

Supplementary Figure Legends

Supplementary Figure 1. The differential expression genes (DEGs) of CRC in three datasets.

a. Workflow diagram for screening the potential tumor markers via different databases.
b-d. Heat maps of differentially expression genes in three CRC datasets (GSE4183, GSE20916 and TCGA-COAD).

Supplementary Figure 2. The Human Protein Atlas analysis.

The protein expression summary analysis for ETV4 via the Human Protein Atlas.

Supplementary Figure 3. ETV4 enhanced cell proliferation and metastasis in CRC

a-b. Lentivirus-mediated stable ETV4 overexpression or knockdown HCT116, RKO and HT29 cells were selected in the presence of puromycin for 72h. Expression of ETV4 was examined by immunoblot analysis. **c.** Cell viability was determined by CCK8 assays in stable ETV4 overexpression or knockdown HT29 cells. **d.** Cell proliferation was analyzed by colony formation assay in HT29 cells with stable overexpression or knockdown ETV4. Cells were seeded into 6-well plates for two weeks and colony numbers were quantified. Representative images (left) and quantification (right) of colony formation of HT29 cells. **e-f.** Transwell migration and invasion assay of HT29 cells stable overexpression or knockdown ETV4. Representative images (left, scale bar: 100µm) and relative number of migrated or invaded cells (right). **g-h.** Stable ETV4 overexpression and knockdown HCT116, RKO and HT29 were performed to the wound healing assay. Data represent the mean $\pm$ SD. All experiments were performed in triplicates. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Supplementary Figure 4. ETV4 overexpression was correlated with high proliferation of CRC in vivo

a, c. Stable ETV4 overexpression or knockdown HCT116 cells were injected subcutaneously into nude mice, then the expression level of ETV4 and PCNA was detected by immunoblot analysis in the indicated xenograft samples. **b, d.** IHC staining and H&E staining of the indicated xenograft samples as described in (a, c), scale bar: 50 μ m.

Supplementary Figure 5. ETV4 promotes EMT in colorectal cancer cells

a. Gene set enrichment analysis of RNA-seq results showed the enrichment of CDH1 in ETV4 overexpression HCT116 cells. **b-c.** Verification of the expression of E-cadherin, N-cadherin, Snail, Twist1 and Vimentin using RT-qPCR in HT29 cells with stable ETV4 overexpression or knockdown. **d.** Immunoblot analysis of E-cadherin, N-cadherin, Vimentin and Twist1 expression in HT29 cells stable ETV4 overexpression or knockdown. **e.** Immunofluorescence analysis for the expression levels of E-cadherin and Vimentin in stable ETV4 overexpression or knockdown HT29 cells, scale bar: 25 μ m. Data represent the mean $\pm$ SD. All experiments were performed in triplicates. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Supplementary Figure 6. LOXL2 is upregulated and correlated with ETV4 in CRC

a. The LOXL2 mRNA expression were analyzed online in TCGA-COAD, GSE4183 and GSE20916 data sets. **b.** The mRNA expression of LOXL2 was detected via RT-qPCR in CRC cDNA array purchased from OriGene. **c-e.** Kaplan–Meier survival curve

analysis representing probabilities of disease-free survival in CRC patients based on the individual or combined expression of ETV4 and LOXL2. **f.** The nucleotide sequences of putative ETV4 binding sites are underlined within LOXL2 promoter. **g.** PCR analysis of ChIP. Chromatin fragments were prepared from HCT116 cells and immunoprecipitated with anti-ETV4 antibody or control IgG. The precipitated DNA was then amplified by real-time PCR with primers directed to the ETV4 binding sites in the LOXL2 promoter region.

Supplementary Figure 7. LOXL2 promotes cell proliferation, migration and invasion in CRC cells

a-b. HCT116 and RKO cells were transiently transfected with negative control siRNA (siNC) and siRNA against LOXL2 (siLOXL2). After 48h, RT-qPCR (a) and immunoblot analysis (b) were conducted to detect LOXL2 expression. **c-f.** HCT116 and RKO cells were treated as described in (a). Cell proliferation and motility were detected by CCK8 assay (c), transwell migration and invasion assay (d, scale bar: 100µm), and wound healing assay (e). Immunoblotting assay were conducted to determine the expression of E-cadherin, Vimentin and Twist1 (f). Data represent the mean ± SD. All experiments were performed in triplicates. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Supplementary Figure 8. Silencing LOXL2 in ETV4-overexpression CRC cells

a-b. siRNA mediated silencing of LOXL2 in stable ETV4 overexpression HCT116 and RKO cells. The LOXL2 expression was detect by RT-qPCR (a) and immunoblot analysis (b).

Supplementary Figure 9. LOXL2 is required for ETV4-induced malignant

phenotypes in CRC cells

a, c, d, f, h. Silencing of LOXL2 in stable ETV4 overexpression RKO and HCT116 cells. Cells were subjected to CCK8 assay (a), colony formation assay (c), transwell migration and invasion assay (d, scale bar: 100µm), and wound healing assay (f, h). **b, e, g, i.** Stable ETV4 knockdown CRC cells were transiently transfected with pcDNA3.1 empty or pcDNA3.1-HA-LOXL2 expression constructs. Then cell proliferation and motility were determined by CCK8 assay (b), transwell migration and invasion assay (e, scale bar: 100µm), and wound healing assay (g, i). **j.** Immunoblot analysis of E-cadherin, N-cadherin and Twist1 expression in RKO cells treated as described in (a). Data represent the mean $\pm$ SD. All experiments were performed in triplicates. * $p < 0.05$, ** $p < 0.01$.

Supplementary Figure 10. LOXL2 regulated NID1 expression in CRC cells.

a-b. HCT116 and RKO cells were transiently transfected with pcDNA3.1 empty or pcDNA3.1-HA-LOXL2 expression constructs. After 48h, RT-qPCR (a) and immunoblot analysis (b) was conducted to detect NID1 expression.

Supplementary Figure 11. NID1 is one of the 15 nearest neighbors with LOXL2 based on tissue protein expression

Expression clustering & correlation for LOXL2 via the Human Protein Atlas database.

Supplementary Figure 12. NID1 is upregulated and may serve as a potential prognostic marker in CRC.

a. The NID1 mRNA expression were analyzed online in three CRC datasets (TCGA-COAD, GSE4183 and GSE20916). **b-e.** Kaplan–Meier survival curve analysis

representing probabilities of disease-free survival in CRC patients based on NID1 (b), as well as the combined expression of ETV4 and NID1 (c), LOXL2 and NID1 (d), and the triple combination of ETV4, LOXL2 and NID1 (e). **f.** Determination of association between EMT markers and NID1 via GSE4183 and GSE20916.

Supplementary Figure 13. NID1 play promoted role in cell growth, migration and invasion in CRC cells.

a-b. HCT116 and RKO cells were transiently transfected with negative control siRNA (siNC) and siRNA against NID1 (siNID1). After 48h, RT-qPCR (a) and immunoblot analysis (b) were conducted to detect NID1 expression. **c-f.** HCT116 and RKO cells were treated as described in (a). Cell proliferation and motility were detected by CCK8 assay (c), colony formation assay (d), transwell migration and invasion assay (e, scale bar: 100µm) and wound healing assay (f). Data represent the mean $\pm$ SD. All experiments were performed in triplicates. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Supplementary Figure 14. NID1 is essential for ETV4-mediated aggressive phenotypes in CRC cells.

a-b,c,e. Stable ETV4 overexpression RKO cells were transiently transfected with negative siRNA or NID1 siRNA, or stable ETV4 knockdown RKO cells were transiently transfected with empty or pcDNA3.0-NID1. The cell viability and motility were determined by CCK8 assay (a-b), wound healing assay (c), transwell migration and invasion assay (e, scale bar: 100µm). **d.** Stable ETV4 overexpression HCT116 cells were transiently transfected with negative siRNA or NID1 siRNA, then cells were subjected to wound healing assay. Data represent the mean $\pm$ SD. All experiments were

performed in triplicates. **p < 0.01, *** p < 0.001.

