

Table S1. List of genes differentially expressed in Caco-2-FOXK2 versus Caco-2-control cells using a human EMT PCR array

Symbol	Caco-2-FOXK2 vs Caco-2-control	Description
EGFR	6.12	Epidermal growth factor receptor
ZEB1	5.57	Zinc finger E-box binding homeobox 1
VIM	4.98	Vimentin
FN1	4.49	Fibronectin 1
MMP2	4.39	Matrix metalloproteinase 2
MMP9	3.87	Matrix metalloproteinase 9
ERBB3	3.32	V-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)
MSN	3.31	Moesin
VCAN	3.11	Versican
ITGAV	2.98	Integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51)
KRT7	2.81	Keratin 7
AKT1	2.78	V-akt murine thymoma viral oncogene homolog 1
CALD1	2.67	Caldesmon 1
ZEB2	2.67	Zinc finger E-box binding homeobox 2
COL5A2	2.45	Collagen, type V, alpha 2
IGFBP4	2.29	Insulin-like growth factor binding protein 4
ITGA5	2.29	Integrin, alpha 5
COL1A2	2.27	Collagen, type I, alpha 2
TWIST1	2.27	Twist homolog 1 (Drosophila)
MAP1B	2.21	Microtubule-associated protein 1B
SNAI1	2.01	Snail homolog 1 (Drosophila)
PTK2(FAK)	1.99	PTK2 protein tyrosine kinase 2
BMP7	1.99	Bone morphogenetic protein 7
ILK	1.91	Integrin-linked kinase
SNAI2(Slug)	1.89	Snail homolog 2 (Drosophila)
ITGB1	1.81	Integrin, beta 1
COL3A1	1.78	Collagen, type III, alpha 1
FGFBP1	1.78	Fibroblast growth factor binding protein 1
CAV2	1.78	Caveolin 2
NODAL	1.71	Nodal homolog
FOXC2	1.71	Forkhead box C2
TGFB1	1.67	Transforming growth factor, beta 1
GSC	1.55	Goosecoid homeobox
PDGFRB	1.34	Platelet-derived growth factor receptor, beta polypeptide
STAT3	1.34	Signal transducer and activator of transcription 3
BMP2	1.23	Bone morphogenetic protein 2
WNT5A	1.22	Wingless-type MMTV integration site family, member 5A
SNAI3	1.22	Snail homolog 3 (Drosophila)
RAC1	1.21	Ras-related C3 botulinum toxin substrate 1
GSK3B	1.16	Glycogen synthase kinase 3 beta

RGS2	1.13	Regulator of G-protein signaling 2, 24kDa
WNT11	1.12	Wingless-type MMTV integration site family, member 11
TGFB2	1.12	Transforming growth factor, beta 2
BMP1	1.12	Bone morphogenetic protein 1
OCLN	1.12	Occludin
MMP3	1.12	Matrix metalloproteinase 3
F11R	1.1	F11 receptor, junction adhesion molecule A
CTNNB1(b-catenin)	1.09	Catenin (cadherin-associated protein), beta 1, 88kDa
FZD7	1.09	Frizzled family receptor 7
VPS13A	1.08	Vacuolar protein sorting 13 homolog A (<i>S. cerevisiae</i>)
TCF3	1.08	Transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)
MST1R(ROn)	1.05	Macrophage stimulating 1 receptor (c-met-related tyrosine kinase)
STEAP1	1.06	Six transmembrane epithelial antigen of the prostate 1
WNT5B	1.03	Wingless-type MMTV integration site family, member 5B
PLEK2	1.03	Pleckstrin 2
HPRT1	1.02	Hypoxanthine phosphoribosyltransferase 1
SERPINE1	1.02	Serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1
AHNAK	1.01	AHNAK nucleoprotein
B2M	1.01	Beta-2-microglobulin
TSPAN13	-1.01	Tetraspanin 13
GEMIN2	-1.04	Survival of motor neuron protein interacting protein 1
SMAD2	-1.09	SMAD family member 2
GNG11	-1.06	Guanine nucleotide binding protein (G protein), gamma 11
TGFB3	-1.06	Transforming growth factor, beta 3
TCF4	-1.07	Transcription factor 4
TMEM132A	-1.14	Transmembrane protein 132A
DESI1	-1.16	PPPDE peptidase domain containing 2
ESR1(ERa)	-1.21	Estrogen receptor 1
NUDT13	-1.22	Nudix (nucleoside diphosphate linked moiety X)-type motif 13
IL1RN	-1.22	Interleukin 1 receptor antagonist
KRT19	-1.23	Keratin 19
TFPI2	-1.33	Tissue factor pathway inhibitor 2
TMEFF1	-1.33	Transmembrane protein with EGF-like and two follistatin-like domains 1
PTP4A1(PRL1)	-1.54	Protein tyrosine phosphatase type IVA, member 1
KRT14	-1.55	Keratin 14
SOX10	-1.56	SRY (sex determining region Y)-box 10
JAG1	-1.76	Jagged 1
DSP	-1.78	Desmoplakin
NOTCH1	-1.81	Notch 1
CDH2	-2.11	Cadherin 2, type 1, N-cadherin (neuronal)
TIMP1	-2.21	TIMP metalloproteinase inhibitor 1
CAMK2N1	-2.29	Calcium/calmodulin-dependent protein kinase II inhibitor 1
DSC2	-2.32	Desmocollin 2
SPP1(Osteopontin)	-2.87	Secreted phosphoprotein 1

SPARC	-3.32	Secreted protein, acidic, cysteine-rich (osteonectin)
CDH1	-3.78	Cadherin 1, type 1, E-cadherin (epithelial)

Table S2. Correlation between nuclear p65 expression and clinicopathological characteristics of CRCs in two independent cohorts of human CRC tissues

Clinicopathological variables		Cohort I (n=363)			Cohort II (n=390)		
		Tumor p65 expression		P Value	Tumor p65 expression		P Value
		Low (n=203)	High (n=160)		Low (n=218)	High (n=172)	
Age		66.75(11.03)	65.72(11.75)	0.727	68.44(10.52)	66.28(12.67)	0.067
Sex	female	79	82	0.02	89	88	0.052
	male	124	78		129	84	
Tumor location	right colon	88	85	0.001	84	81	0.066
	left colon	66	62		95	73	
	rectum	49	13		39	18	
Tumor size	<5cm	92	62	0.24	78	71	0.294
	≥5cm	111	98		140	101	
Tumor differentiation	well or moderate	171	79	<0.001	152	72	<0.001
	poor	32	81		66	100	
Tumor invasion	T1	6	1	0.014	11	5	0.031
	T2	22	8		18	6	
	T3	133	109		151	122	
	T4	42	42		38	39	
Lymph node metastasis	absent	158	41	<0.001	178	48	<0.001
	present	45	119		40	124	
Distant metastasis	absent	189	109	<0.001	208	110	<0.001
	present	14	51		10	62	
AJCC stage	Stage I	28	7	<0.001	15	5	<0.001
	Stage II	128	31		161	38	
	Stage III	33	71		32	69	
	Stage IV	14	51		10	60	

Table S3. Correlation between ZEB1 expression and clinicopathological characteristics of CRCs in two independent cohorts of human CRC tissues

Clinicopathological variables	Cohort I (n=363)			Cohort II (n=390)			
	Tumor ZEB1 expression		P Value	Tumor ZEB1 expression		P Value	
	Low (n=219)	High (n=144)		Low (n=217)	High (n=173)		
Age	66.61(11.39)	65.82(11.29)	0.977	68.06(10.67)	66.76(12.57)	0.197	
Sex	female	89	72	0.085	92	85	0.219
	male	130	72		125	88	
Tumor location	right colon	99	74	0.015	81	84	0.058
	left colon	71	57		102	66	
	rectum	49	13		34	23	
Tumor size	<5cm	93	61	1	88	61	0.296
	≥5cm	126	83		129	112	
Tumor differentiation	well or moderate	182	68	<0.001	150	74	<0.001
	poor	37	76		67	99	
Tumor invasion	T1	4	3	0.157	15	1	0.001
	T2	21	9		15	9	
	T3	149	93		153	120	
	T4	45	39		34	43	
Lymph node metastasis	absent	157	42	<0.001	177	49	<0.001
	present	62	102		40	124	
Distant metastasis	absent	201	97	<0.001	211	107	<0.001
	present	18	47		6	66	
AJCC stage	Stage I	25	10	<0.001	19	1	<0.001
	Stage II	130	29		157	42	
	Stage III	46	58		35	66	
	Stage IV	18	47		6	64	

Table S4. Correlation between EGFR expression and clinicopathological characteristics of CRCs in two independent cohorts of human CRC tissues

Clinicopathological variables	Cohort I (n=363)			Cohort II (n=390)			
	Tumor EGFR expression		P Value	Tumor EGFR expression		P Value	
	Low (n=197)	High (n=166)		Low (n=225)	High (n=165)		
Age	66.32(10.94)	66.27(11.83)	0.434	67.90(10.91)	66.92(12.39)	0.408	
Sex	female	77	84	0.034	99	78	0.538
	male	120	82		126	87	
Tumor location	right colon	100	73	0.378	87	78	0.104
	left colon	51	77		99	69	
	rectum	46	16		39	18	
Tumor size	<5cm	80	74	0.457	88	61	0.675
	≥5cm	117	92		137	104	
Tumor differentiation	well or moderate	165	85	<0.001	152	72	<0.001
	poor	32	81		73	93	
Tumor invasion	T1	5	2	0.137	12	4	0.026
	T2	18	12		15	9	
	T3	133	109		161	112	
	T4	41	43		37	40	
Lymph node metastasis	absent	142	57	<0.001	181	45	<0.001
	present	55	109		44	120	
Distant metastasis	absent	185	113	<0.001	215	103	<0.001
	present	12	53		10	62	
AJCC stage	Stage I	22	13	<0.001	15	5	<0.001
	Stage II	118	41		165	34	
	Stage III	45	59		35	66	
	Stage IV	12	53		10	60	

Table S5. Primer sequences used in the study.

