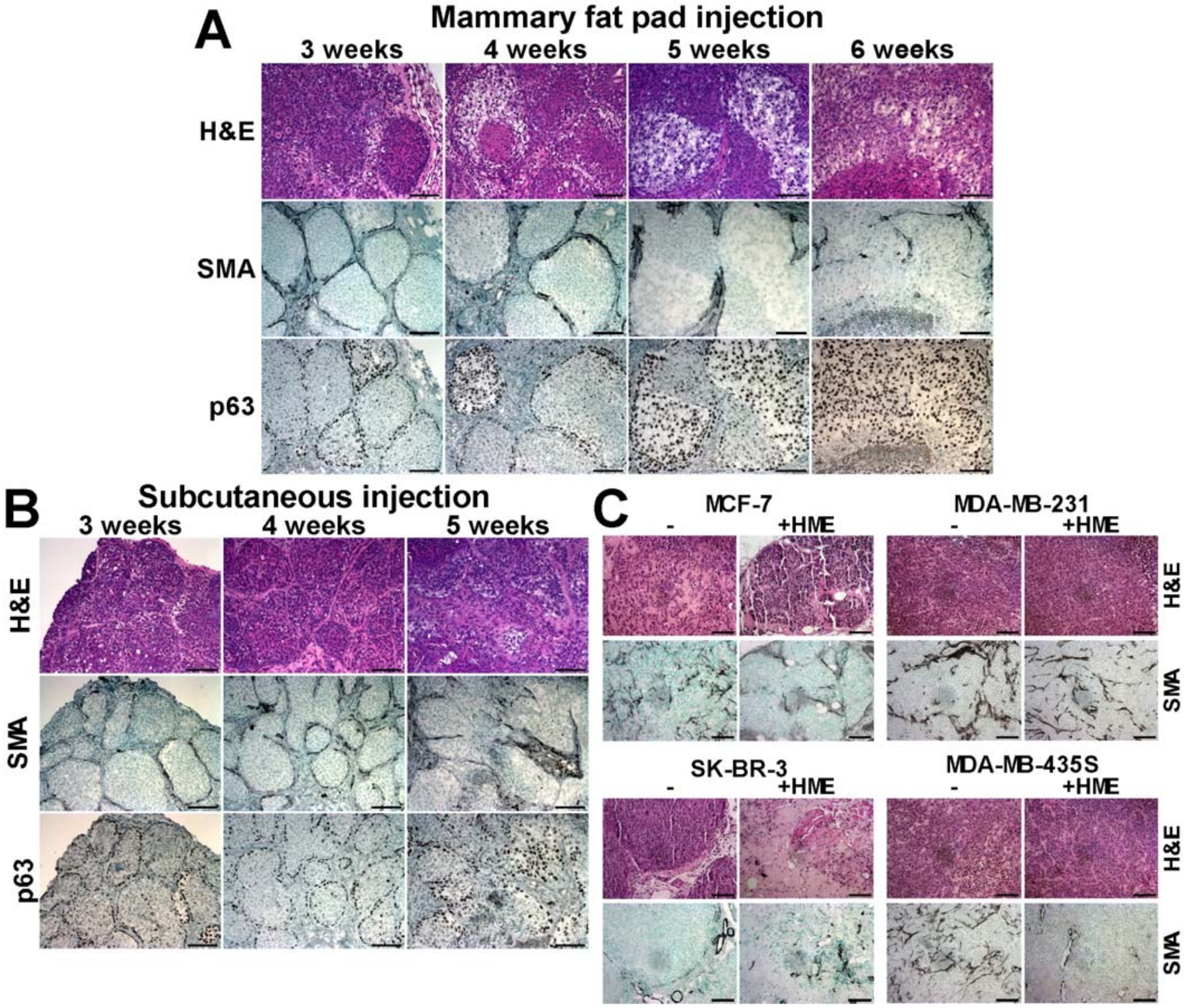


Figure S2. Characterization of the progenitor properties of MCFDCIS cells. Scale bars correspond to 100 μm in all panels. **A:** Confirming the human origin of cytokeratin positive cells in MCFDCIS xenografts. Immuno-FISH analysis of the time course experiment demonstrates that cytokeratin (CTK) positive cells (blue cytoplasm) are of human origin. **B:** FACS analysis of cultured cells for the indicated cell surface markers. With the exception of MUC1 the cells are homogenously positive (ITGA6, ESA, ITGB6, and CD44) or negative (CD10 and CD24) for the proteins analyzed. The expression of ITGB6 is further increased following TGFβ1 treatment. **C:** FACS sorting for MUC1+ and MUC1- cells and histology of xenografts resulting from them. Immunocytochemical (**D)** and immunoblot (**E**) analysis of CK14 and MUC1 expression in single-cell clones. **F:** Immunohistochemical analyses of the indicated markers in xenografts derived from single-cell clones. All clones gave rise to DCIS-like tumors, but clone 7 had a more invasive component. *In vivo,* the expression pattern of all proteins is the same in all clones regardless of their expression *in vitro*.

