

Supplementary Figure 1. (A-E) GSEA of microarray data from S18, S26, CNE2 WT and CNE2 OVOL2-KO cells. (F) qPCR analysis of EMT makers in S18 and S26 cells. (G) GSEA plot of microarray data from CNE2 WT and OVOL2-KO cells reveals an enrichment of gene signatures associated with invasion.

Supplementary Figure 2. (A) T7E1 assay to determine the cleavage efficiency of sgRNAs targeting the human *OVOL2* locus. (B and C) CRISPR-induced indels in HNE1 and CNE2 OVOL2-knockout cells were identified by direct Sanger sequencing.

Supplementary Figure 3. (A) Correlation analysis of OVOL2 and EMT-related proteins in 1036 cell lines from the CCLE database (two-tailed Pearson correlation), $P < 0.001$. (B) Pearson correlation between ZEB1 and OVOL2 mRNA levels in 1036 cancer cell lines from the CCLE database. (C) CNE2 WT, CNE2 OVOL2-KO and S18 cells were transfected with control or ZEB1-specific siRNA; after 72 h, the cells were harvested to analyze the expression of EMT-related proteins. (D) qPCR of EMT-related markers in the cells in c. (E) Overexpression of OVOL2 in NIH-3T3 cells was not able to restore E-cadherin expression. NIH-3T3 cells were infected with control lentivirus or lentivirus carrying OVOL2, and EMT marker expression and cell morphology were examined 72 h post-infection. GFP was used to indicate the transduction efficiency.

Supplementary Figure 1

A

Gene Set Enrichment Analysis in S18 vs. S26: Hallmarks

RANK	NAME	NES	FDR q-val	FWER p-val
1	HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	2.5778787	0	0
2	HALLMARK_TNFA_SIGNALING_VIA_NFKB	2.1917176	0	0
3	HALLMARK_HYPOXIA	2.0717003	0	0
4	HALLMARK_UV_RESPONSE_DN	2.0534444	0	0
5	HALLMARK_IL6_JAK_STAT3_SIGNALING	2.0296838	0	0

B

Gene Set Enrichment Analysis in CNE2 OVOL2 KO vs. WT: EMT

RANK	NAME	NES	FDR q-val	FWER p-val
1	ANASTASSIOU_MULTICANCER_INVASIVENESS_SIGNATURE	1.873348	0.003653846	0.003
2	JECHLINGER_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_UP	1.8166465	0.003044456	0.005
3	LIEN_BREAST_CARCINOMA_METAPLASTIC_VS_DUCTAL_UP	1.8087697	0.002029637	0.005
4	LEF1_UP.V1_DN	1.7329576	0.006408124	0.022
5	LEF1_UP.V1_UP	1.6925118	0.007035764	0.029

C

Gene Set Enrichment Analysis in CNE2 OVOL2 KO vs. WT: Hallmarks

RANK	NAME	NES	FDR q-val	FWER p-val
1	HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	2.3770862	0	0
2	HALLMARK_INTERFERON_ALPHA_RESPONSE	2.3119333	0	0
3	HALLMARK_TNFA_SIGNALING_VIA_NFKB	2.1534803	0	0
4	HALLMARK_HYPOXIA	2.119335	0	0
5	HALLMARK_INTERFERON_GAMMA_RESPONSE	2.0551329	0	0

D

Gene Set Enrichment Analysis in S18 vs. CNE2 OVOL2 KO: Hallmarks

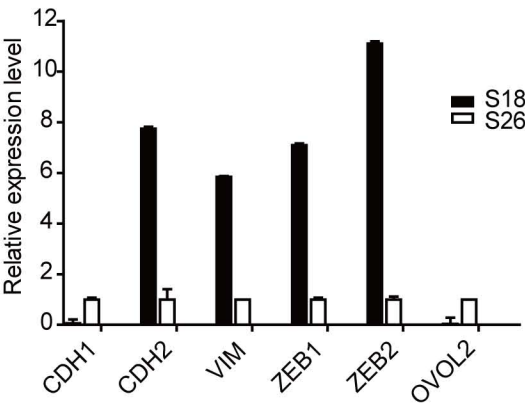
RANK	NAME	NES	FDR q-val	FWER p-val
1	HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	2.079954	0	0
2	HALLMARK_TNFA_SIGNALING_VIA_NFKB	2.0743127	0	0
3	HALLMARK_TGF_BETA_SIGNALING	2.0352209	0	0
4	HALLMARK_INFLAMMATORY_RESPONSE	1.9631493	3.13E-04	0.001
5	HALLMARK_UV_RESPONSE_DN	1.9050014	2.50E-04	0.001

