PNAS
www.pnas.org

Supplementary Information for

Acquisition of a hybrid E/M state is essential for tumorigenicity of basal breast cancer cells

Cornelia Kröger[1], Alexander Afeyan[1, 2], Jasmin Mraz[1, 3], Elinor Ng Eaton[1], Ferenc Reinhardt[1], Yevgenia L. Khodor[4], Prathapan Thiru[1], Brian Bierie[1], Xin Ye[1, 5], Christopher B. Burge[4], Robert A. Weinberg[1,6,7]

Corresponding author: Robert A. Weinberg
Email: weinberg@wi.mit.edu

This PDF file includes:

www.pnas.org/cgi/doi/10.1073/pnas.1812876116

Supplementary Information Text

SUPPLEMENTARY RESULTS

Generation of the xE and xM cell populations.

For these studies we used the well-characterized human mammary epithelial (HME) cell culture model, involving human mammary epithelial cells immortalized by forced expression of hTERT (1, 2). We proceeded to transform these cells by introduction of SV40 early-region genes and the *HRAS* V12 oncogene, resulting in HMLER cells (1), and labelled them for easy detection with either Clover or tdTomato reporters, both membrane-bound fluorescent proteins, (3, 4).

In order to enrich for tumor cell populations with differing epithelial (E) versus mesenchymal (M) traits, we initially FACS-sorted HMLER cells using the frequently employed CD24/CD44 cell-surface markers into epithelial (E) $CD24^{hi}CD44^{low}$ (HMLER44L) and more mesenchymal (M) $CD24^{low}CD44^{hi}$ (HMLER44H) populations (Supp. Fig. 1A) (5, 6). We then further stratified the heterogeneous HMLER44H cells using the marker CD104, as described recently (7), in order to generate both a $CD104^{+}CD44^{hi}$ population of hybrid epithelial-mesenchymal (E/M) phenotype and a more mesenchymal (M) $CD104^{-}CD44^{hi}$ cell population. Importantly, this particular M population remained capable of regenerating substantial numbers of more epithelial CD104+ cells over extended periods of passaging time *in vitro*, which eventually represented ~30% of the overall cell population (Supp. Fig. 1B). Because of this ongoing plasticity, we employed seven sequential rounds of FACS sorting in order to further enrich for more homogeneous cell subpopulations, which resulted in the highly pure E ($CD104^{+}CD44^{low}$), the E/M ($CD104^{+}CD44^{hi}$) and xM ($CD104^{-}CD44^{hi}$) tumor cell populations (Fig. 1A of main text).

Control experiments revealed that proliferation rates in monolayer culture of the cells in these three populations were not significantly different (Supp. Fig. 1 D). Moreover, levels of the introduced RAS oncoprotein, as determined by western immunoblots, were comparable in the E and E/M cells and up to 2-fold elevated in the xM cells (Fig. 1C of main text).

As cited above, the parental, generally mesenchymal, more heterogeneous CD44/CD104-derived M cell population from which we isolated the fully mesenchymal xM cells displayed a ~30% subpopulation of $CD104^{+}$ E/M cells, relative to such a corresponding $CD104^{+}$ subpopulation of only 5% in the xM cells (Fig. 1A, Supp. Fig. 1B). This raised the question of which specific property distinguished the xM and M cells from one another. In contrast to the xM tumors, M tumors weighed on average 0.55 g per tumor (Fig. 2), closely resembling in size the E/M tumors (~ 0.7 g) and were, by contrast, about 10-fold larger in size than the xM-derived tumors. This result indicated that the M cells exhibited a degree of phenotypic plasticity that enabled them to generate a significant $CD104^{+}$ E/M subpopulation while the xM cells lacked this plasticity and thus an ability to generate such E/M cells. The ability to generate E/M cells, in turn, was apparently responsible for the relatively high tumor-forming and stem cell frequencies relative to the M cells, which was similar to that of the E/M cells and was confirmed by stem cell frequency analysis conducted by implantations at limiting dilutions (Supp. Fig. 1C). Such behavior confronted us with two alternative mechanistic hypotheses: (i) that residence in the E/M state,

independent of plasticity, was important for stemness or (ii) that plasticity and an ability to move into and out of the E/M state was critically important for display of this trait.
To determine whether cell plasticity played a critical role in the elevated tumor-initiating and outgrowth capabilities of M cells relative to those of the xM cells, we performed FACS analysis using the cell-surface CD44 and CD104 markers of the cells prepared from individual tumors. Prior to injection, the majority of M and xM cells (70% and 95%, respectively) resided in a predominantly $CD104^-$ state. To gauge the degree of plasticity exhibited by these cells while growing within tumors, we monitored the percentage of cells that shifted from their initial more mesenchymal $CD104^-$ state to a more epithelial $CD104^+$ state. We found that both M cells and xM tumors, showed high degrees of plasticity *in vivo* with a high variability between individual tumors. Thus, 39% to 89% of the cells within the M-initiated tumors had gained epithelial properties and thus shifted into a $CD104^+CD44^{hi}$ E/M state *in vivo* (Supp. Fig. 1F). Similarly, 26-69% of xM cells shifted from an xM to a mixed E/M cell state as determined by their CD104/44 FACS profile (main Fig1 H, I). This revealed that, while the xM carcinoma cells generally resided stably in an xM state, those few cells that successfully grew out to form small tumors exhibited phenotypic plasticity as evidenced by their ability to generate E/M progeny. Taken together, these various observations suggested an association between the generation of subpopulations of E/M cells and acquisition of tumor-initiating ability. Indeed, this was compatible with the notion that residence in the E/M state was required for tumor-initiating ability and might even suffice, in the absence of plasticity, to generate tumor-initiating cells.
As a corollary of this, we presumed that the observed tumorigenic powers of the more heterogeneous mesenchymal (M) cell population depended on their ability to generate E/M subpopulations, doing so in a way that depended on their intrinsic plasticity.
As mentioned above, we note that the E cells bore only a minute subpopulation of tumor-initiating cells (Supp. Fig. 1C). We reasoned that if residence in the E/M state was crucial to tumor-initiating ability, as argued immediately above, we should observe, in the few tumors seeded by the E cells, a shift toward an E/M state *in vivo* with associated gain of certain mesenchymal properties. Confirming results by Chaffer et al. (8), we found that indeed 15% of the tumors generated by the implanted E cells expressed certain mesenchymal markers *in vivo*, as determined by CD44 marker analysis (main Fig. 1H, I, Supp. Fig. 1F), compared to around 2% of $CD44^{hi}$ E cells *in vitro*, suggesting additionally that gain of certain mesenchymal traits enabled concomitant acquisition of tumor-initiating properties (9, 10).