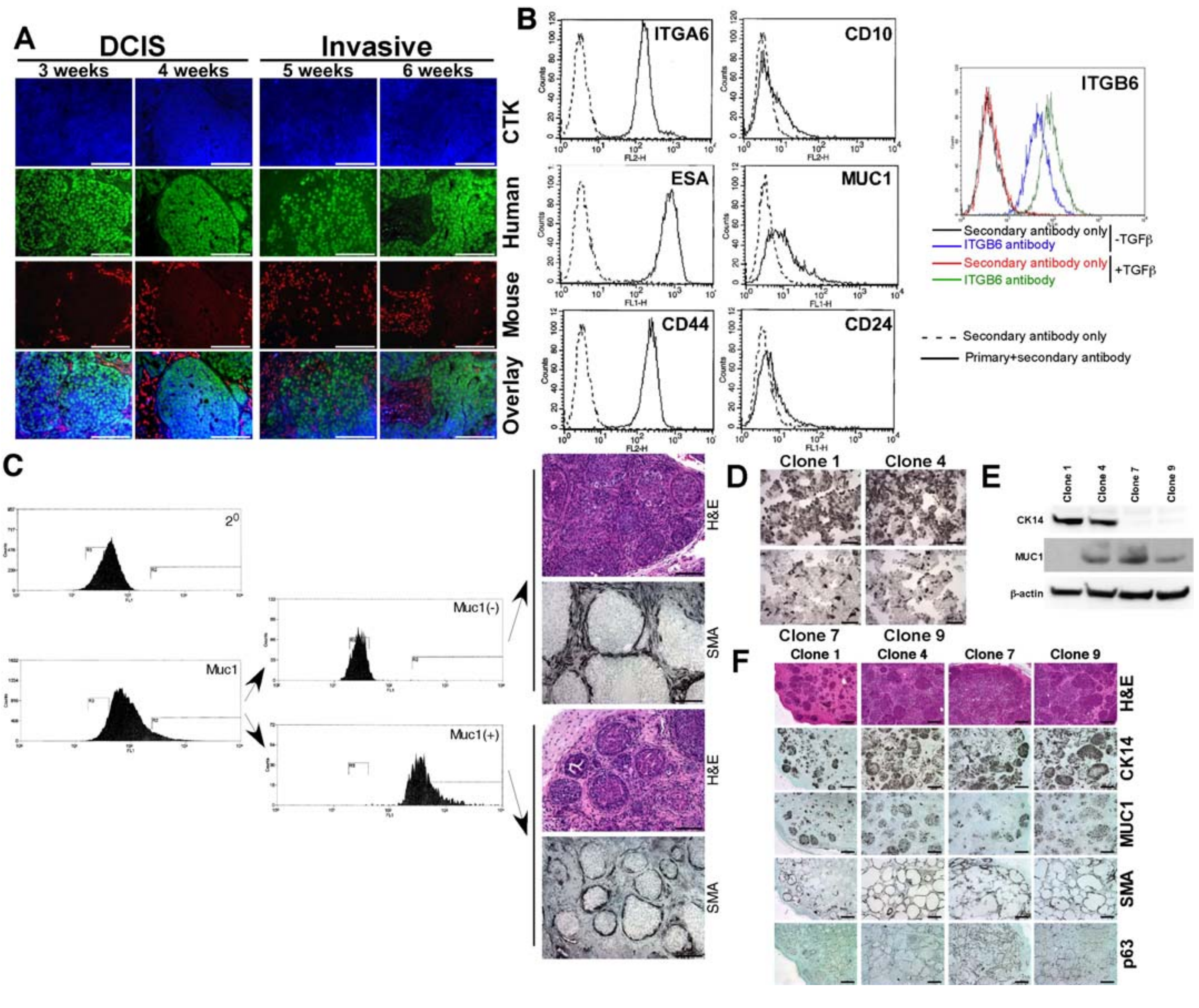


Figure S3. Analysis of genetic changes in MCF10A series cells and MCFDCIS derived xenografts. A: 11K SNP array analysis of the indicated cells and xenografts for copy number changes. Copy number gain of 8q24, including *MYC,* and homozygous deletion of *CDKN2A* at 9p21 are indicated. **B:** FISH confirmation of 8q24 copy number gain and insertion to 10q22. Upper panel: *MYC* probe (red) gave three hybridization signals, two on chromosome 8 (green) and one on chromosome 10 (aqua). Lower panel: *MYC* break apart probe (5' red and 3' green) shows two intact copies (yellow fused signal including both 5' and 3' portions) on chromosome 8 (aqua) while the signal on chromosome 10 contains only the 5' portion of the probe (red). **C:** 11K SNP array analysis of the indicated cells and xenografts. An inferred LOH (loss of heterozygosity) map including all chromosomes indicates that, with the exception of two tumors, none of the samples demonstrated copy number changes compared to MCFDCIS cells. Blue and yellow colors indicate LOH and retention of both alleles, respectively. **D:** 250K SNP array analysis of purified ITGB6+ and MUC1+ cells and their comparison to parental MCFDCIS cells. **E:** Enlarged view of chromosome 8 highlighting the focal copy number gain at 8q24.