Primer name	Primer sequences	Enzyme
Primers for ZEB1 promoter construct:		
(-1947/+66)ZEB1 sense:	5'-TATAGGTACCATGTGTGATAGAGCTGGA-3'	KpnI
(-1655/+66)ZEB1 sense:	5'-TATAGGTACCGATGAATGCAGATATATA-3'	KpnI
(-1262/+66)ZEB1 sense:	5'-TATAGGTACCCTCTATCAATAACTGCTA-3'	KpnI
(-687/+66)ZEB1 sense:	5'-TATAGGTACCCAATCTCATTGAAGTCAC-3'	KpnI
Antisense:	5'-ATATAAGCTTCATGATCCTCTCGCTTGT-3'	HindIII
Primers for ZEB1 promoter site-directed mutagenesis:		
binding site 4 mutation sense:	5'-CAGCTCAGCTTcgcgATTCCATAAAG-3'	
binding site 4 mutation antisense:	5'-CTTTATGGAATcgcgAAGCTGAGCTG-3'	
binding site 3 mutation sense:	5'-TTCCGGAAGGTacgcAAAAGATAATT-3'	
binding site 3 mutation antisense:	5'-AATTATCTTTTgctACCTTCCGGAA-3'	
binding site 2 mutation sense:	5'-CATGGACCAATcgcgAACGCATTTAC-3'	
binding site 2 mutation antisense:	5'-GTAAATGCGTTcgcgATTGGTCCATG-3'	
binding site 1 mutation sense:	5'-AAGGACACCGTAcgtACATACACAAG-3'	
binding site 1 mutation antisense:	5'-CTTGTGTATGTAcgtACGGTGTCTT-3'	
Primers used for ChIP in the ZEB1 promoter:		
binding site 4 sense:	5'-TATCAATAACTGCTACAT-3'	
binding site 4 antisense:	5'-AGACCATACAGATAACCA-3'	
binding site 2,3 sense:	5'-TAGACATAATGTATAGCA-3'	
binding site 2,3 antisense:	5'-TTGTCACCTCTATGGCAG-3'	
binding site 1 sense:	5'-CATAGTCTATGACCTGAT-3'	
binding site 1 antisense:	5'-TGCATTCATCCTCAATGC-3'	
Primers for EGFR promoter construct:		
(-2000/+150)EGFR sense:	5'-TATAGGTACCAAGCTCAAAGGAGAACTCAA-3'	KpnI
(-1113/+150)EGFR sense:	5'-TATAGGTACCACCAGGCCTGAAGGTCCTAGTG-3'	KpnI
(-757/+150)EGFR sense:	5'-TATAGGTACCTTAAGGAGGCCTGTCTCTGC-3'	KpnI
Antisense:	5'-ATATAAGCTTGCACAGGGGGGGCTCTCTGA-3'	HindIII
Primers for EGFR promoter site-directed mutagenesis:		
binding site 2 mutation sense:	5'-CTGGTTGCAATcgctATTAAGGAGGCC-3'	
binding site 2 mutation antisense:	5'-GGCCTCCTTAATAGCGATTGCAACCAG-3'	
binding site 1 mutation sense:	5'-AGATCAGGGTTAcgtAACCAGGCCTG-3'	
binding site 1 mutation antisense:	5'-CAGGCCTGGTTACGTAACCCTGATCT-3'	
Primers used for ChIP in the EGFR promoter:		
binding site 2 sense:	5'-CGGGAGCTACAGGGGCAGTG-3'	
binding site 2 antisense:	5'-GAAATGAGGGCACCAACTC-3'	
binding site 1 sense:	5'-GCAAAGGGCAGGTCTGTAGC-3'	
binding site 1 antisense:	5'-CACTGTTTCCTTCTCCTGCAG-3'	
Primers for FOXK2 promoter construct:		
(-5048/+150)FOXK2 sense:	5'-TATAGGTACCCTTCTTTCTTTTATCTTGGAA-3'	KpnI
(-1834/+150)FOXK2 sense:	5'-ATATGGTACCTCCCCAGCAGAGACTAAATTGG-3'	KpnI
(-746/+150)FOXK2 sense:	5'-ATATGGTACCCATGTTGGCTAGGCTGGTCTC-3'	KpnI
(-176/+150)FOXK2 sense:	5'-ATATGGTACCCGGCGCCCGGAGTCGGGACTGC-3'	KpnI

antisense: 5'-ATATAAAGCTTCAACCCGCTGTCCCACTGGCAG-3' HindIII

Primers for FOXK2 promoter site-directed mutagenesis:

NF- kappa B binding site

binding site 2 mutation sense: 5'-TTAGTAGAGCTtatcgcgagCCATGTTGGCT-3'

binding site 2 mutation antisense: 5'-AGCCAACATGGctcgcgataAGCTCTACTAA-3'

binding site 1 mutation sense: 5'-GACAATTCAGGtataAGTCCCTCCCC-3'

binding site 1 mutation antisense: 5'-GGGGAGGGACTtataCCTGAATTGTC-3'

SP1 binding site

binding site 2 mutation sense: 5'-CTTACTGCAACagtacatgtCCAGGTTGAAG-3'

binding site 2 mutation antisense: 5'-CTTCAACCTGGacatgtactGTTGCAGTAAG-3'

binding site 1 mutation sense: 5'-CTTACTGCAATcgataactCCGCGTTTAAAG-3'

binding site 1 mutation antisense: 5'-CTTAAACGCGGagtatatcgATTGCAGTAAG-3'

Primers used for CHIP in the FOXK2 promoter:

binding site 3 sense: 5'-GCTCCCGCGCCCGGCCAACC-3'

binding site 3 antisense: 5'-CACCAAGCTGCCGACGTCC-3'

binding site 2 sense: 5'-TTGAAGCAATTCTCCTGCCTCC-3'

binding site 2 antisense: 5'-CCAGCACTTTGGGAGGCCGAG-3'

binding site 1 sense: 5'-CACCATGCTGAAAGAAGTGAG-3'

binding site 1 antisense: 5'-TACAACCTCTTAAACAGGAAG-3'

Primers for real-time PCR:

FOXA1 sense: 5'-ACTCGCCTTACGGCTCTACG-3'

FOXA1 antisense: 5'-TGTTTAGGACGGGTCTGGAATA-3'

FOXA2 sense: 5'-AGTATGCTGGGAGCGGTGAA-3'

FOXA2 antisense: 5'-CGGCTTTGCGTGCGTGTA-3'

FOXA3 sense: 5'-ACGCCAAGCCACCGTATT-3'

FOXA3 antisense: 5'-GGTAGCAGCCATTCTCAAACA-3'

FOXB1 sense: 5'-ATGCCTCGGCCCGGCCGAACACGT-3'

FOXB1 antisense: 5'-ACGTGTTGCGGCCGGCCGAGGCAT-3'

FOXB2 sense: 5'-TGCCGCTGAGCGACATCTAC-3'

FOXB2 antisense: 5'-TTCCCGCGTGCAAGTGAGTA-3'

FOXC1 sense: 5'-CAGCATCCGCCACAACCTCT-3'

FOXC1 antisense: 5'-GCAGCCTGTCCTTCTCCTCCT-3'

FOXC2 sense: 5'-AGAACAGCATCCGCCACAACC-3'

FOXC2 antisense: 5'-GGGCTCAGCGTCTCCACCTT-3'

FOXD1 sense: 5'-GGGCGAGGACGAAGAAGACGA-3'

FOXD1 antisense: 5'-GAGGTTGTGGCGGATGCTGTT-3'

FOXD2 sense: 5'-AGGCCGACGCAGACATAGACGT-3'

FOXD2 antisense: 5'-GCCGCTGATGAACTCGCAGAT-3'

FOXD3 sense: 5'-CCCCATCACGGACAGCCTCA-3'

FOXD3 antisense: 5'-GCGGGTCCAGGGTCCAGTAGTT-3'

FOXD4 sense: 5'-CAGCCCTCACTTGGTCTCA-3'

FOXD4 antisense: 5'-CCTCTGCCACCGCCTGATA-3'

FOXE1 sense: 5'-GGTGCTGGCTACCGTGAAGGA-3'

FOXE1 antisense: 5'-GAGGTTGTGGCGGATGCTGTT-3'

