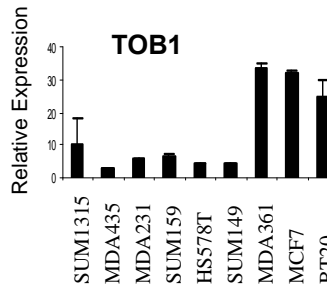
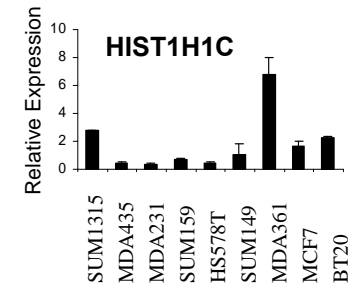
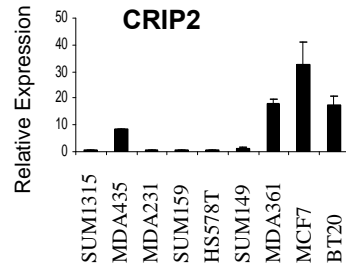
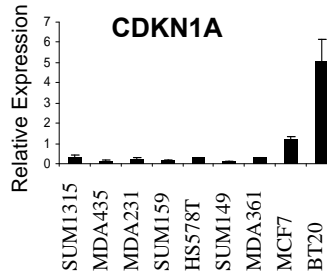
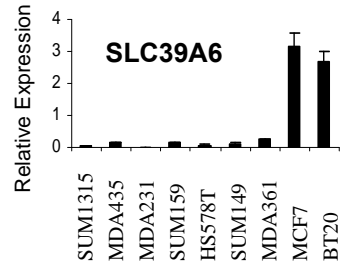
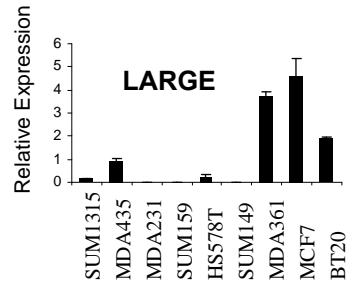
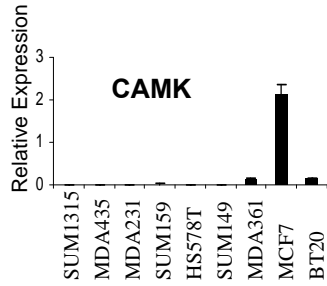
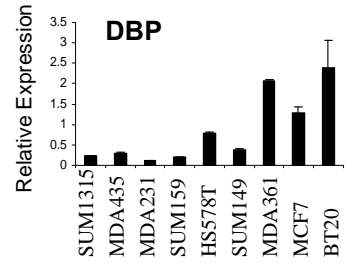
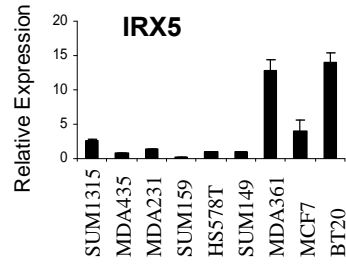
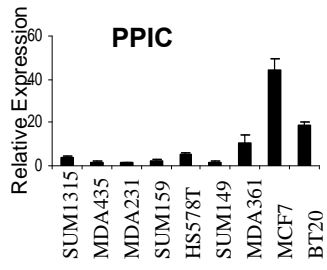
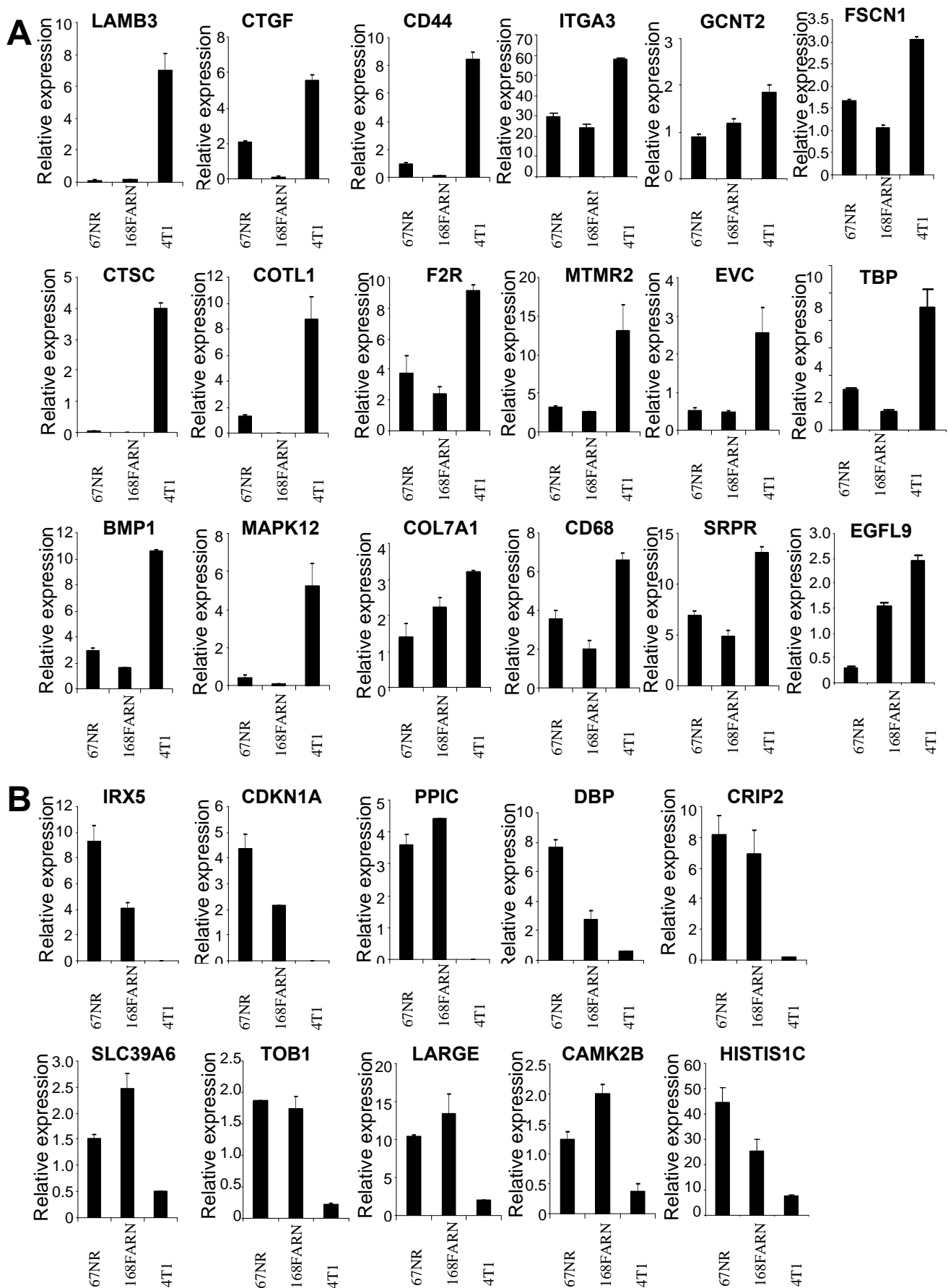