Generation of xE Zeb1 KO cell lines
In order to trap cells entirely in an xE state, we used the CRISPR/Cas9 technology to completely eliminate Zeb1 expression in a population of E cells (Supp. Fig. 1A). We began by comparing the *ZEB1* CRISPR knockout (KO) cell clones and Zeb1-overexpressing clones with their respective isogenic controls, wishing also to exclude possible pre-existing genetic differences among the engineered single cell clones. To do so, we first isolated epithelial and mesenchymal single-cell clones (E-SCC and M-SCC, respectively) from the HMLER parental population. Furthermore, in order to minimize any cell-intrinsic disposition for plasticity, we screened E-SCC and M-SCC clones for those displaying the lowest level of plasticity as determined by either a gain of CD44 or CD104 marker

expression, respectively (Supp. Fig. 2A). Both epithelial and mesenchymal SCCs were then targeted with small guide (sgRNA) for *ZEB1*, subjected subsequently to single-cell cloning, and validated by sequencing of the appropriate DNA sequences, thereby generating the E-SCC-Zeb1KO or M-SCC-Zeb1KO cells (Supp. Fig. 2B-D). Mesenchymal cells were included in the targeting to function as controls, allowing us to calibrate the effectiveness of the *ZEB1* CRISPR knockout. Zeb1 KO was validated by western blotting or IF staining. Because the E cells do not naturally express Zeb1, we induced an EMT in these cells by EMT-TF expression or treatment with TGF-β to validate the efficiency of the Zeb1KO (Supp. Fig. 2B, 3B, 5B, C). These E-SCC-Zeb1KO cells upon TGF-β did not express Zeb1 and were thus termed xE cells. As described in the main text, these cells progressed into an E/M state upon TGF-β treatment and were not able to complete the EMT process and move into an xM state (Supp. Fig. 3B, 5A).

Zeb1 KO in xM cell lines

In light of our demonstration that Zeb1 was required for completion of the entire EMT program, we asked whether, conversely, Zeb1 loss in M cells would result in a mesenchymal-epithelial transition (MET). To this end, we examined M-Zeb1KO cells *in vitro*, in which *ZEB1* had been knocked-out in fully mesenchymal M $CD104^{-}CD44^{hi}$cells. *ZEB1* KO in the mesenchymal cells led to the reactivation of expression of the epithelial marker CD104, yielding cells that stably resided in a $CD44^{hi}CD104^{+}$ E/M state as detected by CD104/CD44 marker analysis of xenografts. Moreover, these cells re-expressed the epithelial marker E-cadherin, which could be detected at low levels at the cell membrane by IF (Supp. Fig. 3D). Similar to certain other E/M cell populations described here, the residence of these cells in this mixed E/M phenotype was stable with low plasticity maintained *in vivo*, as demonstrated by FACS analysis of digested tumors that had previously been growing in mammary fat pads (which revealed uniform expression of the $CD104^{+}CD44^{hi}$ marker state in all the cells) (Supp. Fig. 3C). Hence, loss of Zeb1 expression in these M cells led to a partial MET, with reversion of cells from a mesenchymal state to the hybrid E/M state, which once again displayed low plasticity along the E-to-M axis *in vivo*.

Canonical Wnt signaling in the E/M cells

To further test the hypothesis that the E/M cell population is marked by elevated canonical Wnt signaling, we transfected cells with the 7xTCF-GFP-reporter (7TGP), which acts as a transcriptional reporter of β-catenin-mediated canonical Wnt signaling (11). Indeed, previous data in our lab suggested that canonical Wnt3a and TGF-β1 work synergistically to induce an EMT program (12). To validate this reporter assay, we therefore overexpressed in these cells a strong canonical Wnt inducer, Wnt3a (Supp. Fig. 6C). We then undertook to monitor canonical Wnt signaling as cells passed progressively through various E - M states (E, E/M, xM) by treating these E-7TGP-Wnt3A-expressing cells with TGF-β1. Interestingly, we observed that GFP induction was observed in E cells and was twice as high in the E/M cell populations; however, the levels of GFP^{+} cells were 5-fold lower in the xM cells compared to E/M cells (Supp. Fig. 6D-F). These observations confirmed that canonical Wnt signaling was highly active in E/M cells as accounted for by a strong GFP signal, but was downregulated as cells moved into an xM state.

Canonical Wnt signaling and Snail are both negatively regulated by GSK-3β (13, 14). Using the GSK-3β inhibitor BIO (6-bromoindirubin-3'-oxime) (15) on E cells expressing Wnt3a and the canonical Wnt signaling reporter 7TCF-GFP, we could validate in HMLER cells that inhibition of GSK-3β did indeed lead to an increase in the percentage of GFP+ cells and canonical Wnt signaling (Supp. Fig. 6G). Moreover, we found that GSK-3β inhibition induced an EMT in the HMLER-E cells, as measured by CD104/CD44 marker expression by FACS, similar to an induction by EMT-TFs such as Snail or Zeb1, which drive cells into the $CD44^{hi}$ state (Supp. Fig. 6G). These combined findings are consistent with the notion that the shift from an E to an E/M state could be driven by stabilizing canonical autocrine Wnt signaling.

The alliance between β-catenin and Snail is underscored by the fact that both are down-regulated by GSK-3β (14, 16, 17). Hence, it is not surprising that by blocking GSK-3β using the GSK-3β inhibitor BIO (6-bromoindirubin-3'-oxime) (15), we could not only induce canonical Wnt signaling (Supp. Fig. 6G) but also elicit a partial EMT in the epithelial HMLER cells.

SUPPLEMENTARY MATERIALS AND METHODS

Cell lines and culture conditions

HMLER cells were cultured in 500 ml MEGM media (Lonza Bullet kit), supplemented with 250 ml DME, 250 ml F12, insulin (10µg/ml), EGF (10 ng/ml), hydrocortisone (1µg/ml), 1x Pen/Strep. SUM149 cell lines were cultured in F12 with 5% inactivated calf serum (IFS), 1ug/ml hydrocortisone, insulin (5ug/ml), and pen/strep. All cell lines were passaged every two to three days to keep cells in sub-confluent conditions. If stated cells were treated with BIO (6-bromoindirubin-3'-oxime) compound (Sigma) at a concentration of 1:20.000. Media was exchanged every two days.

Plasmid constructs and virus construction

HME (1) cells were transformed with pENTR1A SV40 LgT (Addgene 22297) cloned into pLenti CMV Neo DEST and pWZL-HRASg12v-Blast (18). To color-code cells, pLV-tdTomato-membrane was used and pLV-YFP-membrane (19) was subcloned and substituted with Clover (3) using the amplification primers (CTGGACGGCGACGTGAAC/ GCCCATATGCCTTACTTGTAC) into the BmgBI/NdeI site, creating pLV-Clover-membrane; and eBFP2 synthesized by IDT, creating pLV-EBFP-membrane. pLenti-CRISPR-Cas9 V2 (Addgene 52961) constructs were produced as previously described (20). Spacer guide sequences used for the constructs were: sgZEB1 (GAGCACTTAAGAATTCACAG) sgSNAIL (GGGACTCTCCTGGAGCCGAA); non-cutting controls sgNC (ACGGAGGCTAAGCGTCGCAA) (kindly provided by Yun Zhang, Whitehead Institute for Biomedical Research, Cambridge, MA). For SNAIL, TWIST and SLUG over-expression pLV IRES Tom expression vector was used (kindly provided by Leonardo Rodriguez, Whitehead Institute for Biomedical Research, Cambridge, MA). ZEB1 over-expression was induced using a pLenti CMV Puro DEST (w118-1) gateway recombined (LR) with ZEB1 V2 from Genecopoeia (GC-F0876-B, Addgene 17452). Blast resistance sequence was substituted in pLenti CMV Blast DEST with EBFP-membrane or tdTomato-membrane creating pLenti CMV tdTomM DEST and pLenti CMV EBFPM DEST. The pLV EBFP Zebmt was constructed exchanging the sgRNA targeting sequence within, with a gene fragment containing the substitute sequence GAACATTTGAGGATACATAG in the Zeb1 vector described above and then subloned into pLenti CMV EBFPM DEST. Fzd8-CRD (21) was subloned into pENTR1A using EcoRI/ApaI into EcoRI/EcoRV restriction side, respectively and further subloned into the Dox inducible gateway expression vector pCW57.1, a gift from David Root (Addgene 41393). Active Wnt3A-V5 (Addgene 43810), Active Wnt5A-V5 (Addgene 43813), Active Wnt7A-V5 (Addgene 43816), Active Wnt7B-V5 (Addgene 43817) (22) were subcloned into pENTR1A and further subcloned into pLenti CMV tdTomM DEST. 7TGP was a gift from Roel Nusse (Addgene 24305). pLenti-based constructs were packaged with pMD2.G (VSVG) and psPAX2 plasmids (Addgene 12259 and 12260, respectively). H-Ras V12 blast was packaged with pUMVC (Addgene 8449) and pMD2.G. Viral infections were performed using 5 µg/ml Polybren (EMD Millipore) for 16 h.

