

# High Expression of ELF1 Associates With Poor Prognosis And Immune Cells Infiltration In Colon Adenocarcinoma: A Bioinformatic Analysis

Zhaopeng Li (✉ [lizhaopengdoctor@163.com](mailto:lizhaopengdoctor@163.com))

Kunming Medical University First Affiliated Hospital

Lin Zhu

Kunming Medical University First Affiliated Hospital

Yang Yang

Kunming Medical University First Affiliated Hospital

Song Bai

Kunming Medical University First Affiliated Hospital

---

## Research

**Keywords:** ELF1, Colon adenocarcinoma, Tumor-infiltrating immune cells, Bioinformatics analysis

**Posted Date:** September 15th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-861154/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

## Abstract

**Background** The human ELFN1 encodes a functional protein belonging to leucine-rich repeat (LRR) family neuronal adhesion proteins. However, there remains a lack of research on ELFN1 in tumors, especially in colon adenocarcinoma (COAD). The paper aimed to explore ELFN1 expression and its potential biological function in COAD through bioinformatics analysis.

**Results** ELFN1 expression was significantly upregulated in COAD tissues from The Cancer Genome Atlas (TCGA) database ( $P < 0.01$ ). The high ELFN1 mRNA expression level in tumor samples was also verified in GSE39582 and GSE146009 cohorts from Gene Expression Omnibus (GEO) website ( $P < 0.01$ ). Besides, high ELFN1 expression was positively correlated with advanced tumor stage (I vs. IV, II vs. IV), perineural invasion and vascular invasion ( $P < 0.05$ ). Moreover, high ELFN1 expression was correlated with worse overall survival (OS) ( $P < 0.05$ ), which was also verified in GSE17536 and GSE29621 datasets ( $P < 0.05$ ). Multivariate Cox regression analysis suggested ELFN1 was an independent prognostic predictor in COAD ( $P < 0.01$ ). Furthermore, the high infiltration levels of M0 Macrophages and Regulatory T cells (Tregs) were found in the high ELFN1 expression COAD tissues, while the infiltration levels of activated CD4 memory T cells, resting or activated Dendritic cells were low in those samples ( $P < 0.05$ ). The correlation analysis revealed that ELFN1 expression had a positive correlation with the infiltrating levels of Tregs and M0 Macrophages, but inversely correlated with that of activated Dendritic cells and both resting and activated CD4 memory T cells in COAD ( $P < 0.05$ ). Gene Set Enrichment Analysis (GSEA) exhibited that ELFN1 was enriched in many immunological and cancer-related signaling pathways, like Hedgehog, ECM receptor interaction, focal adhesion, Notch, and MAPK signaling pathways.

**Conclusions** ELFN1 is overexpressed in COAD and which is correlated with worse clinical outcomes. Higher expression of ELFN1 could influence tumor immune cells infiltrating in COAD. ELFN1 could promote progression of COAD through several immunological and cancer-related signaling pathways. ELFN1 may be a new prognostic predictor and a latent therapeutic target for COAD.

## Background

Colorectal cancer (CRC) is the third most prevalent malignancy worldwide and the second cause of cancer-related deaths(1). Patients with CRC usually develop to an advanced stage at the first diagnosis (2), about 20–25% of them with distant metastasis (3). Multiple factors can contribute to colorectal carcinogenesis, but the successive accumulation of genetic and epigenetic alterations remains the fundamental pathogenesis (4). It is reported that adenocarcinoma makes up 98% of newly diagnosed colon cancer (CC) (5). Although significant advances have been made in multimodal anticancer strategies, the 5-year OS of metastatic CC is only 13.7% (<https://seer.cancer.gov>, February 1, 2021). This is mainly due to our limited knowledge of the etiology and pathogenesis of CC, as well as being the shortage of early diagnostic indicators and effective therapeutic targets.

Immunotherapy like PD-1 blockade has been shown a promising response to metastatic CRC with DNA deficient mismatch repair (6). Tumor immune microenvironment (TIME) is tightly associated with

tumorigenesis, prognosis, and the efficacy of immunotherapy (7). Tumor-infiltrating immune cells (TIICs), like macrophages, T cells, neutrophils, play a crucial role in modulating anticancer response in the TIME (8). It is primarily because TIICs can secrete cytokines and chemokines, which will further activate or inactivate several signaling pathways related to immune response. Therefore, if the nexus between oncogene and TIICs could be entirely delineated, it would supply a novel perception into the pathogenesis of COAD.

Extracellular leucine-rich repeat and fibronectin type III domain-containing 1 (ELFN1) encodes a functional protein belonging to leucine-rich repeat (LRR) family neuronal adhesion proteins (9), which has been validated to perform a crucial role in synapse formation through the coordination of both presynaptic and postsynaptic machinery (10, 11). Besides, it has been identified in interneurons of the hippocampus, cerebral cortex (12), and cone photoreceptors (13). However, there remains a lack of research on ELFN1 in tumors, especially in COAD.

In this study, it is the first time to investigate the expression level of ELFN1 in COAD and uncover its potential biological roles, prognostic value, as well as its relationship with the TIICs. We observed that ELFN1 was overexpressed in COAD, which was correlated with worse clinical outcomes. ELFN1 could also influence TIICs infiltrating in the TIME, particularly Tregs and M0 Macrophages. Moreover, GSEA revealed ELFN1 was linked to many immunological and oncogenic signaling pathways, like the Hedgehog, ECM receptor interaction, focal adhesion, Notch and MAPK signaling pathways. Our work suggests ELFN1 could be a new prognostic predictor and a latent target for COAD therapy.

## Results

### Expression levels of ELFN1 in COAD

The mRNA expression profiles of ELFN1 in both tumor tissues and normal colon tissues from the TCGA and GEO websites were explored. It was found the ELFN1 mRNA expression level was significantly upregulated in COAD tissues from the TCGA database ( $P < 0.01$ ) (Fig. 1a). A similar result could be seen in 41 matched tumor and normal colon tissues from a TCGA cohort ( $P < 0.01$ ) (Fig. 1b). The high ELFN1 mRNA expression level in tumor samples was also verified in GSE39582 and GSE146009 cohorts ( $P < 0.01$ ) (Fig. 1c, d). Simultaneously, ELFN1 was also highly expressed in several types of cancers, such as esophageal cancer, stomach cancer, and breast cancer by bioinformatic analysis for the resting 32 cancer types (Additional File. 1a).

### Association between ELFN1 mRNA expression and clinicopathological characteristics

The ELFN1 mRNA expression levels of 446 COAD samples from the TCGA database in different clinicopathological parameters were presented (Table. 1). The association between ELFN1 mRNA expression level and clinicopathological parameters was analyzed. The outcomes manifested that high ELFN1 mRNA in COAD had a positive association with the advanced Tumor stage (IV vs. I; IV vs. II), T stage (T4 vs. T2), N stage (N2 vs. N0; N2 vs. N1), M stage (M1 vs. M0), Lymphatic metastasis (YES vs.

NO), Vascular invasion (YES vs. NO) and Perineural invasion (YES vs. NO) ( $P < 0.05$ ) (Fig. 2a-g). However, no association was seen between ELFN1 mRNA expression and the patient's Age, Gender ( $P > 0.05$ ) (Fig. 2h-i).

Table1. Clinicopathological characteristics of 446 samples from the TCGA cohort.

		ELFN1 expression		
Characteristic	Type	High	Low	Total
Age(years)	<=65	94(42.15%)	88(39.46%)	182(40.81%)
	>65	129(57.85%)	135(60.54%)	264(59.19%)
Gender	Female	113(50.67%)	99(44.39%)	212(47.53%)
	Male	110(49.33%)	124(55.61%)	234(52.47%)
Vital status	Alive	171(76.68%)	187(83.86%)	358(80.27%)
	Dead	52(23.32%)	36(16.14%)	88(19.73%)
pTumor Stage	Stage I	33(14.8%)	41(18.39%)	74(16.59%)
	Stage II	83(37.22%)	93(41.7%)	176(39.46%)
	Stage III	64(28.7%)	60(26.91%)	124(27.8%)
	Stage IV	36(16.14%)	25(11.21%)	61(13.68%)
	Unknown	7(3.14%)	4(1.79%)	11(2.47%)
pT Stage	T1	5(2.24%)	5(2.24%)	10(2.24%)
	T2	34(15.25%)	42(18.83%)	76(17.04%)
	T3	152(68.16%)	152(68.16%)	304(68.16%)
	T4	32(14.35%)	24(10.76%)	56(12.56%)
pN Stage	N0	124(55.61%)	141(63.23%)	265(59.42%)
	N1	47(21.08%)	55(24.66%)	102(22.87%)
	N2	52(23.32%)	27(12.11%)	79(17.71%)
pM Stage	M0	159(71.3%)	171(76.68%)	330(73.99%)
	M1	36(16.14%)	25(11.21%)	61(13.68%)
	Unknown	28(12.56%)	27(12.11%)	55(12.33%)
Lymphatic invasion	YES	96(43.05%)	63(28.25%)	159(35.65%)
	NO	107(47.98%)	139(62.33%)	246(55.16%)
	Unknown	20(8.97%)	21(9.42%)	41(9.19%)
Vascular invasion	YES	58(26.01%)	37(16.59%)	95(21.3%)
	NO	134(60.09%)	159(71.3%)	293(65.7%)
	Unknown	31(13.9%)	27(12.11%)	58(13%)

Perineural invasion	YES	26(11.66%)	19(8.52%)	45(10.09%)
	NO	58(26.01%)	73(32.74%)	131(29.37%)
	Unknown	139(62.33%)	131(58.74%)	270(60.54%)
CEA level(ng/L)	<=5	92(41.26%)	94(42.15%)	186(41.7%)
	>5	53(23.77%)	43(19.28%)	96(21.52%)
	Unknown	78(34.98%)	86(38.57%)	164(36.77%)

### The potential of ELFN1 to be a diagnostic and prognostic indicator in COAD

Due to the high expression of ELFN1 in COAD, its diagnostic and prognostic value for patients with COAD was analyzed. Generally, the high ELFN1 expression of COAD patients in the TCGA-COAD cohort exhibited worse OS (Fig. 3a) and Progression-free survival (PFI) (Fig. 3b) than that in the low ELFN1 expression patients ( $P < 0.05$ ). Simultaneously, the low expression of ELFN1 in the GSE29621 cohort exhibited favorable OS ( $P < 0.01$ ) (Fig. 3c), while ELFN1 expression had no effect on disease-free survival (DFS) of patients with COAD ( $P > 0.05$ ) (Fig. 3d). Then, low ELFN1 expression in the GSE17536 cohort correlated with a favorable OS (Fig. 3e) and DFS ( $P < 0.05$ ) (Fig. 3f). It was also revealed that high ELFN1 expression in stomach cancer, ocular melanomas and cervical cancer patients were correlated with a poor OS ( $P < 0.05$ ) (Additional File. 1b-d).

The clinicopathological parameters, including ELFN1 expression level, age, T, N, M stage, and tumor stage, were thought to influence patients' OS in COAD by univariate Cox regression analysis ( $P < 0.05$ ) (Fig. 4a). Also, multivariate Cox regression analysis revealed that ELFN1 expression, age and T stage were independent prognostic predictors for COAD patients ( $P < 0.05$ ) (Fig. 4b). Furthermore, ROC curve analysis showed that ELFN1 ( $AUC=0.900$ ) had a more satisfactory diagnostic value than that of carcinoembryonic antigen (CEA) for COAD ( $AUC=0.547$ ) (Fig. 4c). The ROC curves of ELFN1 for OS at 3, 5, and 10 years were also plotted (Fig. 4d).