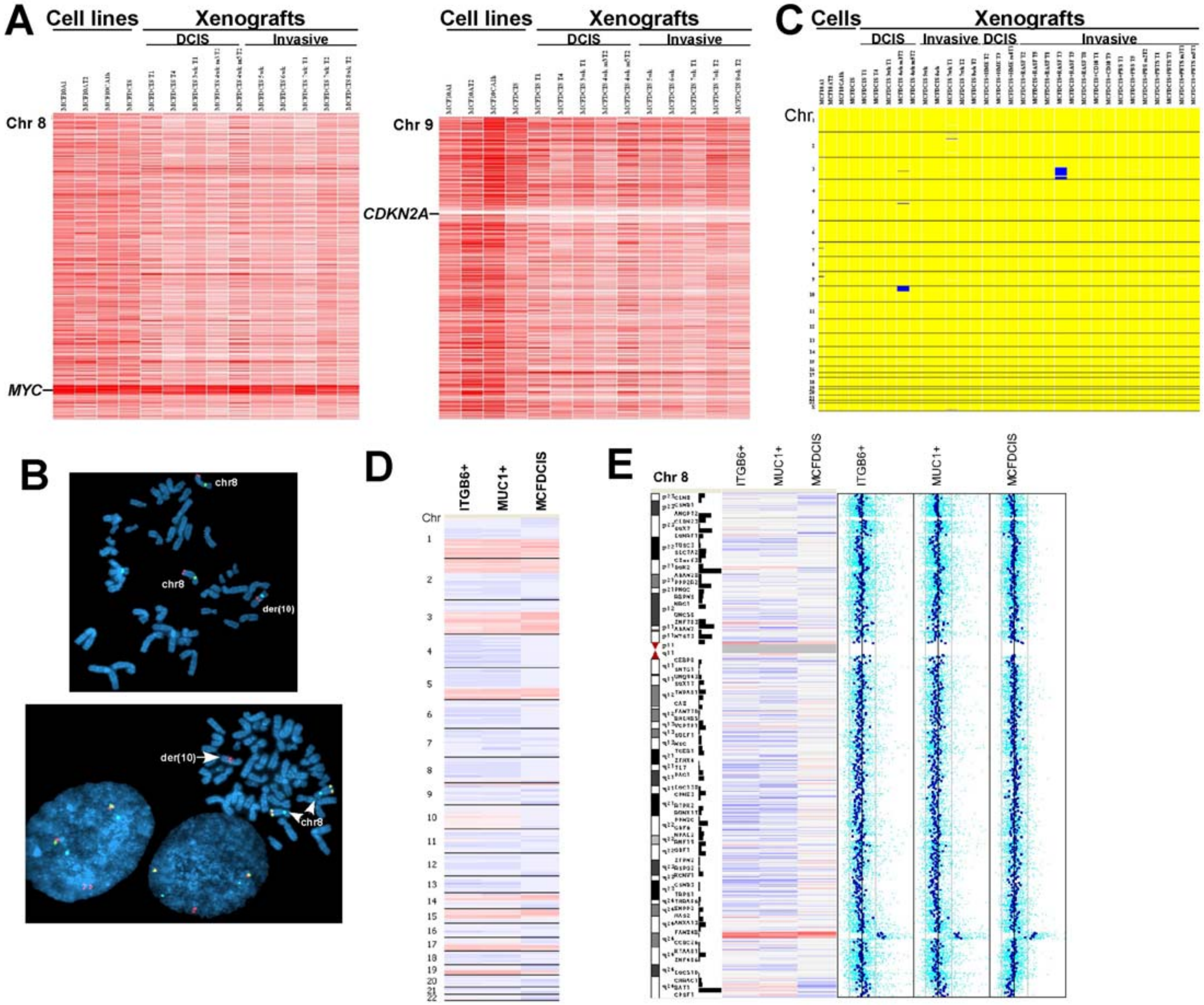
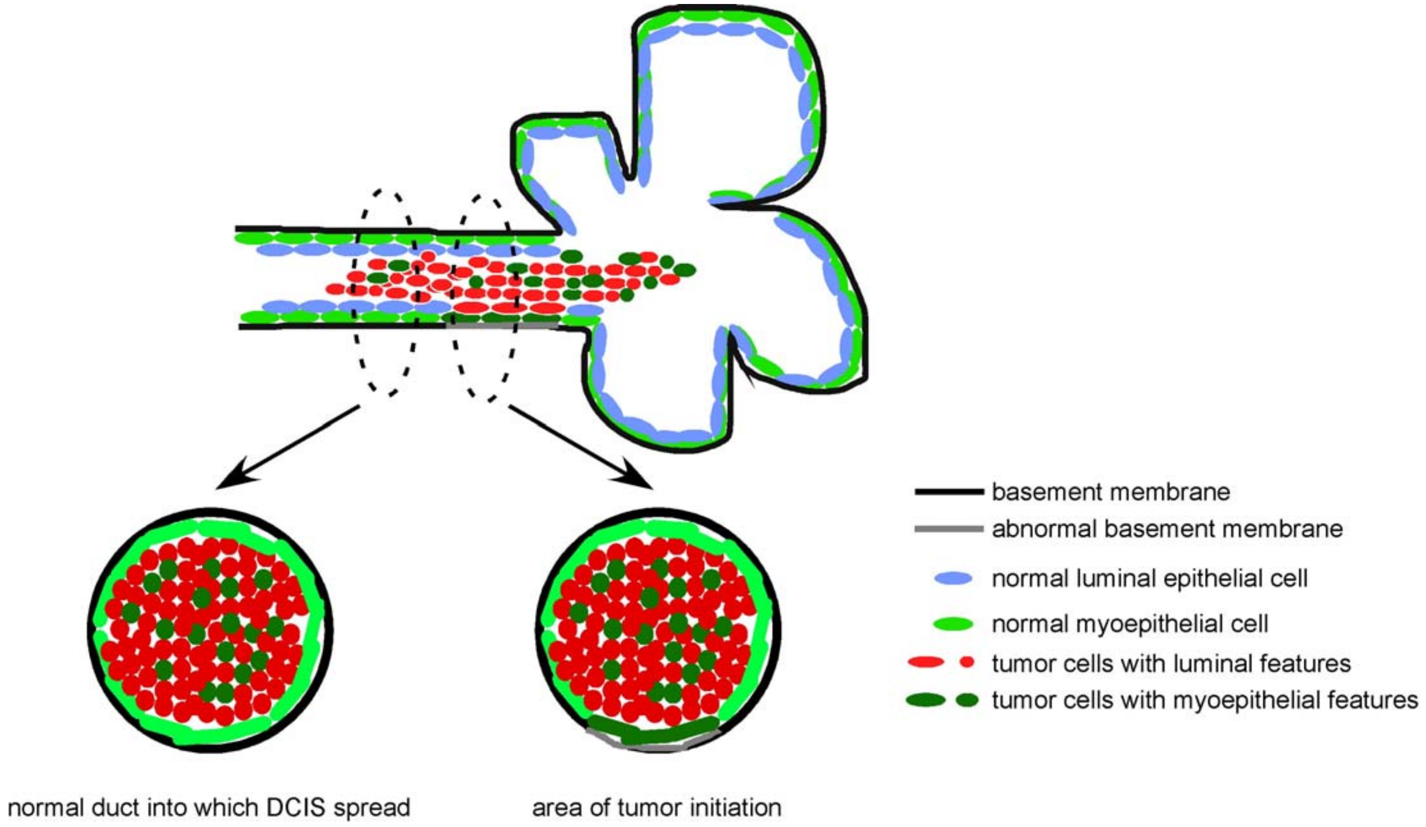


