Cell Reports, Volume *38*

Supplemental information

Coordinate control of basal epithelial cell fate and stem cell maintenance by core EMT transcription factor Zeb1

Yingying Han, Alvaro Villarreal-Ponce, Guadalupe Gutierrez Jr., Quy Nguyen, Peng Sun, Ting Wu, Benjamin Sui, Geert Berx, Thomas Brabletz, Kai Kessenbrock, Yi Arial Zeng, Kazuhide Watanabe, and Xing Dai

SUPPLEMENTARY FIGURES

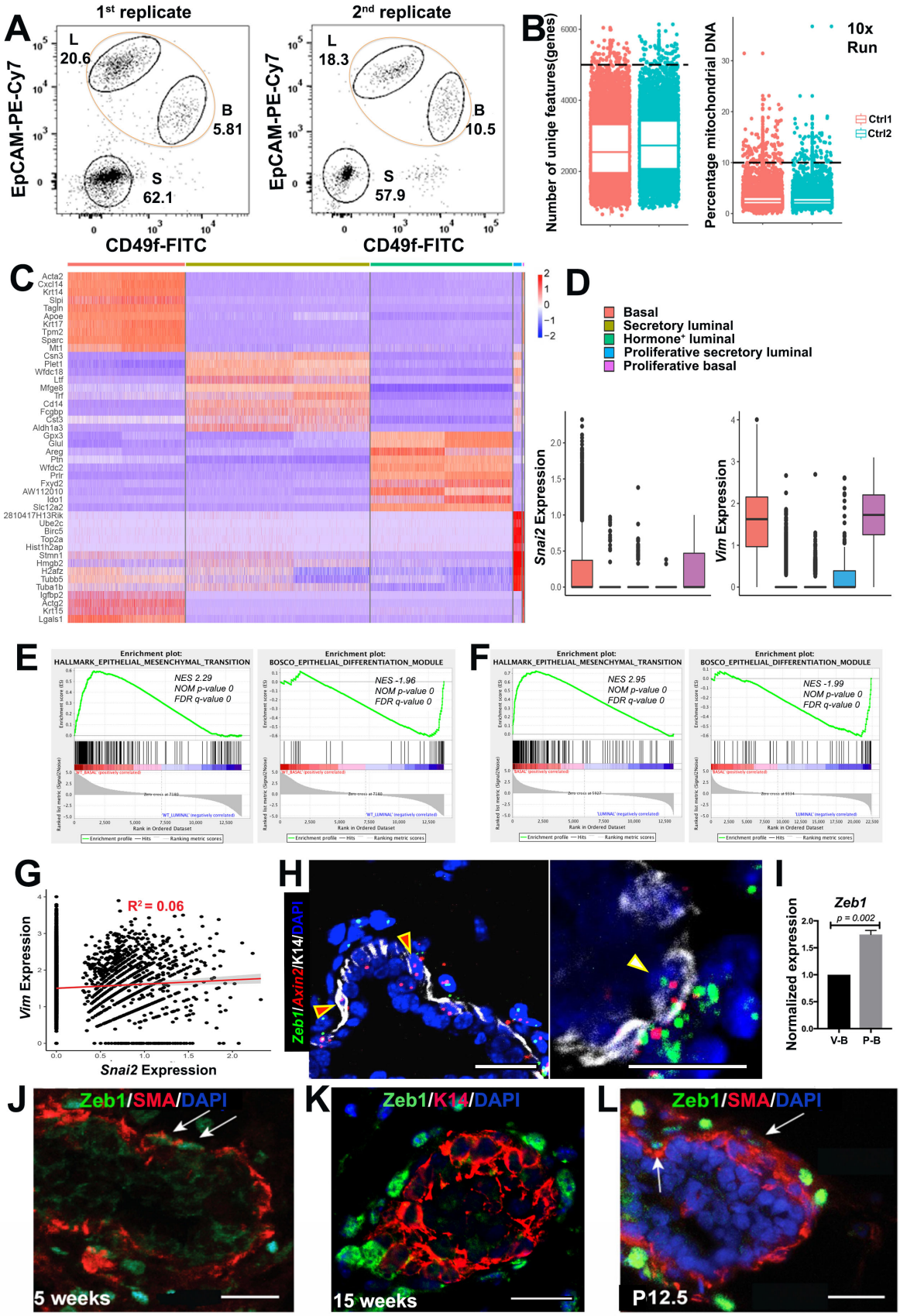

Figure S1. Related to Figure 1.

A. FACS profile and gating (orange circles) strategy used to sort basal (B) and luminal (L) epithelial cells from Lin/EpCAM/CD49f-stained mammary cell populations for scRNA-seq analysis. Mammary stromal cells (S) were excluded by gating.

