Am J Cancer Res 2023;13(6):2572-2587
www.ajcr.us /ISSN:2156-6976/ajcr0150341

Brief Communication

ELL associated factor 2 is a potential diagnostic and prognostic indicator: evidence from the in silico and in vitro experiments

Xiao Yu[1*], Jing-Quan Teng[1*], Jia Chen[1*], Chao Peng[1*], Tong-Rui Guo[2], Meng-Wei Dong[2], Xiao-Lu Hao[1], Kai-Xin Li[1], Xiang-Xiang Huo[3], Cui-Cui Ren[4]

[1]Gansu University of Chinese Medicine, Lanzhou, Gansu, China; [2]Hebei Provincial Hospital of Traditional Chinese Medicine, Shijiazhuang, Hebei, China; [3]Shaanxi University of Chinese Medicine, Xianyang, Shaanxi, China; [4]Xi'an First Hospital, Xi'an, Shaanxi, China. [*]Equal contributors.

Received March 24, 2023; Accepted May 28, 2023; Epub June 15, 2023; Published June 30, 2023

Abstract: Due to the lack of sensitive biomarkers, cancer disease kill 9.6 million individuals each year around the globe. The present study aimed to explore the association between ELL Associated Factor 2 (EAF2) expression and its diagnostic and prognostic landscape across different human cancers using an *in silico* and *in vitro* approach. To achieve the defined goals of this study, we used the following online sources: UALCAN, KM plotter, TNMplot, cBioPortal, STRING, DAVID, MuTarget, Cytoscape, and CTD. In addition to this, we also used additional The Cancer Genome Atlas (TCGA) datasets via TIMER2, GENT2, and GEPIA to confirm the expression of EAF2 on additional cohorts. Finally, we performed RNA sequencing (RNA-seq) and targeted bisulfite sequencing (bisulfite-seq) techniques-based analysis using A549, ABC-1, EBC-1, LK-2 lung cancer cell lines, and MRC-9 normal control lung cell line for further validation of the results. On balance, EAF2 was elevated in 19 types of human cancers and its up-regulation was significantly correlated with shorter overall survival (OS), relapse-free survival (RFS), and metastasis in Liver Hepatocellular Carcinoma (LIHC) and Lung Squamous Cell Carcinoma (LUSC) patients. We further evaluated that EAF2 expression was also elevated across LIHC and LUSC patients belonging to different clinicopathological features. Through pathway analysis, EAF2 associations were observed with four important pathways. Moreover, some worth noticing correlations were also documented between EAF2 expression and its promoter methylation level, genetic alterations, other mutant genes, tumor purity, and different immune cells infiltration. The higher EAF2 expression contributes significantly to the tumorigenesis and metastasis of LIHC and LUSC. Therefore, it can be used as a common biomarker in these cancers.

Keywords: Neoplasm, biomarkers, metastasis

Introduction

Cancer is a condition in which affected cells of the body grow abnormally, depriving normal cells of nutrients and proper functioning [1]. According to a fact sheet provided by the World Health Organization (WHO) in 2021, cancer has been declared the 2nd main cause of death worldwide, causing approximately 19.3 million new cancer cases and 10.0 million cancer-related deaths that year [2]. It is further estimated that in the United States of America (USA) alone, nearly 1,898,160 new cancer cases and 608,570 cancer-related deaths will be registered in 2021 [3] and most of these patients will have to combat the disease for life. According to a survey conducted in Pakistan to record the country-wide incidence of cancer, a total of 63,415 males and 85,590 females were diagnosed with cancer in 2012 [4]. Moreover, breast cancer is ranked first with the highest number of confirmed deaths in Pakistan [5]. There are several internal risk factors for cancer development, including hereditary, hormonal imbalances, metabolic abnormalities, and autoimmune disorders. Moreover, a few external risk factors are also there, including alcohol consumption, dietary imbalance, smoking, exposure to radiation, etc. [6-8]. Despite advancements in cancer diagnosis and therapy, metastases mostly account for the low survival rate in individuals with advanced dis-

ease stages [9, 10]. Therefore, the identification of new diagnostic biomarkers and new treatment targets in cancer patients is urgently needed.