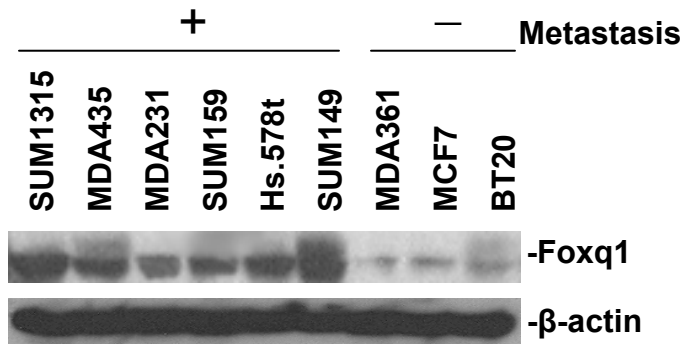
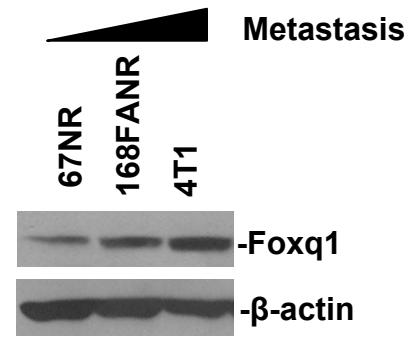
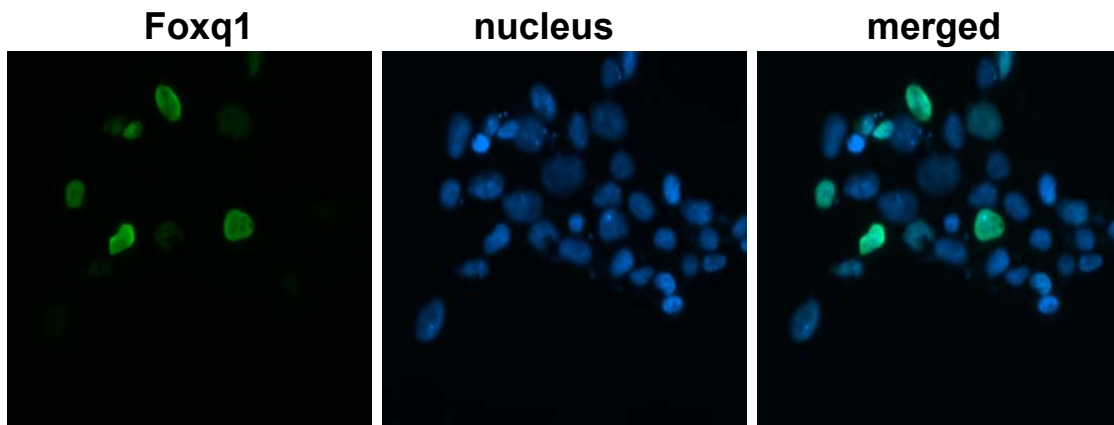
Supplementary figure 1. Zhang et al.

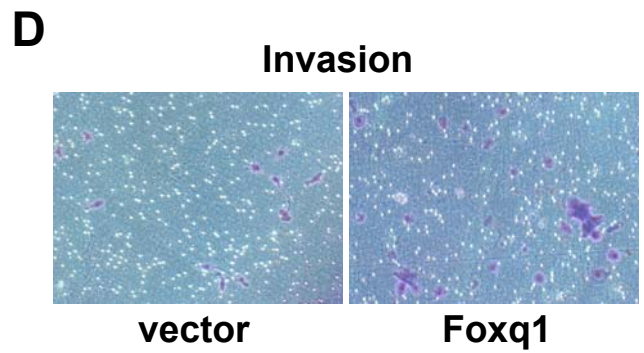
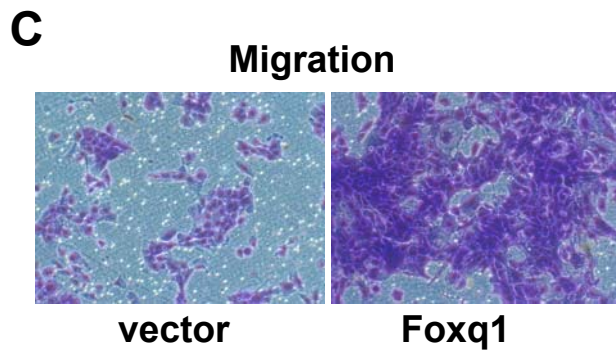
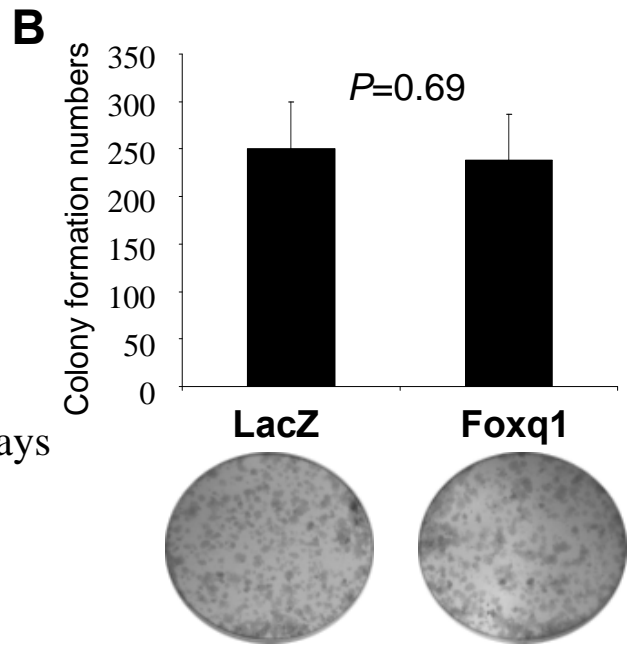
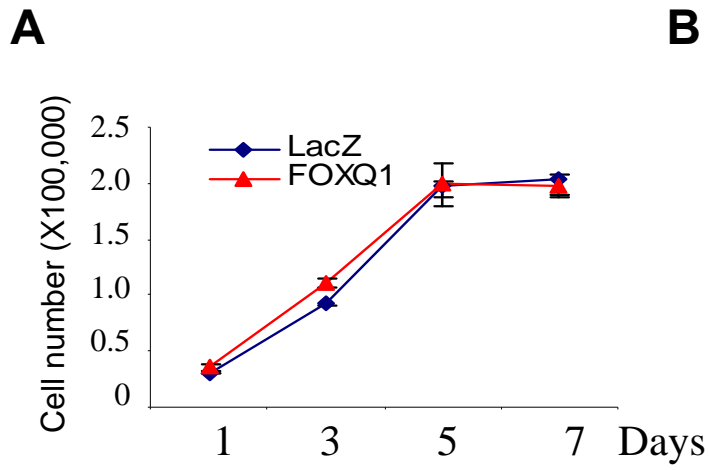


Supplementary figure 2. Zhang et al.