FOXE3 sense: 5'-GCAGAACAGCATCCGCCACA-3'
 FOXE3 antisense: 5'-CGGGGTCCAGCGTCCAGTAG-3'
 FOXF1 sense: 5'-CCGCCCTATTCCTACATCGC-3'
 FOXF1 antisense: 5'-CGCCTGGCATTTCCTTCG-3'
 FOXF2 sense: 5'-AGAGCTACTTGCACCAGAACGC-3'
 FOXF2 antisense: 5'-CCCATTGAAGTTGAGGACGAAA-3'
 FOXG1 sense: 5'-ATGCTGGACATGGGAGATAGGAAAG-3'
 FOXG1 antisense: 5'-CTTTCCTATCTCCCATGTCCAGCAT-3'
 FOXH1 sense: 5'-TCCCTCCCGCAGACTGAA-3'
 FOXH1 antisense: 5'-GGAAAGTTGTGGCGAATG-3'
 FOXI1 sense: 5'-GCCACAACCTGTCGCTCAA-3'
 FOXI1 antisense: 5'-CGCTCACATAGGCTGTCATAGA-3'
 FOXI2 sense: 5'-GCCACAACCTGTCGCTCAAC-3'
 FOXI2 antisense: 5'-CCGCCTCCCGAAAGAAAA-3'
 FOXI3 sense: 5'-GGCGACAACCTTCGGAGTGTATT-3'
 FOXI3 antisense: 5'-GCCGCACCATCTTCATCAG-3'
 FOXJ1 sense: 5'-CCTACTCGTATGCCACGCTCAT-3'
 FOXJ1 antisense: 5'-CTTGCCTGGTTCGTCCTTCTC-3'
 FOXJ2 sense: 5'-ACAGGCAGAGCAGAAGAAC-3'
 FOXJ2 antisense: 5'-AGTCGAAGTCATCAGGGATC-3'
 FOXJ3 sense: 5'-ACTCCTTGACCCAAATAC-3'
 FOXJ3 antisense: 5'-CTTCCCGAAATAATACTACT-3'
 FOXK1 sense: 5'-TCCAGTTCACGTCGCTCTATCA-3'
 FOXK1 antisense: 5'-GGCGGCTTTGACTCATCCTT-3'
 FOXK2 sense: 5'-TGATAGTTCAGGCGATTACGA-3'
 FOXK2 antisense: 5'-TGGATGACAGCGAGTTTGG-3'
 FOXL1 sense: 5'-CATCCTGGCGGAAAGCA-3'
 FOXL1 antisense: 5'-GAGTACCGTGTCCGGAAACCTG-3'
 FOXL2 sense: 5'-TCGCACAGTCAAGGAGCCAGAA-3'
 FOXL2 antisense: 5'-CGGGTCCAGCGTCCAGTAGTT-3'
 FOXM1 sense: 5'-AACTCCATCCGCCACAAC-3'
 FOXM1 antisense: 5'-TCGCCACTAAAGAACTTACTCA-3'
 FOXN1 sense: 5'-CCTGGGTTTCAGAGGTCAAAG-3'
 FOXN1 antisense: 5'-GGAAAGTGCTCCGTCATAAAA-3'
 FOXN2 sense: 5'-TGTATGACATAGAGGGAGAT-3'
 FOXN2 antisense: 5'-GTGCCTGGATAAGATTGG-3'
 FOXN3 sense: 5'-ATTGGGAAAGGGTTCGTTG-3'
 FOXN3 antisense: 5'-TGAGGACAGGTGGGAGGT-3'
 FOXN4 sense: 5'-CGCCGCATCTGTACTCACC-3'
 FOXN4 antisense: 5'-CGCTCATCTGTTCTCCACCT-3'
 FOXO1 sense: 5'-CGCCGTGCTACTCGTTTG-3'
 FOXO1 antisense: 5'-AGCATGTCCAGGGTGGGT-3'
 FOXO3 sense: 5'-ACGGTCACTCTGTCCCA-3'
 FOXO3 antisense: 5'-TGCTGGCGTTAGAATTGGT-3'

FOXO4 sense:	5'-AGCAGGATGGAAGAACTCG-3'
FOXO4 antisense:	5'-TGGCAGCACAGATGGTTT-3'
FOXP1 sense:	5'-GAAGTCTACAGAACCCAAAG-3'
FOXP1 antisense:	5'-GTCGGAAGTAAGCAAACA-3'
FOXP2 sense:	5'-GGAAAGCAAGCGAAAGAG-3'
FOXP2 antisense:	5'-ATGGAGATGAGTCCCTGA-3'
FOXP3 sense:	5'-GAAGGGCAGGGCACAATG-3'
FOXP3 antisense:	5'-GATGAGCGTGGCGTAGGT-3'
FOXP4 sense:	5'-GGACGAGCGGGAGTATCAG-3'
FOXP4 antisense:	5'-GGAGCCAGGGTTCAGCAT-3'
FOXQ1 sense:	5'-AGGGCGACGGCGAACAGAGT-3'
FOXQ1 antisense:	5'-AAGGTTGTGGCGCACGGAGTT-3'
FOXR1 sense:	5'-CGGAAGGCTGGAAGAATA-3'
FOXR1 antisense:	5'-AAGAGGAAGGGCATCACAT-3'
FOXR2 sense:	5'-GAATGAGTTATTTCTGCCTTGT-3'
FOXR2 antisense:	5'-GGGAATGGATAACCTGCT-3'
FOXS1 sense:	5'-TGGACGCTGGACCCTGACT-3'
FOXS1 antisense:	5'-TGCTCCCGATGCCTGATT-3'
ZEB1 sense:	5'-GCCAACAGACCAGACAGTGTT-3'
ZEB1 antisense:	5'-TCTTGCCCTTCCTTTCCTG-3'
E-cadherin sense:	5'-TCACATCCTACACTGCCCAG-3'
E-cadherin antisense:	5'-AGTGTCCCTGTTCCAGTAGC-3'
Vimentin sense:	5'-AATAAGATCCTGCTGGCCGA-3'
Vimentin antisense:	5'-GGTGTTCGCTTCCTCTC-3'
EGFR sense:	5'-GAGCCTCTGGATGGTGCAAT-3'
EGFR antisense:	5'-GAGCCTCTGGATGGTGCAAT-3'
GAPDH sense:	5'-GCACCGTCAAGGCTGAGAAC-3'
GAPDH antisense:	5'-TGGTGAAGACGCCAGTGGA-3'

Table S6. Knockdown shRNA sequences used in this study.

TRC number	Sequence
FOXK2	
TRCN0000016291	CCGGCCAGCCTCTGAAAGCAAATTACTCGAGTAATTTGCTTTCAGAGGCTGGTTTTT
TRCN0000016290	CCGGCCCAGACACAAACATCAAGATCTCGAGATCTTGATGTTTGTGCTCGGGTTTTT
ZEB1	
TRCN0000369266	CCGGCTGAACCTCAGACCTAGTAATCTCGAGATTACTAGGTCTGAGGTTTCAGTTTTTG
TRCN0000369267	CCGGTGTCTCCATAAGTATCAATTCTCGAGAATTGATACTTATGGGAGACATTTTTG
EGFR	
TRCN0000295969	CCGGGCCACAAAGCAGTGAATTTATCTCGAGATAAATTCAGTCTTTGTGGCTTTTTG
TRCN0000121068	CCGGGCCACAAAGCAGTGAATTTATCTCGAGATAAATTCAGTCTTTGTGGCTTTTTG

Table S7. List of genes differentially expressed in Caco-2-FOXK1 versus Caco-2-control cells using a human EMT PCR array

Symbol	Caco-2-FOXK1 vs Caco-2-control	Description
FN1	4.83	Fibronectin 1
TWIST1	4.59	Twist homolog 1 (Drosophila)
SNAI2(Slug)	4.28	Snail homolog 2 (Drosophila)
SNAI1	4.23	Snail homolog 1 (Drosophila)
VIM	4.17	Vimentin
MMP9	4.12	Matrix metalloproteinase 9
BMP1	3.73	Bone morphogenetic protein 1
SNAI3	3.57	Snail homolog 3 (Drosophila)
TGFB2	2.96	Transforming growth factor, beta 2
ITGB1	2.93	Integrin, beta 1
FZD7	2.69	Frizzled family receptor 7
TGFB1	2.66	Transforming growth factor, beta 1
ITGA5	2.47	Integrin, alpha 5
WNT5B	2.36	Wingless-type MMTV integration site family, member 5B
STAT3	2.36	Signal transducer and activator of transcription 3
FGFBP1	2.35	Fibroblast growth factor binding protein 1
IGFBP4	2.32	Insulin-like growth factor binding protein 4
VCAN	2.27	Versican
CTNNB1(b-catenin)	2.21	Catenin (cadherin-associated protein), beta 1, 88kDa
MMP3	2.08	Matrix metalloproteinase 3
PLEK2	1.99	Pleckstrin 2
ERBB3	1.98	V-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)
GSK3B	1.97	Glycogen synthase kinase 3 beta
KRT7	1.93	Keratin 7
ITGAV	1.86	Integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51)
MMP2	1.81	Matrix metalloproteinase 2
CALD1	1.79	Caldesmon 1
ILK	1.74	Integrin-linked kinase
ZEB2	1.74	Zinc finger E-box binding homeobox 2
F11R	1.72	F11 receptor, junction adhesion molecule A
B2M	1.72	Beta-2-microglobulin
TCF3	1.59	Transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)
PDGFRB	1.59	Platelet-derived growth factor receptor, beta polypeptide
ZEB1	1.57	Zinc finger E-box binding homeobox 1
AKT1	1.56	V-akt murine thymoma viral oncogene homolog 1
OCLN	1.55	Occludin
MAP1B	1.53	Microtubule-associated protein 1B
FOXC2	1.52	Forkhead box C2
BMP7	1.52	Bone morphogenetic protein 7
WNT5A	1.48	Wingless-type MMTV integration site family, member 5A