ELL associated factor 2 (EAF2) is a key regulator of transcription elongation and gene expression in both mammals and eukaryotes [11]. Earlier, EAF2 depletion was documented to increase cell proliferation and thus be strongly correlated with cancer in several mouse tissues, demonstrating that it may be a growth inhibitory factor and function as a tumor suppressor [12, 13]. EAF2's contribution to cancer occurrence and progression has already been identified in leukemia, prostate cancer, colorectal cancer, glioblastoma, and gastric cancer [14-17]. However, EAF2 roles in the pathogenesis of other cancers are still unknown. To the best of our knowledge, the EAF2 gene has not yet been investigated in human cancers using pan-caner analysis and *in silico* methods. Therefore, in this manuscript, we analyzed and validated EAF2 expression and prognostic values across various cancers using a variety of *in silico* and *in vitro* experiments.

Materials and methods

EAF2 expression analysis and its cross-validation

The UALCAN database has user-friendly access and easy-to-use features that provide fast access to the cancer multi-omics data (TCGA and MET500) collected from more than 30 different cancer types [18]. We utilized this database for the pan-cancer differential expression analysis of EAF2 across 24 major human cancer types. In UALCAN, the mRNA expression level was normalized as transcript per million (TPM) reads, and box whisker plots were utilized to present the acquired expression values (low to high) from the analyzed sample groups. This tool computes median values from the box whisker plots and compares them to identify the down or up-regulation via student t-test based on the differences in these values between normal and cancerous groups [18]. In the cancerous group, the mRNA with a lower median value than the normal control group is considered down-regulated, while the mRNA with a higher median value than the normal control group is considered up-regulated. In addition to expression data, UALCAN also encompasses the promoter methylation data from TCGA projects, in which every CpG island present in the promoter region of every gene was analyzed to obtain the methylation levels as beta (β) values [18].

Moreover, another tool, MEXPRESS has been developed to visualize The Cancer Genomic Atlas (TCGA) expression data and identify the correlation between promoter methylation and expression level [19]. This tool also encompasses the methylation data from TCGA projects, in which every CpG island present in the promoter region of every gene was analyzed to obtain the methylation levels as beta (β) values [19]. We used this tool to validate the correlation between EAF2 transcription expression and promoter methylation levels in different cancers using new independent cohorts.

The TIMER, GENT2, and GEPIA databases are the cancer transcriptomics data analysis webservers [20-22]. These databases collect and curate large-scale transcriptomic data from different species, including humans. The TIMER, GENT2, and GEPIA provide users with a wide variety of tools and platforms to analyze gene expression profiles under different conditions and across various tissues. Herein, we performed TIMER, GENT2, and GEPIA for the mRNA expression validation of EAF2 across independent cohorts.

Estimation of prognostic and metastatic potentials

KM plotter tool [23] was used for the correlation analysis of EAF2 transcription expression with overall survival (OS) and relapse-free (RFS) durations of the patients suffering from different cancers. KM Plotter is a free online tool that enables users to easily generate Kaplan-Meier plots and perform log-rank survival analysis. It is commonly used by researchers and healthcare professionals to analyze and visualize survival data, allowing them to identify factors that affect survival rates for a specific population of interest. Moreover, the findings of the KM plotter tool were also further validated using the GEPIA tool. In addition to this, we also performed a multivariate cox regression analysis to find a correlation between EAF2 expression and worse OS in patients of different clinical variables with LIHC and LUSC using Gene Expression Omnibus (GEO) datasets.

The TNMplot [24] is used herein to analyze EAF2 expression in metastatic tissues of different cancers relative to primary cancer tissues and corresponding control samples. TNMplot is a web-based tool that offers comprehensive visualizations of cancer patient data [24]. It uses a variety of statistical and analytical approaches to provide interactive, customizable plots of cancer patient data, including survival curves, heatmaps, and principal component analyses.

The cBioportal based analysis

To assess the EAF2-associated genetic mutations, copy number variations (CNVs), and mutational hotspots in LIHC and LUSC, we carried out the genomics data analysis through the cBioPortal platform [25] using the TCGA, Firehose legacy dataset of LIHC (having 442 tumor samples) and the TCGA, Firehose legacy dataset of LUSC (having 511 tumor samples). The cBioPortal in-house sequencing data over more than 30 different types of cancer [25]. The term "EAF2" was entered in the search bar of cBioPortal, and a summary of EAF2 genetic mutations, CNVs, and EAF2 protein architecture showing mutational hotspots was generated and analyzed using default settings.