B. Boxplots denoting number of unique features and total percent of mitochondrial DNA captured per cell per run. Dashed line indicates threshold cutoff per feature per cell.

C. Heatmap depicting the top 10 differentially-expressed marker genes per epithelial cluster (color coded across the top to match cluster colors in Figure 1A). See Table S1 for a complete list of marker genes for each cluster.

D. Boxplots of *Snai2* and *Vim* expression across the epithelial clusters.

E. GSEA of RNA-seq data collected from Lin/CD24/CD29-sorted basal (left) and luminal (right) MECs using Hallmark EMT gene signature and Bosco_Epithelial_Differentiation_Module (Bosco et al., 2010; Liberzon et al., 2015) (see Table S3 for a complete list of genes). NES, nominal enrichment score. NOM, nominal p-value. FDR, false discovery rate.

F. GSEA of microarray data collected from Lin/CD24/CD29-sorted basal (left) and luminal (right) MECs using Hallmark EMT gene signature and Bosco_Epithelial_Differentiation_Module. See E above.

G. Scatter plot showing lack of correlation (Pearson's) between *Snai2* and *Vim* expression in basal MECs.

H. RNAScope images from co-analysis of *Zeb1* and *Axin2* mRNAs in MGs from 15-week-old mice. K14 antibody highlights basal MECs. DAPI stains the cell nuclei. Arrowheads point to *Zeb21/Axin2* double-positive basal cells. Scale bar: 50 μm.

I. RT-qPCR analysis of *Zeb1* expression in $Lin^{-}CD49f^{high}EpCAM^{+}$ basal (B) MECs from age-matched virgin (V, 12-week-old) and pregnant (P13.5) mice. n=3 each.

J-L. Indirect immunofluorescence staining fails to detect robust Zeb1 protein expression in basal MECs (marked by SMA or K14 protein expression) of 5-week-old (J), 10-week-old (K), and P12.5 pregnant (L) mice. Scale bar: 50 μm.