### Relationship between TIICs and ELFN1 expression in COAD

The relative abundances of 22 TIICs in each COAD tumor sample were calculated using the CIBERSORT algorithm. The high infiltration levels of M0 Macrophages and Tregs were found in COAD tissues with the high expression of ELFN1, while the infiltration levels of activated CD4 memory T cells, resting or activated Dendritic cells were low in the high ELFN1 expression COAD tissues ( $P < 0.05$ ) (Fig. 5a). The correlation between the relative abundance of TIICs and ELFN1 expression level was also investigated. The outcomes revealed that ELFN1 expression had a positive correlation with the infiltrating levels of Tregs and M0 Macrophages, but inversely correlated with that of activated Dendritic cells and both resting and activated CD4 memory T cells in COAD ( $P < 0.05$ ) (Fig. 5b-f).

The relationship between infiltrating Tregs, M0 Macrophages, and clinicopathological parameters was analyzed. The higher relative abundance of Tregs in COAD had a positive association with the advanced

Tumor stage (III vs. II) ( $P < 0.05$ ) (Fig. 6a), whereas its abundance changes had no influence on the T stage, N stage, and M stage (Fig. 6b-d) ( $P > 0.05$ ). The relative abundance of M0 Macrophages was positively correlated with the early N stage ( $P < 0.05$ ), surprisingly, it revealed the opposite trend in the advanced N stage (Fig. 6g) ( $P > 0.05$ ). Besides, the altered abundance of M0 Macrophages had no effect on Tumor stage, T stage and M stage (Fig. 6 e, f, h) ( $P > 0.05$ ). The changes of Tregs infiltration levels in COAD had no influence on OS (Fig. 6i) ( $P > 0.05$ ). However, lower Tregs' relative abundance was associated with better PFI (Fig. 6j) ( $P < 0.05$ ). The changes of relative abundances of M0 Macrophages had no effect on OS and PFI (Fig. 6 k, l) ( $P > 0.05$ ).

Besides, there was a weak to moderate association among the relative abundance of different TIICs (Fig. 7). For example, the infiltration level of M0 macrophages in COAD negatively correlated with that of M1 Macrophages ( $r = -0.45$ ) and resting Dendritic cells ( $r = -0.35$ ).

### **Association of ELFN1 methylation with clinicopathological features in COAD**

11 CpG sites of ELFN1 were found by analyzing the methylation database of TCGA-COAD (Fig. 8a). Further analyses showed that the ELFN1 hypermethylation was associated with earlier tumor stage, N stage, and M stage (Fig. 8b-d) ( $P < 0.05$ ). Next, the associations of ELFN1 methylation with OS were evaluated. However, the methylation of ELFN1 did not affect the patient's survival (Additional File. 2a-k) ( $P > 0.05$ ).

### **Analysis of ELFN1 mutation in COAD**

The mutation frequency of ELFN1 was too low in TCGA-COAD and ICGC-COAD cohorts (Fig. 9a, b).

### **Gene set functional enrichment analysis**

Differentially activated signaling pathways related to the high ELFN1 expression in COAD datasets were identified by GSEA software (Table. 2). Many signaling pathways were enriched in the ELFN1 high expression phenotype, such as focal adhesion, pathways in cancer, the intestinal immune network for IGA production, Cell adhesion molecules (CAMs), MAPK, Hedgehog, VEGF, ECM receptor interactions, and Notch signaling pathway (Fig. 10).

Table2. Differentially activated signaling pathways related to the high ELFN1 expression in COAD by GSEA.

Name	ES	NES	Nominal p-value	FDR q-value
KEGG_CELL_ADHESION_MOLECULES_CAMS	0.699	2.212	<0.001	0.004
KEGG_HEDGEHOG_SIGNALING_PATHWAY	0.583	2.081	<0.001	0.013
KEGG_ECM_RECECTOR_INTERACTION	0.712	2.052	0.001	0.014
KEGG_NOTCH_SIGNALING_PATHWAY	0.630	2.044	0.001	0.014
KEGG_FOCAL_ADHESION	0.569	2.013	0.007	0.013
KEGG_INTESTINAL_IMMUNE_NETWORK	0.722	1.938	0.010	0.020
_FOR_IGA_PRODUCTION				
KEGG_MAPK_SIGNALING_PATHWAY	0.433	1.849	0.003	0.029
KEGG_PATHWAYS_IN_CANCER	0.435	1.823	0.010	0.032
KEGG_VEGF_SIGNALING_PATHWAY	0.453	1.805	0.008	0.038

FDR: false discovery rate; ES: enrichment Score; NES: normalized enrichment Score.

### Co-expression analysis of ELFN1

The Pearson correlation coefficients between expression profiles of ELFN1 and PCGs were calculated to determine the co-expression relationships of the ELFN1 and PCGs. Then the Gene Ontology (GO) enrichment analysis was conducted. These genes encoded proteins playing roles mainly in transcription coregulator activity, GTPase regulator activity, growth factor binding, collagen binding crosstalk, extracellular matrix structural constituent, modification-dependent protein binding and transcription corepressor activity (Fig. 11a). The Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis reflected the enrichment of ELFN1-related signatures associated with many immunological and cancer-related signaling pathways, like Human papillomavirus infection, PI3K-AKT, CRC, breast cancer, Hippo, ECM receptor interaction, and focal adhesion, Notch, MAPK, mTOR signaling pathways, and so on (Fig. 11b). The STRING and Cytoscape 3.8.2 software were applied to construct the Protein-protein Interaction (PPI) networks among ELFN1-related co-expressed genes. Proteins interaction data in COAD revealed 189 kinds of proteins could interact with ELFN1 protein, among which 21 kinds of proteins had a negative correlation with the expression of ELFN1, whereas the expression of ELFN1 positively correlated with the other 168 kinds of proteins (Fig. 12a). Based on  $|r| \geq 0.6$ , 20 co-expression genes were identified to be correlated with ELFN1 expression (Fig. 12b, c).

## Discussion

CC is one of the most prevalent malignancies globally, with increasing incidence (14). Despite the rapid progress having been made in screening and therapeutic measures of CC over the years, the patients' clinical outcomes with the advanced CC are still poor (15). That mainly because we have a superficial

knowledge of the pathogenesis of CC (16). Thus, it is necessary to uncover novel biomarkers and enunciate the latent molecular mechanism of CC to refine the patients' prognosis. This study focuses on ELFN1, an adhesion protein of the leucine-rich repeat (LRR) family. ELFN1 is expressed among several neurons and confers target-specific synaptic properties (17). Given the barely reported correlation of ELFN1 and clinicopathological features of tumors, we aim to dissect the biological functions of ELFN1 in COAD and explore its latent mechanisms of action.

Data analysis in TCGA and GEO databases revealed the expression of ELFN1 was significantly upregulated in COAD. The analysis of clinicopathological characteristics manifested that ELFN1 expression in COAD was a positive relationship with the degree of tumor aggression, lymphovascular invasion, distant metastasis, and perineural invasion, which indicated ELFN1 might act as an oncogene to facilitate invasion and migration of COAD. Besides, the ROC curves illustrated that ELFN1 had good diagnostic and prognostic value. Also, Kaplan–Meier analysis displayed that higher ELFN1 expression correlated with worse OS and PFI. Moreover, multivariate Cox regression analysis revealed high expression of ELFN1 might be an independent predictor correlated with worse clinical outcomes of patients in COAD. These results mentioned above suggested that ELFN1 could be a biomarker of diagnosis and prognosis for COAD.

To explore the potential molecular mechanism, we performed GSEA and co-expression analyses. KEGG analysis showed that ELFN1 was associated with several types of tumor signaling pathways, embracing PI3K-AKT, Notch, mTOR signaling pathways, and so on. Also, the results of GSEA exhibited that ELFN1 was enriched in many immunological and cancer-related signaling pathways, like Hedgehog, ECM receptor interaction, focal adhesion, Notch, and MAPK signaling pathways, which indicated that ELFN1 could promote the progression of COAD by diverse biological processes. Additionally, numerous genes encoding several extracellular matrix proteins, like SPON2(18), COL18A1 (19), and SNAI1 (20), were also identified to affect the tumor progression. The extracellular matrix proteins encoded by the genes mentioned above could be interacted with the protein encoded by ELFN1. These results implied that ELFN1 can promote tumor progression through various pathways and it might act as a new and potential therapeutic target for COAD.

A growing number of proofs have focused on the interrelationship between tumor and immune stromal cells (21, 22). Immune escape (23) and functional inhibition of TIICs in the tumor microenvironment(TME) (24) are two significant barriers to cancer immunotherapy.

We utilized CIBERSORT to explore the association of 22 TIICs with ELFN1 in the TME of COAD from the TCGA database. Our results revealed showed a negative nexus between the mRNA expression of ELFN1 and infiltration levels of activated dendritic cells, activated and resting CD4 memory T cells in COAD. Dendritic cells can capture tumor antigens and present them to cytotoxic T lymphocytes (25). Tumor cells can prevent T cells' recruitment and activation via the activated oncogenic pathways in the TME(26). Besides, we observed a positive association between Tregs' infiltration level and ELFN1 expression. We also observed that the high abundance of Tregs is related to the advanced tumor stage. That might imply

ELFN1 and Tregs are implicated in the progression of COAD. Additionally, we observed a positive association between M0 infiltration level and ELF1 expression. We also found that the high relative abundance of M0 Macrophages was positively correlated with the early N stage (N1 vs. N0). However, it revealed the low relative abundance of M0 Macrophages in the advanced N stage (N2 vs. N1). These results imply that different infiltration level of M0 Macrophages in different stages may play different roles in tumor progression. Luo et al. found the increase of Tregs and M0 macrophages in COAD was correlated with worse clinical outcomes (27). In contrast, Forssell et al. found that high macrophages in CC was associated with better patients' survival (28). The reason for this discrepancy may be attributed to cell differentiation with time specificity. M0 macrophages would polarize to M1 and M2 macrophages under different stimulations (29). M1 macrophages act as a protector and appear in the early stage of tumorigenesis (30). However, M2 macrophages play a crucial role in the invasion and immune escape of tumors and weaken the cell attack capacity of natural killer cells and T cells (31). Besides, M2 macrophages reach the peak in the period of tumor progression (32). We all know that most CC belongs to advanced-stage tumors at the diagnosis, so that M0 macrophages might polarize to more M2 macrophages to promote tumor progression. Surprisingly, there was no correlation between M2 macrophages' infiltration level and ELF1 expression. The number of tumor tissues in our study is not large enough, which may be one of the reasons why there is no relation between ELF1 expression and the infiltration level of M2 macrophages in the TME. But a positive trend can be observed between ELF1 expression and M2 macrophages infiltration. Furthermore, we observed the infiltration level of M0 macrophages in COAD was negatively associated with that of M1 Macrophages, activated dendritic cells, and activated CD4 memory T cells, which robustly affirmed the potential biological function of M0 macrophages in the progression of COAD. Among immune cell types, we saw the appreciable nexus between ELF1 and M0 macrophages, Tregs, which indicated that ELF1 could regulate immune cell infiltration in the TME and be a latent anticancer target.

This paper has its own limitations. Firstly, our results originated from the analysis of RNA sequencing data of 514 COAD cases from TCGA and which were verified in a few GEO databases; thus, that should be verified in more and larger cohorts in the future. Secondly, algorithm analysis depended upon RNA-seq, which might not be absolutely accurate to draw the conclusions; so, more corresponding clinical and basic experiments are indispensable to uncover the latent biological mechanisms of ELF1 in COAD in future period.

## Conclusions

ELFN1 is overexpressed in COAD, which is correlated with worse clinical outcomes. Higher expression of ELF1 could influence immune cells infiltrating in the TME. ELF1 could promote progression of COAD through several immunological and cancer-related signaling pathways. ELF1 may be a new prognostic predictor and a latent therapeutic target for COAD.