COL3A1	1.46	Collagen, type III, alpha 1
RGS2	1.45	Regulator of G-protein signaling 2, 24kDa
GNG11	1.45	Guanine nucleotide binding protein (G protein), gamma 11
COL5A2	1.45	Collagen, type V, alpha 2
TCF4	1.43	Transcription factor 4
WNT11	1.42	Wingless-type MMTV integration site family, member 11
MST1R(ROK)	1.41	Macrophage stimulating 1 receptor (c-met-related tyrosine kinase)
STEAP1	1.39	Six transmembrane epithelial antigen of the prostate 1
COL1A2	1.39	Collagen, type I, alpha 2
BMP2	1.39	Bone morphogenetic protein 2
PTK2(FAK)	1.35	PTK2 protein tyrosine kinase 2
VPS13A	1.33	Vacuolar protein sorting 13 homolog A (<i>S. cerevisiae</i>)
EGFR	1.32	Epidermal growth factor receptor
SMAD2	1.29	SMAD family member 2
NODAL	1.29	Nodal homolog
SERPINE1	1.14	Serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1
HPRT1	1.08	Hypoxanthine phosphoribosyltransferase 1
CAV2	-1.04	Caveolin 2
KRT14	-1.05	Keratin 14
AHNAK	-1.11	AHNAK nucleoprotein
SOX10	-1.23	SRY (sex determining region Y)-box 10
TSPAN13	-1.25	Tetraspanin 13
MSN	-1.31	Moesin
GSC	-1.34	Goosecoid homeobox
TIMP1	-1.37	TIMP metalloproteinase inhibitor 1
DESI1	-1.45	PPPDE peptidase domain containing 2
KRT19	-1.56	Keratin 19
ESR1(ERa)	-1.58	Estrogen receptor 1
DSP	-1.65	Desmoplakin
NUDT13	-1.66	Nudix (nucleoside diphosphate linked moiety X)-type motif 13
PTP4A1(PRL1)	-1.82	Protein tyrosine phosphatase type IVA, member 1
RAC1	-1.88	Ras-related C3 botulinum toxin substrate 1
CAMK2N1	-1.94	Calcium/calmodulin-dependent protein kinase II inhibitor 1
TMEM132A	-1.98	Transmembrane protein 132A
TGFB3	-2.13	Transforming growth factor, beta 3
IL1RN	-2.18	Interleukin 1 receptor antagonist
JAG1	-2.19	Jagged 1
TFPI2	-2.24	Tissue factor pathway inhibitor 2
GEMIN2	-2.37	Survival of motor neuron protein interacting protein 1
SPARC	-2.55	Secreted protein, acidic, cysteine-rich (osteonectin)
CDH2	-2.66	Cadherin 2, type 1, N-cadherin (neuronal)
SPP1(Osteopontin)	-2.69	Secreted phosphoprotein 1
NOTCH1	-2.93	Notch 1
DSC2	-3.17	Desmocollin 2

TMEFF1	-3.28	Transmembrane protein with EGF-like and two follistatin-like domains 1
CDH1	-3.75	Cadherin 1, type 1, E-cadherin (epithelial)

Figure S1. Real-time PCR analysis of the expression of FOX family members in normal colorectal epithelial tissues (n = 29) and colorectal cancer tissues (n = 250).

Figure S2. Real-time PCR analysis of the expression of FOX family members in colorectal cancer tissues and paired normal colorectal epithelial tissues (n = 50).

Figure S3. Expression of FOXK2 in publicly available cancer datasets

(A) Representative data extracted from The Cancer Genome Atlas (TCGA) dataset showing the relative mRNA expression of FOXK2 in multiple cancer versus normal tissues. The box-and-whisker plots indicate the median (horizontal line), interquartile range (box) and minimum and maximum (whiskers) of the data. COAD, colon adenocarcinoma; READ, rectum adenocarcinoma; STAD, stomach adenocarcinoma; LIHC, liver hepatocellular carcinoma; ESCA, esophageal carcinoma; LUAD, lung adenocarcinoma. * $P < 0.05$; ** $P < 0.01$.

(B) The mRNA expression of FOXK2 in multiple cancer and normal tissues in different clinical cohorts. Datasets were extracted from the Oncomine (Skrzypczak Colon, Zou Colon, Hong Colorectal, Gaedcke Colorectal, Cho Gastric, Wurmbach Liver, Su Esophagus, and Landi Lung) databases. The box-and-whisker plots indicate the median (horizontal line), interquartile range (box) and minimum and maximum (whiskers) of the data. COAD, colon adenocarcinoma; READ, rectum adenocarcinoma; STAD, stomach adenocarcinoma; LIHC, liver hepatocellular carcinoma; ESCA, esophageal carcinoma; LUAD, lung adenocarcinoma. * $P < 0.05$; ** $P < 0.01$.

Figure S4. Expression of FOXK2 and EGFR in different CRC cell lines.

(A) Western blotting analysis (n = 3) of the expression of FOXK2 and EGFR in the indicated CRC cells.

(B) Correlation between FOXK2 and EGFR expression in the CRC cell lines described.

Figure S5.

(A) LoVo cells were infected with LV-shFOXK2-1, LV-shFOXK2-2 and lentivirus expressing shRNA-resistant FOXK2-WT. Left, Western blotting analysis was used to analyze endogenous FOXK2 and Flag-tagged FOXK2 levels. Right, Transwell analysis of migration and invasion in the indicated cell lines. n = 3 independent experiments performed in triplicate. The data are presented as the mean \pm s.d.

(B) Caco-2-FOXK2 cells were infected with LV-shZEB1-1, LV-shZEB1-2 and lentivirus expressing shRNA-resistant ZEB1-WT. Left, Western blotting analysis was used to analyze endogenous ZEB1 and Flag-tagged ZEB1 levels. Right, Transwell analysis of migration and invasion in the indicated cell lines. n = 3 independent experiments performed in triplicate. The data are presented as the mean \pm s.d.

(C) Caco-2-FOXK2 cells were infected with LV-shEGFR-1, LV-shEGFR-2 and lentivirus expressing shRNA-resistant EGFR-WT. Left, Western blotting analysis was used to analyze endogenous EGFR and Flag-tagged EGFR levels. Right, Transwell analysis of migration and invasion in the indicated cell lines. n = 3 independent experiments performed in triplicate. The data are presented as the mean \pm s.d.

Figure S6.

(A) Transwell assay analysis of the indicated cell migration and invasion. n = 3 independent experiments performed in triplicate. ** $P < 0.01$ compared with the control. The data are presented as the mean \pm s.d.

(B) Caco-2 and Caco-2-FOXK2 cells were transfected with lentivirus LV-shcontrol and LV-shEGFR. Left, Western blotting analysis was used to analyze endogenous EGFR and Flag-tagged EGFR levels. Right, Transwell analysis of migration and invasion in the indicated cell lines. $n = 3$ independent experiments performed in triplicate. * $P < 0.05$ compared with the control. The data are presented as the mean \pm s.d.

(C) LoVo and LoVo-shFOXK2 cells were transfected with lentivirus LV-control and LV-EGFR. Left, Western blotting analysis was used to analyze endogenous EGFR and Flag-tagged EGFR levels. Right, Transwell analysis of migration and invasion in the indicated cell lines. $n = 3$ independent experiments performed in triplicate. * $P < 0.05$ compared with the control. The data are presented as the mean \pm s.d.

Figure S7.

(A) ChIP-qPCR assays demonstrated that FOXK2 directly binds to the ZEB1 promoter in LoVo cells. $n = 3$ independent experiments performed in triplicate. * $P < 0.05$ compared with the control. The data are presented as the mean \pm s.d.

(B) ChIP-qPCR assays demonstrated that FOXK2 directly binds to the EGFR promoter in LoVo cells. $n = 3$ independent experiments performed in triplicate. * $P < 0.05$ compared with the control. The data are presented as the mean \pm s.d.

(C) Western blotting analysis of FOXK2 protein levels in six cases of CRC and paired adjacent nontumor tissues. The experiments were repeated independently at least three times.

Figure S8. Caco-2 cells were treated with EGF 24 hours after transfection with siRNAs against AKT (siAKT1+siAKT2+siAKT3), JNK (siJNK1+siJNK2+siJNK3), ERK (siERK1+siERK2) and p38 (sip38 α +sip38 β).

(A) Western blotting was used to quantify the protein levels of FOXK2, AKT, JNK, ERK and p38 MAPK. n = 3 independent experiments performed in triplicate.

(B) ChIP assays demonstrated that EGF induced NF- κ B to directly bind to the FOXK2 promoter through the ERK pathway. n = 3 independent experiments performed in triplicate. * $P < 0.05$ compared with the control. The data are presented as the mean \pm s.d.

Figure S9.

(A) FOXK2 downregulation attenuated the enhanced cell migration and invasion induced by EGF overexpression. Left, Western blotting analysis of the expression of EGF and FOXK2 in the indicated CRC cells. Right, Transwell analysis of migration and invasion in the indicated cell lines. n = 3 independent experiments performed in triplicate. * $P < 0.05$ compared with the control. The data are presented as the mean \pm s.d.

(B) FOXK2 overexpression rescued the inhibited cell migration and invasion induced by EGF downregulation. Left, Western blotting analysis of the expression of EGF and FOXK2 in the indicated CRC cells. Right, Transwell analysis of migration and invasion in the indicated cell lines. n = 3 independent experiments performed in triplicate. * $P < 0.05$ compared with the control. The data are presented as the mean \pm s.d.

Figure S10. ZEB1 is a downstream gene of EGFR.

(A) Cells were infected with indicated lentivirus. Western blotting analysis was used to analyze EGFR and ZEB1 protein levels 48h after infection.

(B) Transwell analysis of migration and invasion in the indicated cell lines.

Figure S11.

(A) Overlapping of differentially upregulated genes (log 2 fold change > 1) between FOXK1 and FOXK2 overexpression in CRC cells.

(B) Western blotting was used to quantify the protein levels of FOXK1, FOXK2, E-cadherin, vimentin, Slug and ZEB1. The experiments were repeated independently at least three times.

(C) Transwell analysis of migration and invasion in the indicated cell lines. The experiments were repeated independently at least three times. * $P < 0.05$ compared with the control. The data are presented as the mean \pm s.d.

Figure S12.