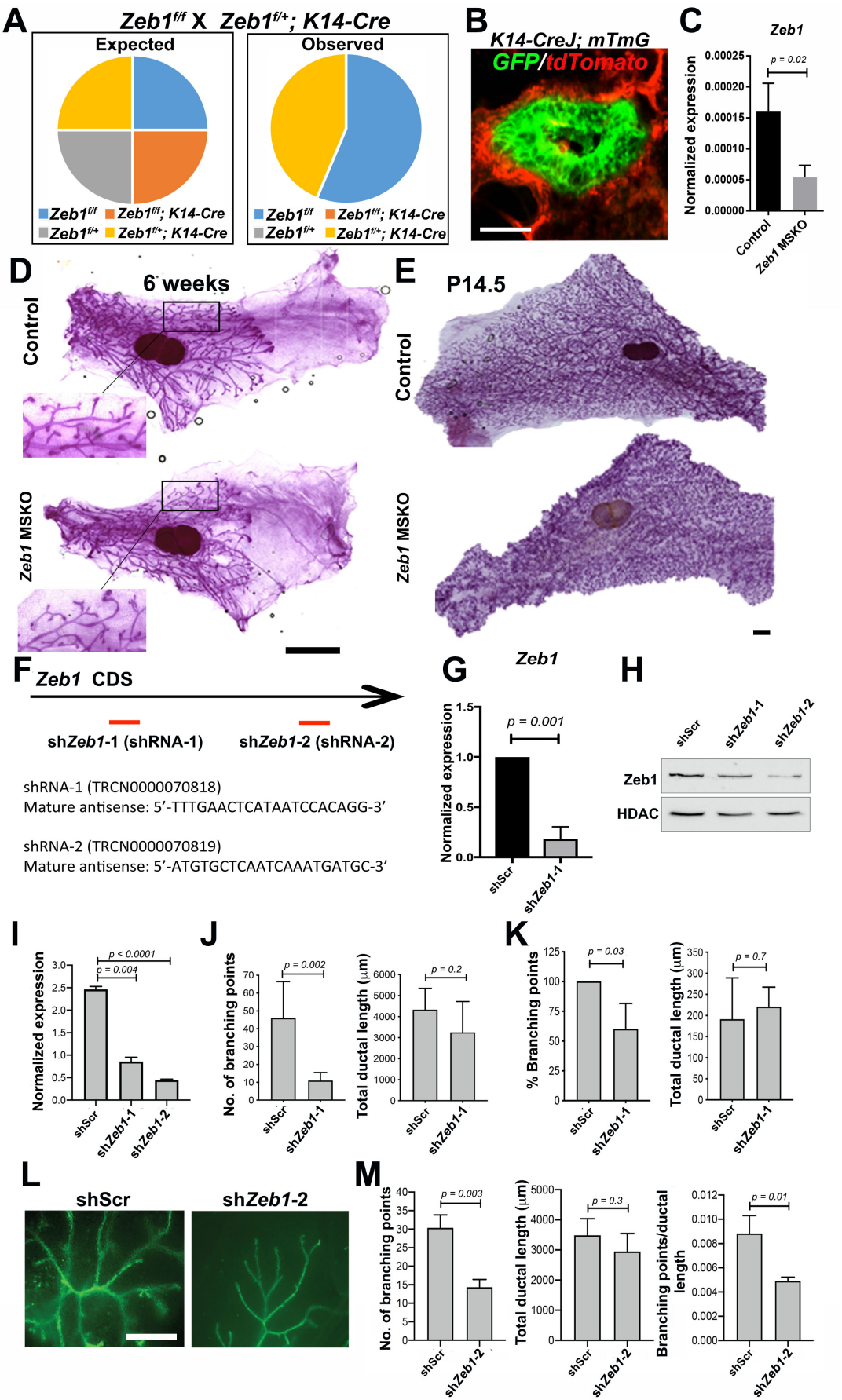
A
Zeb1f/f X Zeb1f/+; K14-Cre
Expected
Observed
Zeb1f/f
Zeb1f/f; K14-Cre
Zeb1f/+
Zeb1f/+; K14-Cre
B
K14-CreJ; mTmG
GFP/tdTomato
C
Zeb1
p = 0.02
Normalized expression
0.00025
0.00020
0.00015
0.00010
0.00005
0.00000
Control
Zeb1 MSKO
D
6 weeks
Control
Zeb1 MSKO
E
P14.5
Control
Zeb1 MSKO
F
Zeb1 CDS
shZeb1-1 (shRNA-1)
shZeb1-2 (shRNA-2)
shRNA-1 (TRCN0000070818)
Mature antisense: 5'-TTTGAACTCATAATCCACAGG-3'
shRNA-2 (TRCN0000070819)
Mature antisense: 5'-ATGTGCTCAATCAAATGATGC-3'
G
Zeb1
p = 0.001
Normalized expression
1.5
1.0
0.5
0.0
shScr
shZeb1-1
H
shScr
shZeb1-1
shZeb1-2
Zeb1
HDAC
I
p < 0.0001
p = 0.004
Normalized expression
3.0
2.5
2.0
1.5
1.0
0.5
0.0
shScr
shZeb1-1
shZeb1-2
J
p = 0.002
No. of branching points
70
60
50
40
30
20
10
0
shScr
shZeb1-1
p = 0.2
Total ductal length (μm)
6000
5000
4000
3000
2000
1000
0
shScr
shZeb1-1
K
p = 0.03
% Branching points
125
100
75
50
25
0
shScr
shZeb1-1
p = 0.7
Total ductal length (μm)
350
300
250
200
150
100
50
0
shScr
shZeb1-1
L
shScr
shZeb1-2
M
p = 0.003
No. of branching points
40
35
30
25
20
15
10
5
0
shScr
shZeb1-2
p = 0.3
Total ductal length (μm)
5000
4000
3000
2000
1000