## Materials And Methods

## Data sources

Clinical data and Level 3 RNA sequencing data of 514 COAD cases (normal samples, 41 cases; tumor samples, 473 cases) came from the TCGA site (<https://portal.gdc.cancer.gov>, December 14, 2020). Gene mutation data were extracted from the TCGA project. DNA methylation data (n=347) were acquired from the GDC TCGA-COAD cohort on the UCSC Xena browser (<https://xenabrowser.net>). Pan-Cancer data of TCGA were obtained from the Xena browser. The whole data for this study came from public sources, and no specific ethical approval is required.

## The mRNA expression analysis of ELFN1

The "limma" (33) package was applied to analyze the mRNA expression profile of ELFN1 in TCGA-COAD via the R 4.0.3 software(34). The variations of ELFN1 mRNA expression between tumor tissue and neighboring normal tissue were visualized via the "ggpubr"(35) package. After deleting 27 samples due to incomplete clinical data, 446 samples were obtained for further analyzing the association of ELFN1 expression with clinicopathological features in TCGA-COAD using the "ggpubr" package. For verification, the expression profile of ELFN1 in GSE39582 (36) (tumor, n=443; normal, n=19) and GSE146009 (37) (tumor, n=33; normal, n=32) datasets came from the GEO website (38) (<https://www.ncbi.nlm.nih.gov/geo>). Besides, the mRNA expression data of ELFN1, including tumor and normal tissues in other 32 cancer types from the TCGA, were also analyzed using the same methods described above. Matrixes were filtered, depending on analysis requirements via the Perl language (<https://www.perl.org/>) scripts.

## Survival analysis

Survival information of TCGA samples was obtained from TCGA, including PFI and OS. After removing samples with OS less than 30 days, 421 samples were included in the final analysis. Samples were separated into low or high groups depending on the median of ELFN1 mRNA expression level. The Kaplan-Meier (K-M) curves were plotted via the "survminer" (39) and "survival" (40) package. Univariate and multivariate analyses of prognostic factors were conducted with Cox regression by the "survival" package. The ROC curves were plotted via the "timeROC" (41), "survminer," "survival," "pROC" (42) packages. The survival outcomes of GSE17536 (43)(CC samples, n=177), GSE29621(44) (CC samples, n=65) came from the GEO database to verify the results. Additionally, the same method was used for survival analysis in the other 32 cancer types.

## The nexus of ELFN1 expression and TIICs

The "CIBERSORT R script v1.03" (45) was applied to compute the distribution of 22 TIICs in COAD depending upon the transcriptome profiles. Then, the relative abundances of each type of immune cells were calculated separately. Samples were separated into low or high expression groups depending upon the median mRNA expression level of ELFN1 with  $P < 0.05$ . The Spearman method was performed to analyze the correlation between the relative abundance of TIICs and mRNA expression level of ELFN1.

The results were visualized by the correlation test  $P < 0.05$ . Next, the associations of the relative abundances of ELFN1-related TlICs with clinicopathological features and OS in TCGA-COAD were also analyzed. These analyses were performed using the "limma" "ggExtra(46)" "ggpubr" "ggplot2(47)" "corrplot(48)" "vioplot(49)" "survival," and "survminer" packages in the R software.

## ELFN1 methylation analysis

The "plyr(50)", "reshape2(51)", "ggpubr" and "ggplot2" packages were performed to analyze and visualize the methylation of ELFN1. The associations between clinicopathologic variables and the methylation of ELFN1 were analyzed and visualized via the packages as described above. The survival analysis between the methylation of ELFN1 and OS in TCGA-COAD was done via the "survminer" and "survival" packages.

## ELFN1 mutation analysis

The "GenVisR" (52) package was applied to analyze the ELFN1 mutation data from the TCGA-COAD and the International Cancer Genome Consortium website (<https://dcc.icgc.org/releases/current/Projects/COCA-CN>).

## Gene Set Functional Enrichment Analysis

The "c2.cp.kegg.v7.2.symbols.gmt" of the GSEA 4.1.0 software (<http://software.broadinstitute.org/gsea/index.jsp>) was applied for performing GSEA to identify differentially activated signaling pathways in high and low ELFN1 expression of COAD datasets. Gene set permutations were set at 2000. The gene set was considered obviously enrichment with  $Q < 0.05$  and  $P < 0.05$ .

Depending upon the  $|Pearson\ correlation\ coefficient(r)| \geq 0.40$  and  $P < 0.001$ , PCGs were considered ELFN1-related PCGs. The KEGG and GO enrichment analysis of ELFN1-related PCGs and visualization of results were conducted via the "clusterProfiler(53)" "org.Hs.eg.db(54)" "DOSE(55)" packages. The PPI networks were built through the tools called STRING (<https://string-db.org/>, Version: 11.0) and Cytoscape 3.8.2. ELFN1-related co-expressed genes with the  $|r| \geq 0.6$  and  $P < 0.001$  were utilized to construct the correlation heat map and Chord diagram via the "corrplot" package.

## Statistical analysis

The R software (version 4.0.3, <https://www.r-project.org>) was applied for all Statistical analyses. Survival analysis was conducted with the K-M method. The significance of prognostic factors was evaluated via univariate and multivariate Cox regression analyses. Subgroup differential analyses were applied through Wilcoxon rank-sum test and Kruskal-Wallis test. Differences were considered statistically significant with  $P < 0.05$ ; all statistical analyses employed two-tailed tests.

## Abbreviations

Extracellular leucine-rich repeat and fibronectin type III domain-containing 1 ELFN1

Colon adenocarcinoma COAD

Colorectal cancer CRC

Colon cancer CC

Tumor immune microenvironment TIME

Tumor microenvironment TME

Tumor-infiltrating immune cells TIICs

The Cancer Genome Atlas TCGA

Gene Set Enrichment Analysis GSEA

Gene Expression Omnibus GEO

Kyoto Encyclopedia of Genes and Genomes KEGG

Gene Ontology GO

Overall survival OS

Progression-free survival PFI

Disease-free survival DFS

Regulatory T cells Tregs

International Cancer Genome Consortium ICGC

Carcinoembryonic antigen CEA

Protein-coding genes PCGs

Protein-protein Interaction PPI

## Declarations

### Ethics approval and consent to participate

*The Ethics Committee* of the First Affiliated Hospital of Kunming Medical University ruled that no ethics approval and consent to participate were required in this research.

## **Consent for publication**

Not applicable.

## **Availability of data and materials**

The datasets generated and analyzed during the current study are available in the TCGA and GEO repositories, [<https://portal.gdc.cancer.gov>, December 14, 2020; <https://www.ncbi.nlm.gov/geo> ]

## **Competing interests**

The authors declare that they have no competing interests.

## **Funding**

Not applicable.

## **Authors' contribution**

ZP-L edited the manuscript of this article. LZ conducted the database mining, data analysis and charts compiling. YY searched the literature. SB conceived and designed the study. All authors read and approved the final manuscript.

## **Acknowledgments**

The authors appreciate that TCGA, GEO, and ICGC websites share a considerable amount of data. The authors thank the programmers for developing the software tools to acquire the data and prepare the figures.

## **References**

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Erratum. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J Clin.* 2020;70(4):313.
2. Or CR, Chang Y, Lin WC, Lee WC, Su HL, Cheung MW, et al. Obatoclax, a Pan-BCL-2 Inhibitor, Targets Cyclin D1 for Degradation to Induce Antiproliferation in Human Colorectal Carcinoma Cells. *Int J Mol Sci.* 2016;18(1).
3. Suh SS, Kim TK, Kim JE, Hong JM, Nguyen TTT, Han SJ, et al. Anticancer Activity of Ramalin, a Secondary Metabolite from the Antarctic Lichen Ramalina terebrata, against Colorectal Cancer Cells. *Molecules (Basel, Switzerland)*. 2017;22(8).
4. Fawzy MS, Ibrahim AT, AlSel BTA, Alghamdi SA, Toraih EA. Analysis of microRNA-34a expression profile and rs2666433 variant in colorectal cancer: a pilot study. *Scientific reports.* 2020;10(1):16940.

5. Zhao Z, Qin X. MicroRNA-708 targeting ZNF549 regulates colon adenocarcinoma development through PI3K/Akt pathway. *Scientific reports*. 2020;10(1):16729.
6. Liu DX, Li DD, He W, Ke CF, Jiang W, Tang JH, et al. PD-1 blockade in neoadjuvant setting of DNA mismatch repair-deficient/microsatellite instability-high colorectal cancer. *Oncoimmunology*. 2020;9(1):1711650.
7. Wu T, Wu X, Wang HY, Chen L. Immune contexture defined by single cell technology for prognosis prediction and immunotherapy guidance in cancer. *Cancer communications (London England)*. 2019;39(1):21.
8. Maybruck BT, Pfannenstiel LW, Diaz-Montero M, Gastman BR. Tumor-derived exosomes induce CD8(+) T cell suppressors. *J immunother Cancer*. 2017;5(1):65.
9. Cao Y, Sarria I, Fehlhaber KE, Kamasawa N, Orlandi C, James KN, et al. Mechanism for Selective Synaptic Wiring of Rod Photoreceptors into the Retinal Circuitry and Its Role in Vision. *Neuron*. 2015;87(6):1248–60.
10. Dolan J, Mitchell KJ. Mutation of Elfn1 in mice causes seizures and hyperactivity. *PLoS one*. 2013;8(11):e80491.
11. Dunn HA, Patil DN, Cao Y, Orlandi C, Martemyanov KA. Synaptic adhesion protein ELFN1 is a selective allosteric modulator of group III metabotropic glutamate receptors in trans. *Proc Natl Acad Sci USA*. 2018;115(19):5022–7.
12. Tomioka NH, Yasuda H, Miyamoto H, Hatayama M, Morimura N, Matsumoto Y, et al. Elfn1 recruits presynaptic mGluR7 in trans and its loss results in seizures. *Nature communications*. 2014;5:4501.
13. Cao Y, Wang Y, Dunn HA, Orlandi C, Shultz N, Kamasawa N, et al. Interplay between cell-adhesion molecules governs synaptic wiring of cone photoreceptors. *Proc Natl Acad Sci USA*. 2020;117(38):23914–24.
14. Wang D, Rai B, Qi F, Liu T, Wang J, Wang X, et al. Influence of the Twist gene on the invasion and metastasis of colon cancer. *Oncol Rep*. 2018;39(1):31–44.
15. Lu H, Shi T, Wang M, Li X, Gu Y, Zhang X, et al. B7-H3 inhibits the IFN- $\gamma$ -dependent cytotoxicity of V $\gamma$ 9V82 T cells against colon cancer cells. *Oncoimmunology*. 2020;9(1):1748991.
16. He JH, Han ZP, Luo JG, Jiang JW, Zhou JB, Chen WM, et al. Hsa\_Circ\_0007843 Acts as a miR-518c-5p Sponge to Regulate the Migration and Invasion of Colon Cancer SW480 Cells. *Frontiers in genetics*. 2020;11:9.
17. Kim HJ, Kim M, Kang B, Yun S, Ryeo SE, Hwang D, et al. Systematic analysis of expression signatures of neuronal subpopulations in the VTA. *Mol Brain*. 2019;12(1):110.
18. Kang HG, Kim WJ, Noh MG, Chun KH, Kim SJ. SPON2 Is Upregulated through Notch Signaling Pathway and Promotes Tumor Progression in Gastric Cancer. *Cancers (Basel)*. 2020;12(6).
19. Guo Z, Zhang T, Wu J, Wang H, Liu X, Tian L. Genetic polymorphisms in COL18A1 influence the development of osteosarcoma. *Int J Clin Exp Pathol*. 2015;8(9):11531–6.