Supplementary Figure 15. ETV4/LOXL2/NID1 activates the ERK signaling pathway through FAK

a. Immunoblotting was performed to determine FAK and p-FAK in stable ETV4 overexpression HCT116 and RKO cells. **b.** Immunoblotting was performed to determine the protein expression of p-ERK, ERK, p-FAK and FAK after PF573228 (FAK inhibitor) treatment in stable ETV4 overexpression HCT116. **c.** Immunoblotting was performed to explore the protein expression of p-FAK and FAK in HCT116 and RKO cells transiently transfected with negative siRNA or NID1 siRNA.

Supplementary Figure 16. ETV4 and LOXL2 promote NID1 transcription through promoter demethylation

a. The association of NID1 promoter methylation status with ETV4 and LOXL2 via UCSC Xena database (<https://xena.ucsc.edu/public/>). **b.** The association of NID1 methylation with NID1 or LOXL2 expression via TCGA database. **c.** Verification of NID1 promoter methylation status in stable ETV4 overexpression HCT116 and RKO cells using methylation specific PCR. **d.** Schematic illustration of 7 CpG sites located between nucleotides -370 and -330 in the NID1 promoter. **e.** The correlation of CpG sites in NID1 promoter region with ETV4 and LOXL2 via TCGA databases.

Supplementary Figure 17. ETV4/LOXL2 can induce NID1 transcriptional expression by promoter demethylation

a. The nucleotide sequences of putative ETV4 binding sites are underlined within NID1 promoter. **b.** Stable ETV4 overexpression HCT116 and RKO cells were transiently

transfected with negative siRNA or LOXL2 siRNA, and cells were treated with DAC (5-aza-2'-deoxycytidine, DNA methyltransferase inhibitor), then the methylation level of NID1 promoter was determined by MSP. **c-d.** HCT116 and RKO cells were treated as described in **(b)**, then the expression of NID1 was determined by RT-qPCR (c) and immunoblot analysis (d). Data represent the mean $\pm$ SD. All experiments were performed in triplicates. **p < 0.01, *** p < 0.001.