0
shScr
shZeb1-2
p = 0.01
Branching points/ductal length
0.012
0.010
0.008
0.006
0.004
0.002
0.000
shScr
shZeb1-2

Figure S2. Related to Figure 2.

A. Initial attempt to ablate *Zeb1* with our previously used *K14-Cre* line (Andl et al., 2004; Gu et al., 2009, 2013; Watanabe et al., 2014) failed to produce MSKO offspring at the expected Mendelian ratios. Shown are the expected and observed genotypes of offspring from *K14-Cre;Zeb1*$^{f/+}$ X *Zeb1*$^{f/f}$ crosses. The significant deviation in genotyping distribution suggests this particular *K14-Cre* transgene may be integrated near or at the *Zeb1* locus.

B. The *K14-Cre* line (Jackson Laboratory) used in this study directs efficient and specific recombination in MECs as previously reported (Cai et al., 2017), evident through robust GFP expression in the epithelium and tdTomato expression in the stroma of *K14-Cre*;*mTmG* mice. Scale bar: 100 μm.

C. RT-qPCR analysis of *Zeb1* expression in basal MECs from 15-week-old control and MSKO mice.

D. Whole-mount carmine staining of MGs from control and *Zeb1* MSKO mice at 6-weeks of age. Boxed areas are shown at higher magnification. Scale bar: 5 mm.