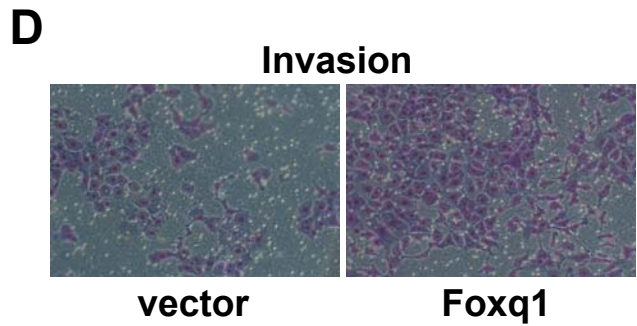
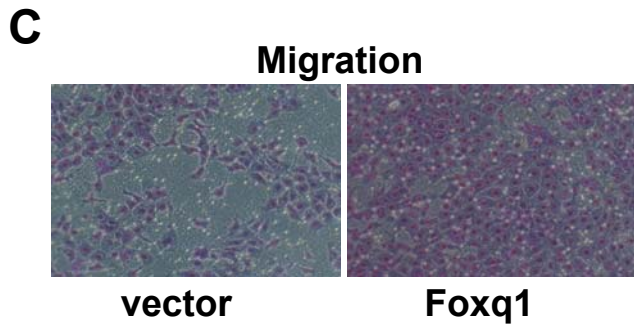
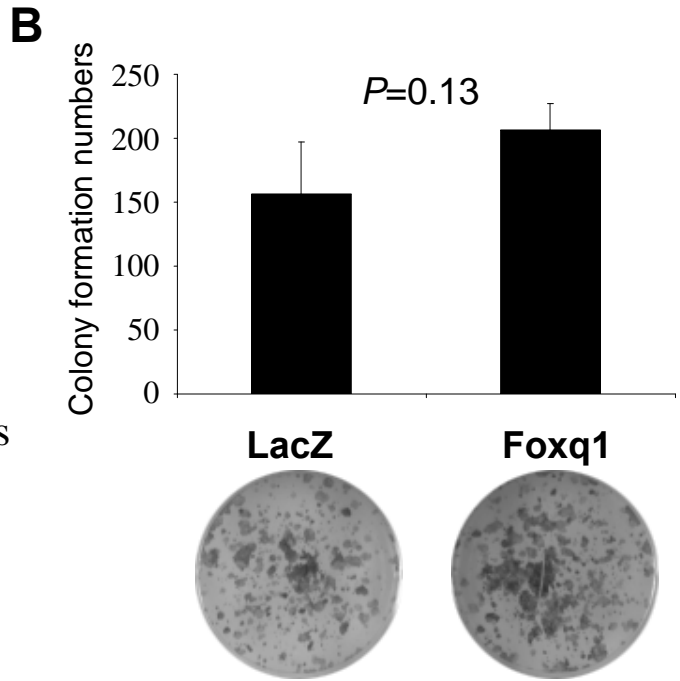
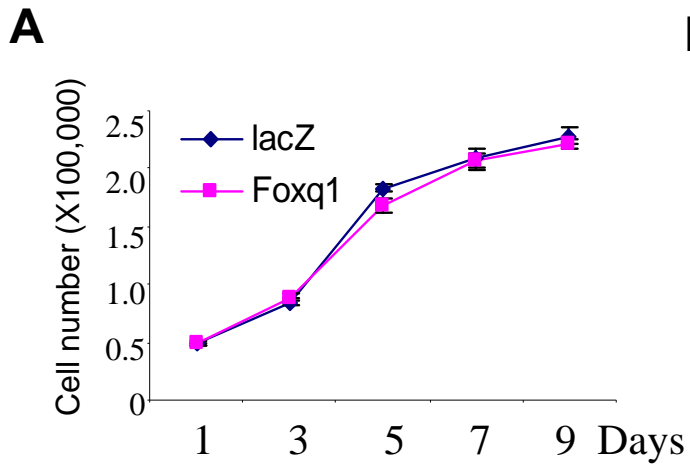