Figure S4. Schematic model of basal-like DCIS. The majority of histologic sections in DCIS represent normal ducts and lobules into which the tumor cells spread and only at the site of tumor initiation we may detect genetically abnormal bipotential tumor initiating cells. A prediction of this model is that some DCIS may already be invasive and spread outside of the ducts at a very early stage. This hypothesis is supported by recent data in a mouse model of DCIS describing systemic spread at early steps of tumorigenesis (Husemann et al., 2008).


basement membrane
abnormal basement membrane
normal luminal epithelial cell
normal myoepithelial cell
tumor cells with luminal features
tumor cells with myoepithelial features

normal duct into which DCIS spread

area of tumor initiation

Figure S5. Maps of signaling pathways differentially activated between epithelial and myoepithelial cells both in primary human breast tissue and in the MCFDCIS xenograft model. A: WNT, TGFβ, and cytoskeletal remodeling. **B:** Lipid metabolism. Red rectangles mark differentially expressed genes.

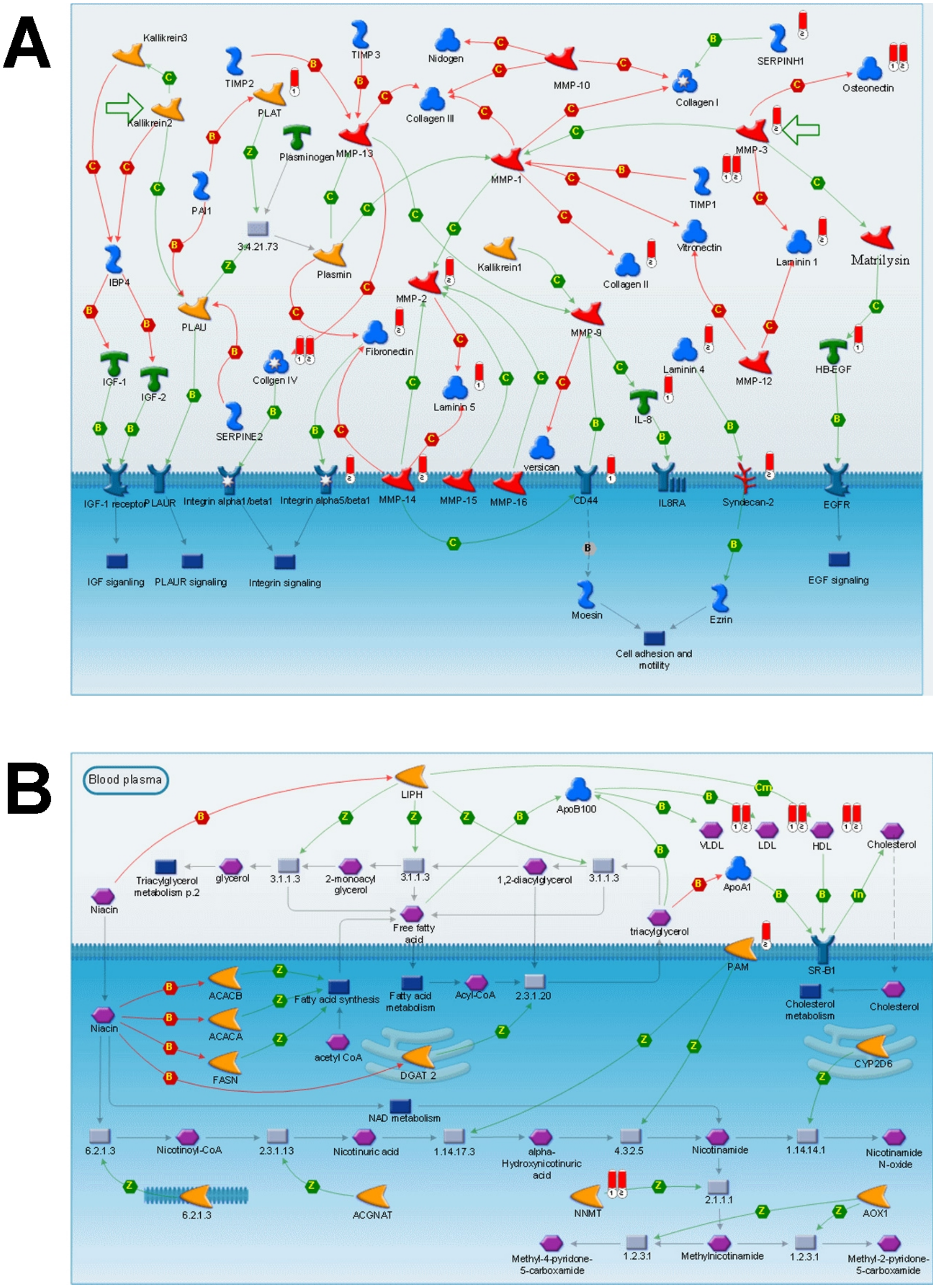


Figure S6. Characterization of xenografts derived from MCFDCIS cells and co-injected non-epithelial cells. Scale bars correspond to 100 μm in all panels. **A:** Immuno-FISH using pan-cytokeratin antibody (panCK-blue) identifies epithelial cells. The overlay demonstrates the human origin of cytokeratin positive epithelial cells and the mouse origin of stromal cells. **B:** Immuno-FISH using smooth muscle actin antibody (SMA-blue) identifies myoepithelial cells or myofibroblasts. The overlay demonstrates the human origin of SMA-positive myoepithelial cells (yellow arrows) in DCIS structures and the mouse origin of SMA-positive myofibroblasts (yellow stars).