E. Whole-mount carmine staining of MGs from control and *Zeb1* MSKO mice during pregnancy (P14.5). Scale bar: 2 mm.

F. Diagram showing the positions and sequences of *shZeb1*-1 and *shZeb1*-2. CDS, coding sequence.

G. RT-qPCR analysis of *Zeb1* expression in control (shScr) and *Zeb1* knockdown (sh*Zeb1*-1) MECs. Data were obtained from 3 independent experiments and are represented as the mean ± SD.

H. Western blotting for Zeb1 protein in nuclear extracts from 3T3 mouse fibroblasts to test shRNA-mediated knockdown efficiency. HDAC was used as a loading control.

I. Quantification of Zeb1 protein expression in (G) (n=3). Data are represented as the mean ± SD. One-way ANOVA with multiple comparisons was used to analyze statistical significance between groups.

J-M. Additional measurements of cleared fat pad transplantation of basal MECs infected with lentiviruses expressing either sh*Zeb1*-1 (J-K) or sh*Zeb1*-2 (L-M). Panels (J), (L), and (M) describe data from virgin hosts, while (K) is from pregnant hosts. (L) shows representative images of data in (M). n=5 for (J), n=3 for (K) and (M). Data are represented as the mean ± SD. Scale bar: 1 mm in (L).

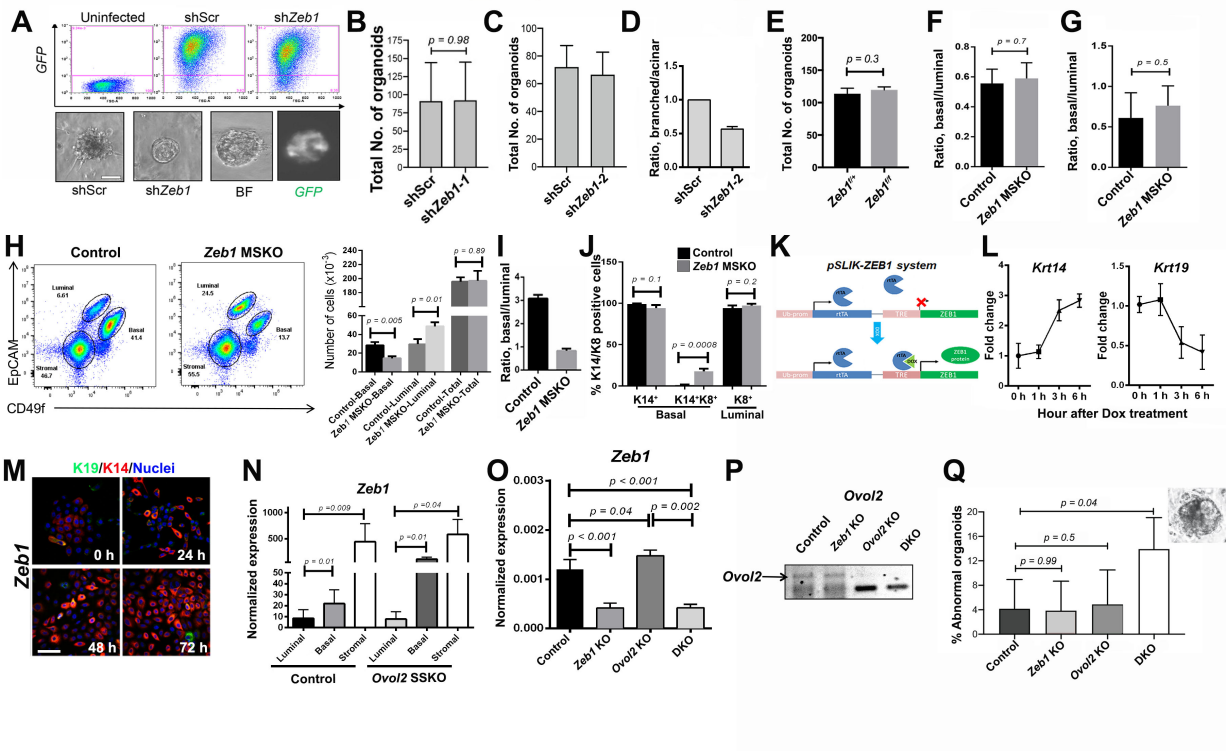

Figure S3. Related to Figure 3.

A. Top: FACS plots of basal MECs infected with lentiviruses expressing GFP and either sh*Zeb*-1 or shScr. Bottom: bright-field (BF) and fluorescent images of GFP-positive organoids produced by control and *Zeb1*-depleted MECs. Scale bar: 50 μm.