Generation of CRISPR/Cas9 HMLER KO and overexpression clones

To compare CRISPR clones to isogenic controls, single-cell clones (SCCs) were isolated from E and M populations. In brief, single-cells that had been sorted by FACS were seeded into 96-well plates and 24 SCCs were grown up. Clones were selected after FACS CD104/CD44 marker analysis, namely clone M-SSC #12 and E-SCC #5. CRIPSR targeting with sgRNAs for Zeb1 or Snail were performed (Supp. Fig. 2B). Out of 3 different sgRNAs tested, the most efficient was chosen for subsequent experiments. Homozygous frame-shift mutations in the genomic region targeted by the CRISPR-Cas9 construct were confirmed by sequencing using the following sequencing primers: Zeb1 (CAAAACAACCATCAGGCTCA/ TTCCATTCCCCTGCTAATTG) Snail (ATTGAGAATCGGCCCCACC/ CCAAACGTTTCCAGAGAGCC). Clones were analyzed for CRISPR/Cas9 levels (determined by FLAG expression) and similar tumorigenic potential. To exclude sub-clone specific results clones were analyzed first individually and then 3 were pooled in equal ratios (Supp. Fig. 2 C and D). In similar fashion over-expression cell clones were generated in M-SCC cells using pLVX Zeb1.

Quantitative Real-Time-PCR and Primers and RNA-Seq

Q-PCR and RNA-Seq. were performed as described before (7). In brief total RNA was isolated using TRIzol followed by RNeasy Micro Kits (Quiagen). Reverse transcription was performed with cDNA Synthesis Kits, including RNAse Inhibitor (Applied Biosystems). Quantitative RT-PCR was performed using a Roche Diagnostics LightCycler 480 II and SYBR Green Mastermix (Roche). Primers used for analysis: Krt5 (CAGCCGGAGCCTCTACAACC/ CGCCGAAACCAAATCCACTAC), Krt8 (GATGCTGGAGACCAAGTGGAG/ GCCGCCTAAGGTTGTTGATG); E-Cad (TTGCACCGGTCGACAAAGGAC/ TGGATTCCAGAAACGGAGGCC), Zeb1 (GATGATGAATGCGAGTCAGATGC/ ACAGCAGTGTCTTGTTGTTGT), Slug (TACCGCTGCTCCATTCCACG/ CATGGGGGTCTGAAAGCTTGG), Snail (TCGGAAGCCTAACTACAGCGA/ AGATGAGCATTGGCAGCGAG), GAPDH (GTCTCCTCTGACTTCAACAGCG/ ACCACCCTGTTGCTGTAGCCAA).

Nuclear fractionation and western blotting

Western blotting was performed as follows: Cells were washed in PBS and total protein was extracted in RIPA buffer (Invitrogen) supplemented with Phosphatase Inhibitors (mixture 2 and 3, Sigma) and Complete Protease Inhibitors (Roche) for 30 min on ice. For nuclear fractionations cells were scraped off in ice-cold PBS and pelleted at 1000 g for 3 min. Cells were re-suspended by pipetting and lysed in Nuclear lysis buffer (5 mM PIPES (pH8.0), 85 mM KCl, 0.5% NP40, supplemented with protease and phosphatase inhibitors as stated above) for 10 min on ice. Lysates were microfuged at 2300 g for 5 min at 4°C to pellet nuclei. Supernatant was collected as cytoplasmic fraction. Nuclear pellets were lysed using RIPA buffer as stated above. All protein lysates were microfuged at 13,000 g for 30 min at 4°C before total protein was determined by the BioRad protein quantification kit. Samples were prepared and western blot performed according to manufacturer's

instructions (Thermo Fisher Scientific). Separation of total protein extracts was carried out in 1xMOPS buffer using NuPAGE Novex 4-12% Bis-Tris gels. Proteins were electrotransferred to PVDF membrane by wet blotting in NuPAGE Transfer buffer. Blocking and antibody incubations were performed as described by CST. Secondary antibodies (CST) were used at 1:10,000 dilution detected with Pierce Femto or Dura ECL (Thermo Fisher Scientific) as substrate

Immunofluorescence and histology analysis
Immunofluorescence analysis on cells and tissues were performed as described before (23, 24). Sterilized coverslips were placed directly into culture dishes. For keratin staining, cells were fixed for 5 min in -20°C methanol, 30 s in -20°C acetone, otherwise cells were fixed for 15 min in 4% formalin/PBS at 4°C, washed in TBS, and then permeabilized for 5 min in 0.25% Triton/TBS. After a TBS wash, they were incubated for 1 h or overnight at 4°C with primary antibodies diluted in TBS/1% BSA. Secondary antibodies at 1:400 dilutions were applied and incubated for 30–60 min. Nuclei were counterstained using 1:1,000 diluted DAPI. Coverslips were mounted with Prolong Gold (Invitrogen). For histology, tissues were fixed overnight at 4°C in 4% formalin, sequentially incubated at 4°C overnight in 15% and 30% sucrose, embedded in paraffin, and sectioned (5 μm). Antigens were retrieved with citrate buffer (18 mM citric acid monohydrate, 82 mM sodium citrate, pH 6.0) and the antibodies were diluted in TBS containing 1% BSA and 5% NGS. For Zeb1 staining on tissues an amplification step using the TSA Plus Kit (PerkinElmer) was performed according to manufactures instructions. Slides were mounted in ProLong Gold antifade reagent (Invitrogen).

Immunofluorescence microscopy and data processing. H&E images were acquired using an Axioskop microscope with a Plan Neofluar 10x /0.30 Ph1 objective using an AxioCam 105 color (Zeiss) and fluorescent images were obtained using an Observer.Z1with an ApoTome2 (Zeiss) microscope with Apochromat 20x/ 0.8 M27 or 63x /1.4 oil immersion objectives at room temperature using an AxioCam506 (Zeiss). Image analysis and processing were performed using the Zen2 (Zeiss) and Adobe PhotoShop CC2017 software. LUT (brightness and gamma) was adjusted using Adobe Photoshop.