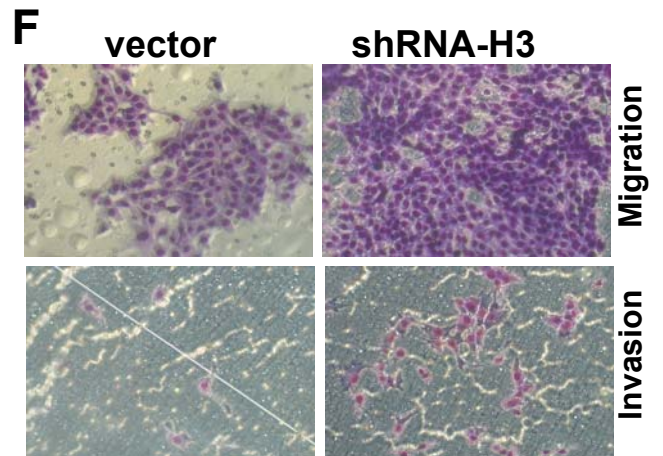
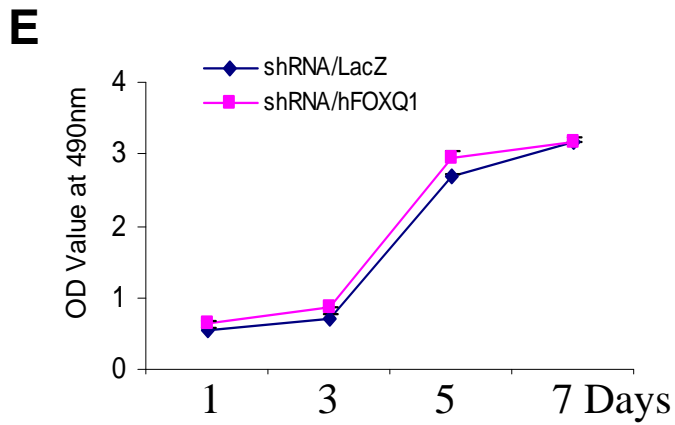
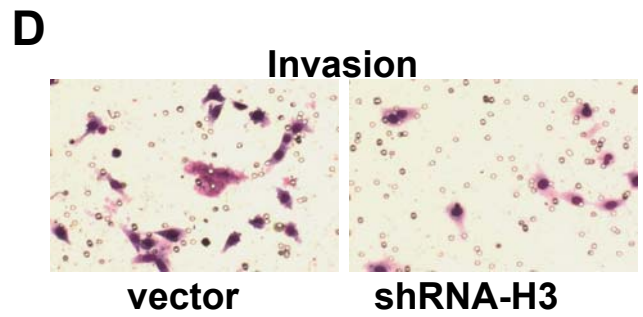
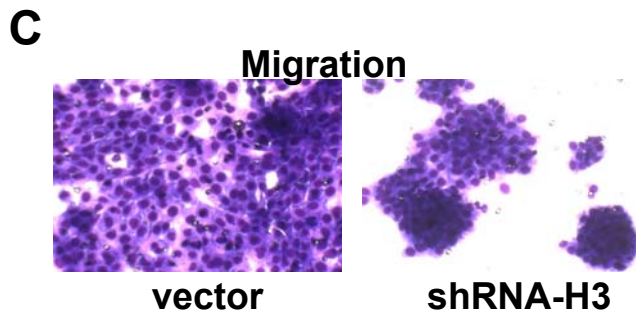
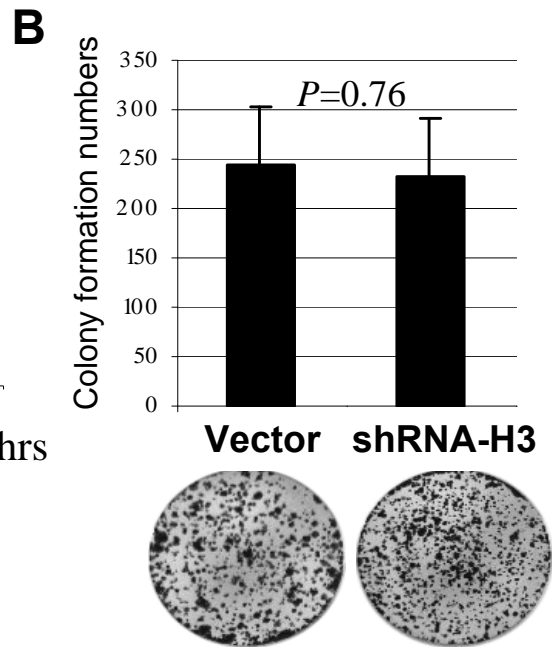
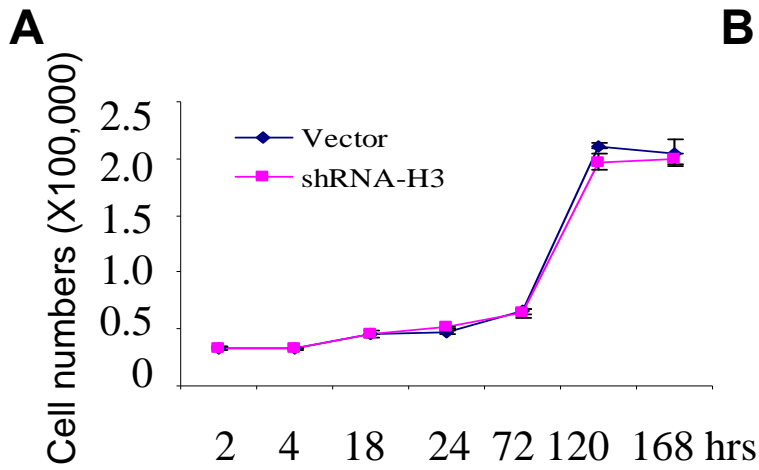
Supplementary figure 3. Zhang et al.

**A****B****C**

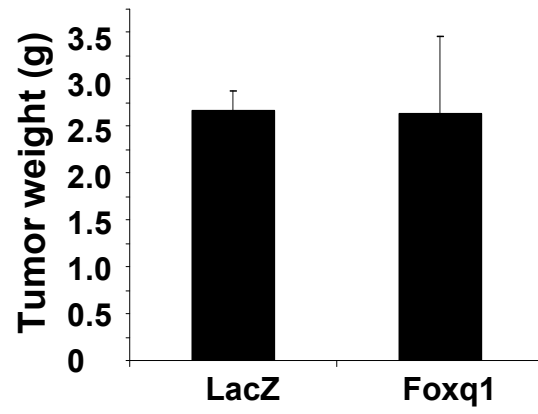


Supplementary figure 5. Zhang et al.



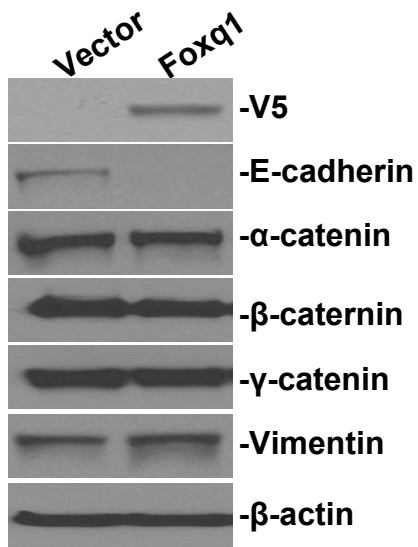
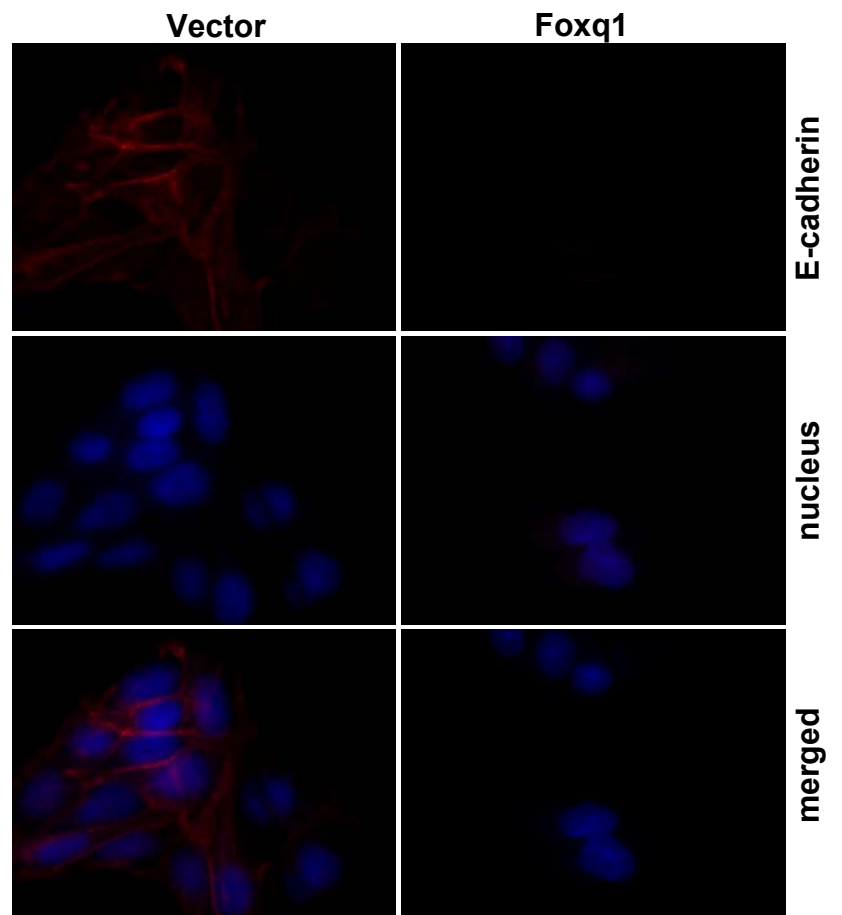