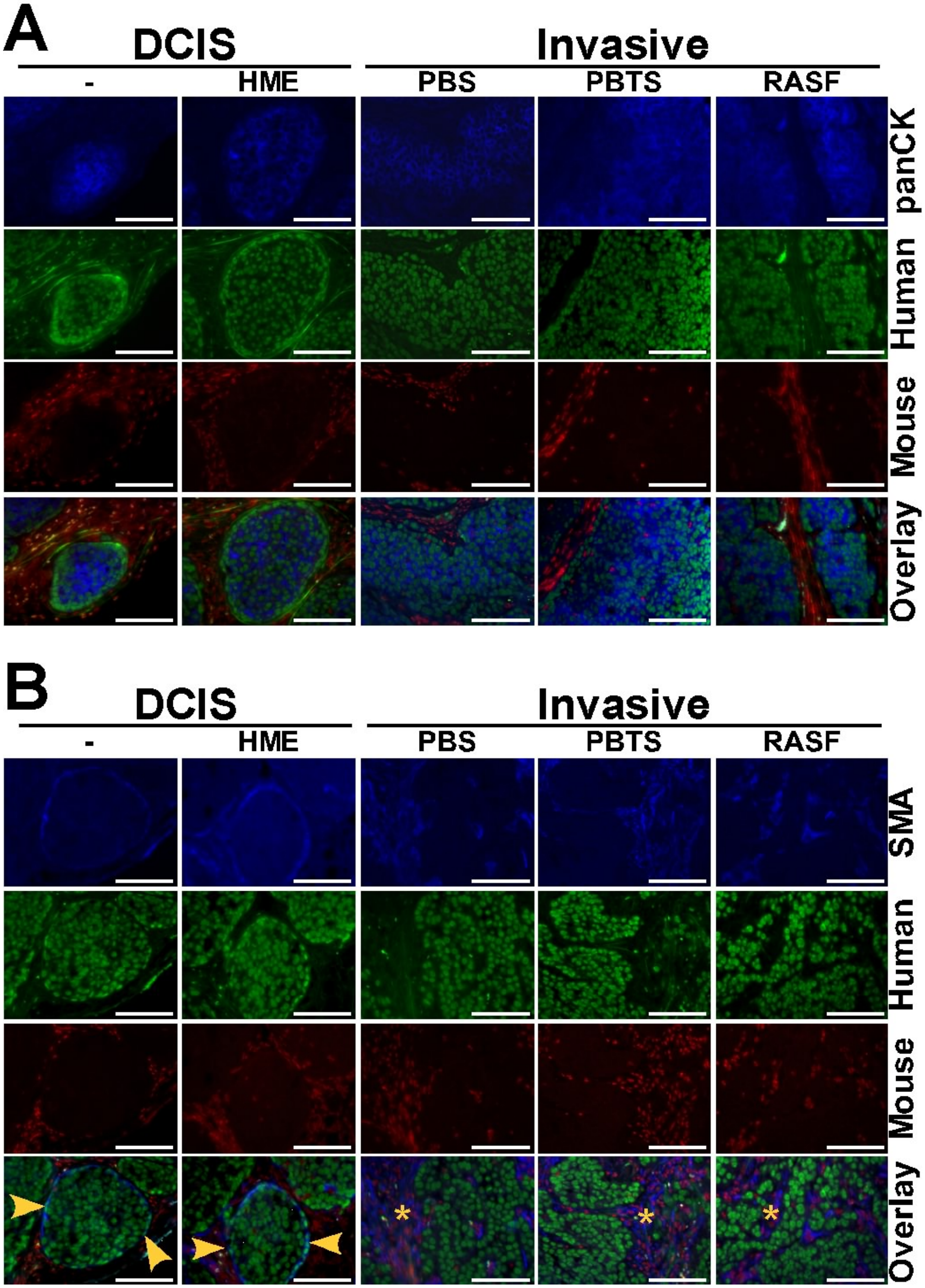


Figure S7. Validation of shRNA clones. shRNA clones highlighted in red were used for xenograft studies. **A:** Quantitative RT-PCR analysis of SMAD4 mRNA levels in control shGFP and SMAD4 shRNA clones. Inset is western blot analysis of SMAD4 protein levels. **B:** Quantitative RT-PCR analysis of TGFBR2 mRNA levels in control shGFP and TGFBR2 shRNA clones. Inset is western blot analysis of TGFBR2 protein levels. **C:** Semi-quantitative RT-PCR analysis of GLI2 mRNA levels in control shGFP and GLI2 shRNA clones. Right panel is western blot analysis of Gli2 protein levels. **D:** Western blot analysis of the indicated proteins in control and Gli2, SMAD4, and TGFBR2 shRNA clones. **E:** TGFβ and Gli activity determined using luciferase reporters in the indicated MCFDCIS derivatives and treatment conditions. Error bars represent mean ± SD.