20. Feng L, Zhang J, Sun M, Qiu F, Chen W, Qiu W. Tumor Suppressor LINC02487 Inhibits Oral Squamous Cell Carcinoma Cell Migration and Invasion Through the USP17-SNAI1 Axis. *Front Oncol.* 2020;10:559808.
21. Pagès F, Mlecnik B, Marliot F, Bindea G, Ou FS, Bifulco C, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet.* 2018;391(10135):2128–39.
22. Restifo NP, Smyth MJ, Snyder A. Acquired resistance to immunotherapy and future challenges. *Nature reviews Cancer.* 2016;16(2):121–6.
23. Di Pace AL, Tumino N, Besi F, Alicata C, Conti LA, Munari E, et al. Characterization of Human NK Cell-Derived Exosomes: Role of DNAM1 Receptor In Exosome-Mediated Cytotoxicity Against Tumor. *Cancers (Basel).* 2020;12(3).
24. Kalvala A, Wallet P, Yang L, Wang C, Li H, Nam A, et al. Phenotypic Switching of Naïve T Cells to Immune-Suppressive Treg-Like Cells by Mutant KRAS. *Journal of clinical medicine.* 2019;8(10).
25. Zhu H, Shan Y, Ge K, Lu J, Kong W, Jia C. Oxaliplatin induces immunogenic cell death in hepatocellular carcinoma cells and synergizes with immune checkpoint blockade therapy. *Cellular oncology (Dordrecht).* 2020;43(6):1203–14.
26. Chen D, Wang H. The clinical and immune features of CD14 in colorectal cancer identified via large-scale analysis. *Int Immunopharmacol.* 2020;88:106966.
27. Luo J, Liu P, Wang L, Huang Y, Wang Y, Geng W, et al. Establishment of an immune-related gene pair model to predict colon adenocarcinoma prognosis. *BMC Cancer.* 2020;20(1):1071.
28. Forsslund J, Oberg A, Henriksson ML, Stenling R, Jung A, Palmqvist R. High macrophage infiltration along the tumor front correlates with improved survival in colon cancer. *Clinical cancer research: an official journal of the American Association for Cancer Research.* 2007;13(5):1472–9.
29. Murray PJ, Allen JE, Biswas SK, Fisher EA, Gilroy DW, Goerdt S, et al. Macrophage activation and polarization: nomenclature and experimental guidelines. *Immunity.* 2014;41(1):14–20.
30. Zhao Y, Ge X, Xu X, Yu S, Wang J, Sun L. Prognostic value and clinicopathological roles of phenotypes of tumour-associated macrophages in colo rectal cancer. *J Cancer Res Clin Oncol.* 2019;145(12):3005–19.
31. Wang Y, Wang Y, Xu C, Liu Y, Huang Z. Identification of Novel Tumor-Microenvironment-Regulating Factor That Facilitates Tumor Immune Infiltration in Colon Cancer. *Molecular therapy Nucleic acids.* 2020;22:236–50.
32. Molinero L, Li Y, Chang CW, Maund S, Berg M, Harrison J, et al. Tumor immune microenvironment and genomic evolution in a patient with metastatic triple negative breast cancer and a complete response to atezolizumab. *J immunother Cancer.* 2019;7(1):274.
33. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research.* 2015;43(7):e47-e.
34. Team RC. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2020.

35. Kassambara A. *ggpubr: "ggplot2" Based Publication Ready Plots* (Version 0.3. 0)[Computer software]. 2020.
36. Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP, Vescovo L, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med.* 2013;10(5):e1001453.
37. Paredes J, Zabaleta J, Garai J, Ji P, Imtiaz S, Spagnardi M, et al. Immune-Related Gene Expression and Cytokine Secretion Is Reduced Among African American Colon Cancer Patients. *Frontiers in oncology.* 2020;10:1498.
38. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic acids research.* 2012;41(D1):D991-D5.
39. Kassambara A, Kosinski M. *survminer: drawing survival curves using "ggplot2."* 2017. 2020.
40. Therneau T. *A Package for Survival Analysis in R.* R package version 3.2-3. Computer software] Rochester, MN: Mayo Clinic Retrieved from <https://CRAN.R-project.org/package=survival>. 2020.
41. Blanche P, Dartigues J-F, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med.* 2013;32(30):5381–97.
42. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S + to analyze and compare ROC curves. *BMC Bioinformatics.* 2011;12:77.
43. Smith JJ, Deane NG, Wu F, Merchant NB, Zhang B, Jiang A, et al. Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. *Gastroenterology.* 2010;138(3):958–68.
44. Chen D-T, Hernandez JM, Shibata D, McCarthy SM, Humphries LA, Clark W, et al. Complementary strand microRNAs mediate acquisition of metastatic potential in colonic adenocarcinoma. *J Gastrointest Surg.* 2012;16(5):905–13.
45. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nature methods.* 2015;12(5):453–7.
46. Attali D, Baker C. *ggExtra: Add marginal histograms to 'ggplot2', and more 'ggplot2'enhancements.* R package version. 2019.
47. Wickham H. *GGPLOT2: Elegant Graphics for Data Analysis* 2016. New York: Springer-Verlag; 2016.
48. Wei T, Simko V. *R package "corrplot": Visualization of a Correlation Matrix (Version 0.84).* 2017.
49. Adler D, Kelly S. *vioplot: Violin plot.* R package version 0.3. 2. 2018. 2020.
50. Wickham H. The split-apply-combine strategy for data analysis. *Journal of statistical software.* 2011;40(1):1–29.
51. Wickham H. Reshaping data with the reshape package. *Journal of statistical software.* 2007;21(12):1–20.
52. Skidmore ZL, Wagner AH, Lesurf R, Campbell KM, Kunisaki J, Griffith OL, et al. GenVisR: genomic visualizations in R *Bioinformatics.* 2016;32(19):3012–4.

53. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology*. 2012;16(5):284–7.
54. Carlson M org. Hs. eg. db: Genome Wide Annotation for Human. R package version 3.2. 3. 2019.
55. Yu G, Wang L-G, Yan G-R, He Q-Y. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics*. 2015;31(4):608–9.

## Figures

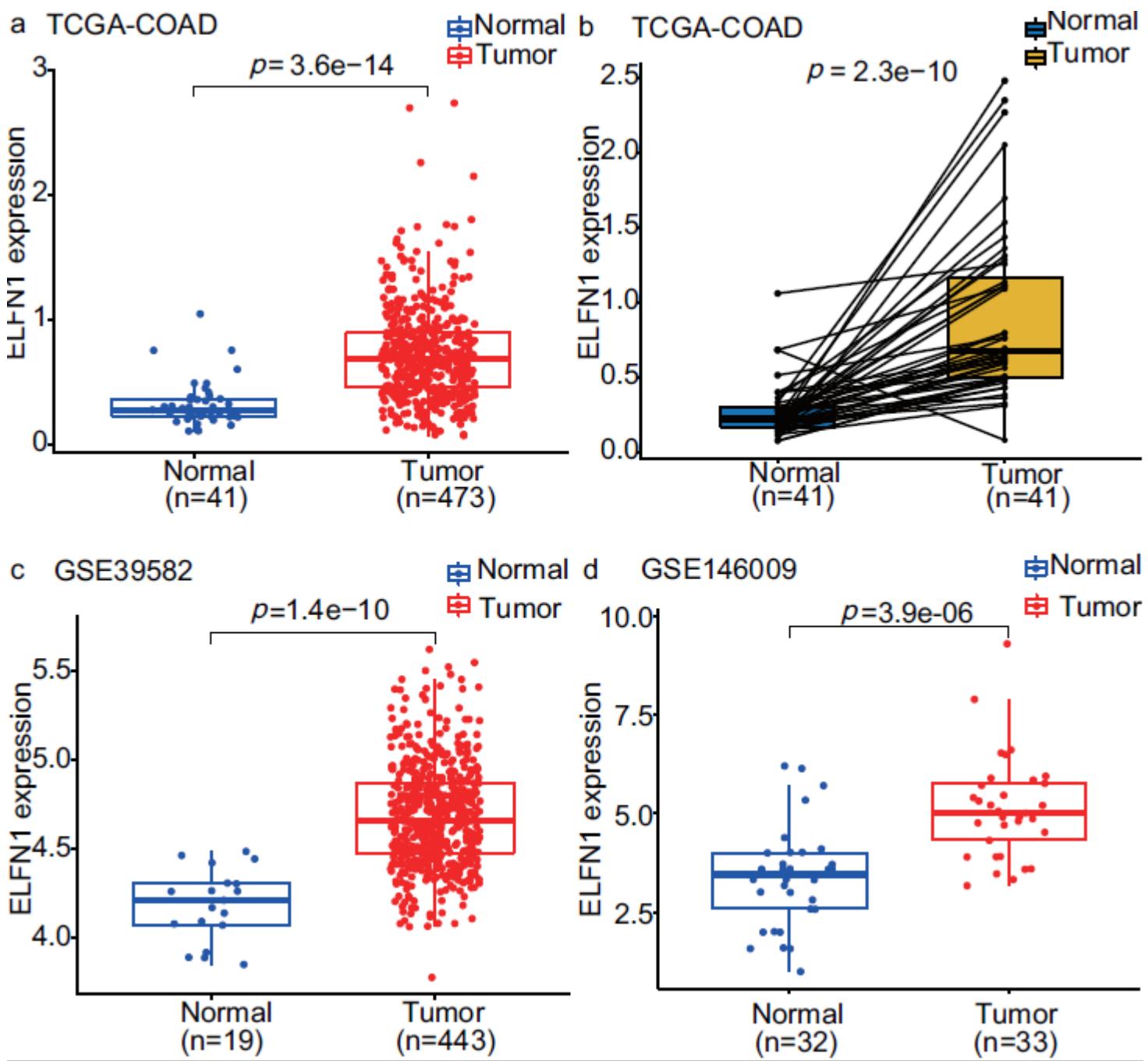
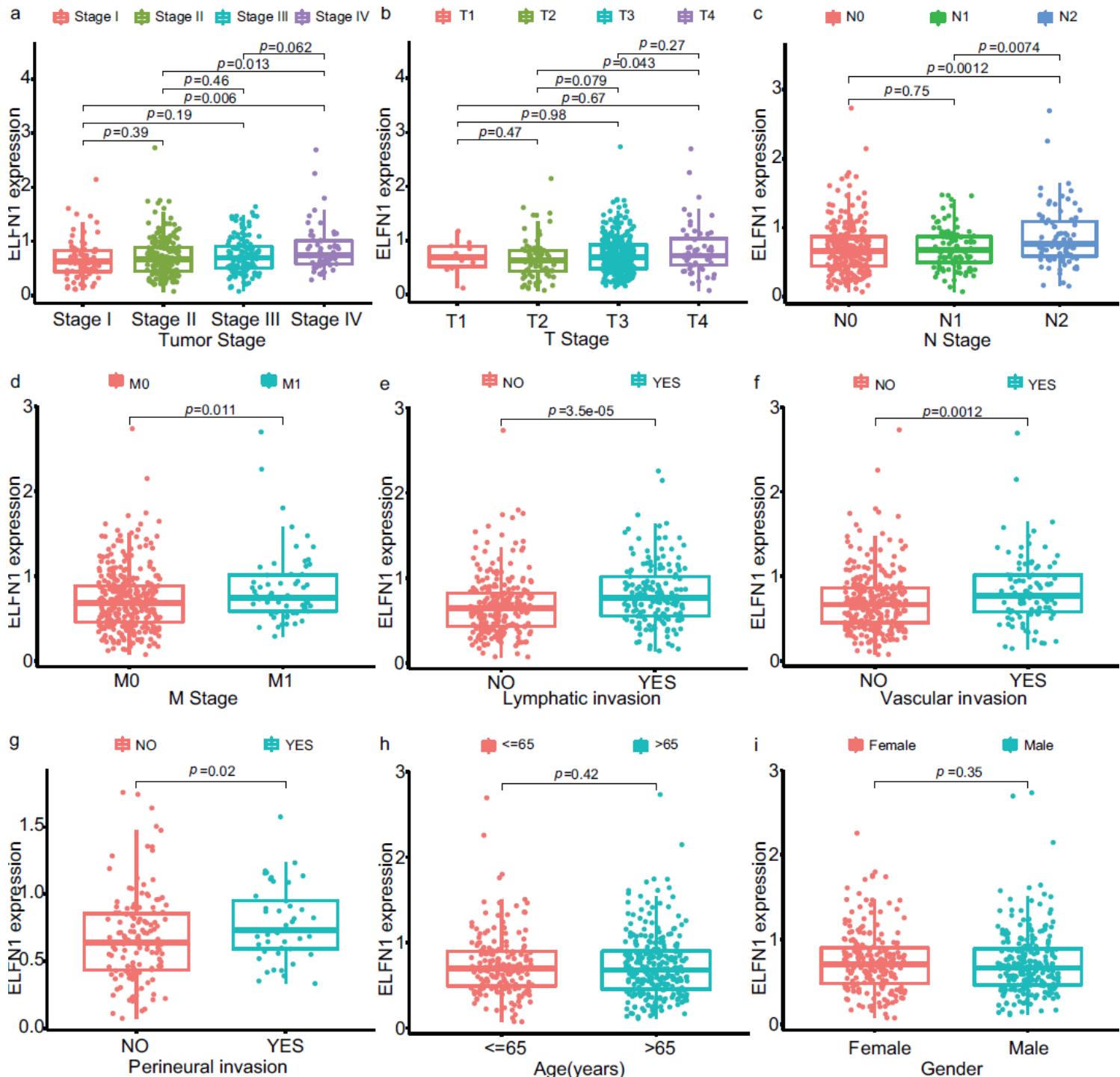