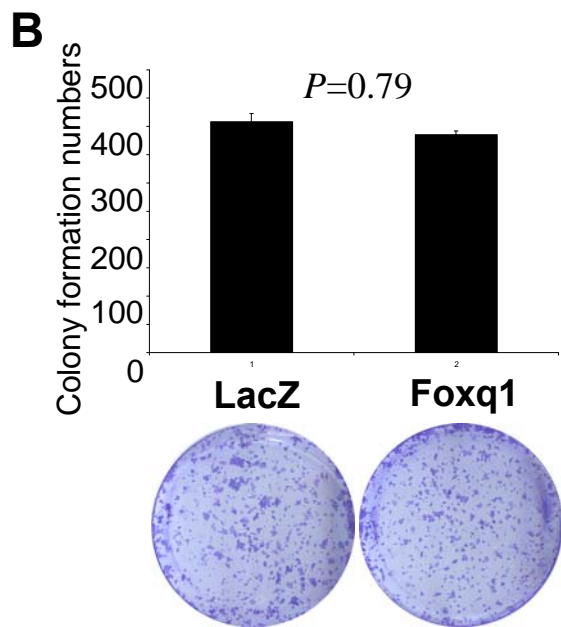
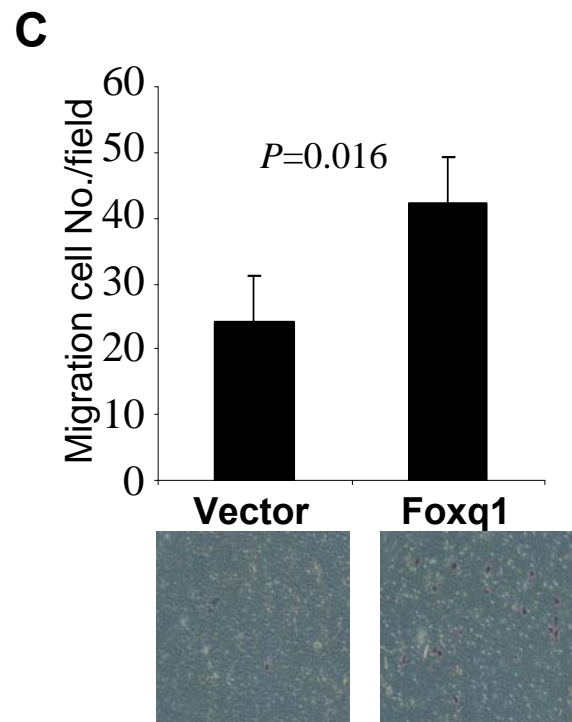
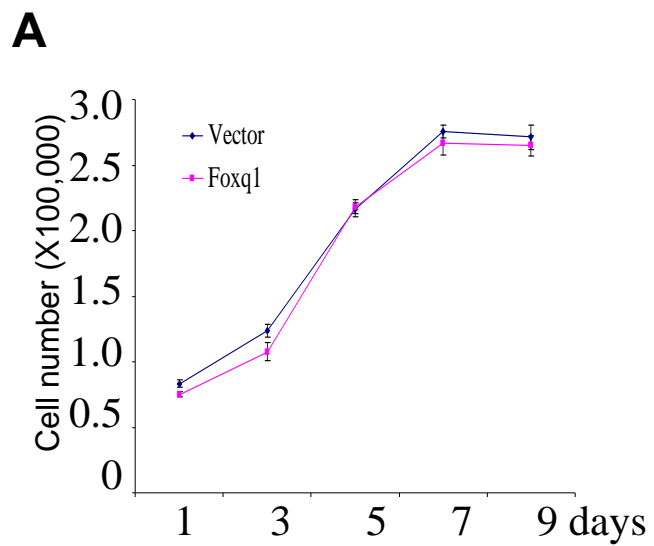
Supplementary figure 7. Zhang et al.

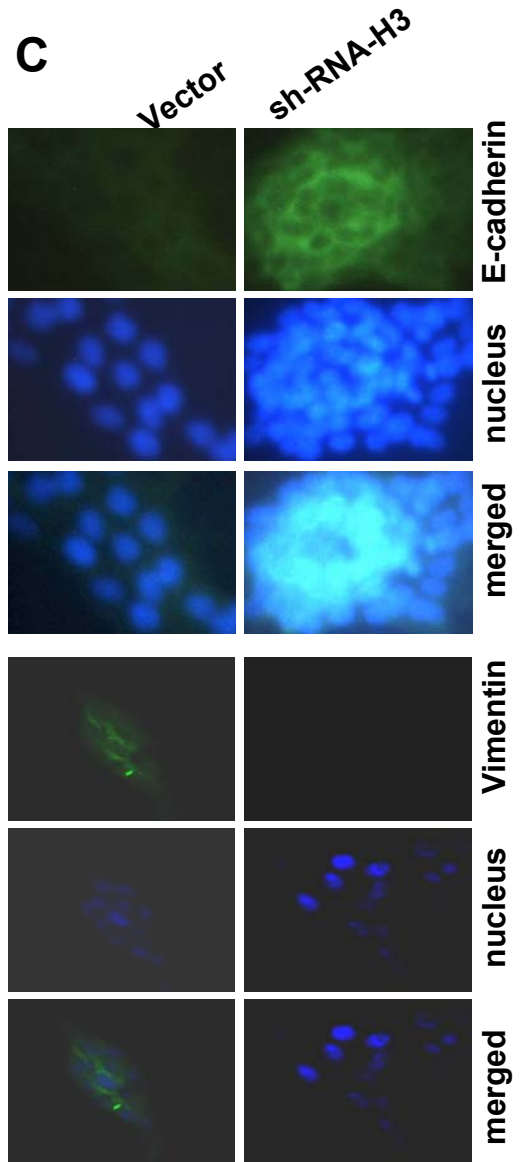
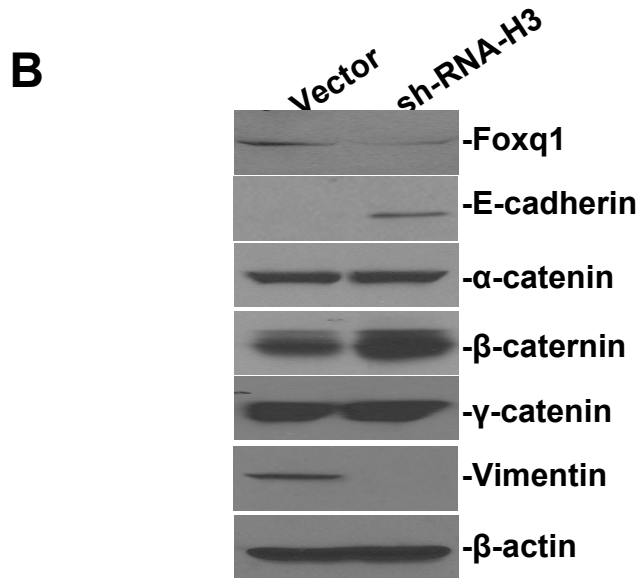
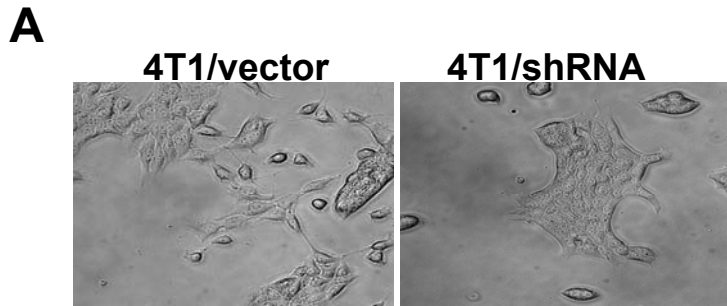