Figure S1

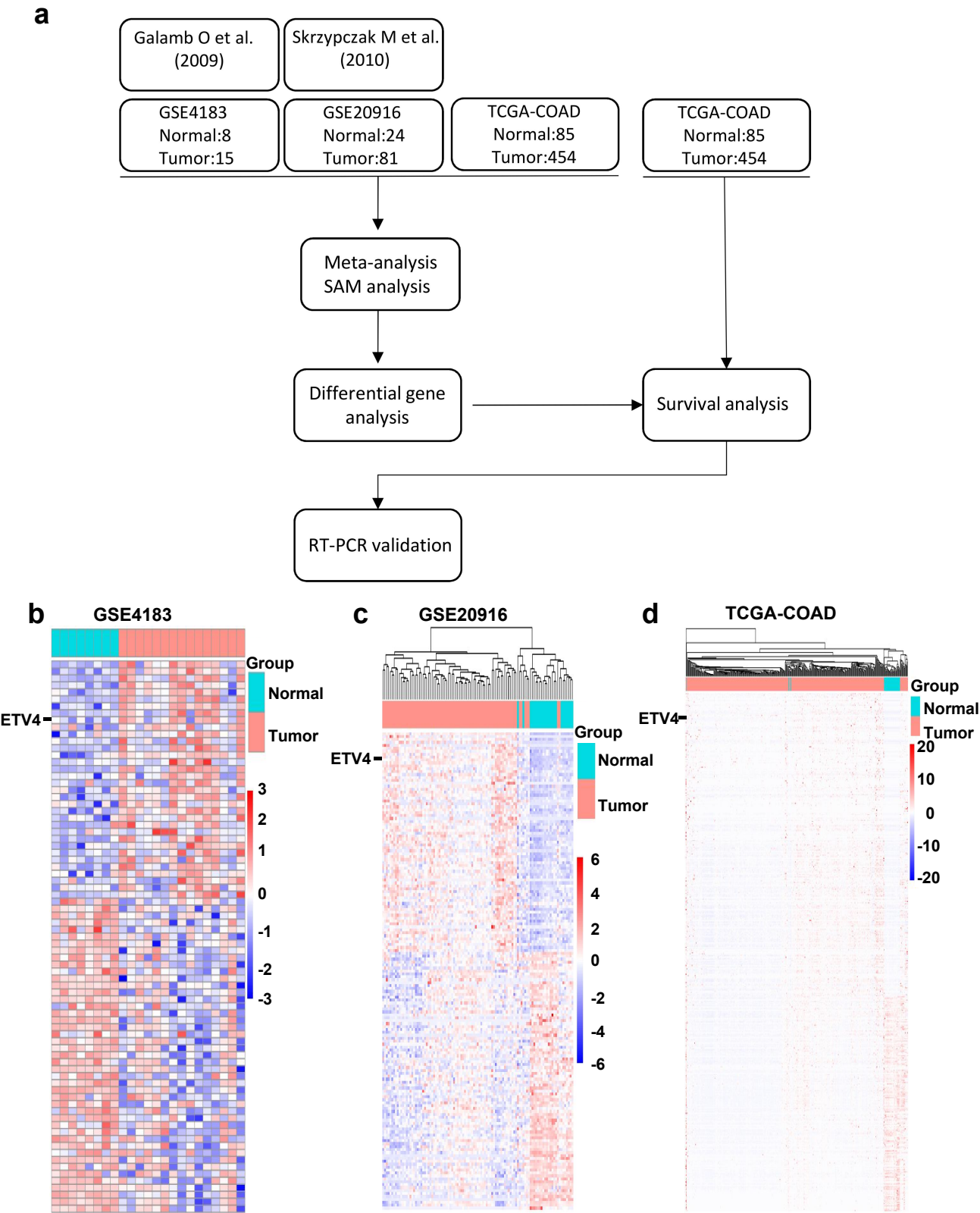


Figure S2

a

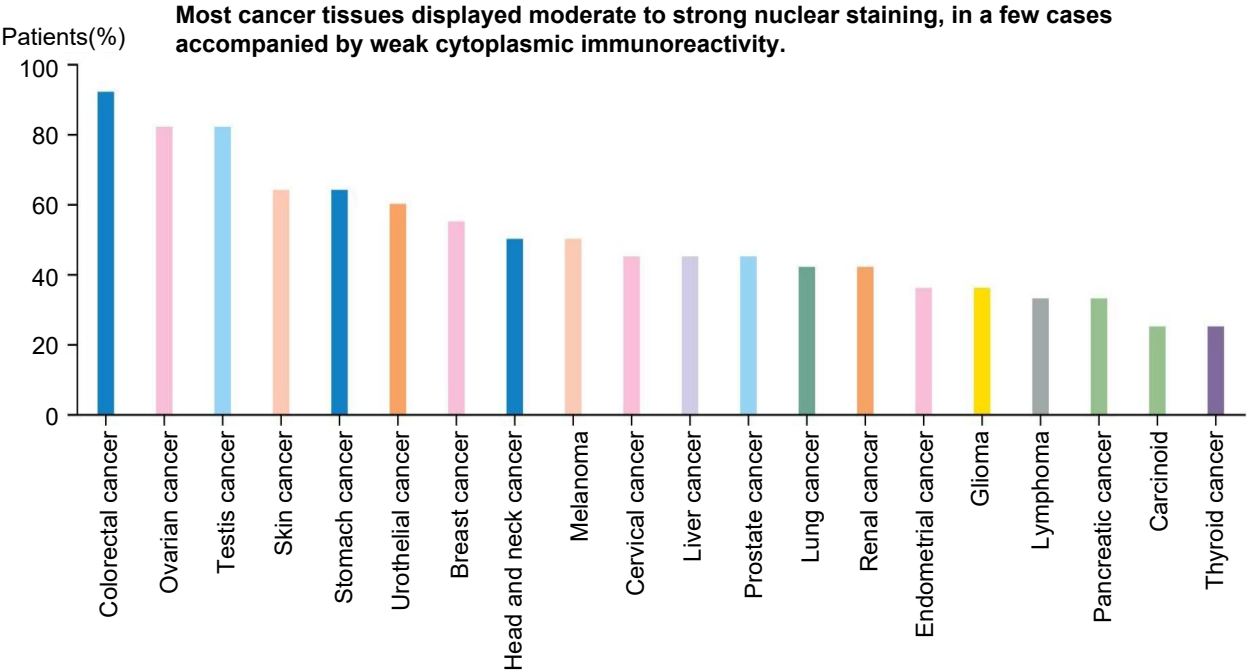


Figure S3

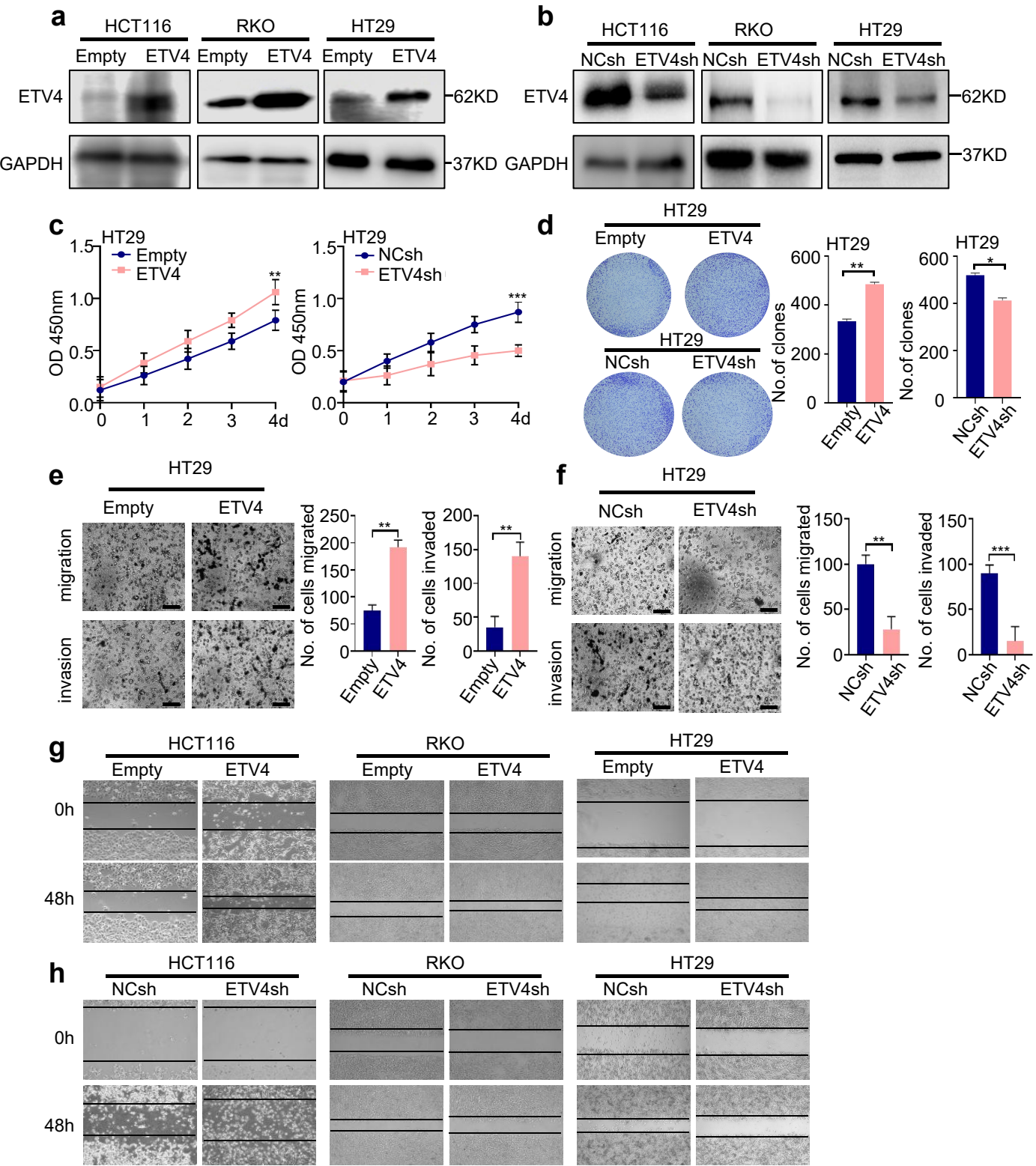


Figure S4

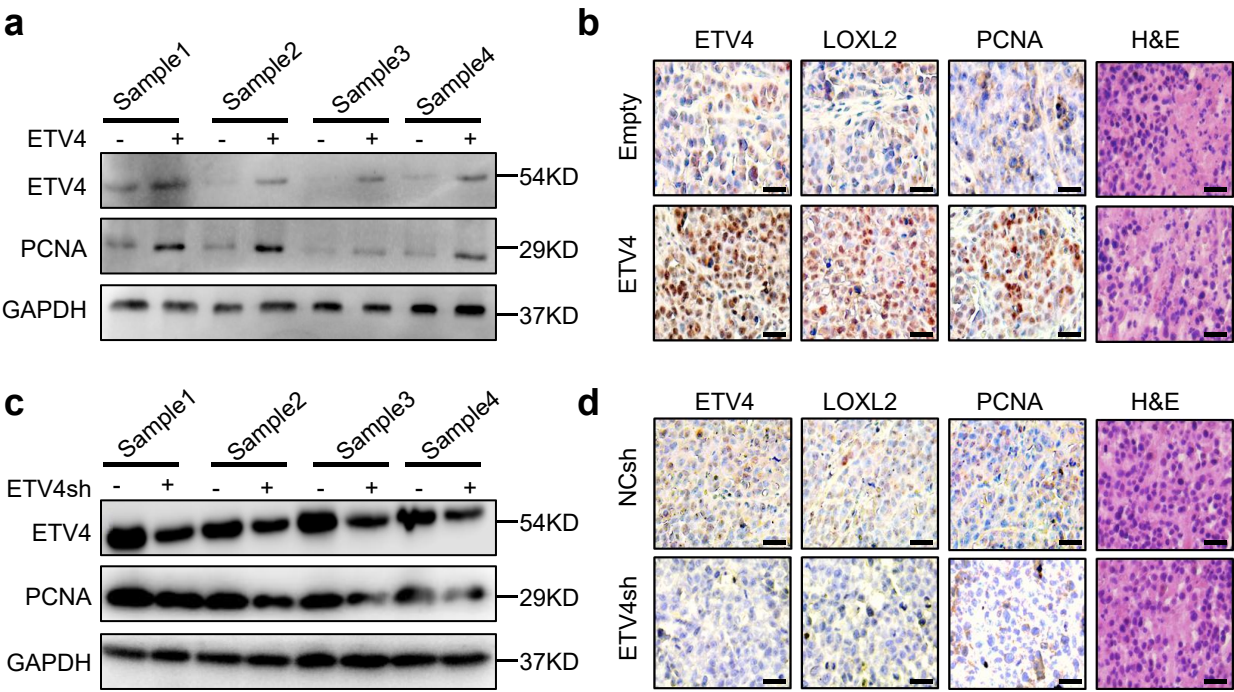


Figure S5

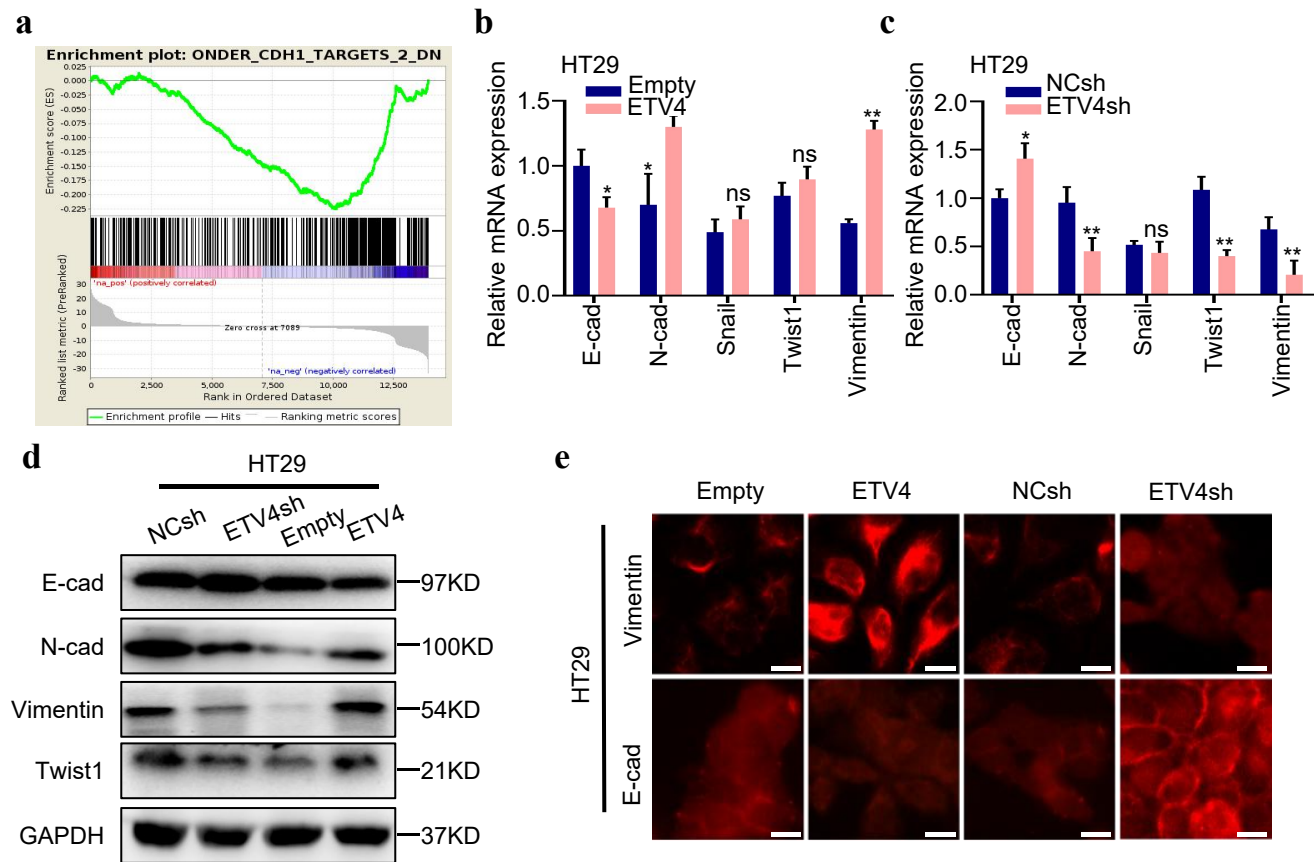


Figure S6

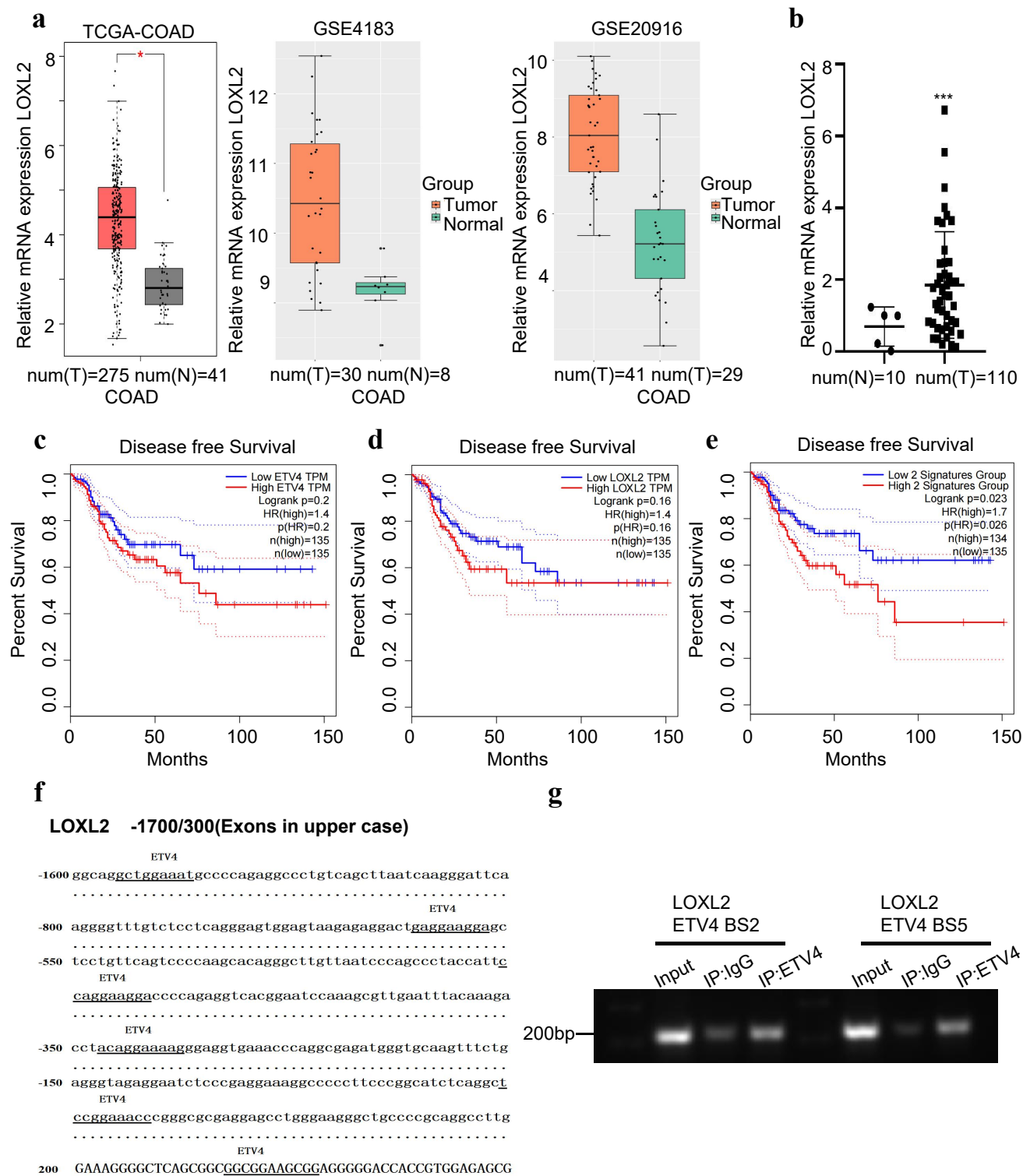


Figure S7