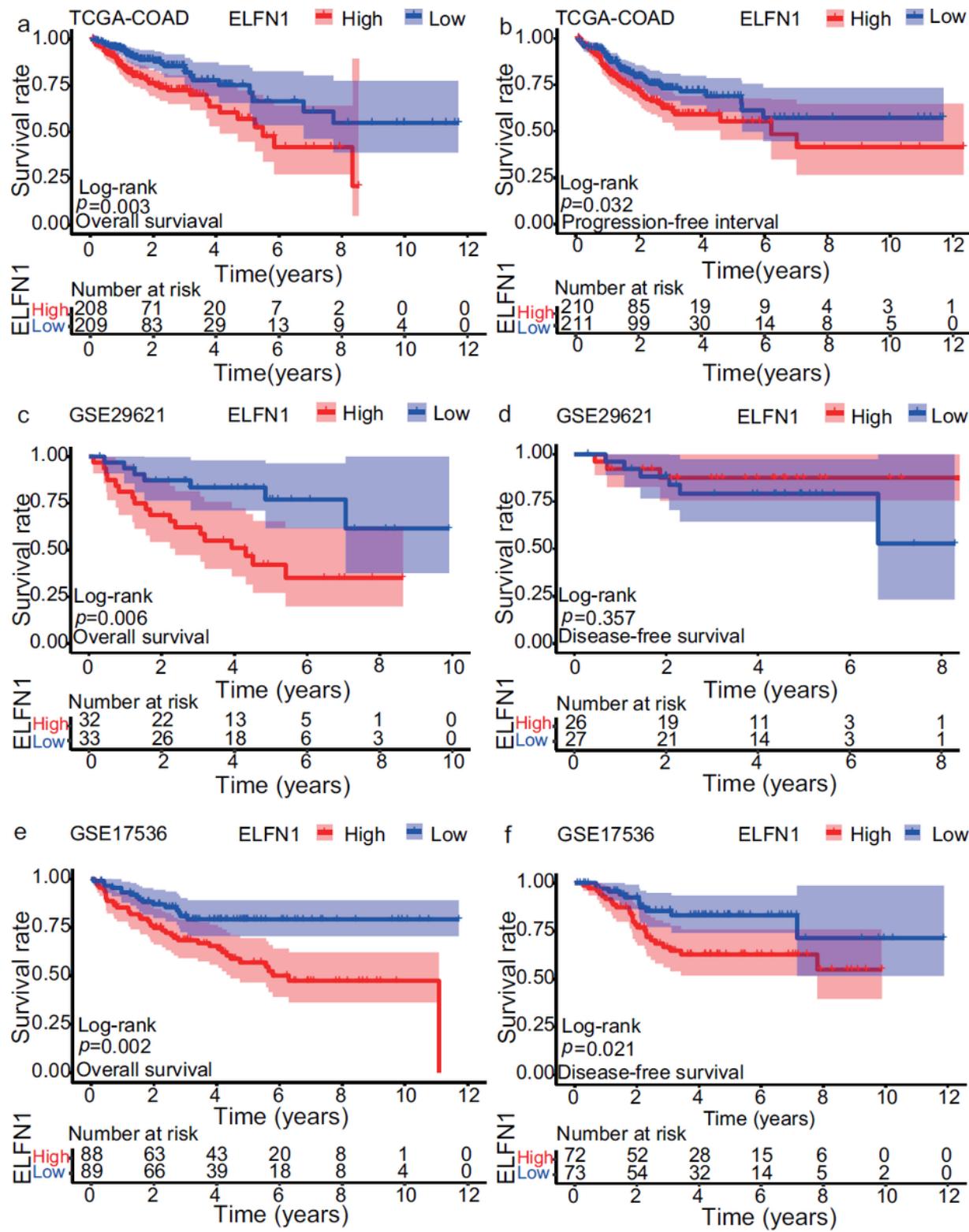
Figure 1

Analysis of ELFN1 expression levels in COAD. a ELFN1 mRNA level in COAD and adjacent normal colon tissues from the TCGA cohort. b ELFN1 mRNA level in matched COAD tissues and normal colon tissues from the TCGA cohort. c ELFN1 mRNA expression in COAD GSE39582 cohort from GEO database. d ELFN1 mRNA level in COAD GSE146009 cohort from GEO database.



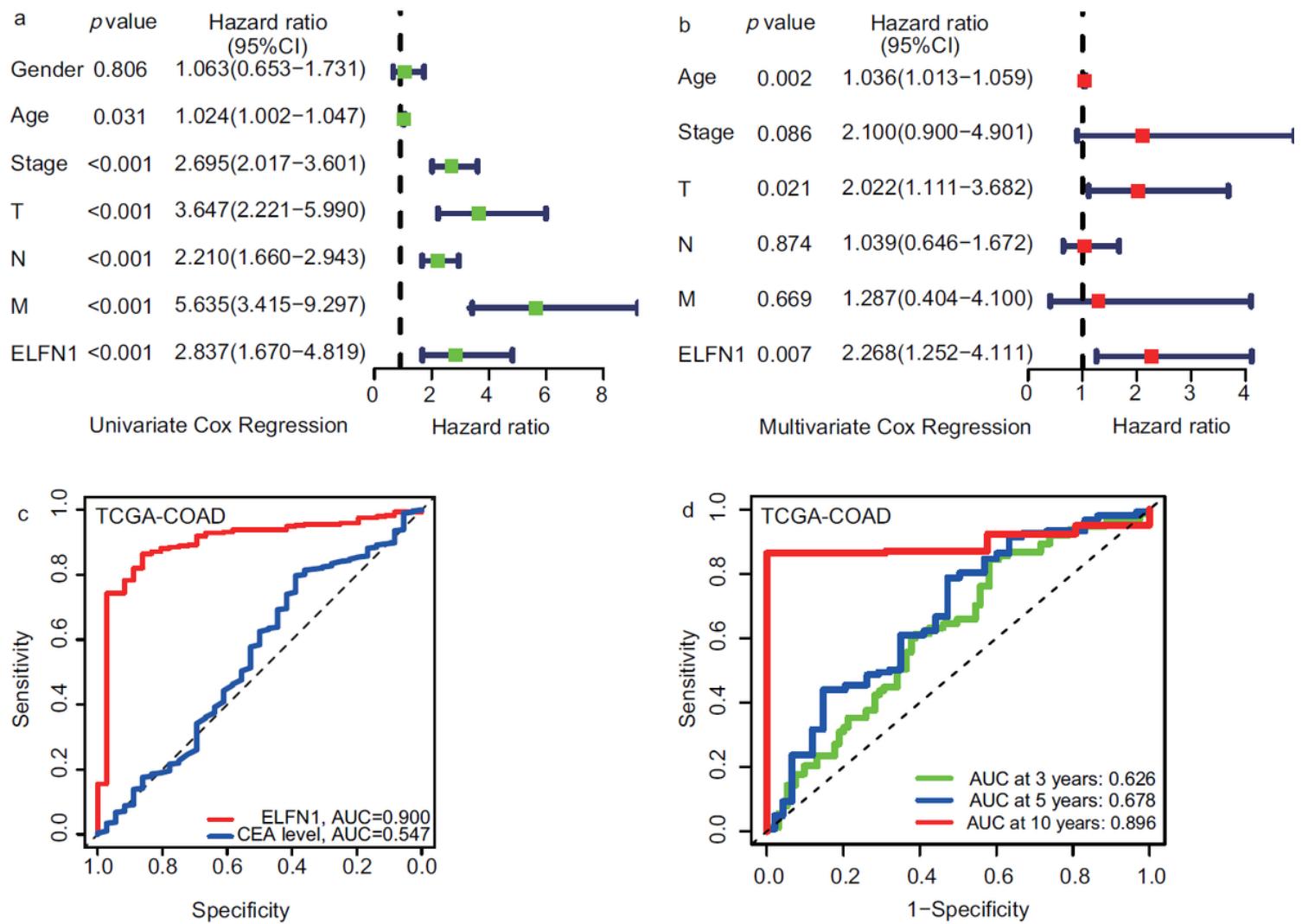
**Figure 2**

Visualization of difference in ELFN1 expression levels among clinicopathological characteristics from the TCGA cohort. a Tumor stage, b T stage, c N stage, d M stage, e Lymphatic invasion, f Vascular invasion, g Perineural invasion, h Age, i Gender.