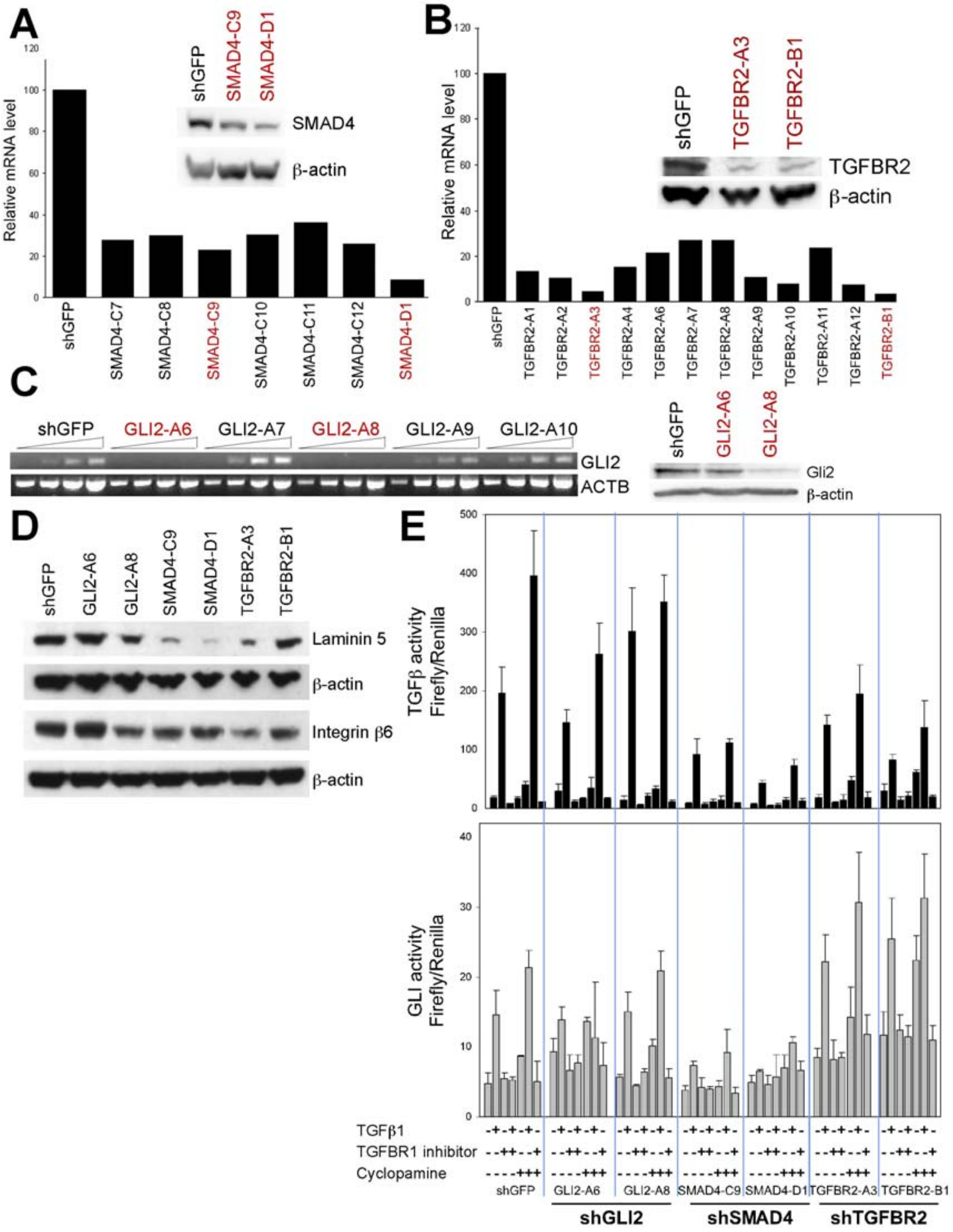