Antibodies
CD44-PECy7 (Biolegends) FACS (F) 1:500; CD44-BV785 (Affymetrix) F 1:200; CD104-efluor660 (Affymetrix) F 1:200; CD104 (Novus) WB 1:500; CD24-PECy7 (Affymetrix) F 1:50; E-Cad (CST) IF 1:200, WB 1:1000; E-Cad (ECCD-2, SantaCruz) IF 1:50; Zeb1(CST) WB 1:1000; Vim (GP59, Progen) IF 1:50: Vim (CST) WB 1:1000; pan-Krt-AL488 (eBioscience) IF 1:50; Krt5 (AF138, BioLegend) WB 1:1000; Krt5 (GP-CK5, Progen) IF 1:50; Krt8 (BioLegend) IF 1:50, WB 1:1000, Krt8/18 (Sigma) IF 1:50; Zeb1 (H10, SantaCruz) IF 1:50; Snail (CST) IF 1:50, WB 1:1000, Ras (SantaCruz) WB 1:1000; CoxIV (CST) WB 1:1000; GAPDH (CST) WB 1:1000; N-Cad (CST) WB 1:1000; Slug (CST) WB 1:1000; Twist (BD Transduction Laboratories) WB 1:1000; FN (BD Transduction Laboratories) WB 1:5000; Flag (Sigma) WB 1:1000; SV40 LgT (Pab101,SantaCruz) IF (1:200), WB 1:1000; SV40 LgT (v-300, SantaCruz) IF (1:50);

Wnt3a (CST) WB 1:1000; Wnt7a (GeneTex) WB 1:1000; Wnt7b (GeneTex) WB 1:1000; Wnt5a (CST) WB 1:1000; β-cat (BD Transduction Laboratories) WB 1:1000; P-β-cat S33/37/T41 (CST) WB 1:1000; GSK3β (CST) WB 1:1000; P- GSK3β S9 (CST) WB 1:1000; TCF1 (CST) WB 1:1000; Lamin A/C (CST) WB 1:1000; MET (CST) WB 1:1000; JNK (CST) WB 1:1000; P-JNK T183/Y185 (CST) WB 1:1000; JunC (CST) WB 1:1000; P-JunC S63(CST) WB 1:1000; PKCα (CST) WB 1:1000; P- PKCα T497 (CST) WB 1:1000; PLCγ1 (CST) WB 1:1000; P- PLCγ1 Y1217 (CST) WB 1:1000; LRP6 (CST) WB 1:1000; P- LRP6 S1490 (CST) WB 1:1000; Rac1 (BD Tranduction Laboratories) WB 1:1000; Alexa secondary antibodies for IF (Jackson ImmunoResearch) 1:400.

Proliferation

Proliferation assays were performed in triplicate with 1000 cells/ well in 96-well plates using CyQuant (ThermoFisher) according to manufactures instructions over the course of 4 d.

Primer

Zeb1 (CRISPR sequencing)	(CAAAACAACCATCAGGCTCA/ TTCCATTCCCCTGCTAATTG)
Snail (CRISPR sequencing)	(ATTGAGAATCGGCCCCACC/ CCAAACGTTTCCAGAGAGCC)
Krt5	(CAGCCGGAGCCTCTACAACC/ CGCCGAAACCAAATCCACTAC)
Krt8	(GATGCTGGAGACCAAGTGGAG/ GCCGCCTAAGGTTGTTGATG)
E-Cad	(TTGCACCGGTCGACAAAGGAC/ TGGATTCCAGAAACGGAGGCC)
Zeb1	(GATGATGAATGCGAGTCAGATGC/ ACAGCAGTGTCTTGTTGTTGT)
Slug	(TACCGCTGCTCCATTCCACG/ CATGGGGGTCTGAAAGCTTGG)
Snail	(TCGGAAGCCTAACTACAGCGA/ AGATGAGCATTGGCAGCGAG)
GAPDH	(GTCTCCTCTGACTTCAACAGCG/ ACCACCCTGTTGCTGTAGCCAA)

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. A. FACS profiles for CD24/ CD44 and CD104/CD44 of HMLER 44L and 44H cell populations. **B.** FACS profiles for CD104 and CD44 of parental HMLER E and M cell populations. **C.** Differences in tumor-initiating ability of E, E/M, M and xM cells were assessed upon transplantation with limiting dilution into NOD/SCID mice. Data are presented as mean ± SEM. **D.** Cell proliferation assay of E, E/M and xM cells monitoring population growth over 4 d. Data are presented as mean ± SEM. **E.** Analysis of E, E/M and xM tumor sections using IF staining for Krt5/E-cad or Vim/αSma. LgT staining was used to differentiate tumor cells from mouse stromal cells. Nucleus is visualized by DAPI staining. **F.** FACS analysis for CD104, CD44 marker expression of dissociated neoplastic cells of E and M tumors. **G.** FACS histogram analysis for CD104 of HMLER E, E/M and xM tumor cell populations. Bars: Tissue IF 2 μm

Supplementary Figure 2. A. FACS profiles for CD104 and CD44 marker expression of SSCs of xM cells were assessed for plasticity, monitoring their respective ability, following passage in culture for a period of 3 weeks, to generate progeny that had developed increased CD104 expression, the latter indicating entrance into the E/M state. **B.** Western blot analyses of Zeb1 and CRISPR *ZEB1* KO M cells to gauge construct efficiency including the knockout of *ZEB1* expression. **C.** Sequencing results of the CRISPR targeted region of *ZEB1* sgRNA for different epithelial ZEB1 KO clones (E-SCC-ZEB1KO #1,2 and 4). **D.** Western blot characterization of E-SCC, E-SCCZeb1KO clones and controls. Flag expression is representative of Cas9 expression. **E.** Tumor growth ability monitored by tumor weight of E-SCC, E-SCCZeb1KO clones and controls. Data are presented as mean ± SEM. **F.** Western blot characterization of M-SCC, xM-SCCZeb1 clones and controls to identify high Zeb1 expressing clones "locked" in the xM state. **G.** Tumor growth ability monitored by tumor weight of M-SCC, xM-SCCZeb1 clones and controls. Data are presented as mean ± SEM. **H.** Analysis E, E-SCC-Zeb1KO, M and xM-SCC-Zeb1 tumor sections using IF staining for Krt5/Zeb1. LgT staining was used to differentiate tumor cells from mouse stromal cells. Nucleus is visualized by DAPI staining. Bars: Tissue IF 2 μm