PPI network and pathway analysis

STRING [26] was conducted to obtain the PPI network of the EAF2 protein with default settings, and Cytoscape was performed to show the PPI of EAF2 and related proteins [27]. STRING is a protein-protein interaction database that comprehensively maps the known and predicted interactions between proteins. Cytoscape, on the other hand, is an open-source software platform used to visualize networks, including protein-protein interaction networks derived from STRING, among other types of data. Together, these tools enable researchers to explore complex biological networks and identify novel molecular biomarkers.

Moreover, KEGG pathway analysis was carried out through the DAVID tool [28]. DAVID serves as a one-stop-shop for functional annotation and enrichment analysis in genomic research. DAVID catalogs thousands of genes and proteins across dozens of organisms and integrates a range of molecular data types like gene ontology, KEGG pathways, and biological function annotations [28].

EAF2 correlation with other genes

The GEPIA database [22] was used in our study to find the pairwise gene correlation analysis between EAF2 and its other enriched genes (obtained via STRING analysis) through the "Correlation Analysis" module.

MuTarget

MuTarget [29] was used in the current study to explore the different mutant genes linked with expression alteration in EAF2 across different cancers with default thresholds. This database mainly focuses on the cross-talk between the target proteins of the multi-functional drugs, and these targets are classified based on their molecular function and pathways [29].

TIMER

The TIMER database is an online application that is used to identify the correlations between tumor purity, infiltration of immune cells (B cells, CD8+ T cells, CD4+ T cells, dendritic cells, neutrophils, and macrophages), and gene expression of particular genes in cancer tissues using RNA-seq based expression data [30]. In the current study, we utilized this database to measure correlations among tumor purity, immune cell infiltration, and EAF2 mRNA expression in LIHC and LUSC.

Screening of EAF2 drugs

An online resource, the Comparative Toxicogenomics Database (CTD), was searched in our study to gather information on therapeutic drugs that can decrease or increase the mRNA expression level of EAF2 [31]. The CTD is a comprehensive database that allows researchers to compare biological responses to toxic substances across species and experiment types. This database can help identify genes and pathways that are involved in toxic response, as well as species differences that may impact toxicity. It can also help to predict toxicity in humans based on data from other species, which is important in drug development and chemical safety testing [31].

Statistical analysis

A t-test was used to evaluate differences in EAF2 expression in normal and cancer tissues via the UALCAN, GEPIA, TIMER, and GENT2 databases. The correlation analyses were per-

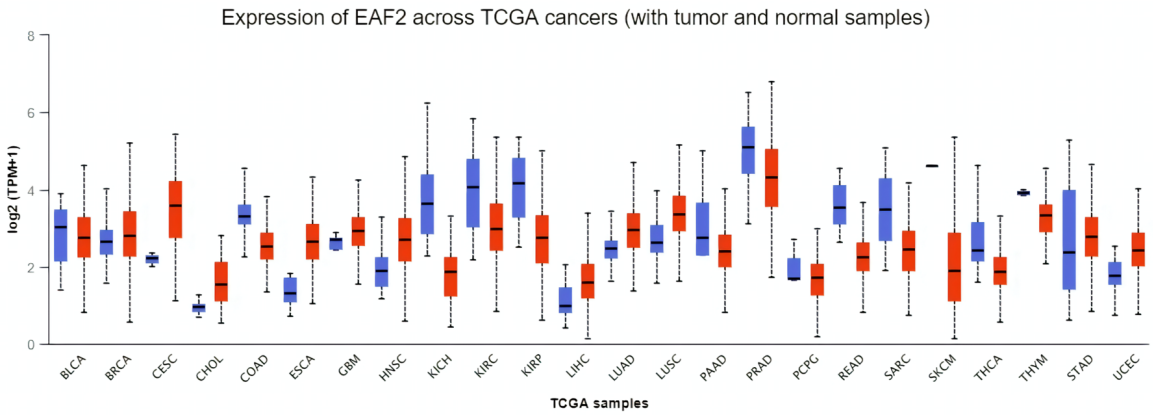


Figure 1. The mRNA based expression profiling of EAF2 across 24 types of cancerous tissues paired with corresponding controls using pan-cancer analysis via UALCAN.

formed using Spearman's correlation. What's more, $P < 0.05$ was considered statistically significant.