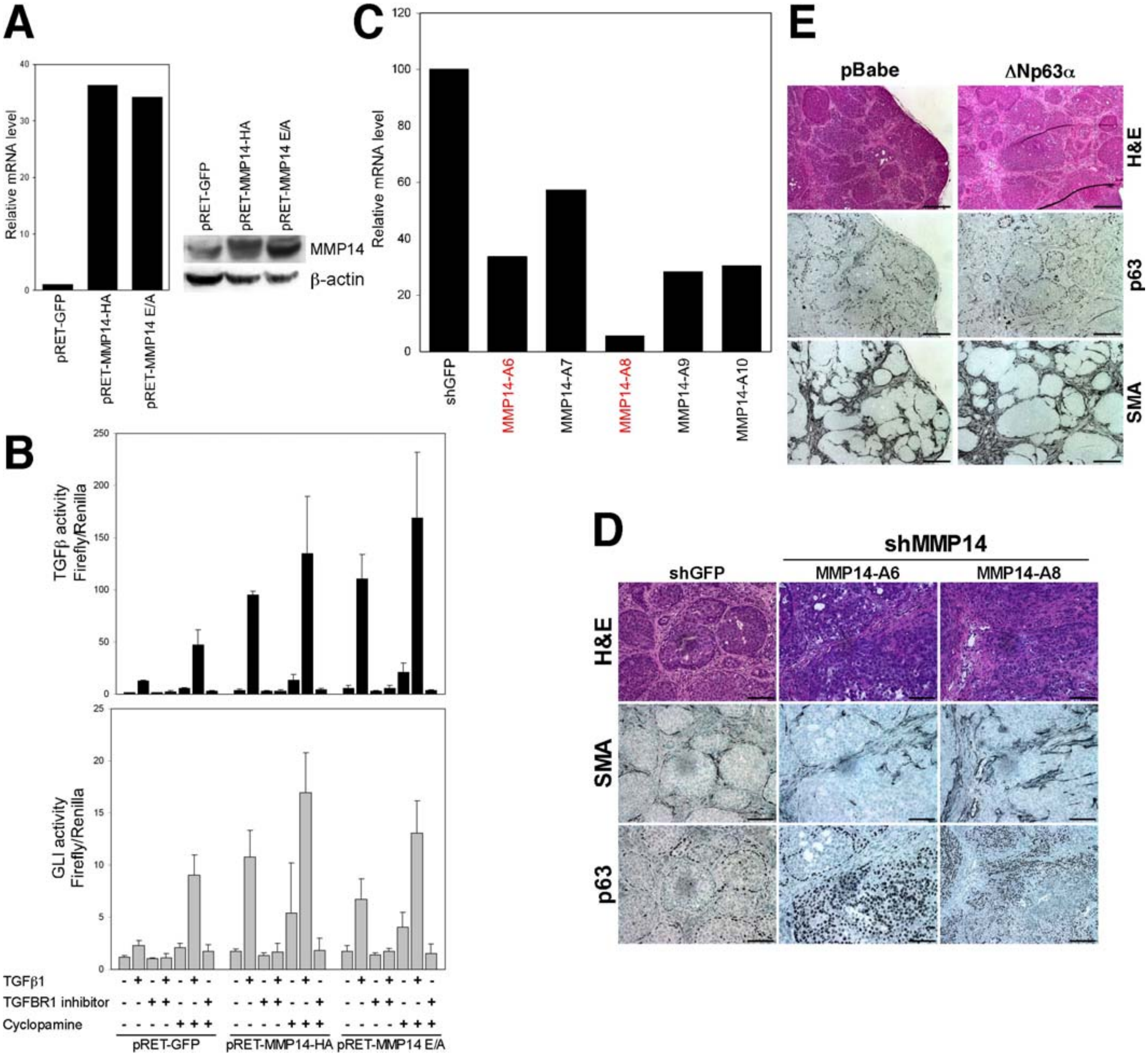