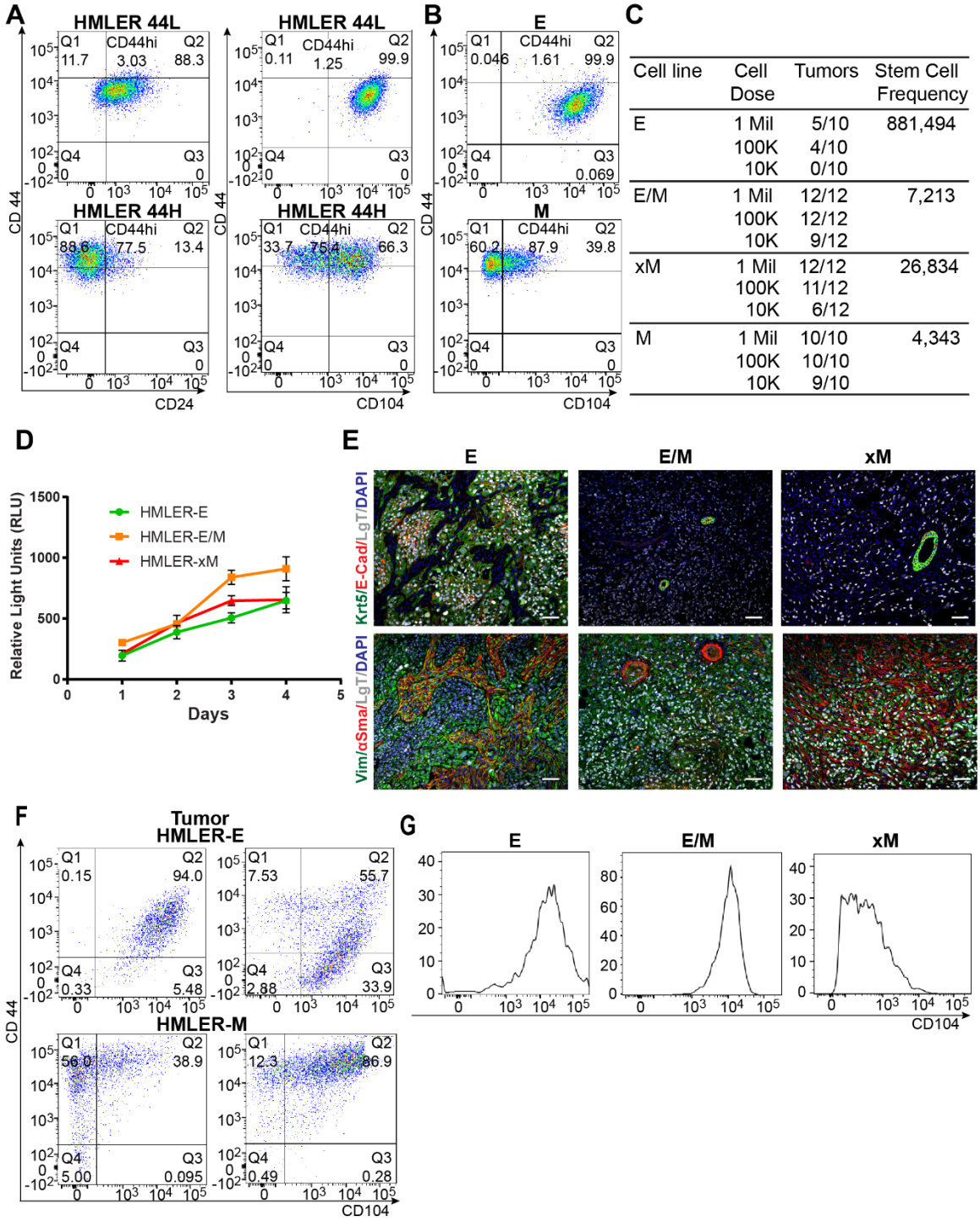
Supplementary Figure 3. A. Representative images of Zeb1 staining in E-SCC-Sn and E-SCC-Zeb1KO clones. Nuclei are visualized by DAPI. **B.** FACS profiles for CD104 and CD44 E-SCC-Sn and E-SCC-Zeb1KO clones. Green populations represent the non-targeted cells and red populations the Snail, tdTom expressing cells. **C.** FACS profiles for CD104 and CD44 of M and M-Zeb1KO cells, as well as M-Zeb1KO cells post FACS sorting (ps.) before injections into NOD/SCID mice and expression of dissociated neoplastic cells of M-Zeb1KO tumors (tum.). **D.** Representative images of E-cad staining in E, M and M-Zeb1KO cells. Nuclei are visualized by DAPI. Bars: IF 1 μm.

Supplementary Figure 4. A. FACS profiles for CD104 and CD44 of SUM149-SCC and SUM149-SCC-Zeb1KO cells treated with TGF-β. **B.** Differences in tumor-initiating ability of M-SCC, M-SCC-Zeb1 and M-SCC-SnKO cells upon transplantation at limiting dilutions into NOD/SCID mice. **C.** FACS profiles for CD104 and CD44 of E-SCC, E-SCCSn, E-SCC-Zeb1KO and E-SCC-Zeb1KOSn cell populations.

Supplementary Figure 5. A. FACS profiles for CD104 and CD44 of E, E-SCC, E-SCC-Zeb1KO and E-SCC-SnKO cells treated with TGF-β1 or with forced expression of the EMT-TFs Slug, Twist or Zeb1mt (the CRISPR KO-resistant mutant form of the *ZEB1* gene). **B.** Western blot analysis of E, E-SCC, E-SCC-Zeb1KO and E-SCC-SnKO treated with TGF-β1 and untreated. **C.** Western blot analysis of E, E-SCC, E-SCC-Zeb1KO and E-SCC-SnKO with forced Zeb1 and Zeb1mt expression.

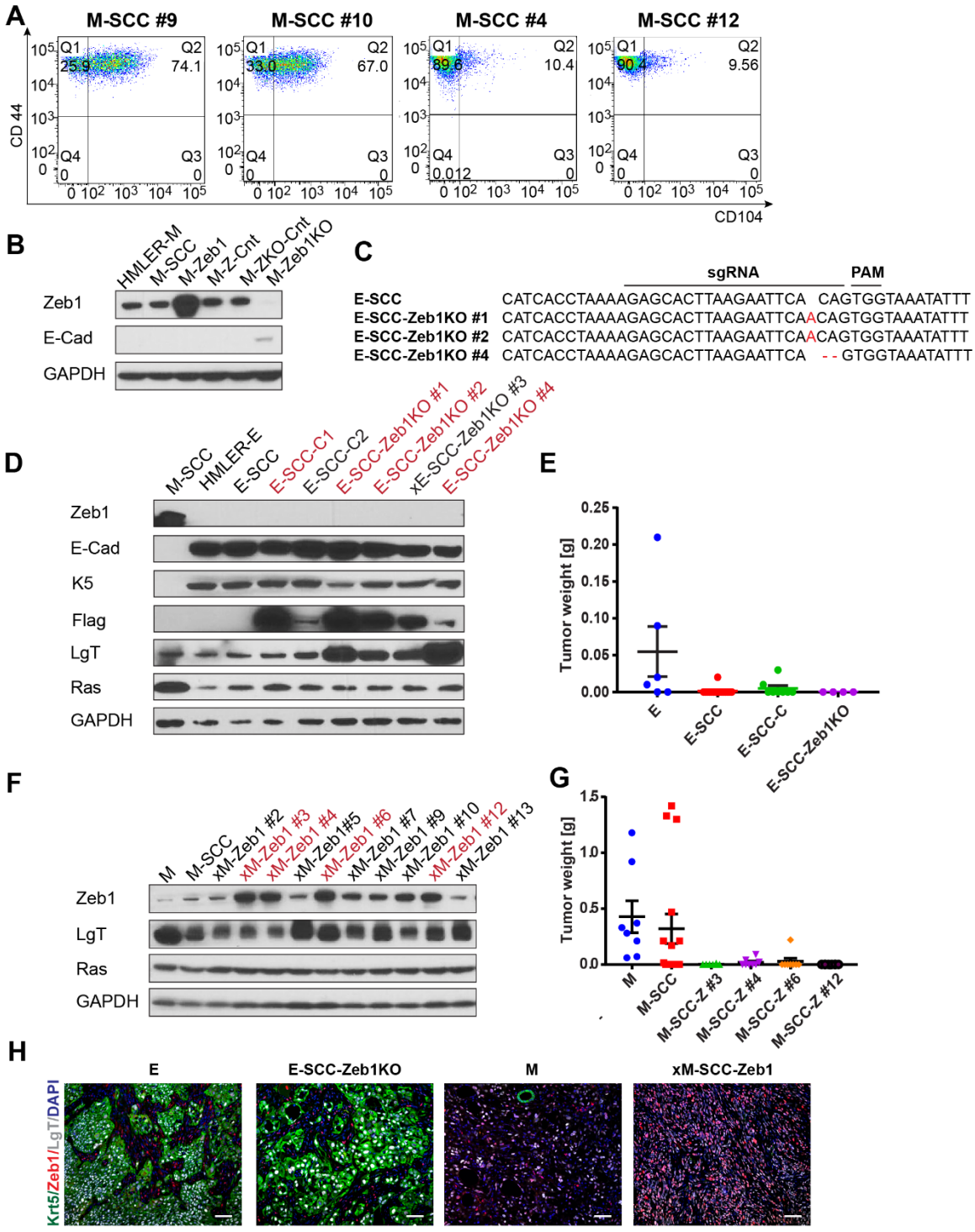
Supplementary Figure 6. A.
A. Transcriptome analysis of E, E/M and xM cells. Hierarchical clustering of differentially expressed genes in E, E/M, and xM states. Gene expression is converted into z-score by row for all samples and averaged for each state, and individual genes are clustered together based on uncentered correlation using Cluster3. **B.** Significantly changing genes in the Wnt ligand and cognate receptor families, across E, E/M, and xM states in HMLER and SUM159 cells (data from ref. 7) with a focus on the EMT-TFs, as well as the canonical and non-canonical Wnt ligands and receptors. Higher expression in red, lower expression in blue. **C.** Western blot analysis of Wnt3a and Wnt5a overexpression and control construct in E cells. **D.** Western blot analysis of E, M, E/M and xM 7-TGP/Wnt3A overexpressing cells for Wnt3a expression. **E.** FACS profiles for CD104 and CD44 of M-7-TGP/Wnt3A expressing cells. **F.** GFP FACS histogram of E ($CD104^{+}CD44^{low}$), E/M ($CD104^{+}CD44^{hi}$) and xM ($CD104^{-}CD44^{hi}$) cell fractions of the total M-7-TGP/Wnt3A cell population analyzed for TCF-GFP activation/canonical-Wnt-signaling activation and quantification of the median GFP expression in the different E, E/M and xM fractions. **G.** GFP FACS histogram of E-7-TGP/Wnt3A cells treated with the Bio compound for 0 d (red), 3 d (blue) and 11 d (orange) to induce TCF-GFP activation/canonical-Wnt-signaling activation. The corresponding FACS profiles for CD104 and CD44 marker analysis demonstrate that cells undergo an EMT upon GSK3β blockage. **H.** FACS profiles for CD104 and CD44 of E/M and M cells as well as WNT5a overexpressing E/M and M show no influence of Wnt5a on CD104/CD44 marker expression.