RNA-seq and targeted bisulfite-seq analysis based in vitro validation of EAF2 expression and methylation status

A total of 4 lung cancer cell lines, including A549, ABC-1, EBC-1, LK-2, and one normal control lung cell line (MRC-9) were purchased from the ATCC (American Type Culture Collection). The purchased cell lines were cultured in DMEM (HyClone), supplemented with 10% fetal bovine serum (FBS; TBD), 1% glutamine, and 1% penicillin-streptomycin in 5% CO_2 at 37°C. Total RNA extraction from all these three cells lines was done using TRIzol® reagent method [32], while total DNA was extracted via the organic method [33]. Finally, RNA and DNA samples were sent to the Beijing Genomics Institute (BGI) company for RNA-seq and bisulfite-seq analysis.

After RNA-seq analysis, the gene expression values of the EAF2 were normalized using fragments per kilo base million reads (FPKM). While, methylation values were normalized as beta values. The obtained FPKM, and beta values against hub genes in lung cancer and normal control cell line were compared to identify differences in the expression and methylation levels.

Results

EAF2 expression

To find out whether expression variations in EAF2 expression have any association with cancer or not, we evaluated EAF2 expression across 24 major tumor types paired with normal tissues through the UALCAN platform. Results demonstrated that EAF2 was significantly overexpressed in the majority of human cancers as compared to the normal controls (**Figure 1**). This abnormal expression pattern of EAF2 gives us clues that overexpressed EAF2 may be linked with the development of several cancers.

EAF2 prognostic and metastatic potentials

To estimate the prognostic potential of overexpressed EAF2 in different cancers (that showed significantly elevated expression levels of EAF2), the km plotter tool was applied. Results revealed that EAF2 with high expression was significantly ($P < 0.05$) linked to the poor OS and RFS durations of the LIHC and LUSC patients with *p* values of 0.002 and 0.0050 in OS and 0.013 and 0.019 in RFS analysis, while in other cancers (BLCA, BRCA, CESC, CHOL, ESCA, GBM, HNSC, LUAD, PRAD, PCPG, SKCM, THYM, UCEC), the overexpressed EAF2 was not found to be associated with the reduced OS and RFS (**Figure 2A**, **2B**). In view of the prognostic analysis results, it is speculated that overexpressed EAF2 is more closely related to the pathogenesis of LIHC and LUSC. Secondly, to further clarify the role of overexpressed EAF2 in the metastasis of LIHC and LUSC, we utilized the TNMplot database. Results of the analysis showed that elevated expression of EAF2 was also associated with metastasis because the metastatic tissues of LIHC and LUSC presented a higher level of EAF2 than the primary cancer tissues and nor-