(A) The normal colonic epithelial cell line NCM460 and the CRC cell line Caco-2 were transfected with FOXK2 (LV-FOXK2) and control (LV-control) lentivirus for 48 hours. Then, luciferase reporter assay, qRT-PCR and Western blotting were performed to detect the expression levels of EGFR. n = 3 independent experiments performed in triplicate. * $P < 0.05$ compared with the control. The data are presented as the mean \pm s.d.

(B) Cells isolated from CRC tissues and normal colonic epithelial tissues were transfected with FOXK2 (LV-FOXK2) and control (LV-control) lentivirus for 48 hours. Then, luciferase reporter assay, qRT-PCR and Western blotting were performed to detect the expression levels of EGFR. n = 3 independent experiments performed in triplicate. * $P < 0.05$ compared with the control. The data are presented as the mean \pm s.d.

(C) The normal colonic epithelial cell line NCM460 and the CRC cell line Caco-2 were incubated with or without EGF (100 ng/mL) for 24 hours. Then, luciferase reporter assay, qRT-PCR and Western blotting were performed to detect the expression levels of FOXK2. n = 3 independent experiments performed in triplicate. * $P < 0.05$ compared with the control. The data are presented as the mean \pm s.d.

(D) Cells isolated from CRC tissues and normal colonic epithelial tissues were exposed to EGF (100 ng/mL) or vehicle for 24 hours. Then, luciferase reporter assay, qRT-PCR and Western blotting were performed to detect the expression levels of FOXK2. $n = 3$ independent experiments performed in triplicate. * $P < 0.05$ compared with the control. The data are presented as the mean \pm s.d.