SUPPLEMENTARY FIGURES

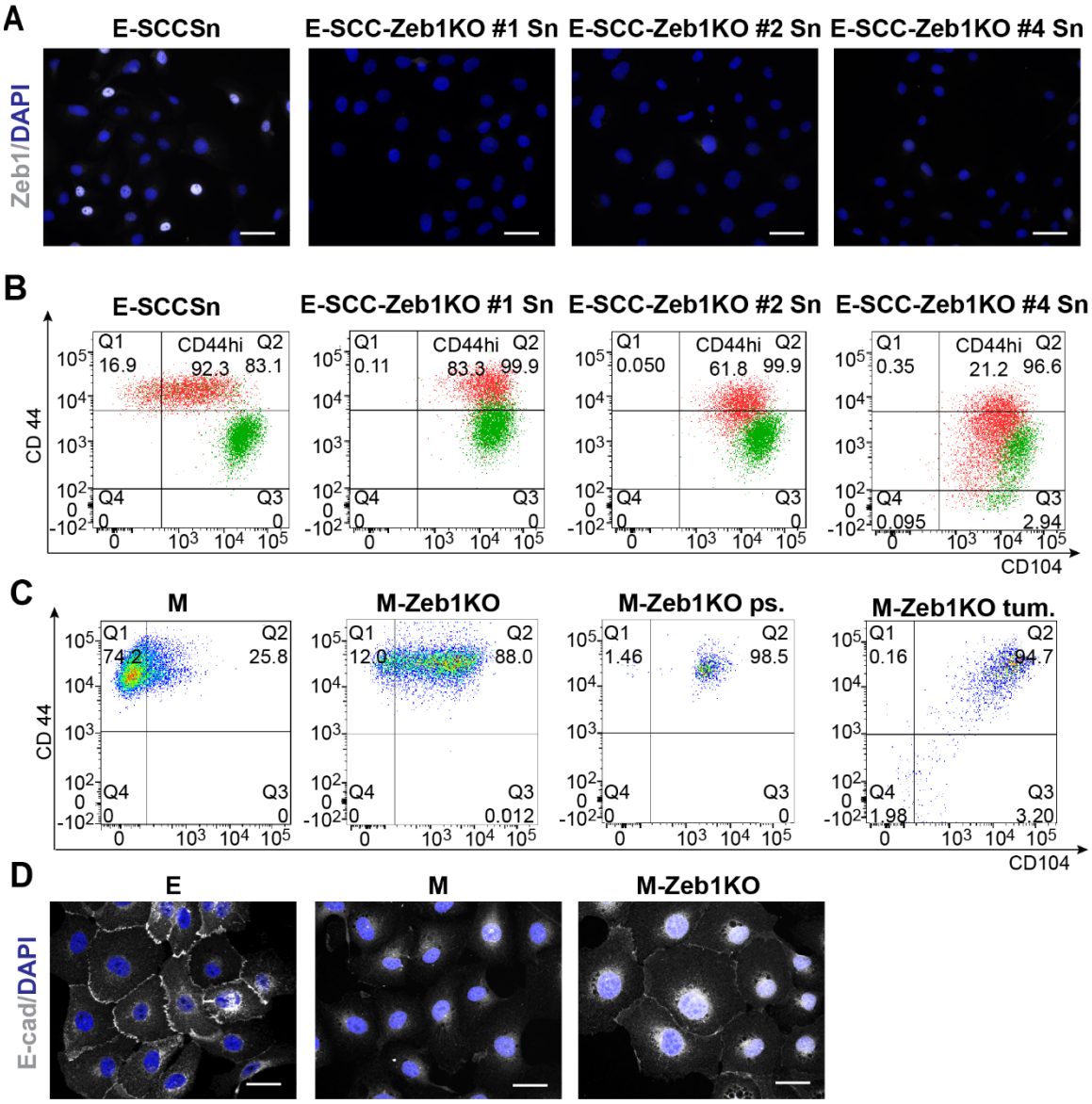


Cell line	Cell Dose	Tumors	Stem Cell Frequency
E	1 Mil	5/10	881,494
	100K	4/10	
	10K	0/10	
E/M	1 Mil	12/12	7,213
	100K	12/12	
	10K	9/12	
xM	1 Mil	12/12	26,834
	100K	11/12	
	10K	6/12	
M	1 Mil	10/10	4,343
	100K	10/10	
	10K	9/10	