E

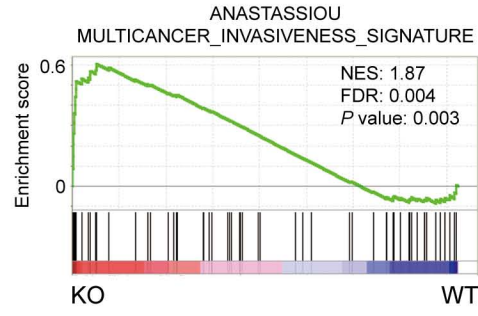
Gene Set Enrichment Analysis in CNE2 vs. S26: Hallmarks

RANK	NAME	NES	FDR q-val	FWER p-val
1	HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	1.663294	0.03123561	0.029
2	HALLMARK_MYC_TARGETS_V2	1.5866247	0.03354433	0.062

F

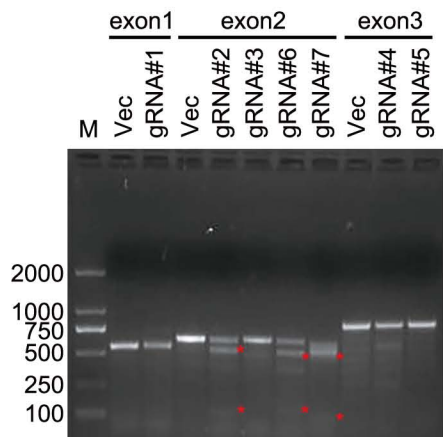


G



Supplementary Figure 2

A



B

PAM

HNE1

WT GCGTGC GGGGGCTCTCGCTT **TGG**GGGCTCGGGCCGCTGCGGCTCCCGGGGC

KO #1

Allele_1 GCGTGC GGGGGCTCTCGCTT **TGG**GGGCTCGGGCCGC **GCGAACCCACCTCTCCGATAAGCCGATACTGACCCGTGTTGTCTGTTA** +188 bp
 GCCGACGAGACTACGGCGGCACAAGGCCGACAGTCGCGTCCCGCGGGCCAAGAAAAACAGTTCTGGCTGGACAGGCCACGGG
 ACTTACTTGACGTTCTGCTCCGTCGCGCCGATAGCACCCGACCGGTGCTGCCCGCAAGTGCGGCTCCCGGGGC

Allele_2 GCGTGC GGGGGCTCTCGCTT **TGG**GGGCTCGGGCCGC **GCGAACCCACCTCTCCGATAAGCCGATACTGACCCGTGTTGTCTGTTA** +188 bp
 GCCGACGAGACTACGGCGGCACAAGGCCGACAGTCGCGTCCCGCGGGCCAAGAAAAACAGTTCTGGCTGGACAGGCCACGGG
 ACTTACTTGACGTTCTGCTCCGTCGCGCCGATAGCACCCGACCGGTGCTGCCCGCAAGTGCGGCTCCCGGGGC

KO #2

WT AGCAGCAGCGCGGGGAGCCTGGAGGAGCAGAGAGCAGCTCGTCCCGCACGCCCCGAGAGCGAAACCCCGAGCCCGGGC
 ACGCCGAGGGCCCCGAT**TGG**ACACCTG

Allele_1 AGCAGCAGCGCGGGGAGCCTGGAGGAGCAGAGAGCAGCTCGTCCCGCA-----AGCGAAACCCCGAGCCCGGGC -10 bp
 ACGCCGAGGGCCCCGAT**TGG**ACACCTG

Allele_2 AGCAGCAGCGCGGGGAGCCTGGAGGAGCA-----CGCCCCGAGAGCGAAACCCCGAGCCCGGGC -20 bp
 ACGCCGAGGGCCCCGAT**TGG**ACACCTG

KO #3

WT GCGGGCGTGCGGGTTGACGCCGAGTGTGTGTGTCGGGGCGGGCCCTCCCTCCACAGTGGGCCTAGGCCGCCTGCTCCACGA
 CCCCCGAGGACTGCCGACGACGGCGCAGCAGCAGCGGCAGCAGCAGCGCGGGGAGCCTGGAGGAGCAGAG
 AGCAGCTCGTCCCGCACGCCCCCGAGAGCGAAACCCCGAGCCCGGCGACGCCGAGGGCCCCGAT**TGG**ACAC

Allele_1 GCG----- -180 bp
 -----GCCCGGAGAGCGAAACCCCGAGCCCGGCGACGCCGAGGGCCCCGAT**TGG**ACAC

Allele_2 GCGGGCGTGCGGGTTGACGCCGAGTGTGTGTGTCGGGGCGGGCCCTCCCTCCACAGTGGGCCTAGGCCGCCTGCTCCACGA
 CCCCCGAGGACTGCCGACGACGGCGCAGCAGCAGCGGCAGCAGCAGCGCGGGGAGCCTGGAGGAGCAGAGA
 GCAGCTCGTCC-----CGAGAGCGAAACCCCGAGCCCGGCGACGCCGAGGGCCCCGAT**TGG**ACAC -10 bp