B. Total number of organoids produced after *Zeb1* knockdown with sh*Zeb1*-1 (F) (n=3).

C. Total number of organoids produced after *Zeb1* knockdown with sh*Zeb1*-2 (n=2).

D. Ratio of branched to acinar organoids after *Zeb1* knockdown with sh*Zeb1*-2 (n=2). For (B-D), only GFP-positive organoids were quantified, and data are represented as the mean ± SD.

E. Total number of organoids from control and *Zeb1* knockout cultures (n=3 pairs).

F. Ratio between basal and luminal MECs from 6-week-old virgin control and *Zeb1* MSKO mice (n=3 pairs).

G. Ratio between basal and luminal MECs from 8-9-week-old virgin control and *Zeb1* MSKO mice (n=3 pairs).

H. Representative flow plots (left) and quantification of the number of basal and luminal MECs (right) in 15-week-old control and *Zeb1* MSKO mice.

I. Ratio between basal and luminal MECs in pregnant control and *Zeb1* MSKO females that were bred at the ages of 8 weeks (n=2 each).

J. Percentage of $K14^+$, $K14^+K8^+$, and $K8^+$ cells in basal and luminal MEC populations isolated from 15-week-old *Zeb1* MSKO virgin females and control littermates (n=3 pairs). Note: $K14^+K8^+$ are included in the $K14^+$ quantification.

K. Diagram of the pSLIK system used to induce *Zeb1* overexpression.

L. Time course of *Krt14* and *Krt19* mRNA expression in MCF10A cells after induction of *Zeb1* overexpression (n=3).

M. Immunostaining using K19 and K14 antibodies in MCF10A cells after induction of *Zeb1* overexpression. DAPI was used for nuclear staining. Scale bar: 50 μm.

N. RT-qPCR analysis of *Zeb1* expression in basal, luminal, or stromal cells from control and *Ovol2* knockout (Watanabe et al., 2014) mice (n=3 pairs).

O. RT-qPCR analysis of *Zeb1* expression in organoids derived from control, *Zeb1* single KO, *Ovol2* single KO, and *Zeb1*/*Ovol2* DKO MECs (n=3 each). See Figure 3K for more information.

P. *Ovol2* mRNA expression in organoids derived from control, *Zeb1* single KO, *Ovol2* single KO, and *Zeb1*/*Ovol2* DKO MECs (n=3 each).

Q. Percent of abnormal organoids that do not show the characteristic branched or acinar morphology (a colony with mixed branched-acinar morphology is shown on the top right as an example). Note that these organoids were excluded from analysis in Figure 3L. Data are represented as the mean ± SD (n=4). One-way ANOVA with multiple comparisons was used to analyze statistical significance between groups.

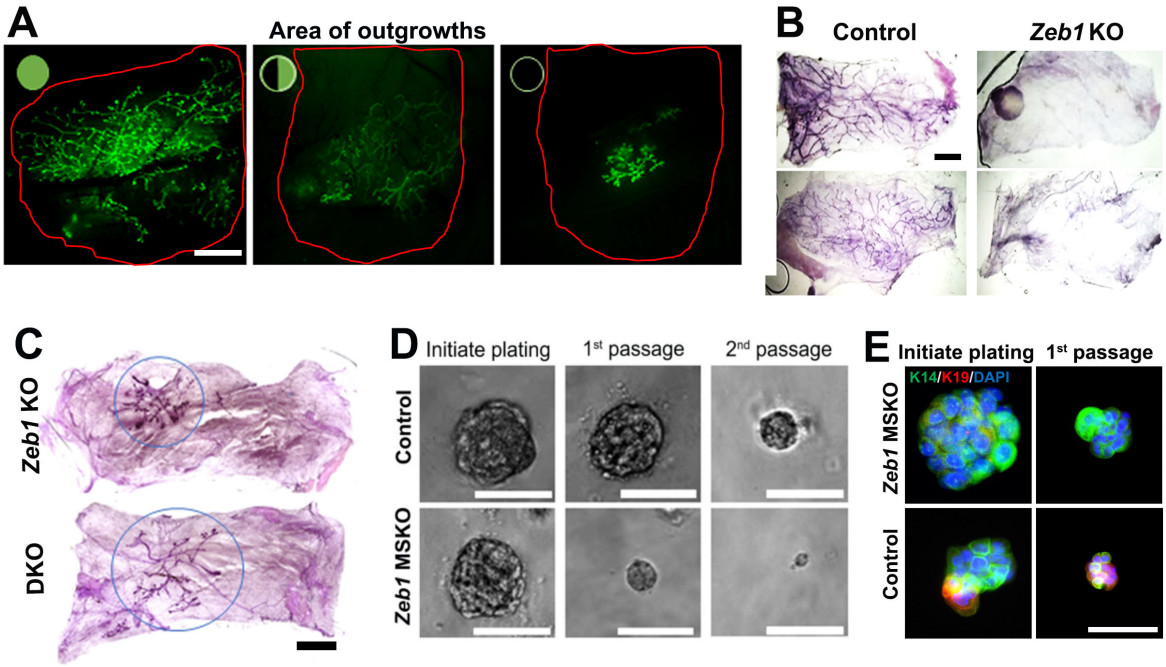

Figure S4. Related to Figure 4.