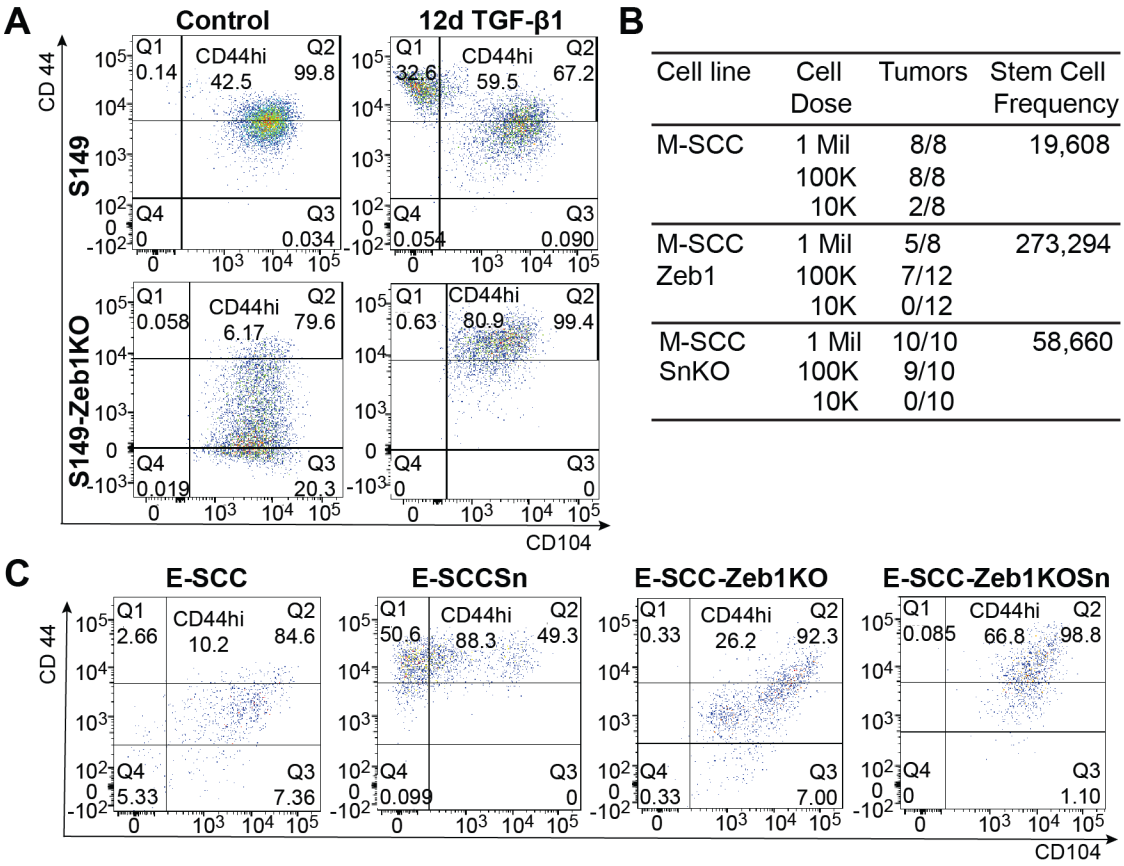
Supplementary Figure 1.



Supplementary Figure 2.

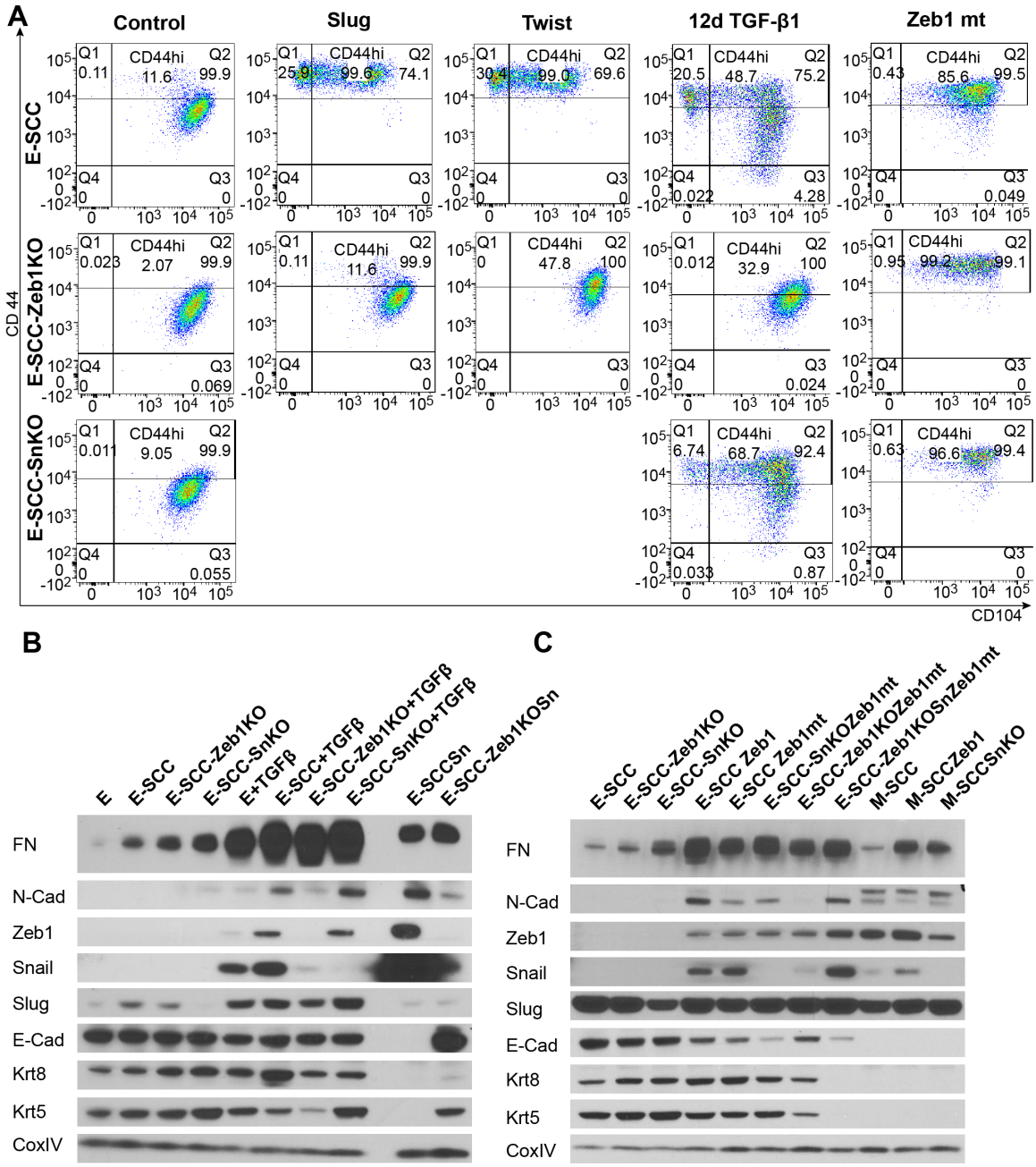


Supplementary Figure 3.