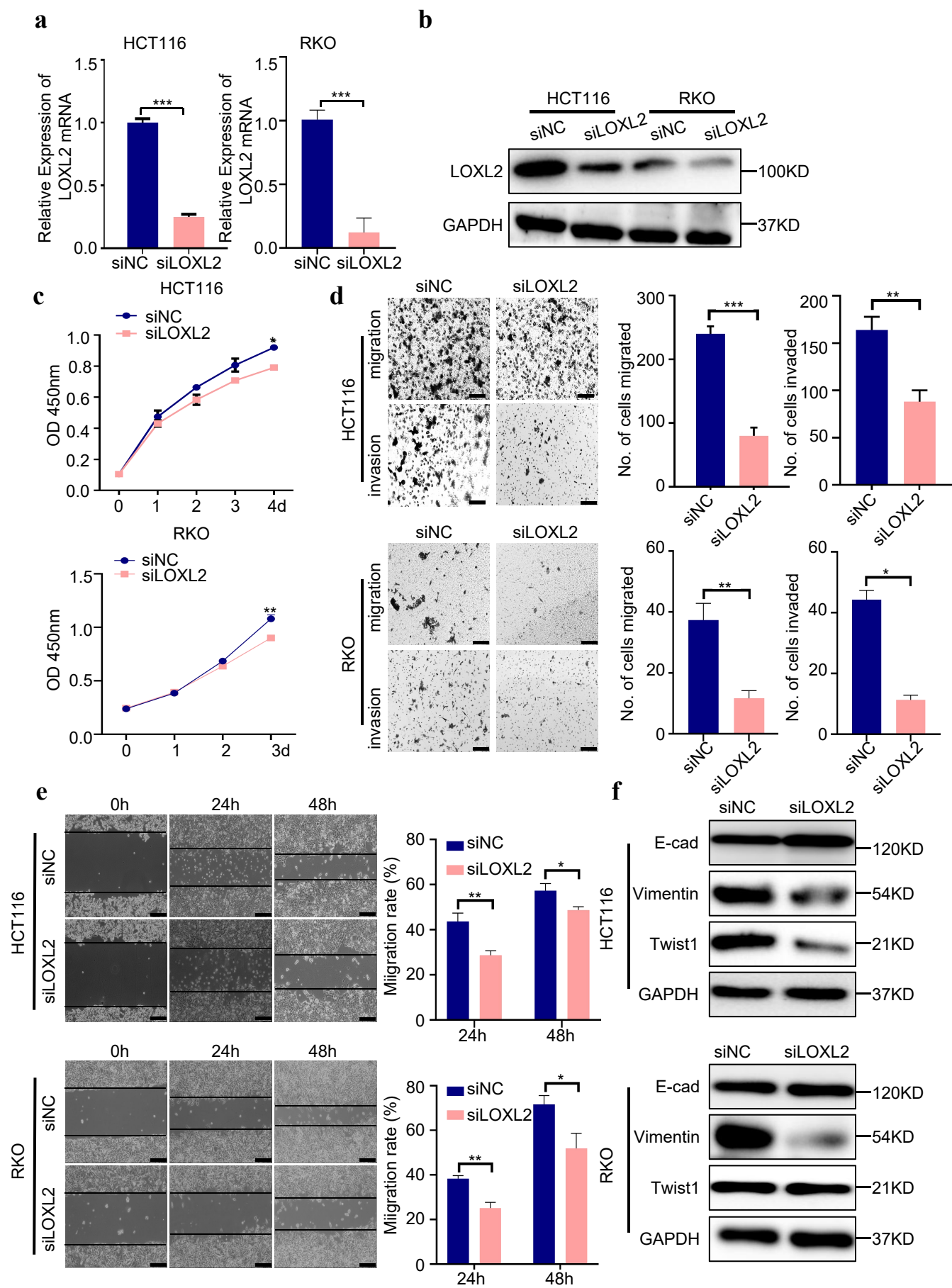


Figure S8

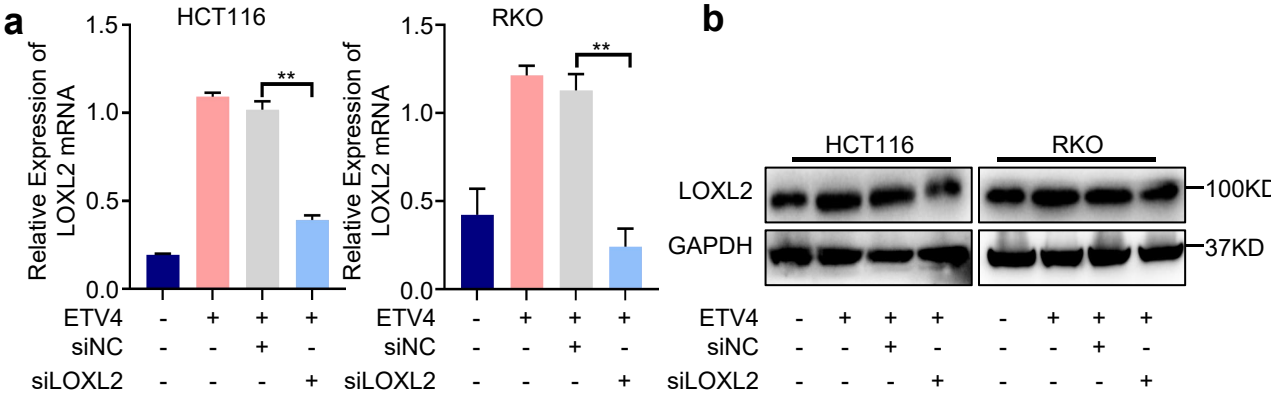


Figure S9

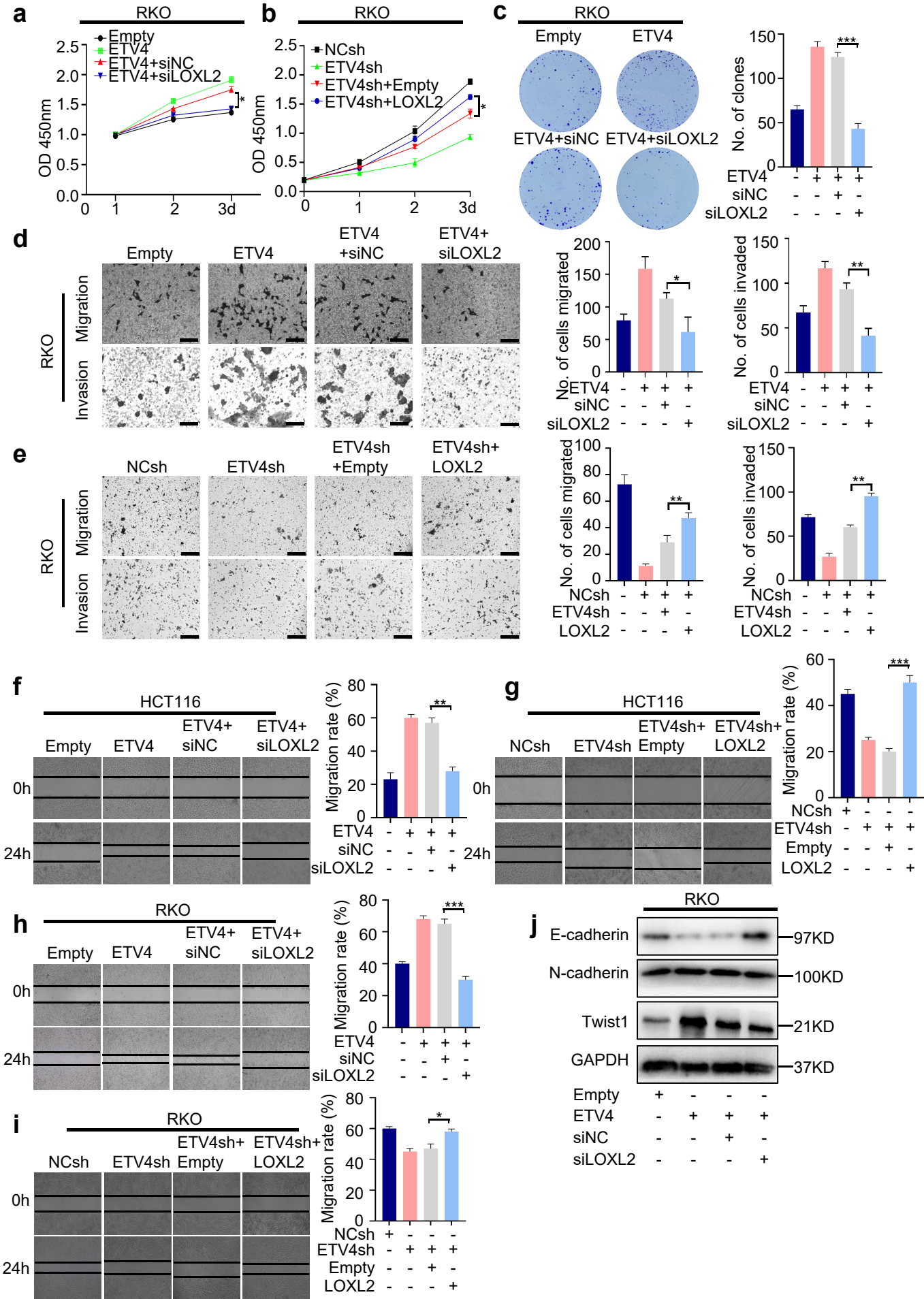


Figure S10

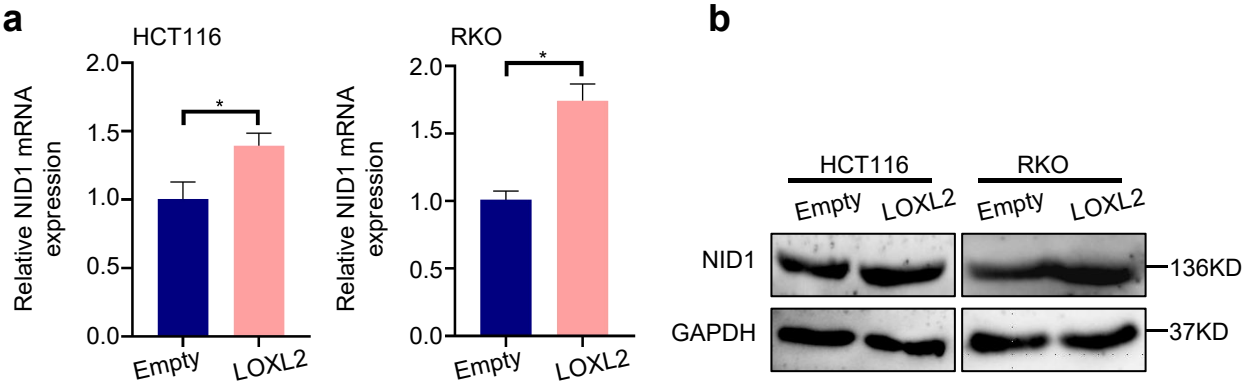


Figure S11

LOXL2 is part of cluster 47 Fibroblasts - ECM organization with confidenceⁱ 0.82
271 genes in cluster



15 nearest neighbours based on tissue RNA expression

Neighbour ⁱ	Description ⁱ	Correlation ⁱ	Cluster ⁱ
CLMP	CXADR like membrane protein	0.9526	47
CTHRC1	Collagen triple helix repeat containing 1	0.9439	69
ITIH5	Inter-alpha-trypsin inhibitor heavy chain 5	0.9351	69
COL5A2	Collagen type V alpha 2 chain	0.9316	47
COL3A1	Collagen type III alpha 1 chain	0.9298	47
HAS2	Hyaluronan synthase 2	0.9281	47
GJC1	Gap junction protein gamma 1	0.9246	47
MMP2	Matrix metalloproteinase 2	0.9228	47
PCDH18	Protocadherin 18	0.9211	47
COL1A2	Collagen type I alpha 2 chain	0.9193	47
NID1	Nidogen 1	0.9193	47
LUM	Lumican	0.9035	47
MMRN1	Multimerin 1	0.9035	47
IL33	Interleukin 33	0.9035	47
FBN1	Fibrillin 1	0.9018	69