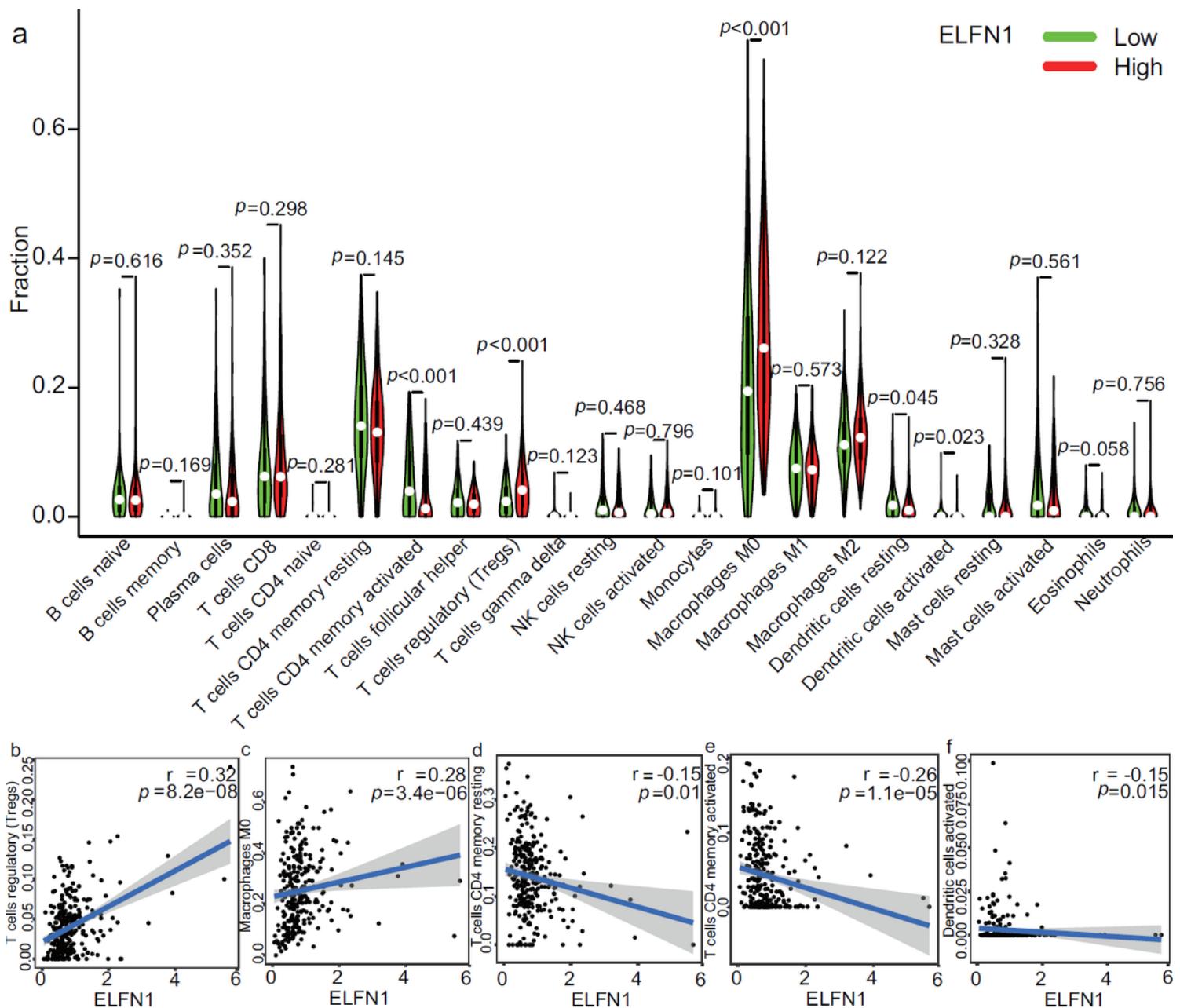
**Figure 3**

Survival analysis of ELFN1 in COAD patients. Kaplan-Meier survival analysis in COAD patients with different ELFN1 mRNA levels: a OS in TCGA COAD dataset, b PFI in TCGA COAD dataset, c OS in GSE29621 dataset, d DFS in GSE29621 dataset, e OS in GSE17536 dataset, f DFS in GSE17536 dataset. OS: Overall survival. PFI: Progression-free interval. DFS: Disease-free survival.



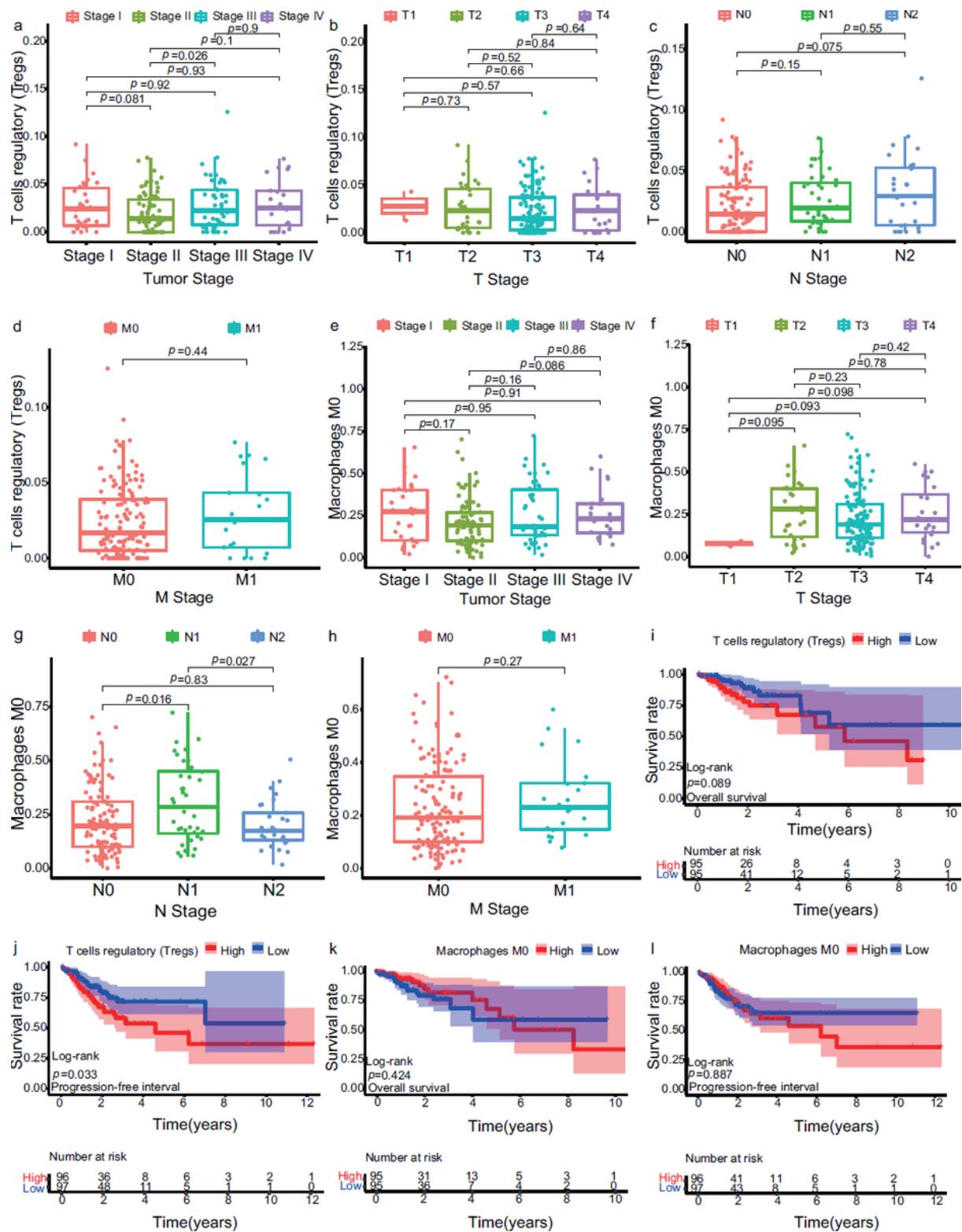
**Figure 4**

Diagnostic and prognostic value of ELFN1 in COAD patients. a Forest plot in univariate Cox regression showed ELFN1 expression could influence the OS for COAD patients. b Forest plot in multivariate Cox regression revealed ELFN1 expression was an independent prognostic indicator for COAD patients. c ROC curve showed ELFN1 had a more satisfactory diagnostic value than that of CEA for COAD patients. d ROC curve revealed the prognostic value of ELFN1 at 3, 5, and 10 years after diagnosis for COAD patients.



**Figure 5**

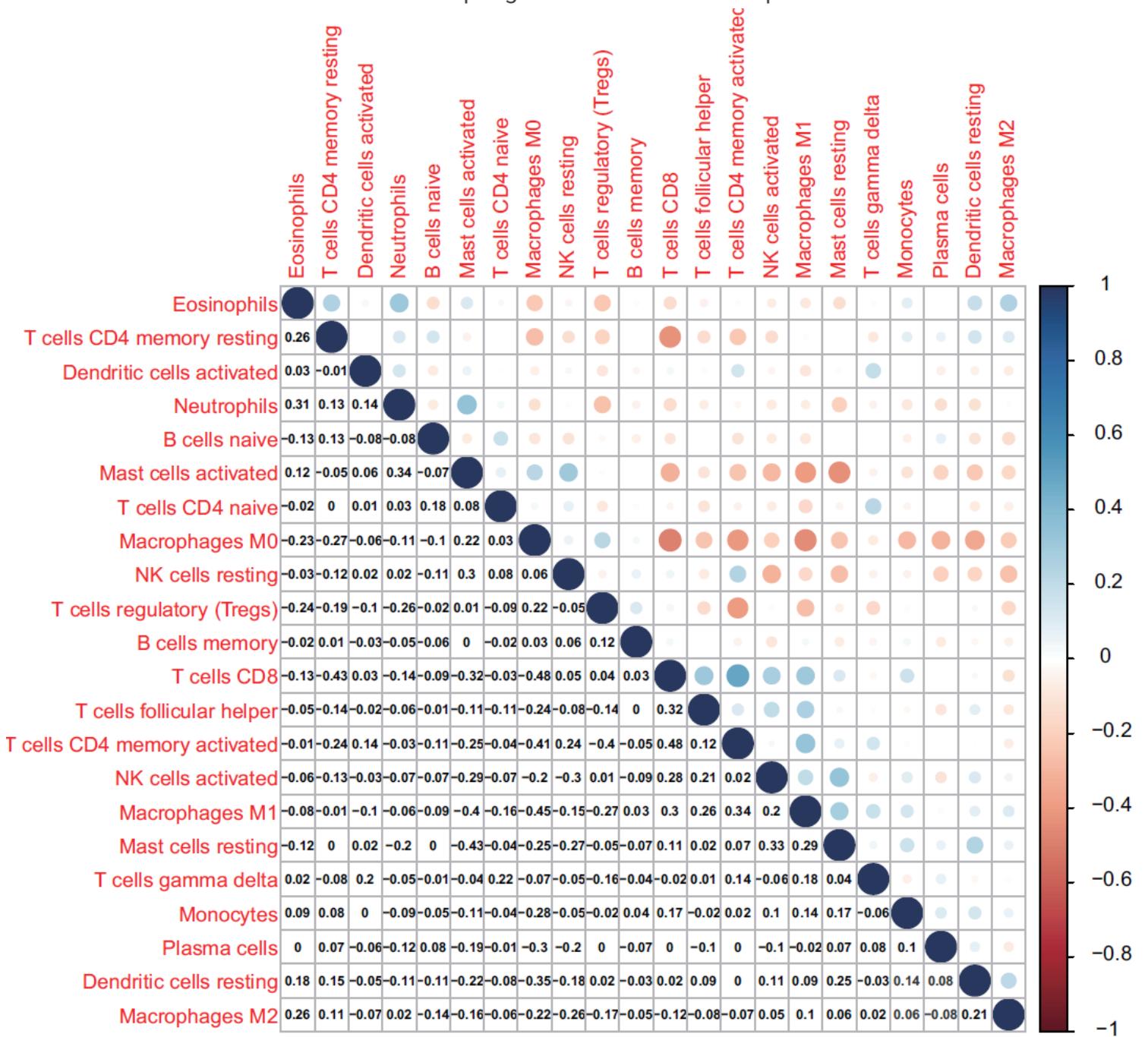
Association between ELFN1 expression and TIIICs in COAD. a The relative abundances of TIIICs in both high ELFN1 and low ELFN1 samples. b-f The correlations between ELFN1 expression level and TIIICs (Tregs, M0 Macrophages, activated Dendritic cells, resting and activated CD4 memory T cells) levels were presented in scatterplots.  $r$  value represented the Spearman correlation coefficient,  $-1 \leq r \leq 1$ ,  $r > 0$  meant a positive correlation, while  $r < 0$  meant a negative correlation.