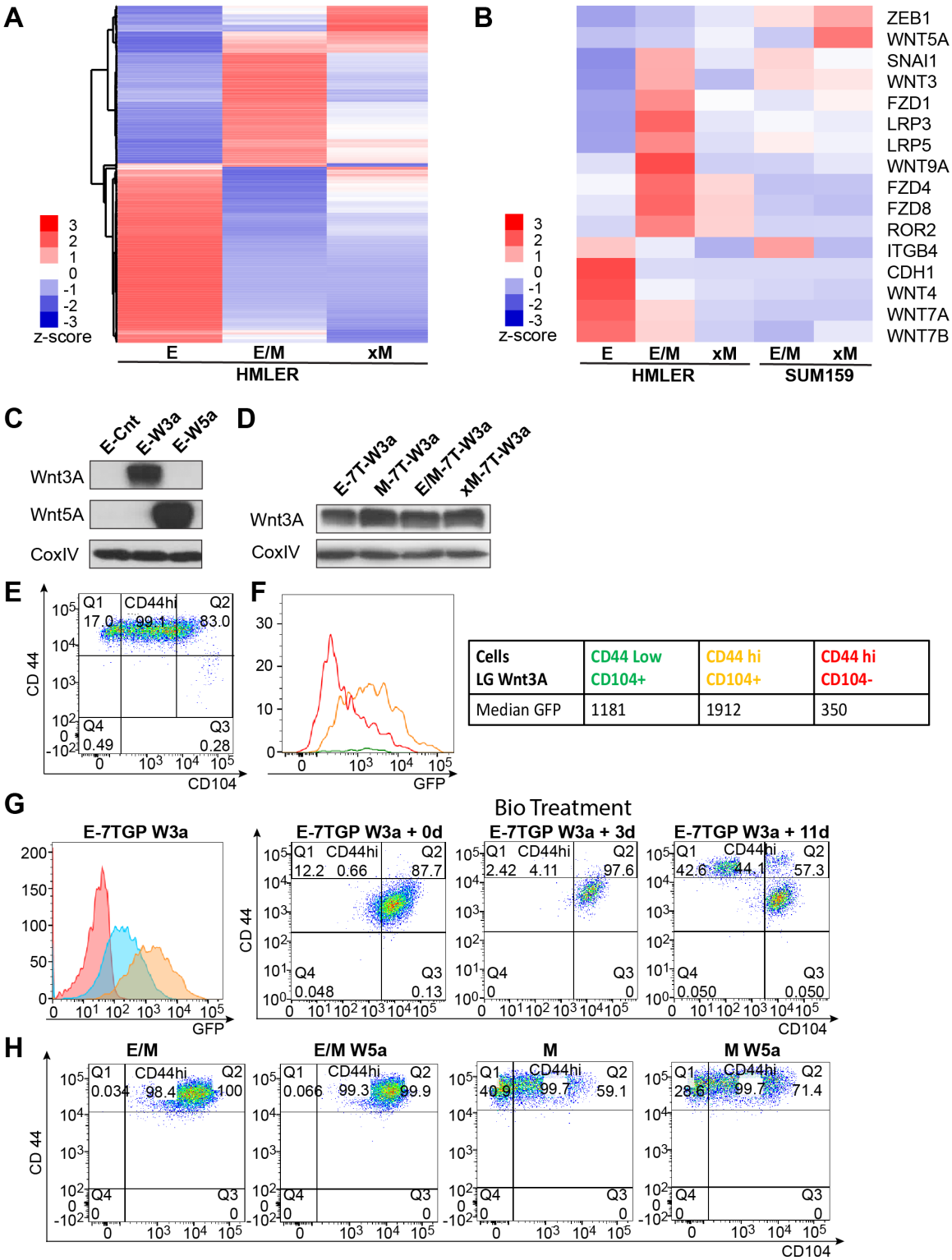


Cell line	Cell Dose	Tumors	Stem Cell Frequency
M-SCC	1 Mil	8/8	19,608
	100K	8/8	
	10K	2/8	
M-SCC Zeb1	1 Mil	5/8	273,294
	100K	7/12	
	10K	0/12	
M-SCC SnKO	1 Mil	10/10	58,660
	100K	9/10	
	10K	0/10	

Supplementary Figure 4.



Supplementary Figure 5.



Supplementary Figure 6.

SUPPLEMENTARY REFERENCES

1. Elenbaas B*, et al.* (2001) Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. *Genes Dev* 15(1):50-65.
2. Hahn WC*, et al.* (2002) Enumeration of the simian virus 40 early region elements necessary for human cell transformation. *Mol Cell Biol* 22(7):2111-2123.
3. Lam AJ*, et al.* (2012) Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat Methods* 9(10):1005-1012.
4. Shaner NC*, et al.* (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat Biotechnol* 22(12):1567-1572.
5. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, & Clarke MF (2003) Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* 100(7):3983-3988.
6. Chaffer CL*, et al.* (2011) Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proc Natl Acad Sci U S A* 108(19):7950-7955.
7. Bierie B*, et al.* (2017) Integrin-beta4 identifies cancer stem cell-enriched populations of partially mesenchymal carcinoma cells. *Proc Natl Acad Sci U S A* 114(12):E2337-E2346.
8. Chaffer CL & Weinberg RA (2011) A perspective on cancer cell metastasis. *Science* 331(6024):1559-1564.
9. Mani SA*, et al.* (2008) The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133(4):704-715.
10. Morel AP*, et al.* (2008) Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PloS one* 3(8):e2888.
11. Fuerer C & Nusse R (2010) Lentiviral vectors to probe and manipulate the Wnt signaling pathway. *PloS one* 5(2):e9370.
12. Scheel C*, et al.* (2011) Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell* 145(6):926-940.
13. Stemmer V, de Craene B, Berx G, & Behrens J (2008) Snail promotes Wnt target gene expression and interacts with beta-catenin. *Oncogene* 27(37):5075-5080.
14. Zhou BP*, et al.* (2004) Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition. *Nature cell biology* 6(10):931-940.
15. Sato N, Meijer L, Skaltsounis L, Greengard P, & Brivanlou AH (2004) Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nature medicine* 10(1):55-63.
16. Caspi M, Zilberberg A, Eldar-Finkelman H, & Rosin-Arbesfeld R (2008) Nuclear GSK-3beta inhibits the canonical Wnt signalling pathway in a beta-catenin phosphorylation-independent manner. *Oncogene* 27(25):3546-3555.
17. Stamos JL & Weis WI (2013) The beta-catenin destruction complex. *Cold Spring Harb Perspect Biol* 5(1):a007898.
18. Tam WL*, et al.* (2013) Protein kinase C alpha is a central signaling node and therapeutic target for breast cancer stem cells. *Cancer cell* 24(3):347-364.
19. Shibue T, Brooks MW, Inan MF, Reinhardt F, & Weinberg RA (2012) The outgrowth of micrometastases is enabled by the formation of filopodium-like protrusions. *Cancer discovery* 2(8):706-721.
20. Sanjana NE, Shalem O, & Zhang F (2014) Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 11(8):783-784.

21. Smallwood PM, Williams J, Xu Q, Leahy DJ, & Nathans J (2007) Mutational analysis of Norrin-Frizzled4 recognition. *J Biol Chem* 282(6):4057-4068.
22. MacDonald BT*, et al.* (2014) Disulfide bond requirements for active Wnt ligands. *J Biol Chem* 289(26):18122-18136.
23. Kroger C*, et al.* (2013) Keratins control intercellular adhesion involving PKC-alpha-mediated desmoplakin phosphorylation. *J Cell Biol* 201(5):681-692.
24. Kroger C*, et al.* (2011) Placental vasculogenesis is regulated by keratin-mediated hyperoxia in murine decidual tissues. *Am J Pathol* 178(4):1578-1590.