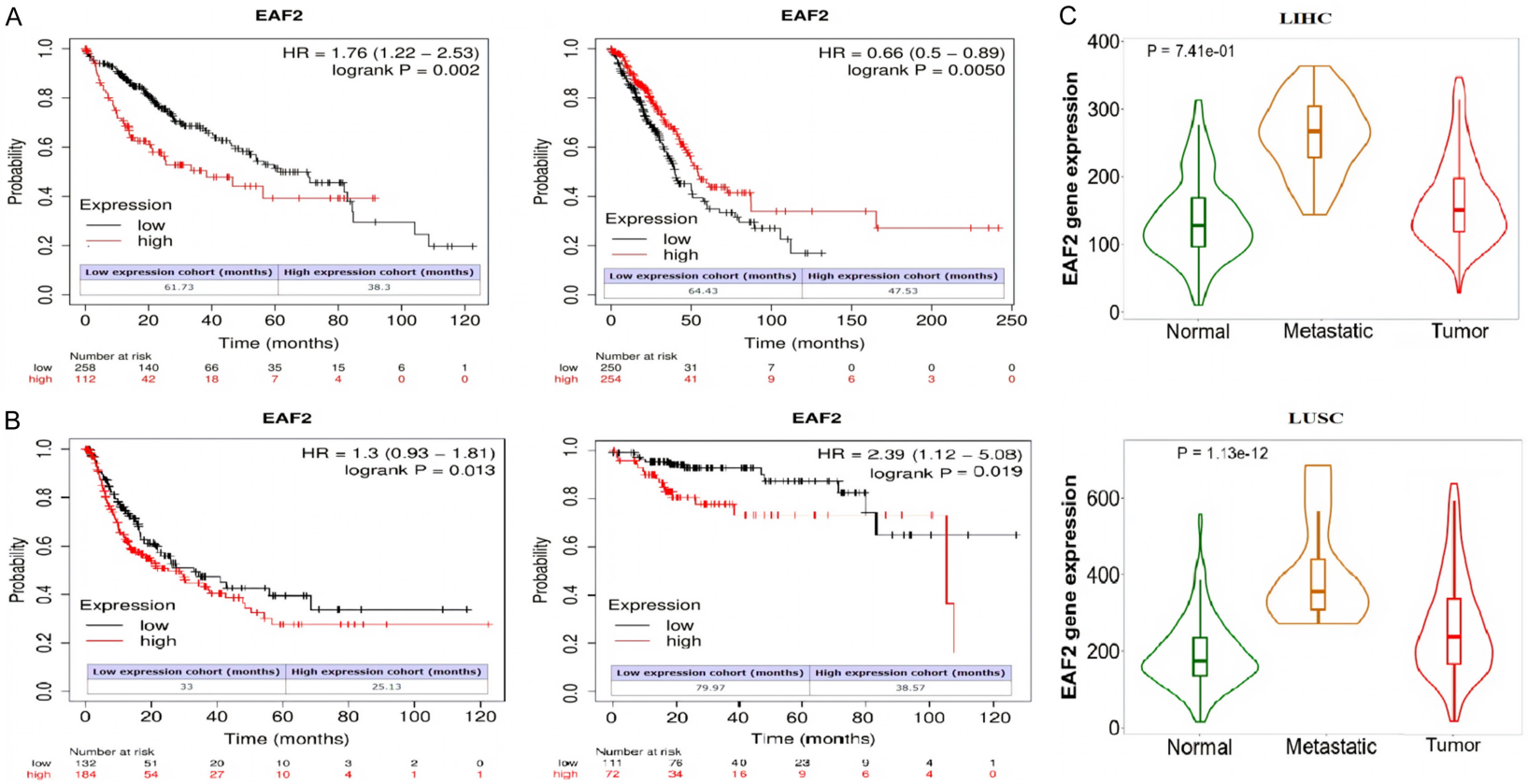


Figure 2. Prognostic values and expression of EAF2. (A) OS values across LIHC and LUSC, (B) RFS values across LIHC and LUSC, and (C) Expression status of EAF2 across primary and metastasis tissues.

mal controls (**Figure 2C**). Taken together, these data suggested that EAF2 might play important roles in the development, and metastasis of LIHC and LUSC. However, further extensive research is required to investigate the link between the overexpression of EAF2 and cancer patient prognosis in other types of cancer.

GEPIA and multivariate cox regression analysis-based EAF2 prognostic potential verification

Cross-validation of the results is an important aspect to consider while performing integrative research. In our study, to validate the EAF2 prognostic potential in new cohorts of LIHC and LUSC, we took advantage of the GEPIA tool. In view of our prognostic potential validation results, initially, it was also observed that higher expression of EAF2 was significantly ($P < 0.05$) linked to the poor OS and RFS of the LIHC patients via GEPIA analysis (Supplementary Figure 1A). Then, a multi cox regression analysis found that higher EAF2 expression was associated with worse overall survival in patients with LIHC and LUSC patients of different stages and age groups (Supplementary Figure 1B). This suggests that EAF2 could serve as a prognostic biomarker LIHC and LUSC patients of different clinical variables. Collectively, the overall data of survival analysis suggested that EAF2's higher expression is a notable factor that affects the survival of LIHC and LUSC cancers significantly.

EAF2 expression level across different clinicopathological variables

Subsequently, following survival analysis, we investigate the relationship between EAF2 expression and different clinicopathological variables, such as cancer stage, patient's race, patient's gender, and age, in patients with LIHC and LUSC. The expression profiling of EAF2 in both LIHC and LUSC samples based on the clinicopathological variables showed the notable overexpression of EAF2 in LIHC and LUSC samples of different clinical variables relative to normal controls, i.e., LIHC and LUSC samples stratified by different cancer stages (stage 1, stage 2, stage 3, and stage 4, p value ≤ 0.05), races (Caucasian, American-African, and Asian, p value ≤ 0.05), genders (male and female, p value ≤ 0.05), and ages (21-40 yrs, 41-60 yrs, 61-80 yrs, and 81-100 yrs, p value ≤ 0.05) (**Figure 3**). Taken together, our results indicated that EAF2 was overexpressed in LIHC and LUSC patients regardless of cancer stage, patient race, patient gender, and age-based clinicopathological variables.