C

PAM

CNE2

WT TCGTCGAGCAGGGGCGTGCGGGGGCTCTCGCTT **TGG**GGGCTC

KO #1

Allele_1 TCGTCGAG-----CGGGGCTCTCGCTT **TGG**GGGCTC -10 bp

Allele_2 TCGTCGAGCAGGGG-----GGCTCTCGCTT **TGG**GGGCTC -8 bp

KO #2

WT CTCTCGCTT **TGG**GGGCTCGGGCCGCTGCGGCTCCCGGGC

Allele_1 CTCTCGCTT **TGG**GGGCTCGGGCCGCTGCGGTAAATGGACTACAAGACCCCTTATATTTACAGTCCGACTCTAATAGTTTTCTAGA +117 bp
 AGTGGATCTAGGAAAAGTGCATCTTTTCGTCAGGCGTCTTTGCCACGACTGGGGTGGGCTCCCGGGC

Allele_2 CTCTCGCTT **TGG**GGGCTCGGGCCGCTGCGGTAAATGGACTACAAGACCCCTTATATTTACAGTCCGACTCTAATAGTTTTCTAGA +117 bp
 AGTGGATCTAGGAAAAGTGCATCTTTTCGTCAGGCGTCTTTGCCACGACTGGGGTGGGCTCCCGGGC

KO #3

WT CTCGTCCCGCACGCCCCGAGAGCGAAACCCCGAGCCCGGCGACGCCGAGGGCCCCGAT**TGG**ACAC

Allele_1 CTC-----GCCCGGAGAGCGAAACCCCGAGCCCGGCGACGCCGAGGGCCCCGAT**TGG**ACAC -10 bp

Allele_2 CTCGTCC-----CCGAGAGCGAAACCCCGAGCCCGGCGACGCCGAGGGCCCCGAT**TGG**ACAC -9 bp

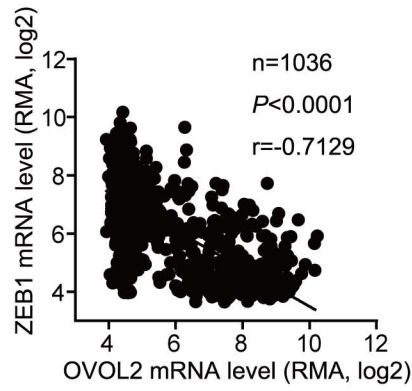
Supplementary Figure 3

A

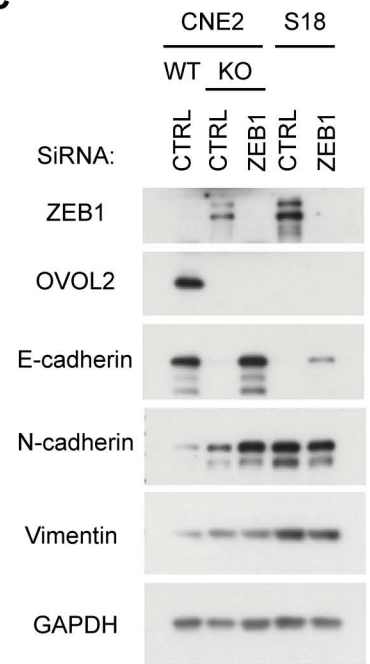
Correlation analysis

gene A	gene B	Pearson r
OVOL2	CDH1	0.7872
	ZEB1	-0.7129
	VIM	-0.7043
	ZEB2	-0.5452
	CDH2	-0.2991
	TWIST1	-0.2930
	TWIST2	-0.2477
	SNAIL1	-0.1943
	SNAIL2	-0.1848

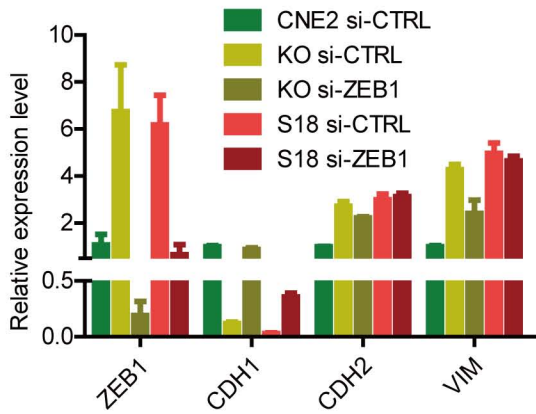
B



C



D



E