Figure S1

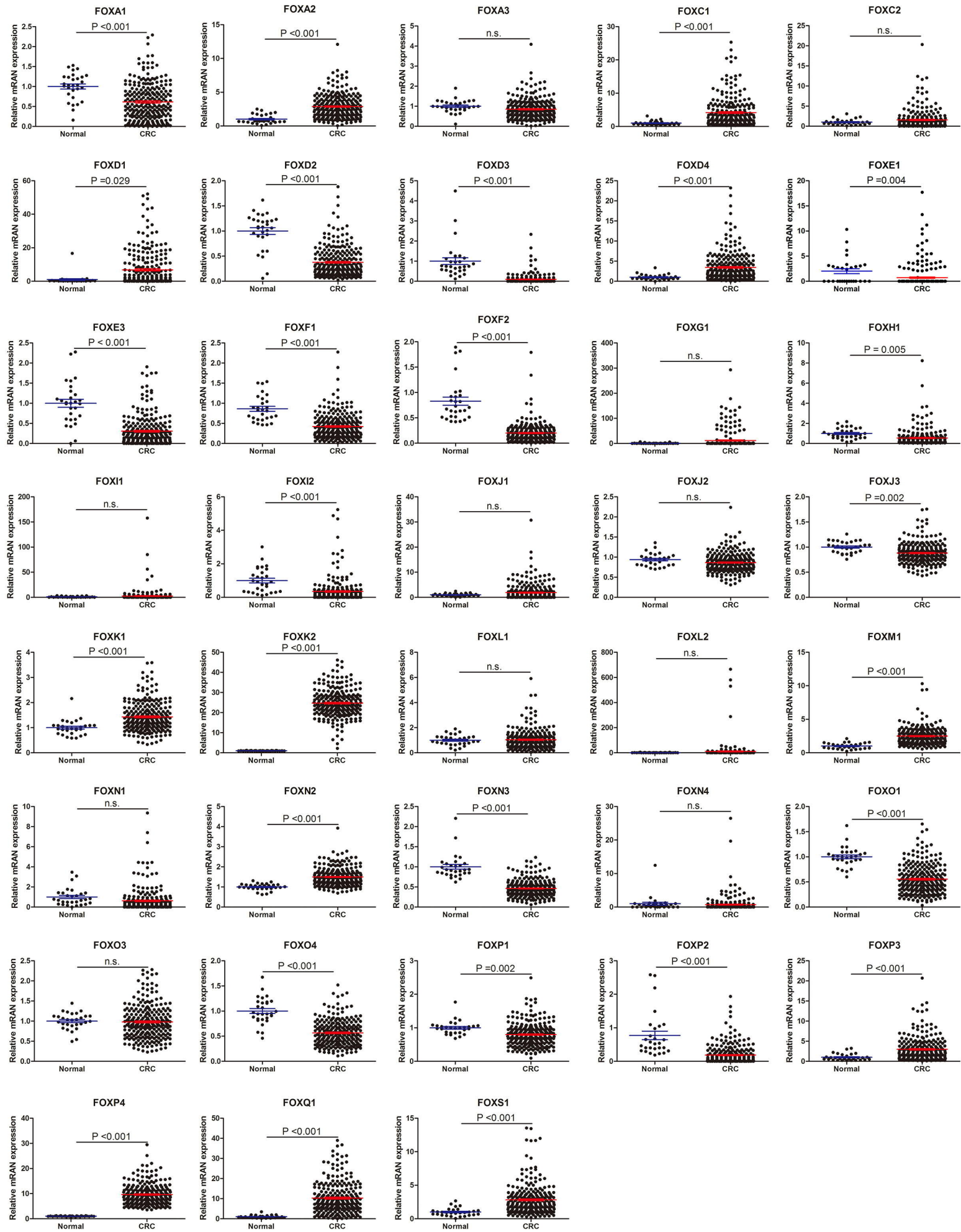


Figure S2

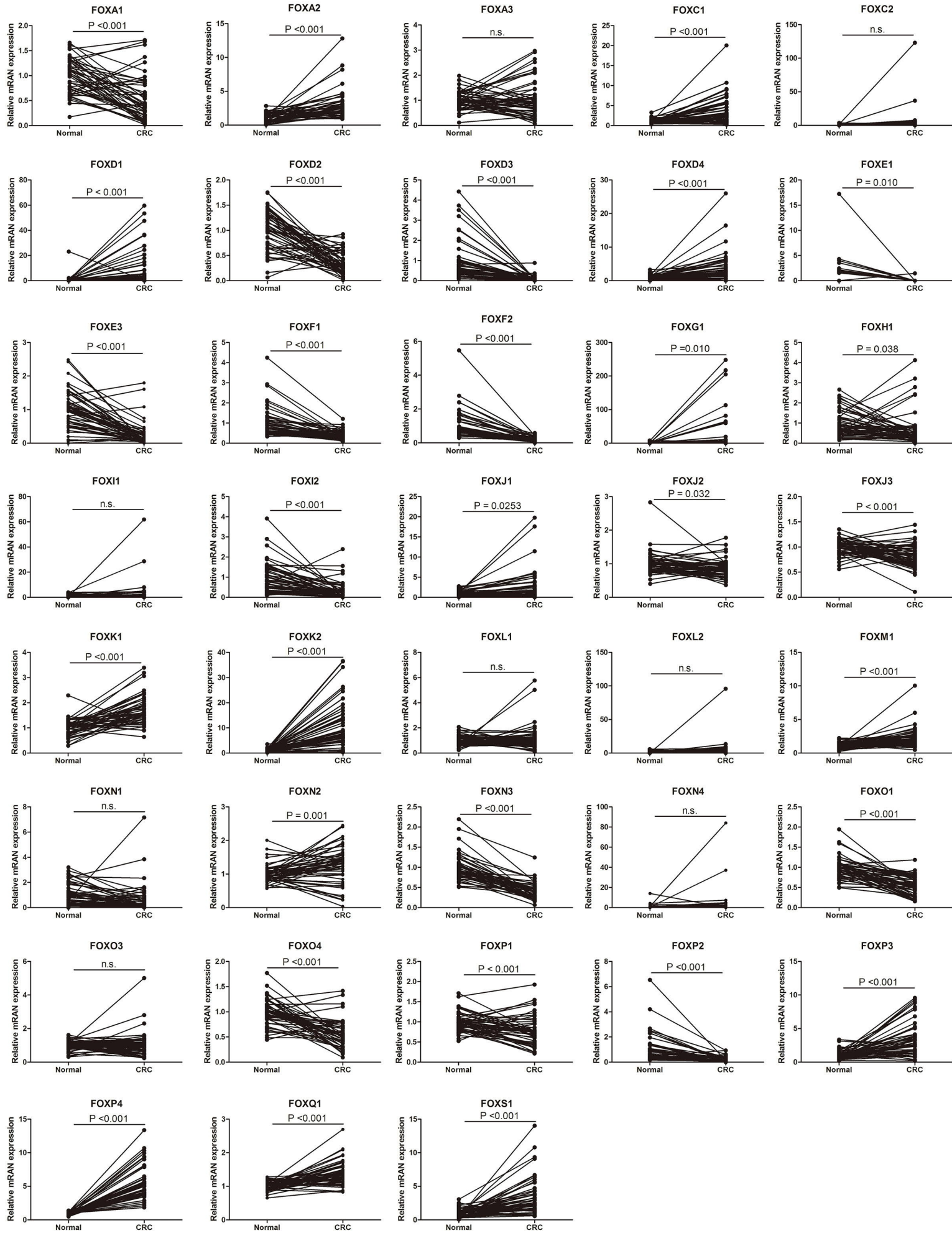


Figure S3

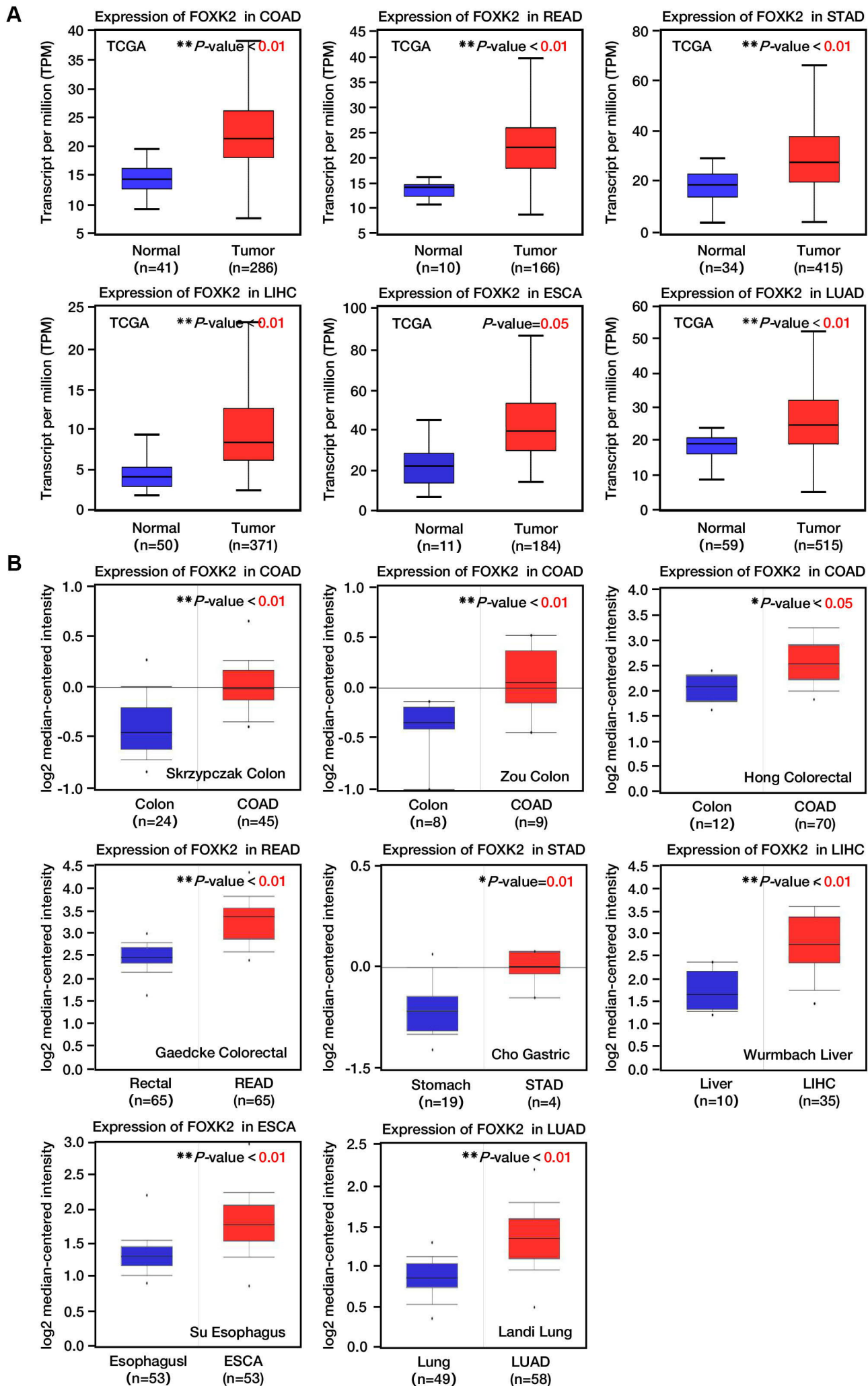


Figure S4

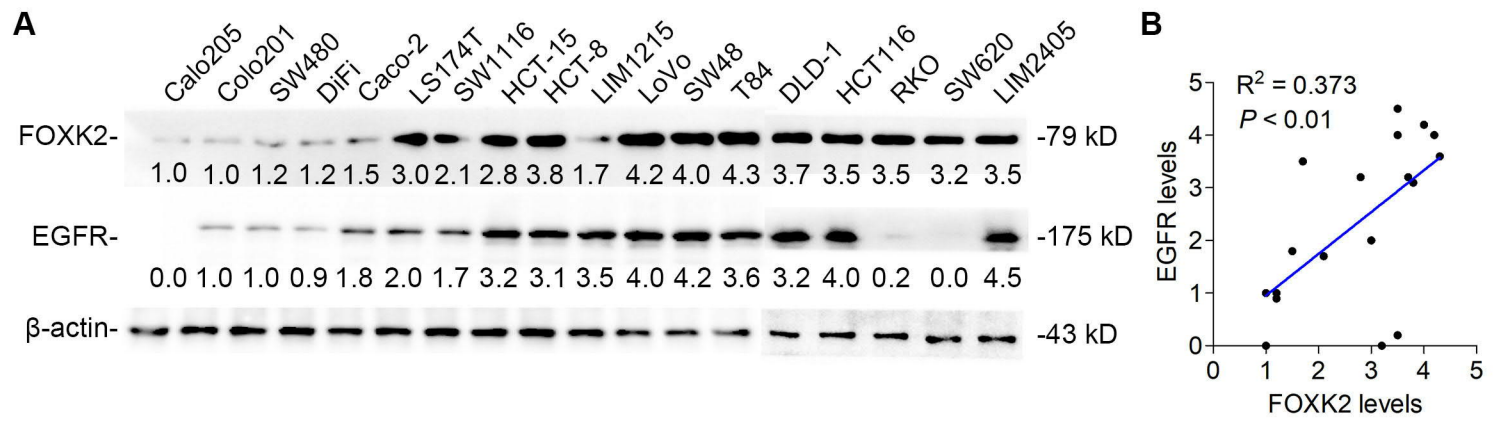
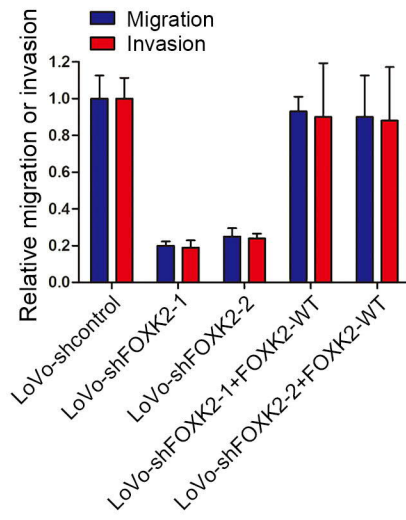
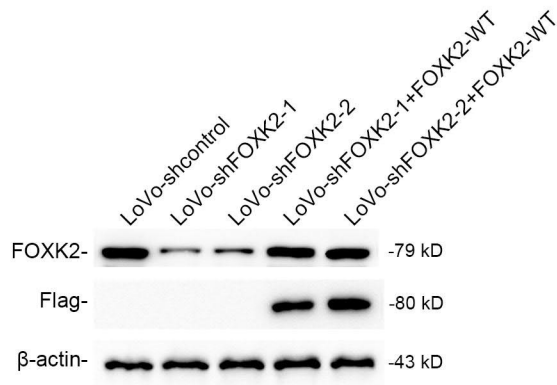
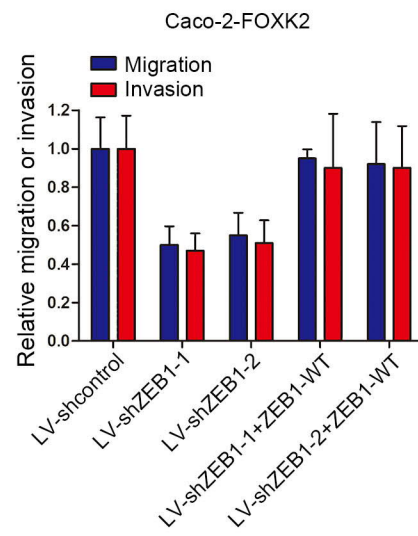
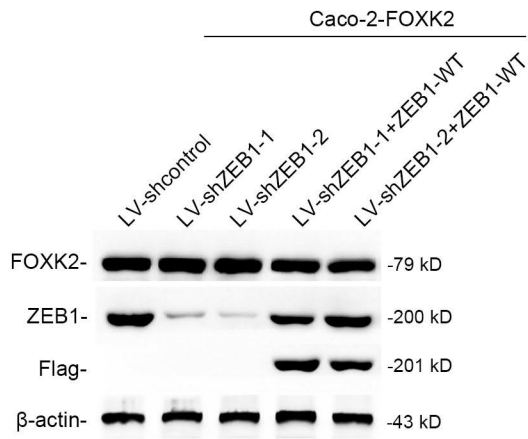


Figure S5

A



B



C

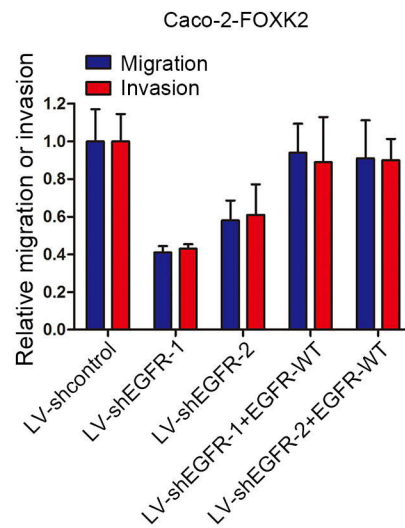
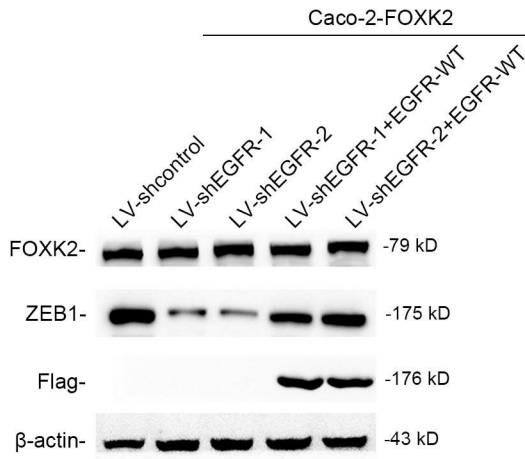