Figure S12

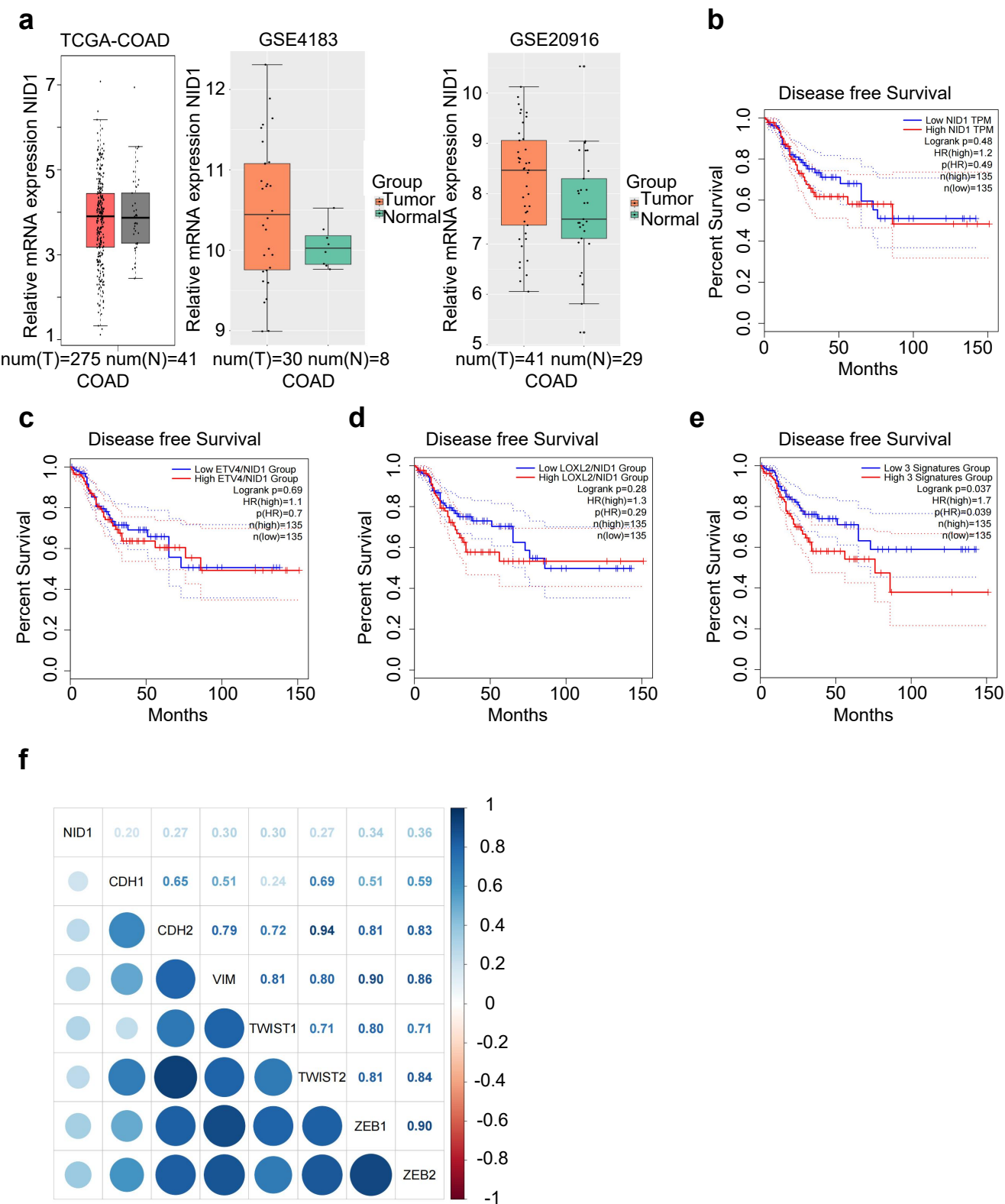


Figure S13

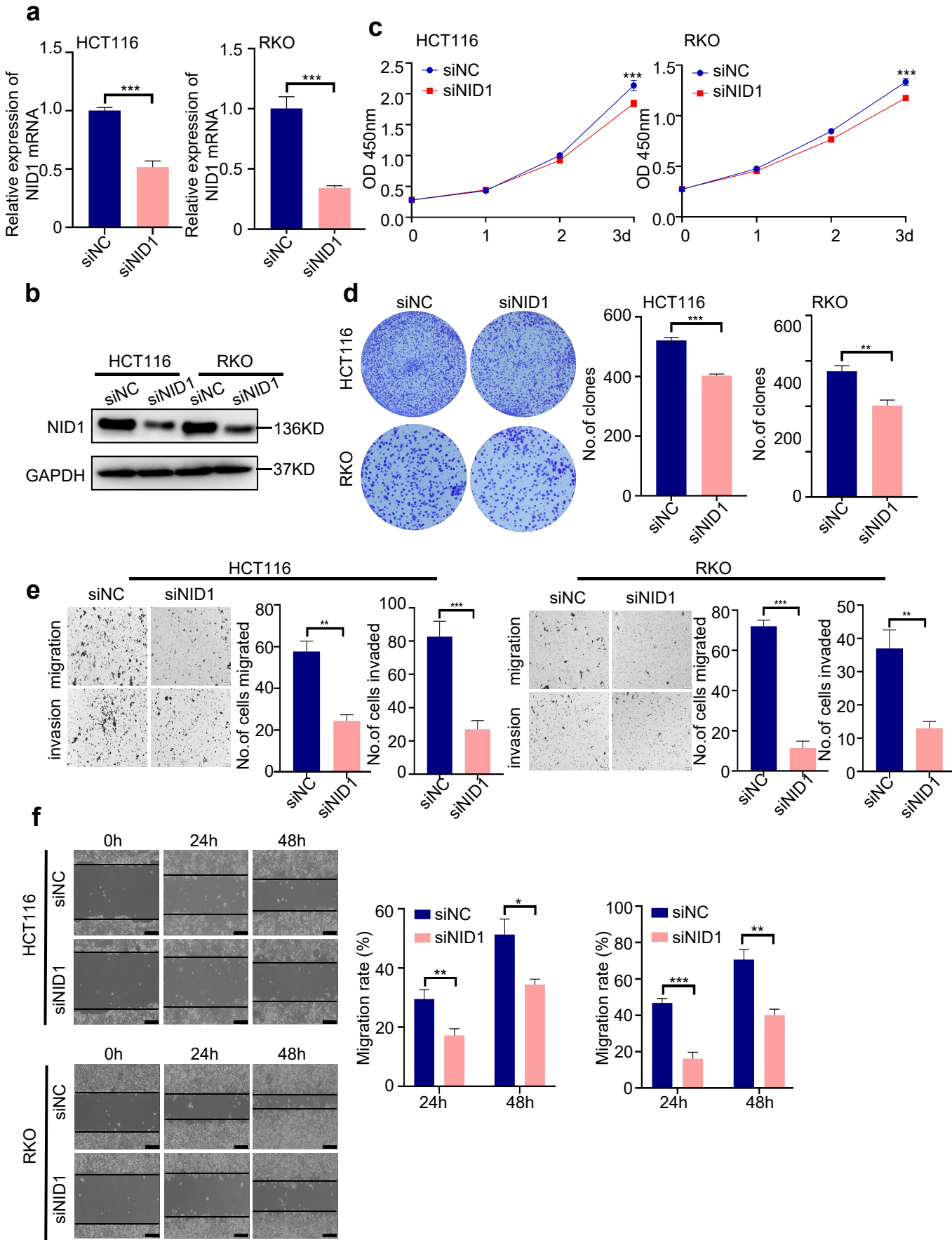


Figure S14