**Figure 6**

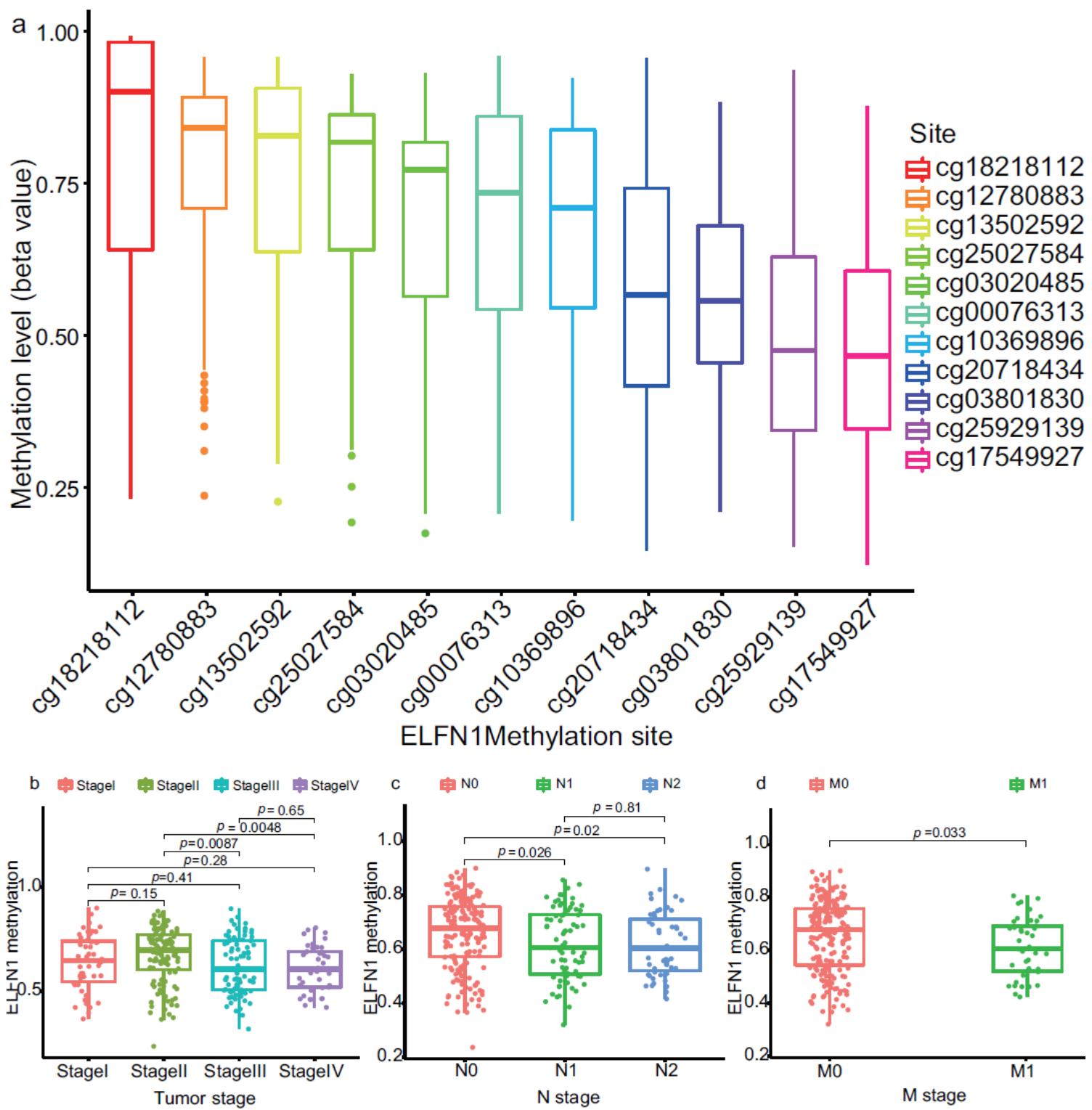
The relationship between TILCs and clinicopathological characteristics. a Higher relative abundance of Tregs in the advanced tumor stage. b-d The relative abundance changes of Tregs did not influence the T, N, M stage. e There was no statistical significance in the Tumor stage with the relative abundance changes of M0 Macrophages. f Relative abundances of M0 Macrophages in different T stages. g Higher relative abundance of M0 Macrophages in the early N stage. h The relative abundance changes of M0

macrophages had no statistical significance on the M stage. i No statistical significance was seen in OS with the relative abundance changes of Tregs. j Lower relative abundance of Tregs with better PFI. k, l The different relative abundances of M0 Macrophages had no influence on patients' OS and PFI.



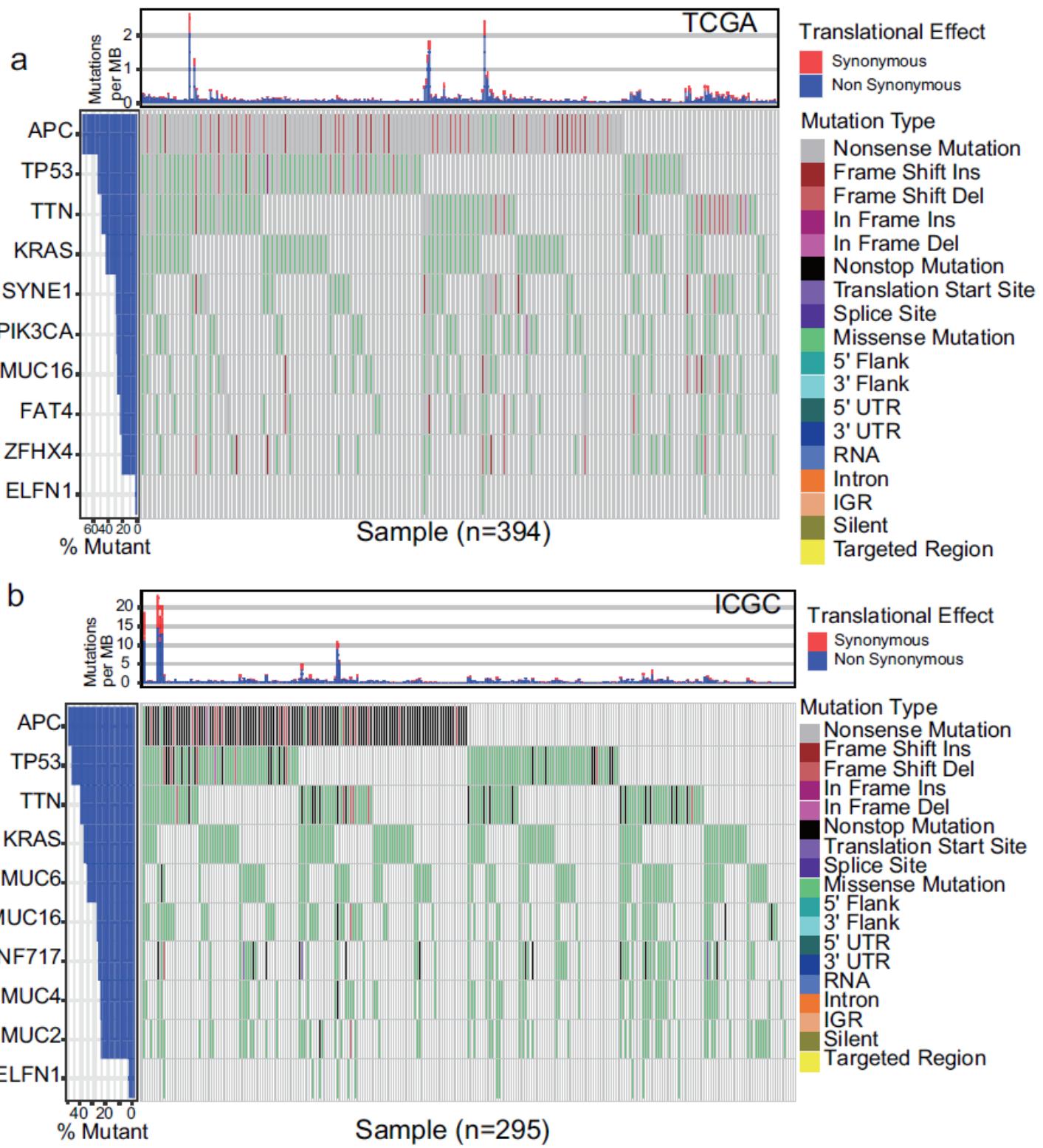
**Figure 7**

Correlations between different TILCs subpopulations in the correlation heatmap. The right ordinate scale reflected r value, which represented the Spearman correlation coefficient,  $-1 \leq r \leq 1$ ,  $r=0$  meant a positive correlation, while  $r=0$  meant a negative correlation.



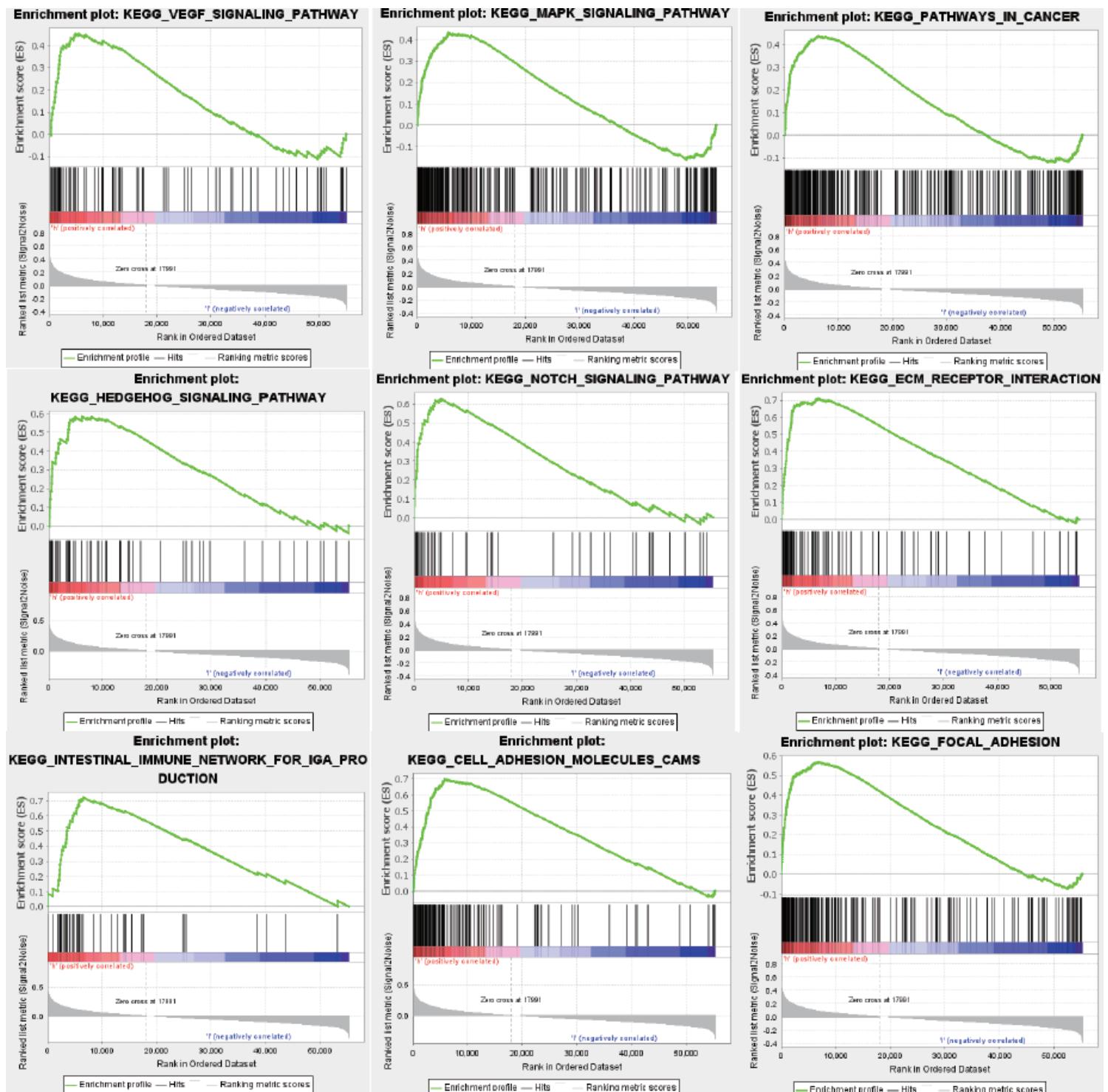
**Figure 8**

Association of ELFN1 methylation with clinicopathological features in COAD from TCGA cohort. a Methylation analysis for 11 CpG sites of the ELFN1 gene in COAD tissues. Visualization of difference in ELFN1 methylation levels among clinicopathological parameters, b Tumor stage, c N stage, d M stage.