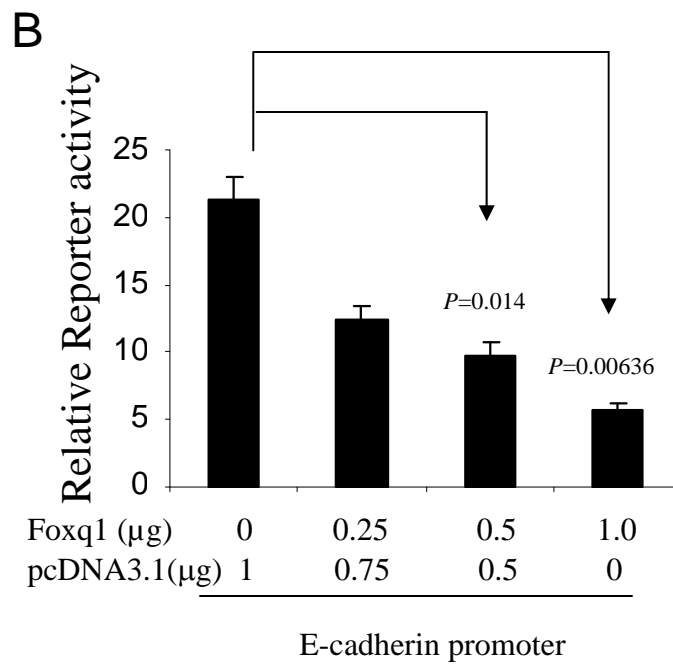
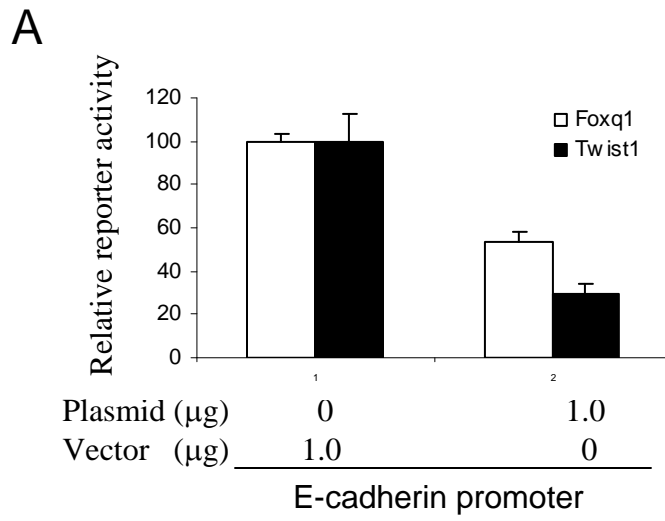
Supplementary figure 8. Zhang et al.



**A****B**



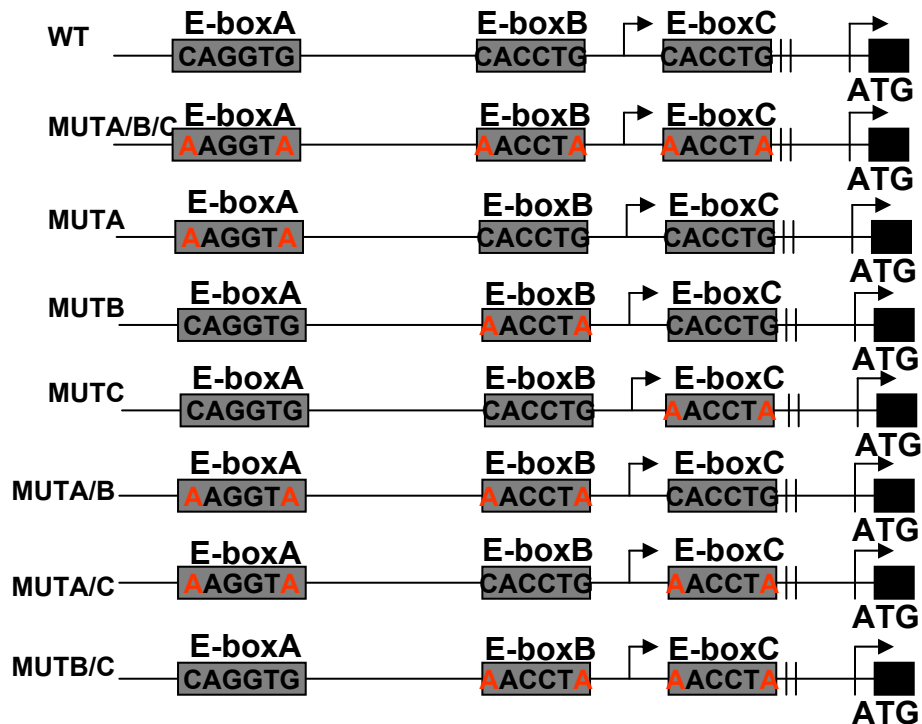




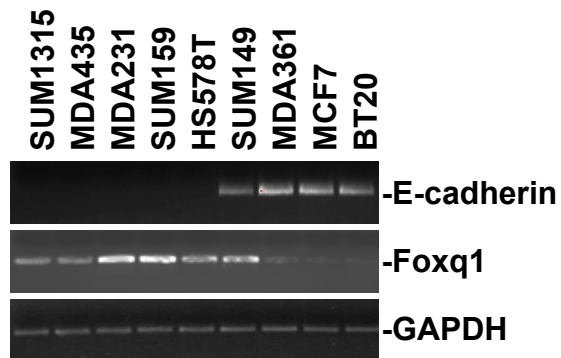
Supplementary figure 12. Zhang et al.