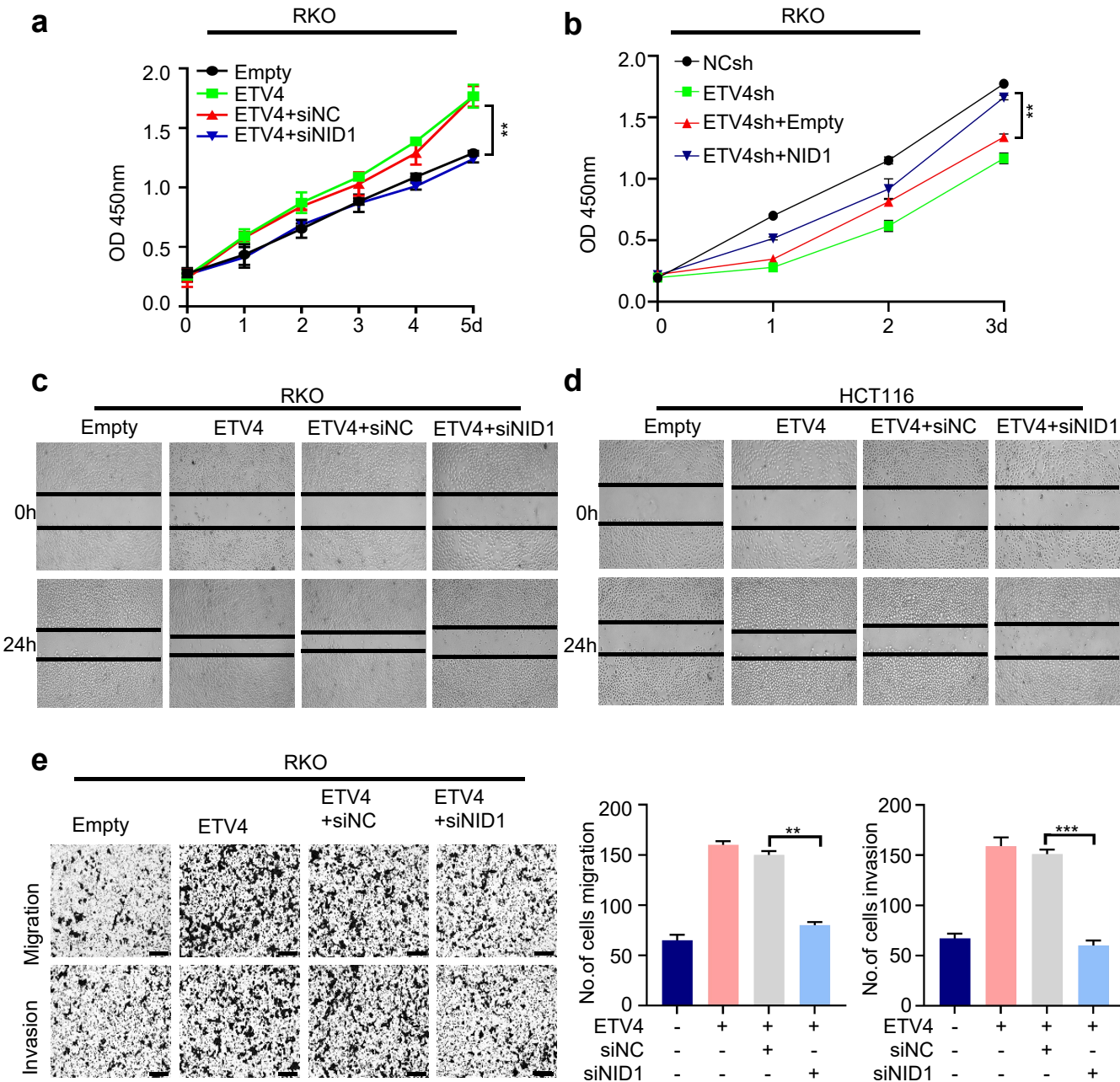


Figure S15

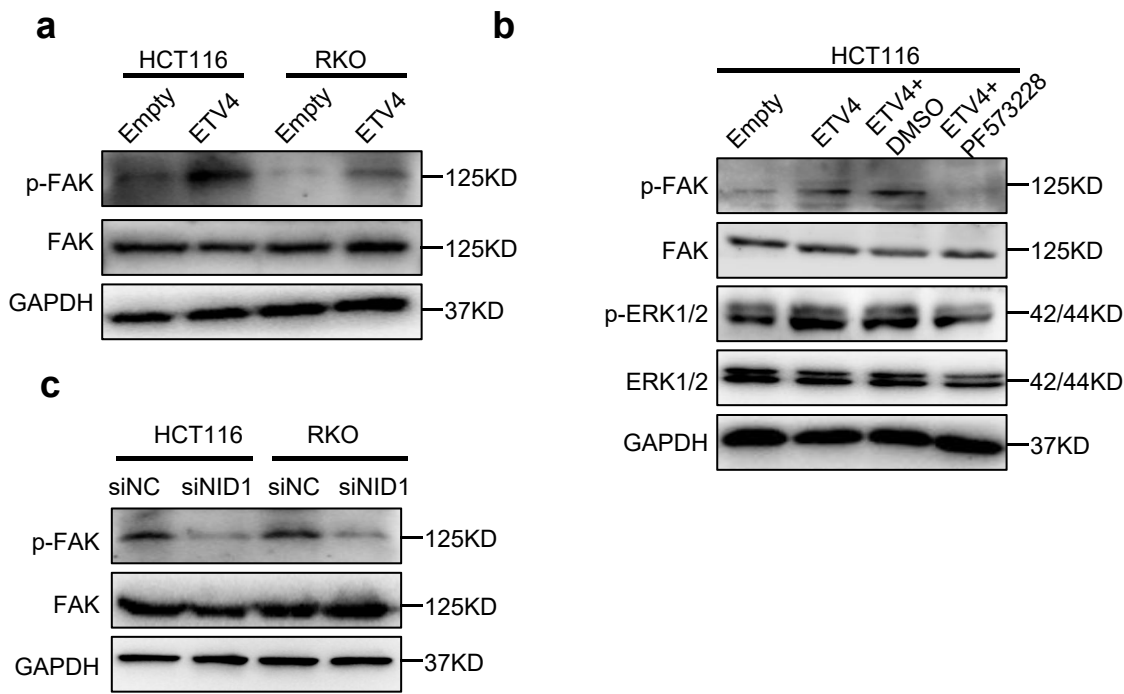


Figure S16

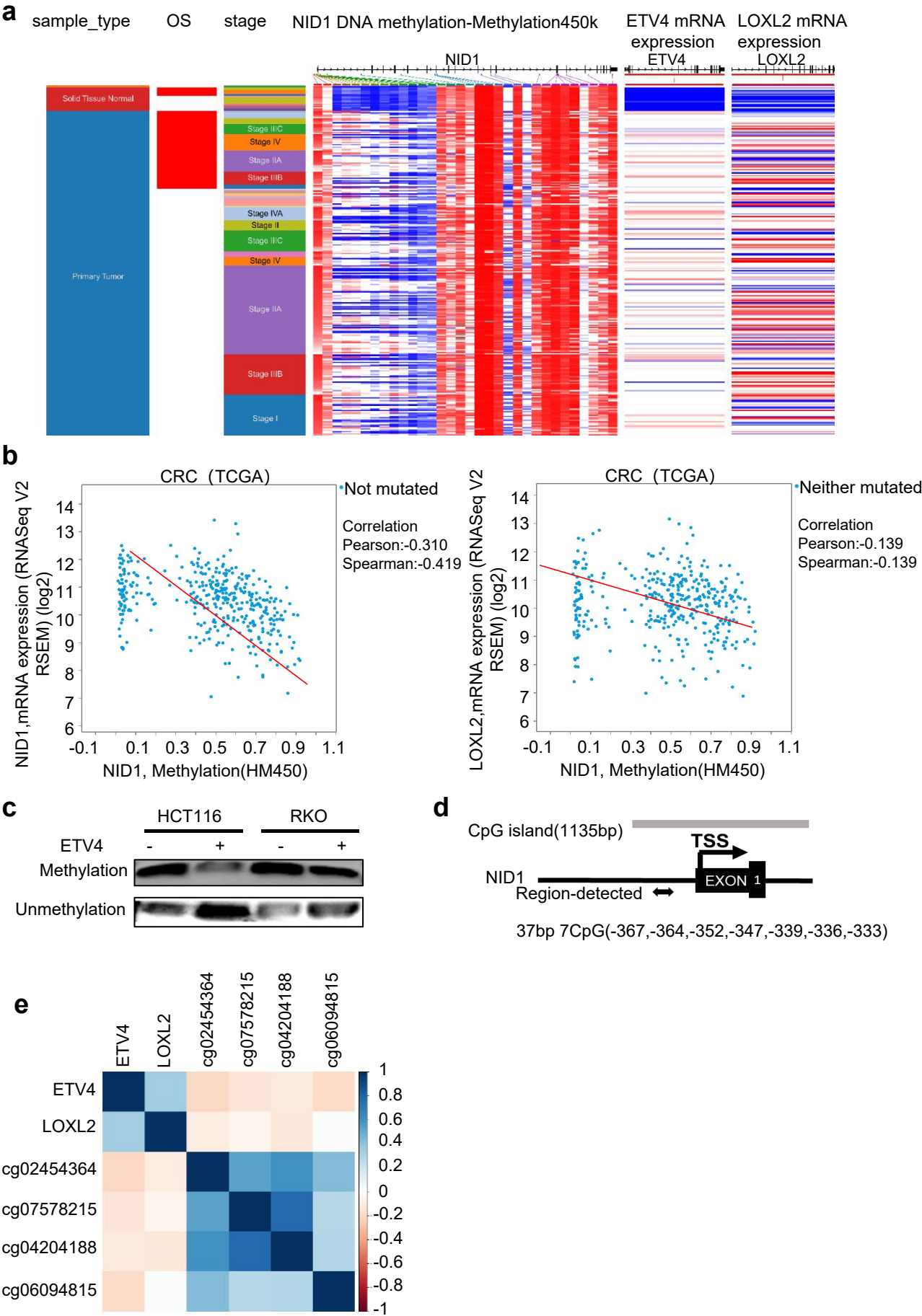


Figure S17