Validation of EAF2 higher expression using additional TCGA datasets

In order to verify the higher level of EAF2 expression, we further carried out the EAF2 expression re-analysis across new independent cohorts of LIHC and LUSC via the TIMER, GENT2, and GEPIA platforms. The results of the re-analysis also revealed a significant ($P < 0.05$) higher expression of EAF2 in LIHC and LUSC patients from new independent cohorts relative to healthy donors (Supplementary Figure 2). Taken together, the results of EAF2 expression analysis and validation expression analysis in LIHC and LUSC, our findings provided reliable validation of EAF2's role in the pathogenesis of LIHC and LUSC.

EAF2 promoter methylation analysis

We computed the level of EAF2 promoter methylation in LIHC and LUSC samples via the UALCAN platform, and later we also employed the MEXPRESS tool to cross-validate the findings of UALCAN. In view of our results via UALCAN, the promoter methylation level of EAF2 was significantly ($P < 0.05$) lower in LIHC and LUSC samples than in normal tissues (Supplementary Figure 3A). Moreover, the results of cross-validation via MEXPRESS also presented similar results to UALCAN (Supplementary Figure 3B). Taken together results of promoter methylation analysis, the obtained values suggested a negative correlation between EAF2 expression and its promoter methylation level in LIHC and LUSC samples as compared to the normal controls. Therefore, it is speculated that the lower promoter methylation level is involved in up-regulating EAF2 expression in LIHC and LUSC.

Gene mutations, CNVs analysis, and mutational hotspots identification

Through the cBioPortal platform, EAF2 genetic alterations, CNVs analysis, and mutational hotspots identification in LIHC and LUSC were carried out using the TCGA LIHC and LUSC datasets via cBioPortal.

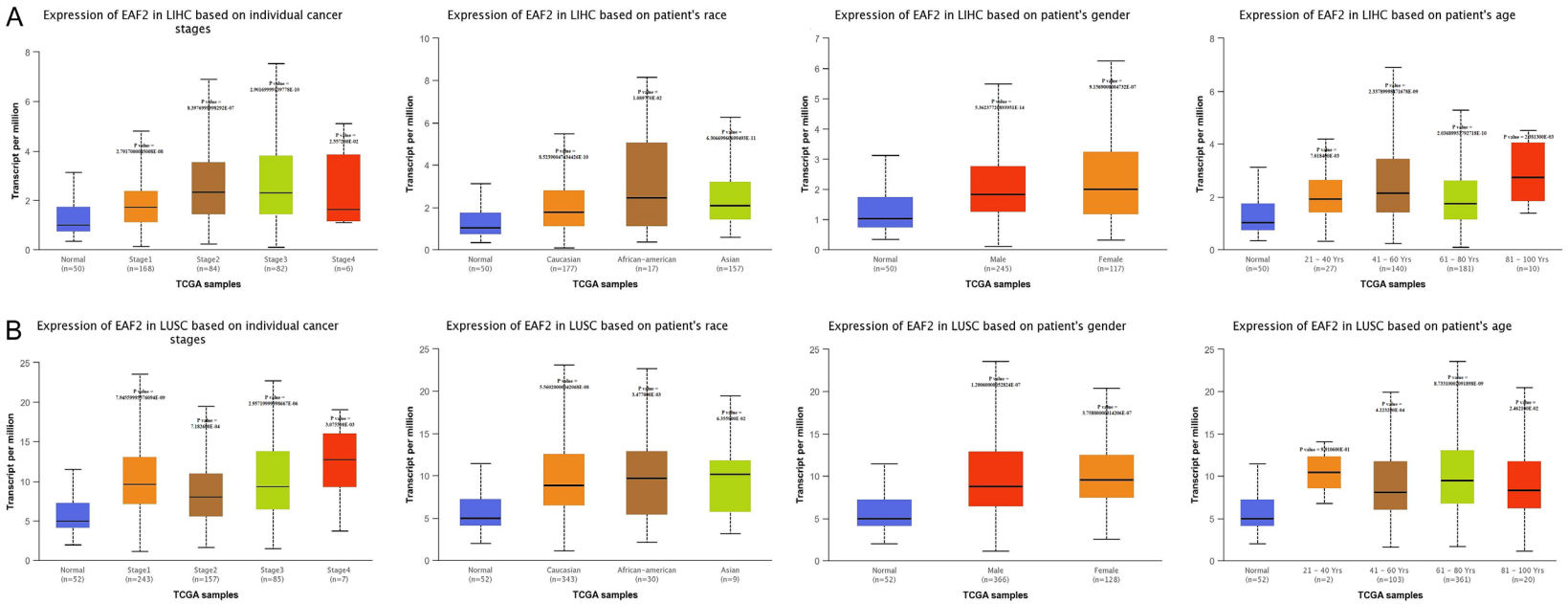


Figure 3. EAF2 and different clinicopathological variables. (A) EAF2 expression in LIHC patients stratified by cancer stages, races, genders, and ages, and (B) EAF2 expression in LUSC patients stratified by cancer stages, races, genders, and ages.