**Figure 9**

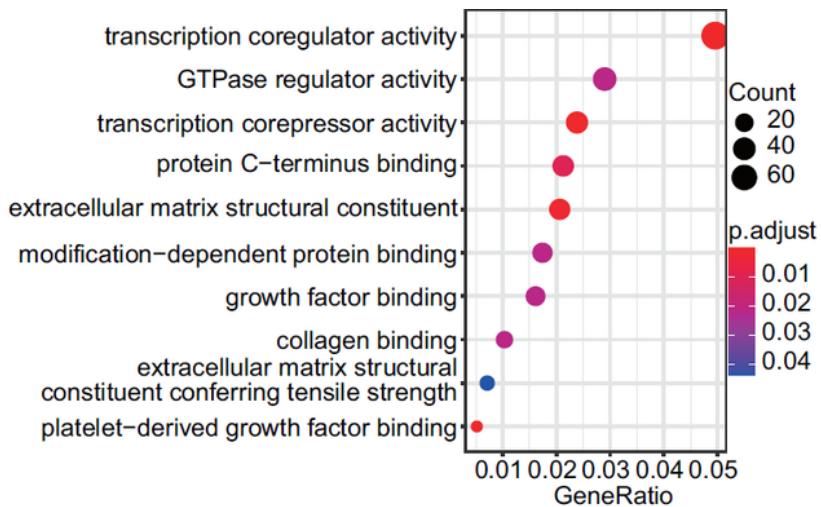
Analysis of ELFN1 mutation in COAD samples from TCGA and ICGC cohorts. a The very low mutation frequency of ELFN1 gene in 394 COAD tissues from TCGA database. b The very low mutation frequency of ELFN1 gene in 295 COAD tissues from ICGC database.



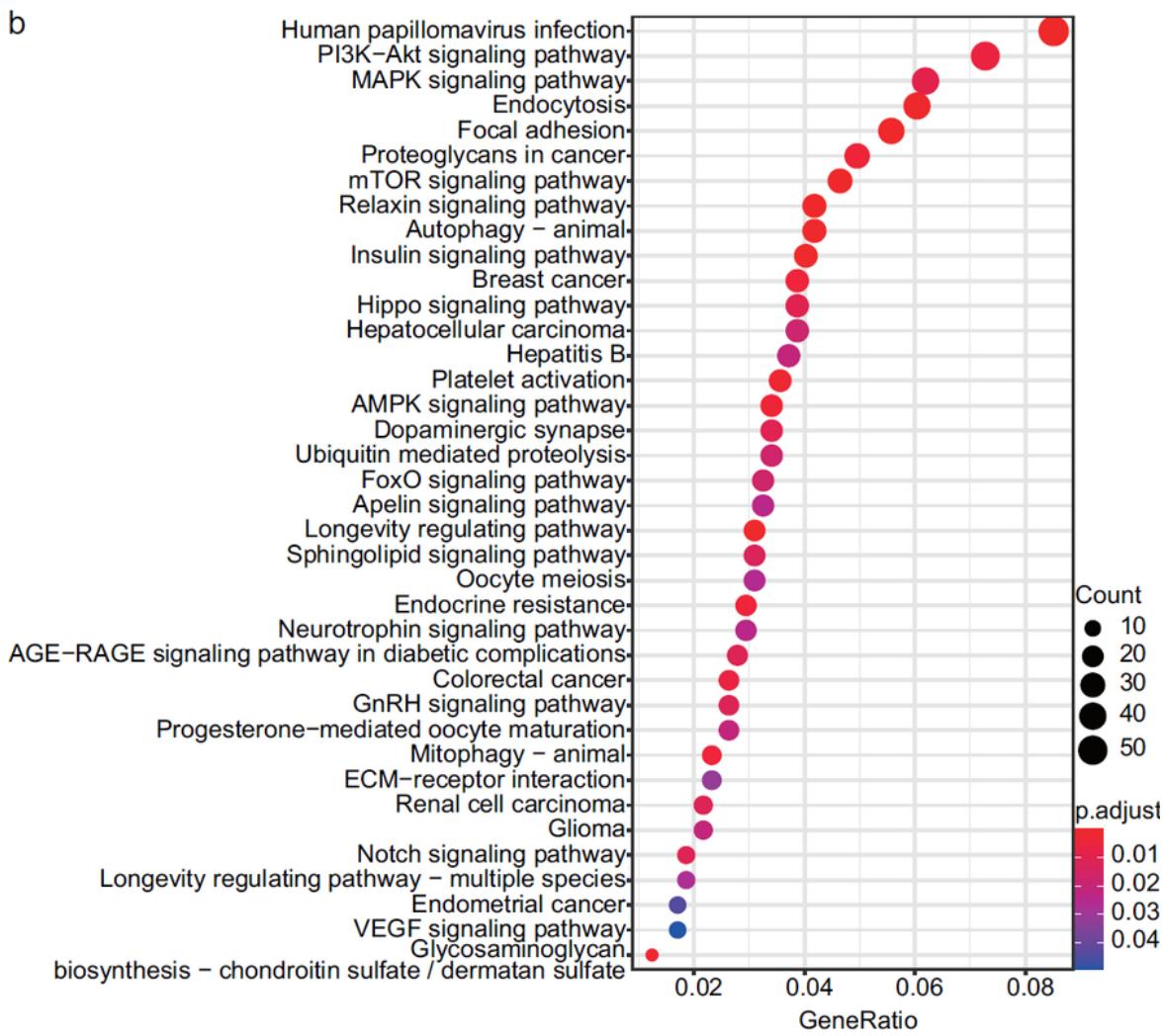
**Figure 10**

Differentially enriched pathways associated with phenotype arising from ELF1 high expression in COAD by GSEA: VEGF, MAPK, pathways in cancer, Hedgehog, Notch, ECM receptor interactions, intestinal immune network for IGA production, CAMs, focal adhesion.

a

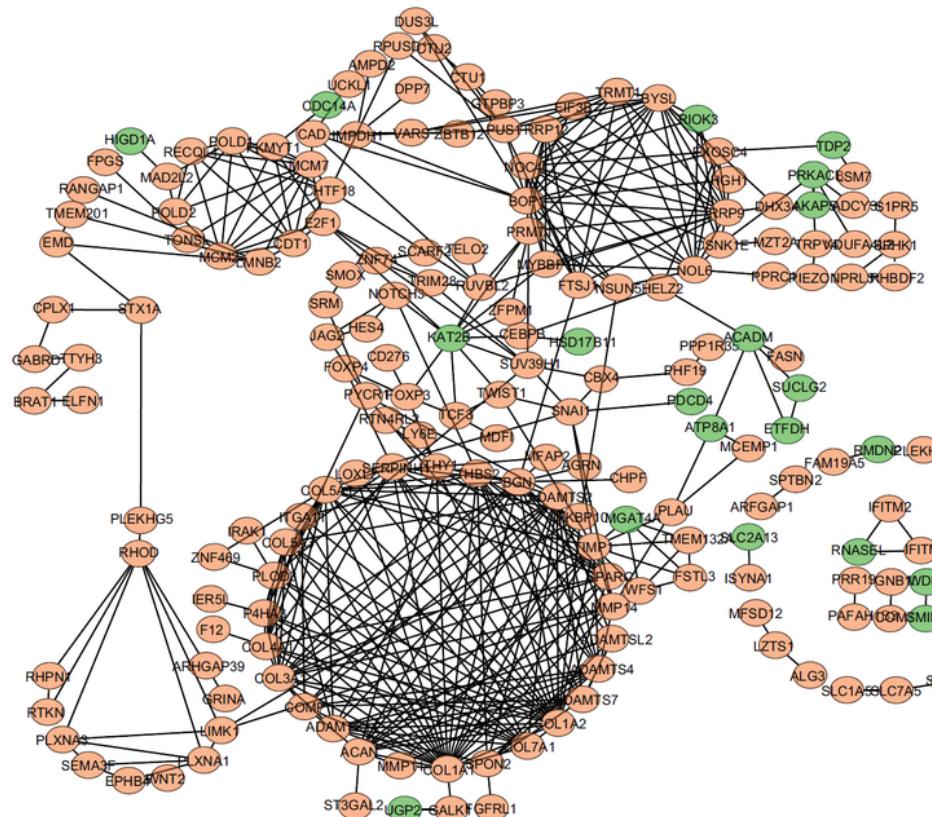


b

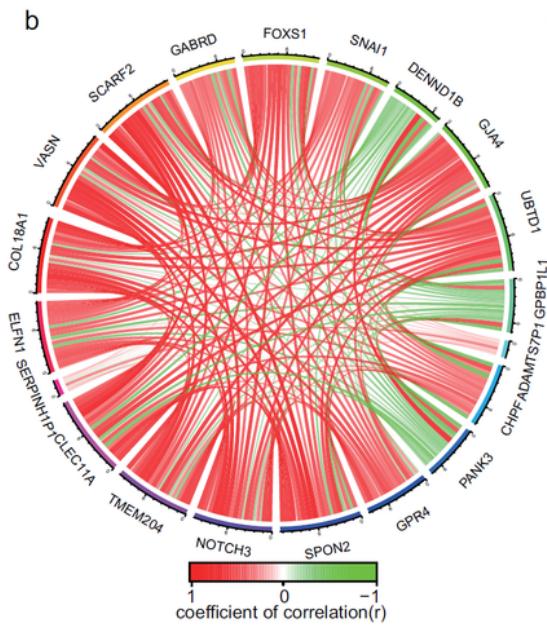
**Figure 11**

GO and KEGG enrichment analyses of ELF1-related PCGs. a GO analysis of positively related biological processes (top ten) with high ELF1 expression in bubble map, b KEGG analysis of many signaling pathways enriched in ELF1-related signatures.

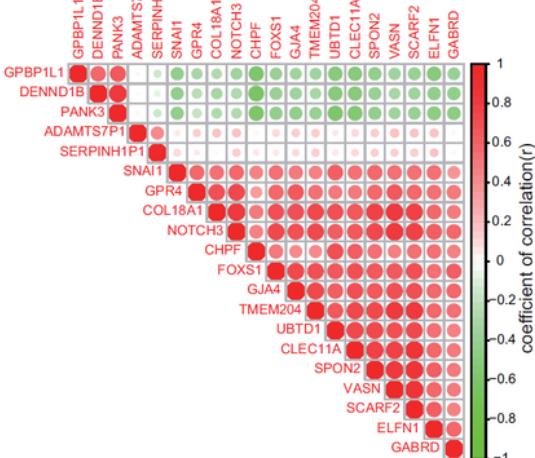
a



b



c

**Figure 12**

Co-expression analysis of ELF11. a The PPI network was constructed with 189 nodes and 432 edges using the STRING and Cytoscape software (note upregulated nodes in brown color and downregulated nodes in green color). b Chord diagram depicting the relationships between ELF11 and co-expression genes, with  $|r| \geq 0.6$ . c Heat map depicting the relationships between ELF11 and co-expression genes, with  $|r| \geq 0.6$ .  $r$  value represented Pearson correlation coefficient,  $-1 \leq r \leq 1$ ,  $r \geq 0$  meant a positive correlation, while  $r \leq 0$

0 meant a negative correlation. The line or circle in red color represented a positive correlation between ELFN1 and co-expression genes, while the line or circle in green color represented a negative correlation between ELFN1 and co-expression genes.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigA1.eps](#)
- [FigA2.eps](#)