A. GFP images of representative transplants described in Figure 4A along with pie chart approximation of fat pad re-epithelialization. The red outlines indicate the border of fat pads. Scale bar: 1mm.

B. Additional replicates for Figure 4B. Scale bar: 5 mm.

C. Whole-mount carmine staining of transplants derived from *Zeb1* single KO and *Zeb1*/*Ovol2* DKO basal MECs. The mammary outgrowths are indicated with a circle. n=2 each. Scale bar: 5 mm.

D. Representative colony images (from 8-week-old control and *Zeb1* MSKO mice) for Figure 4D. Scale bar: 100 μm.

E. Whole-mount immunostaining of colonies from initiate plating and first passage of basal MECs from 15-week-old control and *Zeb1* MSKO mice. Scale bar: 100 μm.

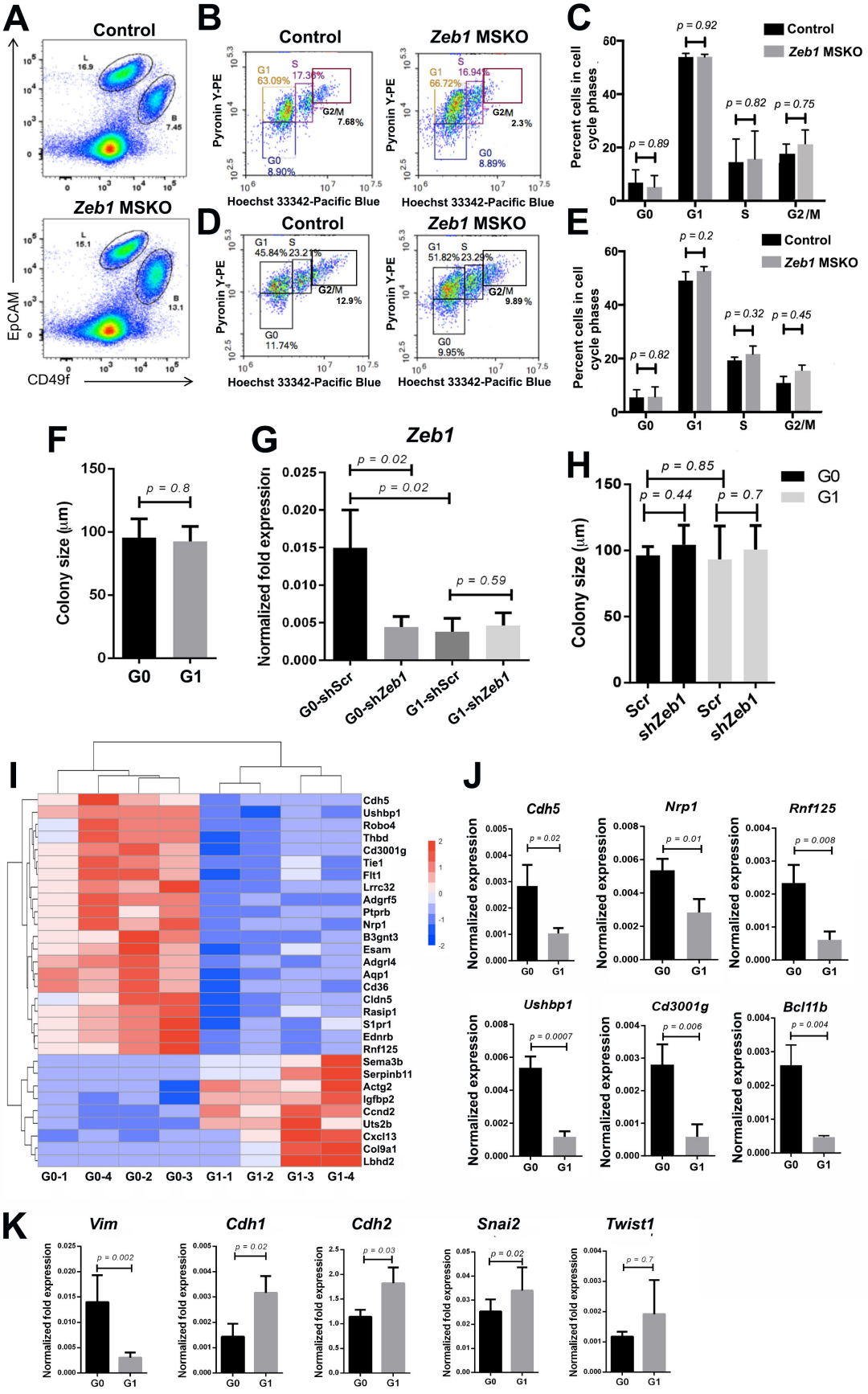
A
Control
Zeb1 MSKO
EpCAM
CD49f
B
Control
Zeb1 MSKO
Pyronin Y-PE
Hoechst 33342-Pacific Blue
G1 63.09%
S 17.36%
G2/M 7.68%
G0 8.90%
G1 66.72%
S 16.94%
G2/M 2.3%
G0 8.89%
C
Control
Zeb1 MSKO
Percent cells in cell cycle phases
p = 0.89
p = 0.92
p = 0.82
p = 0.75
G0
G1
S
G2/M
D
Control
Zeb1 MSKO
G1 45.84%
S 23.21%
G2/M 12.9%
G0 11.74%
G1 51.82%
S 23.29%
G2/M 9.89%
G0 9.95%
E
Control
Zeb1 MSKO
p = 0.82
p = 0.2
p = 0.32
p = 0.45
F
Colony size (μm)
p = 0.8
G0
G1
G
Zeb1
Normalized fold expression
p = 0.02
p = 0.02
p = 0.59
G0-shScr
G0-shZeb1
G1-shScr
G1-shZeb1
H
G0
G1
p = 0.85
p = 0.44
p = 0.7
Colony size (μm)
Scr
shZeb1
Scr
shZeb1
I
Cdh5
Ushbp1
Robo4
Thbd
Cd3001g
Tie1
Flt1
Lrrc32
Adgrf5
Ptprb
Nrp1
B3gnt3
Esam
Adgrl4
Aqp1
Cd36
Cldn5
Rasip1
S1pr1
Ednrb
Rnf125
Sema3b
Serpinb11
Actg2
Igfbp2
Ccnd2
Uts2b
Cxcl13
Col9a1
Lbhd2
G0-1
G0-4
G0-2
G0-3
G1-1
G1-2
G1-3
G1-4
J
Cdh5
Nrp1
Rnf125
Ushbp1
Cd3001g
Bcl11b
Normalized expression
p = 0.02
p = 0.01
p = 0.008
p = 0.0007
p = 0.006
p = 0.004
K
Vim
Cdh1
Cdh2
Snai2
Twist1
Normalized fold expression
p = 0.002
p = 0.02
p = 0.03
p = 0.02
p = 0.7