Figure S6

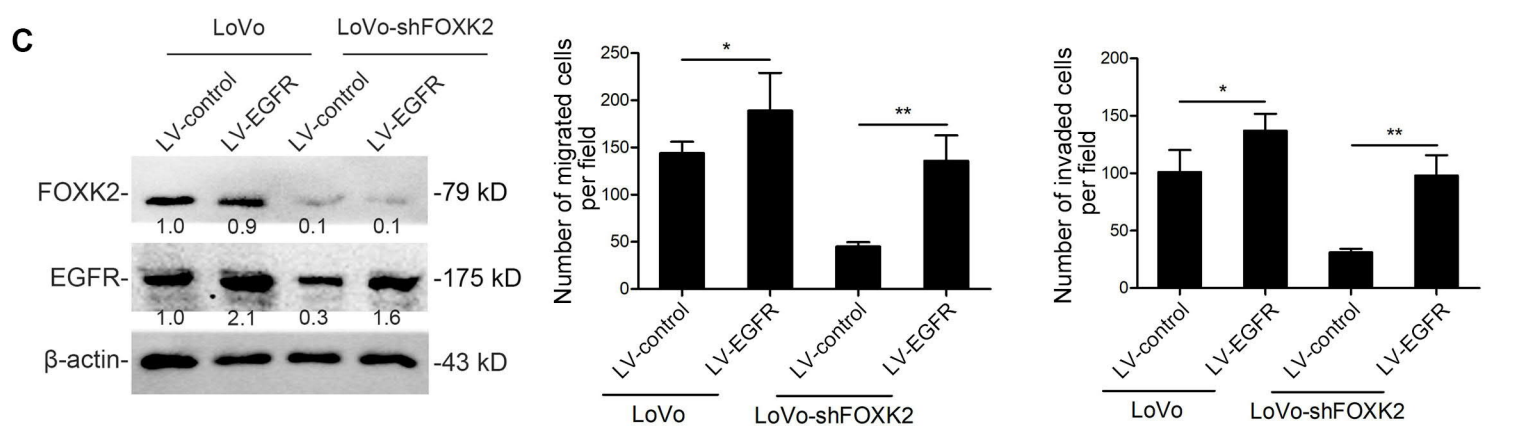
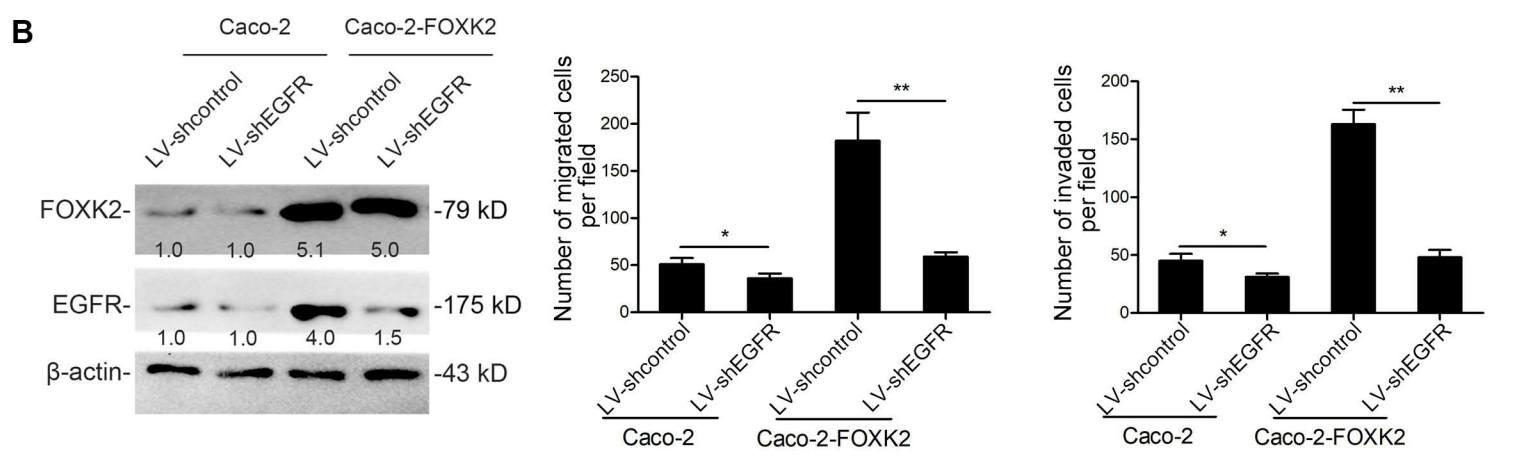
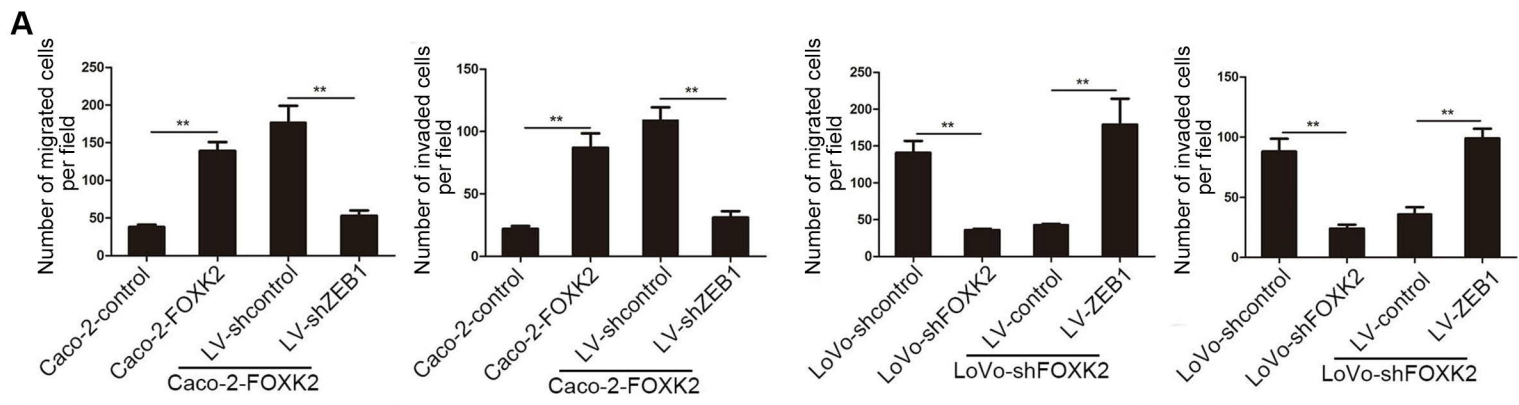


Figure S8

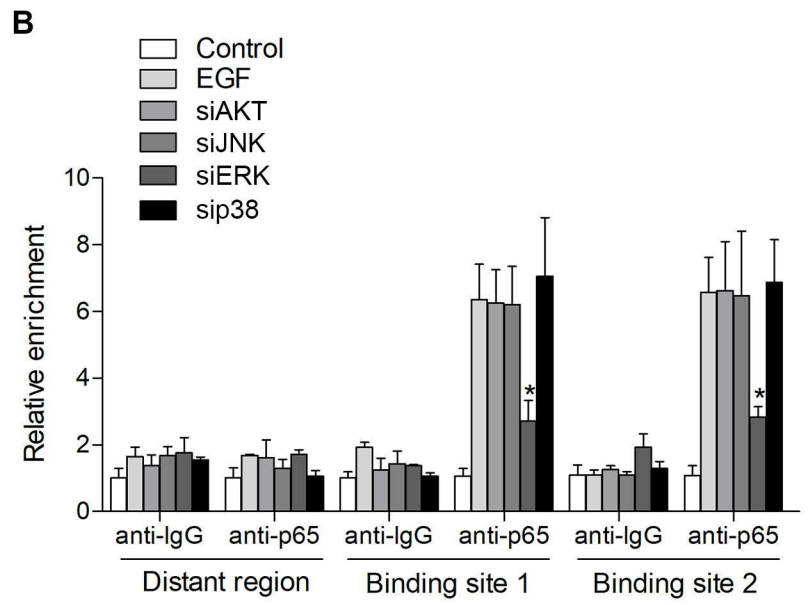
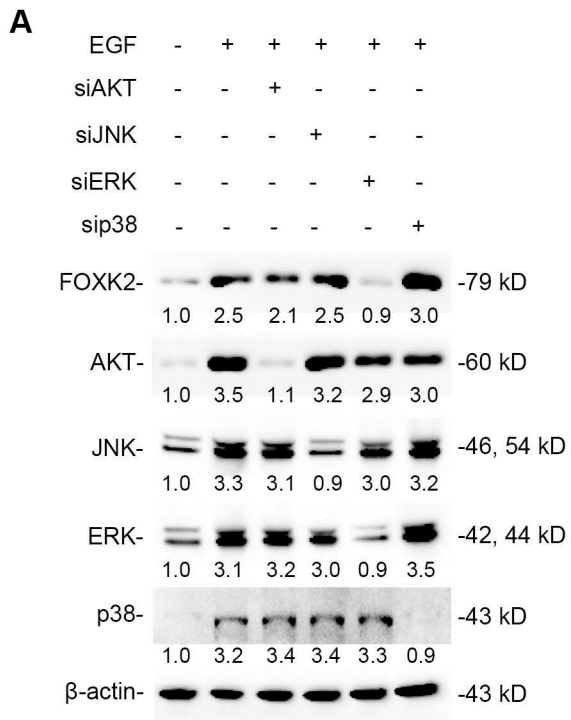
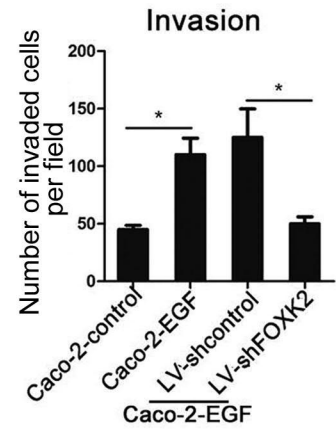
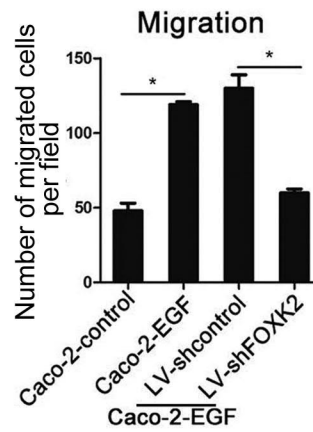
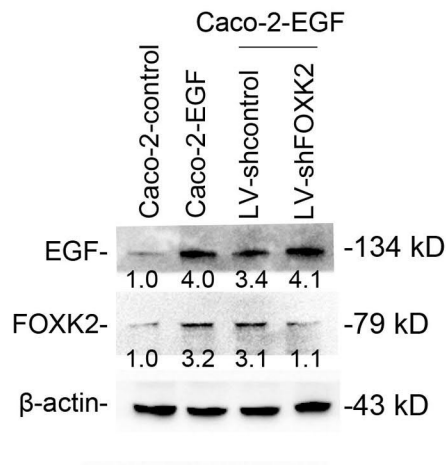