**A**

```
cgagggccgggtgggcegggcccgtcagctccgccctggggaggggtccgcgc  
tgctgattggctgtggccggcaggtgaaccctcagccaatcagcggtaacg  
E-box-A  
gggggcegggtgcctccggggtcacctggtcagccaacgcacccccctctc  
E-box-B  
AGTGGCGTCGGAAGTCAAAGCACCTGTGAGCTTGCGGAAGTCAGTTCAG  
Exon1 E-box-C  
ACTCCAGCCCGCTCCAGCCCGGCCCGACCCGACCCGACCCGGCGCCTGCC  
CTCGCTCGGCGTCCCCGGCCAGCCATGGGCCCTTGGAGCCGCAGCCTCTC  
GGCGCTGCTGCTGCTGCTGCAGgtacccccggatcccctgacttgcgaggg
```

**B**

Supplementary figure 13. Zhang et al.



Supplementary figure 14. Zhang et al.

### **Supplementary figure legends**

#### **Figure S1. Confirmation of the microarray data with quantitative real time RT-PCR.**

The expression patterns of 18 genes in a panel of human breast cancer cell lines were investigated by quantitative real time RT PCR. SUM1315, MDA435, MDA231, SUM159, HS578T, SUM149 are highly-metastatic cell lines. MDA361, MCF7 and BT20 are poorly-metastatic cell lines. These 18 genes showed up-regulation in highly metastatic cell lines.

#### **Figure S2. Confirmation of the microarray data with quantitative real time RT-PCR.**

The expression patterns of 10 genes in a panel of human breast cancer cell lines were investigated by quantitative real time RT PCR. SUM1315, MDA435, MDA231, SUM159, HS578T, SUM149 are highly-metastatic cell lines. MDA361, MCF7 and BT20 are poorly metastatic cell lines. These 10 genes showed down-regulation in highly-metastatic cell lines.

#### **Figure S3. Confirmation of the microarray data with quantitative real time RT-PCR.**

*A*, The expression patterns of 18 genes in a panel of mouse breast cancer cell lines were investigated by quantitative real time RT PCR. These 18 genes showed up-regulation in highly-metastatic mouse cell line 4T1. *B*, The expression patterns of 10 genes in a panel of mouse breast cancer cell lines were investigated by quantitative real time RT PCR. These 10 genes showed down-regulation in highly-metastatic mouse cell line 4T1.

**Figure S4. The expression of *Foxq1* in human and mouse breast cancer cell lines and Subcellular localization of the *Foxq1* gene.** *A* and *B*, Investigation of the protein level of *Foxq1* in human (*A*) and mouse (*B*) model cell lines using western blotting.  $\beta$ -actin is used as a protein loading control. +, highly metastatic. -, poorly metastatic. Triangular bar indicates an increase of metastasis. *C*, Subcellular localization of the *Foxq1* gene. The *Foxq1* gene is V5-tagged and was transfected into 293 cells. An immunofluorescence analysis was performed with an anti-V5 antibody (green, left). HOECHST 33342 was used to stain DNA (blue, center). Merged images were present on the right.

**Figure S5. The effect of ectopic expression of *Foxq1* on cell proliferation and colony formation in HMLE cells.** *A*, MTT assay shows that overexpression of *Foxq1* in HMLE cells doesn't change the cell proliferation in seven days. *B*, Cell colony formation assay shows that ectopic expression of *Foxq1* in HMLE cells doesn't change the colony formation capability of HMLE cells ( $P=0.69$ ). *C* and *D*, ectopic expression of *Foxq1* in HMLE cells promotes migration (*C*) and invasion (*D*).

**Figure S6. The effect of ectopic expression of *Foxq1* on cell proliferation and colony formation in EpRas cells.** *A*, MTT assay shows that overexpression of *Foxq1* in EpRas cells doesn't change the cell proliferation in nine days. *B*, Cell colony formation assay shows that ectopic expression of *Foxq1* in EpRas doesn't change the colony formation capability of EpRas cells ( $P=0.13$ ). *C* and *D*, ectopic expression of *Foxq1* in EpRas cells promotes migration (*C*) and invasion (*D*).



**Figure S7. The effect of inhibition of expression of *Foxq1* on cell proliferation and colony formation in 4T1 cells.** *A*, MTT assay shows that knockdown of *Foxq1* in 4T1 cells doesn't change the cell proliferation in 168 hours (7 days). *B*, Cell colony formation assay shows that knockdown of *Foxq1* in 4T1 doesn't change the colony formation capability of 4T1 cells ( $P=0.76$ ). *C* and *D*, Knockdown of *Foxq1* expression in 4T1 cells inhibits migration (*C*) and invasion (*D*). *E*. Introduction of hFoxq1 to 4T1 cells with *Foxq1* knockdown does not change the cell proliferation. *F*. Introduction of hFoxq1 to 4T1 cells with *Foxq1* knockdown promotes cell migration (top two panels) and invasion (low two panels).

**Figure S8. The effect of *Foxq1* expression on tumor growth *in vivo*.** Comparison of the weights of primary mammary tumors formed by EpRas cells expressing either *Foxq1* or control vector. Same numbers of both cells were injected into nude mice. Each data point represents the mean  $\pm$  SEM.

**Figure S9. The effect of *Foxq1* on EMT program in MDCK cells.** *A*, Western blotting results show the effect of *Foxq1* overexpression on EMT markers in the MDCK cells. The epithelial markers E-cadherin was down-regulated and the other epithelial markers and mesenchymal marker Vimentin were not changed. *B*, Immunofluorescence assays confirmed the results of the western blot in MDCK cells. Panel showed down-regulation of E-cadherin in MDCK cells with *Foxq1* expression. Red signals stand for the stain of the E-cadherin gene. Blue signals stand for the nucleus staining.

**Figure S10. The effect of ectopic expression of *Foxq1* on cell proliferation, colony formation and migration in MDCK cells.** *A*, MTT assay shows that ectopic expression of *Foxq1* in MDCK cells doesn't change the cell proliferation in nine days. *B*, Cell colony formation assay shows that ectopic expression of *Foxq1* in MDCK cells doesn't change the colony formation capability of MDCK cells ( $P=0.79$ ). *C*, Cell migration assay using the BD transwell migration chambers shows that overexpression of *Foxq1* leads to an increase of migration ability of MDCK cells (53%,  $P=0.016$ ).