Figure S8. Luciferase assays and xenograft results. A: Quantitative RT-PCR analysis of MMP14 mRNA levels in control pRET-GFP and clones overexpressing wild-type (MMP14-HA) or inactive mutant (MMP14 E/A) MMP14. Right panel is western blot analysis of MMP14 protein levels. **B:** Luciferase activity assessing TGFβ and Gli activity in control and MMP14 overexpressing MCFDCIS cells following the indicated treatments. Error bars represent mean ± SD. **C:** Quantitative RT-PCR analysis of MMP14 mRNA levels in control shGFP and MMP14 shRNA clones. **D:** The effect of decreased MMP14 levels on the histology of MCFDCIS xenografts. Xenografts derived from shMMP14 clones lack myoepithelial cells and have invasive phenotype. **E:** Histology and SMA and p63 expression in xenografts derived from vector control (pBabe) and ΔNp63α overexpressing MCFDCIS cells. Despite its widespread overexpression in 2D culture, in xenografts only the myoepithelial cells are positive for p63 suggesting its post-transcriptional downregulation in luminal epithelial cells potentially due to their lack of contact with the basement membrane. Scale bars correspond to 100 μm in all panels.



References:

Cai, L., Huang, H., Blackshaw, S., Liu, J. S., Cepko, C., and Wong, W. H. (2004). Clustering analysis of SAGE data using a Poisson approach. Genome Biol *5*, R51.

Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A *95*, 14863-14868.

Husemann, Y., Geigl, J. B., Schubert, F., Musiani, P., Meyer, M., Burghart, E., Forni, G., Eils, R., Fehm, T., Riethmuller, G., and Klein, C. A. (2008). Systemic spread is an early step in breast cancer. Cancer Cell *13*, 58-68.

Ney, P. A., Andrews, N. C., Jane, S. M., Safer, B., Purucker, M. E., Weremowicz, S., Morton, C. C., Goff, S. C., Orkin, S. H., and Nienhuis, A. W. (1993). Purification of the human NF-E2 complex: cDNA cloning of the hematopoietic cell-specific subunit and evidence for an associated partner. Mol Cell Biol *13*, 5604-5612.

Peters, B. A., Diaz, L. A., Polyak, K., Meszler, L., Romans, K., Guinan, E. C., Antin, J. H., Myerson, D., Hamilton, S. R., Vogelstein, B.*, et al.* (2005). Contribution of bone marrow-derived endothelial cells to human tumor vasculature. Nat Med *11*, 261-262.