Figure S9

A



B

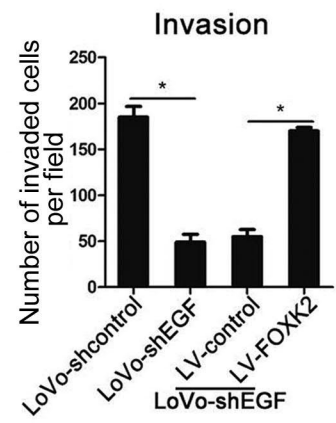
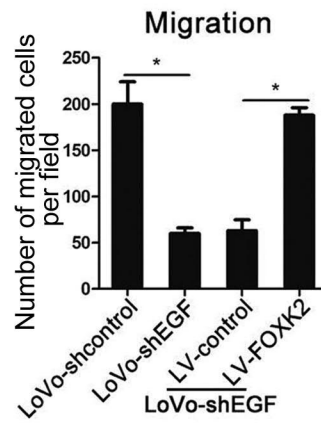
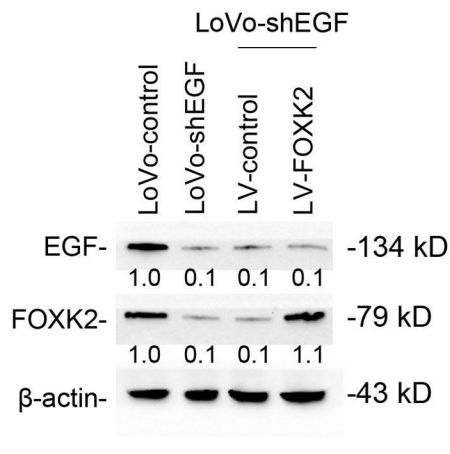


Figure S10

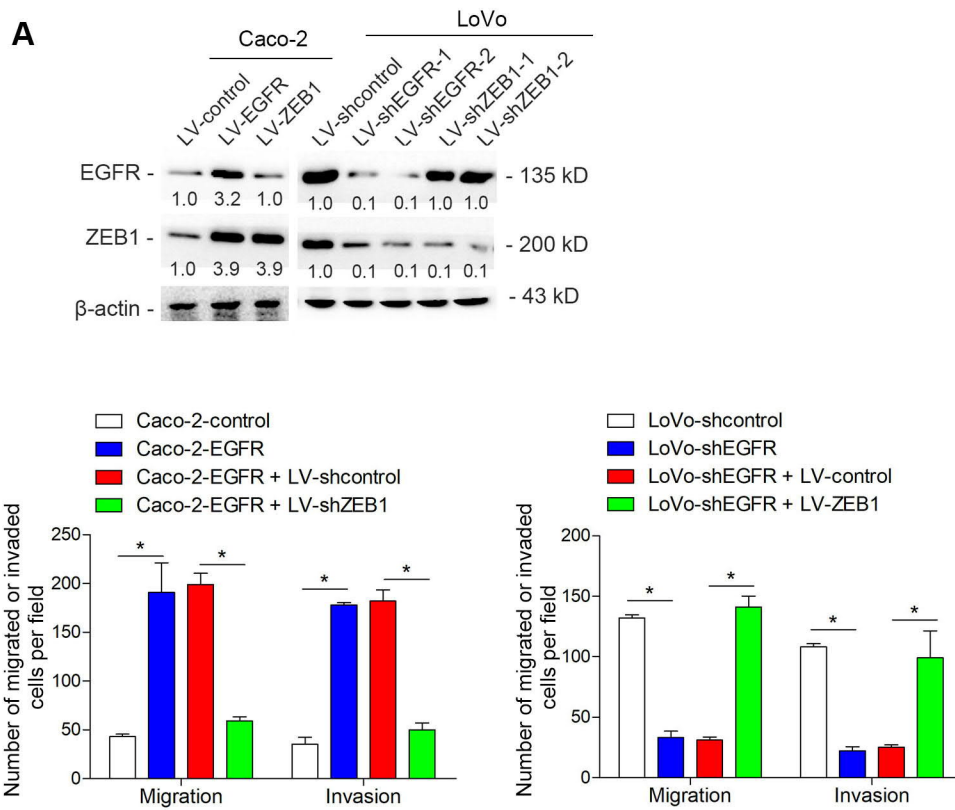


Figure S11

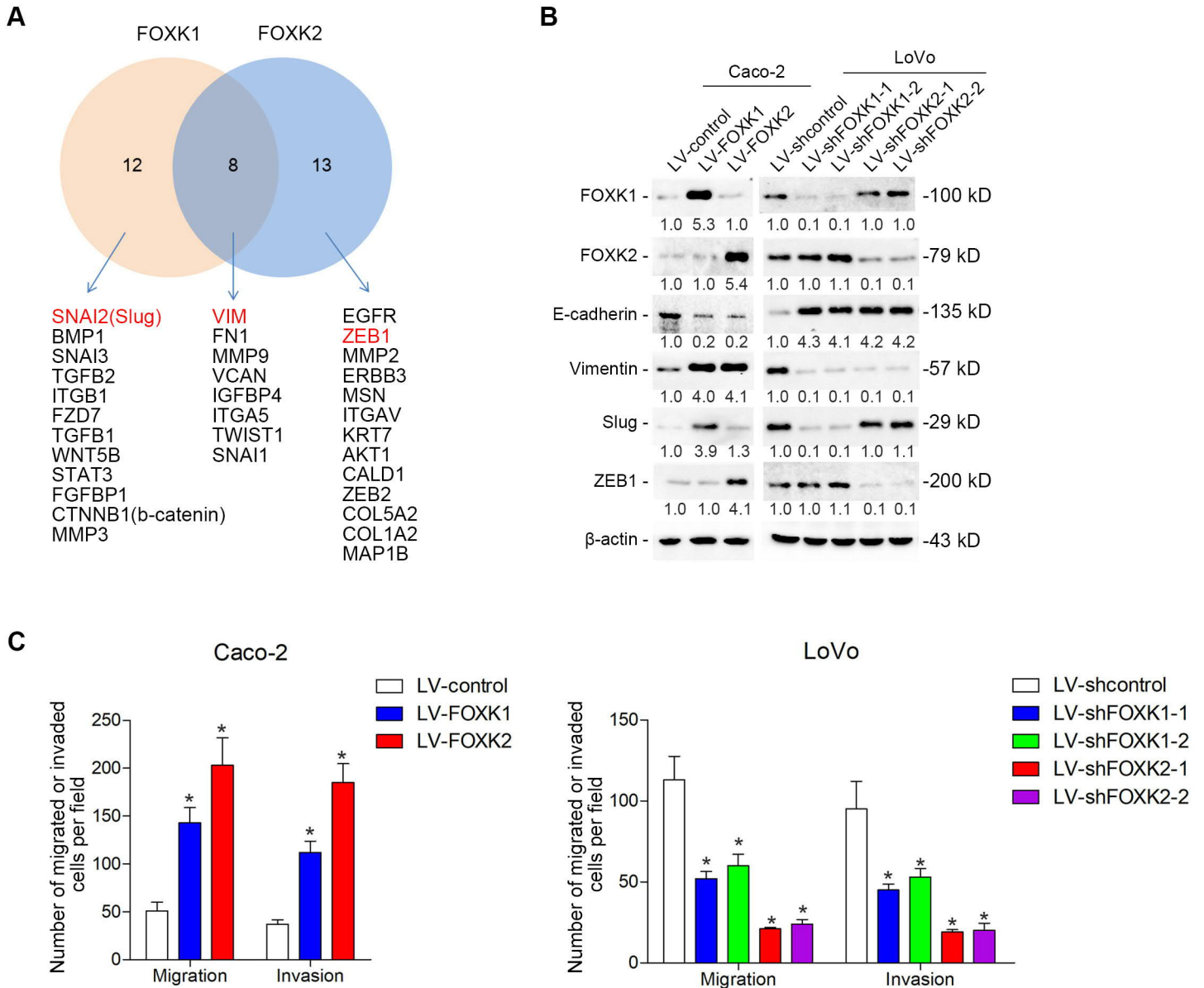


Figure S12