**Figure S11. The effect of *Foxq1* on EMT program in mouse mammary epithelial cell 4T1.** *A*, Representative pictures show the morphological change after inhibition of *Foxq1* expression in 4T1 cell. *B*, Western blotting results show the effect of *Foxq1* inhibition in the 4T1 cells on EMT markers. The epithelial markers E-cadherin and  $\beta$ -catenin were up-regulated and mesenchymal marker Vimentin is inhibited. ShRNA-H3 represents cell lysates of *Foxq1* shRNA-H3 cells. *C*, Immunofluorescence assays confirmed the results of the western blot in 4T1 cells (*B*). Top panels shows up-regulation of E-cadherin in 4T1 cells with knockdown of *Foxq1* expression. The low panels show Vimentin expression was down-regulated due to the knockdown of *Foxq1* expression. ShRNA-H3 represents 4T1 cells with *Foxq1* shRNA-H3.

**Figure S12. Foxq1 repress *E-cadherin* promoter.** *A*, The luciferase assay shows that the twist gene harbors stronger repressive capability on E-cadherin gene promoter than the Foxq1 gene. *B*, The Luciferase assay shows that *Foxq1* down-regulates the *E-cadherin* promoter in a dose-dependent manner in MCF7 cells. Variable amounts of the

*Foxq1* expression construct and pcDNA3.1 vector were co-transfected into MCF7 cells with the *E-cadherin* promoter plasmid. Bars, SD.

**Figure S13. The sequence and constructs of the *E-cadherin* gene promoter.** *A*, The sequence of the *E-cadherin* gene promoter. The three E- box regions, transcriptional start point and start point (ATG) are indicated as labeled. *B*, Schematic diagram of the *E-cadherin* gene promoter cloned into the pGL3 basic reporter construct. A series of mutant constructs are shown. The position of the three E-box regions relative to ATG is shown for each construct. The wild type and mutant sequences of E-box were shown in the shaded boxes.

**Figure S14. Reverse correlation of *Foxq1* and *E-cadherin*.** Reverse correlation of *Foxq1* and *E-cadherin* expression in breast cancer cell lines. RT-PCR showed the expression pattern of *Foxq1* and *E-cadherin* in a set of breast cancer cell lines. GAPDH is an RNA control.

**Table S1. twenty-two up-regulated genes**

Gene	4T1		MDA231	SUM159
	67NR	168FARN	MCF7	
	PLAU	10.86717	4.396752	58.65388
CTGF	2.299719	5.782955	49.72356	10.10379
NT5E	3.816829	3.006329	32.80148	4.837398
FOXQ1	3.473339	3.400526	20.53662	17.81016
CD44	2.482212	2.542759	20.51093	11.31299
CD68	2.284462	5.098053	14.04149	4.337182
LAMB3	22.18648	12.19286	12.25399	5.538799
TGFBR2	2.553168	2.557427	12.10823	16.33052
GCNT2	3.690767	4.180192	8.903633	4.866379
ITGA3	3.756690	2.343037	8.236642	5.767789
CTSC	4.604886	4.925045	7.155259	7.10876
COTL1	2.409051	5.254732	5.849719	5.479707
EGFL9	3.872563	2.254671	4.102571	2.732542
SRPR	2.088897	2.2592	3.743962	3.572755
COL7A1	2.505199	2.441787	3.550176	8.232615
F2R	2.336289	3.299181	3.028073	2.680544
FSCN1	2.800273	3.146898	2.767984	7.555569
BMP1	4.325937	5.831634	2.399728	3.691573
JOSD3	2.46529	2.221018	2.277617	2.733281
MAPK12	6.058898	5.185764	2.272213	3.926462
MTMR2	3.077857	3.435021	2.172874	2.145049
EVC	3.875810	3.764287	2.025917	2.840985

**Table S2. twelve down-regulated genes**

Gene	67NR	168FARN	MCF7	
		4T1	MDA231	SUM159
CXCL12	27.8304	64.65065	7.716821	6.504192
CRIP2	17.51906	14.20295	20.92149	25.19247
IRX5	12.01882	6.735616	3.061279	14.11685
CDKN1A	9.333535	8.908734	2.319016	3.196259
PHLDA3	6.350332	5.793361	6.439936	4.808639
DBP	5.48507	2.847345	2.944285	3.032524
PPIC	5.30236	16.38163	3.41355	3.817891
HIST1H1C	2.22425	3.565479	6.004867	5.954476
SLC39A6	2.213763	2.727612	3.007224	2.251689
TOB1	2.145595	2.462454	6.377567	5.356411
CAMK2B	2.107861	2.47037	7.830648	7.128167
LARGE	2.070512	3.454575	13.29347	13.00762

**Table S3. categories of differentially expressed genes**

Gene	Detailed description or Function	EMT	Metastasis
<i>ECM and adhesion protein</i>			
PLAU	UPA, plasminogen activator, urokinase		X
CTSC	cathepsin C		X
ITGA3	integrin, alpha 3		X
COL7A1	collagen, type VII, alpha1		X
LAMB3	laminin, beta 3		X
<i>ligand, growth factors and receptors</i>			
<b>CXCL12</b>	chemokine (C-X-C motif) ligand 12		X
CTGF	connective tissue growth factor	X	X
BMP1	bone morphogenetic protein1		X
TGFR2	transforming growth factor, beta receptor II	X	X
SRPR	signal recognition particle receptor		
F2R	coagulation factor II (thrombin) receptor		X
<i>antigen/marker</i>			
CD44	CD44 antigen		X
CD68	macrophage marker		
<i>transcription factor</i>			
FOXQ1	forkhead box q1		
TAF1D	TATA box binding protein associated factor		
<b>IRX5</b>	Iroquois homeobox 5		
<b>DBP</b>	D-site of a albumin promoter binding protein		
<i>enzymes</i>			
GCNT2	glucosaminyl(N-acetyl) transferase2, I branching enzyme		
NT5E	ecto-5'-nucleotidase (CD73)		X
<b>PPIC</b>	peptidylprolyl isomerase C (cyclophilin C)		
<b>LARGE</b>	like-glycosyltransferase		
<i>kinase and phosphatase</i>			
MAPK12	Mitogen-activated protein kinase12 ( P38 gamma)		
MTMR2	myotubularin related protein 2		
<b>CAMK2B</b>	calcium/calmodulin-dependent protein kinaseII beta		
<b>CDKN1A</b>	cyclin- dependent kinase inhibitor 1A (p21)		
<i>cytoskeleton</i>			
COTL1	coactosin-like 1(Dictyostelium)		
FSCN1	fascin homolog1, actin -binding protein		
<i>signaling</i>			
<b>TOB1</b>	transducer of ERBB2,1		
<b>PHLDA3</b>	pleckstrin homology-like domain, family, member3		
<b>CRIP2</b>	Cysteine-rich protein 2		
<b>LIV-1</b>	solute carrier family 39 (zinc transporter)	X	X
EGFL9	EGF-like-domain, multiple 9		
<i>other</i>			
EVC	Ellis van creveld Syndrome		
<b>HIST1H1C</b>	histone cluster 1, Hic		

\* Black genes, upregulated. Red genes, downregulated. X, reported previously.