Figure S5. Related to Figure 5.

A. Representative flow plots of mammary cells in 8-week-old control and *Zeb1* MSKO mice.

B. Cell cycle analysis of luminal cells in 8-9-week-old control and *Zeb1* MSKO mice. Shown are represented FACS profiles of one pair of mice.

C. Summary of data from multiple pairs (n=3) as in (B).

D. Cell cycle analysis of luminal cells in 12-week-old control and *Zeb1* MSKO mice. Shown are represented FACS profiles of one pair of mice.

E. Summary of data from multiple pairs (n=3) as in (D).

F. Average size of colonies formed by G0 and G1 basal MECs. n=3 each.

G. RT-qPCR analysis of *Zeb1* expression in G0 and G1 basal MECs following lentivirus-mediated expression of shScr or sh*Zeb1*. n=3 each.

H. Average size of colonies formed by G0 and G1 basal MECs upon *Zeb1* knockdown. n=3 each.

I. Heatmap of genes differentially expressed in G0 and G1 cells revealed by RNA-seq. See Table S7 for additional information.

J. RT-qPCR analysis of G0-enriched genes in G0 and G1 cells. n=3 each.

K. RT-qPCR analysis of EMT-associated genes in G0 and G1 cells. n=3 each.

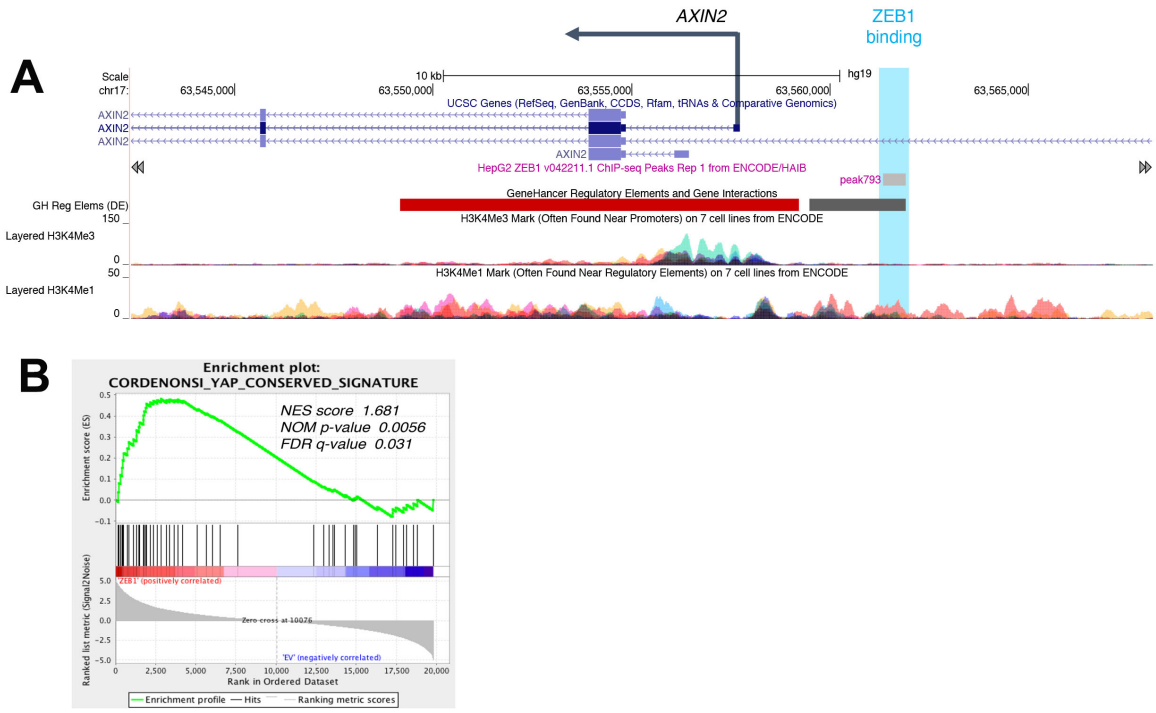

Figure S6. Related to Figure 6.

A. UCSC Genome Browser data showing a ZEB1 binding peak residing in an enhancer region (dark grey; GeneHancer database of human regulatory elements) upstream of the *AXIN2* promoter. Fidelity of the promoter and enhancer regulatory regions is supported by the detected presence of H3K4Me3 and H3K4Me1 histone modifications, respectively, in various cell lines.

B. GSEA of MCF10A RNA-seq data showing *Zeb1* overexpression-induced enrichment of a YAP gene signature (see Table S3 for a complete list of genes).

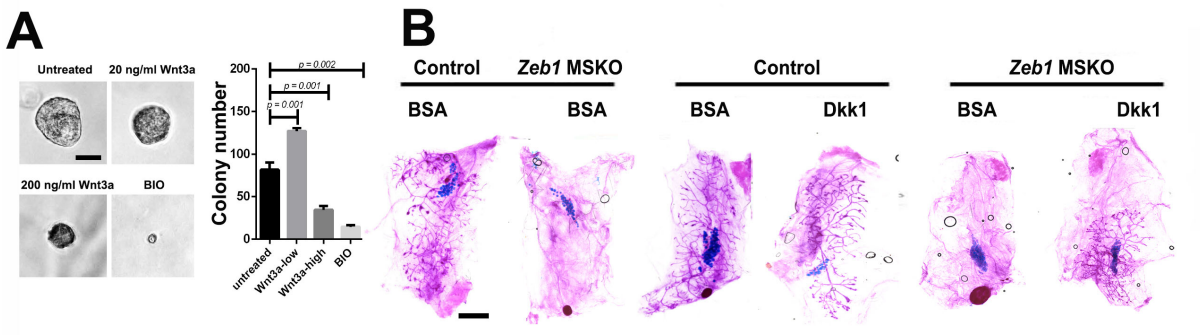

Figure S7. Related to Figure 7.

A. Morphology of representative basal MEC-derived colonies in the absence or presence of Wnt3a or BIO, and quantification of colony number. See Figure 7A. Scale bar: 50 μm.

B. Representative whole-mount images of transplants derived from Dkk1-bead- or BSA-bead-pretreated basal MECs from 15-week-old control and *Zeb1* MSKO mice. See Figure 7F. Scale bar: 5 mm.