Table 1. Detail of Kyoto encyclopedia of genes and genomes pathway

Pathway ID	Pathway Name	Gene count	*P*-value	Gene name
hsa05202	Transcriptional misregulation in cancer	3	< 0.05	CDK9, MLLT1, MLLT3
hsa03020	RNA polymerase	2	< 0.05	POLR2A, POLR2I
hsa00240	Pyrimidine metabolism	2	< 0.05	POLR2A, POLR2I
hsa00230	Purine metabolism	2	< 0.05	POLR2A, POLR2I

In view of the results of this analysis, EAF2 was found to harbor genetic alterations in only 0.3% cases of the queued LIHC samples, and all the observed mutations were truncated mutations in these cases (Supplementary Figure 4A). Similar to LIHC, EAF2 also showed the incidence of genetic alterations in only 12% cases of queued LUSC samples, and in these samples, the deep amplification genetic abnormality was observed as the most frequently reported genetic abnormality (Supplementary Figure 4A).

Next, we also explored that in LIHC, the mutational hotspots of the most frequently observed EAF2 truncated mutation (T112X) lie inside the EAF domain of the encoded protein, while in LUSC, the EAF2 mutational hotspots of the most frequently observed missense mutation (E200K) lie outside the EAF domain (Supplementary Figure 4B). Taken together, we speculated that as genetic alterations were observed in very small proportions (0.3%, and 12%) of the analyzed LIHC and LUSC samples, respectively. Therefore, involvement of these factors in the dysregulation of EAF2 is unlikely. Moreover, different observed EAF2 mutational hotspots in LIHC and LUSC overall suggested a somehow high level of complexity regarding EAF2 mutations in LIHC and LUSC.

A PPI network and pathway analysis

Next, a PPI network of EAF2 was constructed using the STRING database and visualized through Cytoscape software to recognize its associated genes. In total, 10 genes were noticed in the PPI network that were found to be associated with EAF2 (Supplementary Figure 5). We further subjected EAF2-associated genes to pathway enrichment using DAVID. In view of the results of pathway enrichment analysis, EAF2-associated genes were found to be significantly involved in four diverse pathways, including "Transcriptional misregulation in cancer", "RNA polymerase", "Pyrimidine metabolism", and "Pyrimidine metabolism" (Supplementary Figure 5; **Table 1**).

Correlation analysis between EAF2 and its different other associated genes

Results of this study further highlighted that the expression of EAF2 was notably positively correlated with its other related genes expression including ELL, CDK9, MLLT1, TCEB3C, AFF4, POLR2I, MLLT3, TCEA1, POLR2A, and EAF1 in LIHC and LUSC (**Figure 4**). In view of these results, we ultimately speculated that along with EAF2, the aberrant expression of its other associated genes may also exert a tumor-promoting role in LIHC and LUSC.

MuTarget analysis

Via MuTarget, we explored different mutant genes responsible for EAF2 overexpression in LIHC and LUSC. We selected the top 5 mutant genes as shown in Supplementary Figure 6 for EAF2 in LIHC and LUSC, respectively. These genes include TP53, COBLL1, CSMD3, SLC6A11, and TRIM66 in LIHC, and RREB1, AADACL2, CD244, CD1B, and ZFHX3 in LUSC. Taken together, these findings suggest that EAF2 expression has a strong correlation with different other mutant genes acting as possible regulators of EAF2 expression in LIHC and LUSC.

TIMER analysis

Tumor purity and immune cells are the essential regulators of the anticancer immune response and thus act as the backbone of the present cancer immunotherapies [34]. In this study, the Spearman correlations between tumor purity, immune cell infiltration including CD4+ T cells, CD8+ T cells, B cells, macrophages, neutrophils, and EAF2 expression in LIHC and LUSC were carried out via TIMER. As per the tumor purity analysis, we observed negative correlations between EAF2 expression and

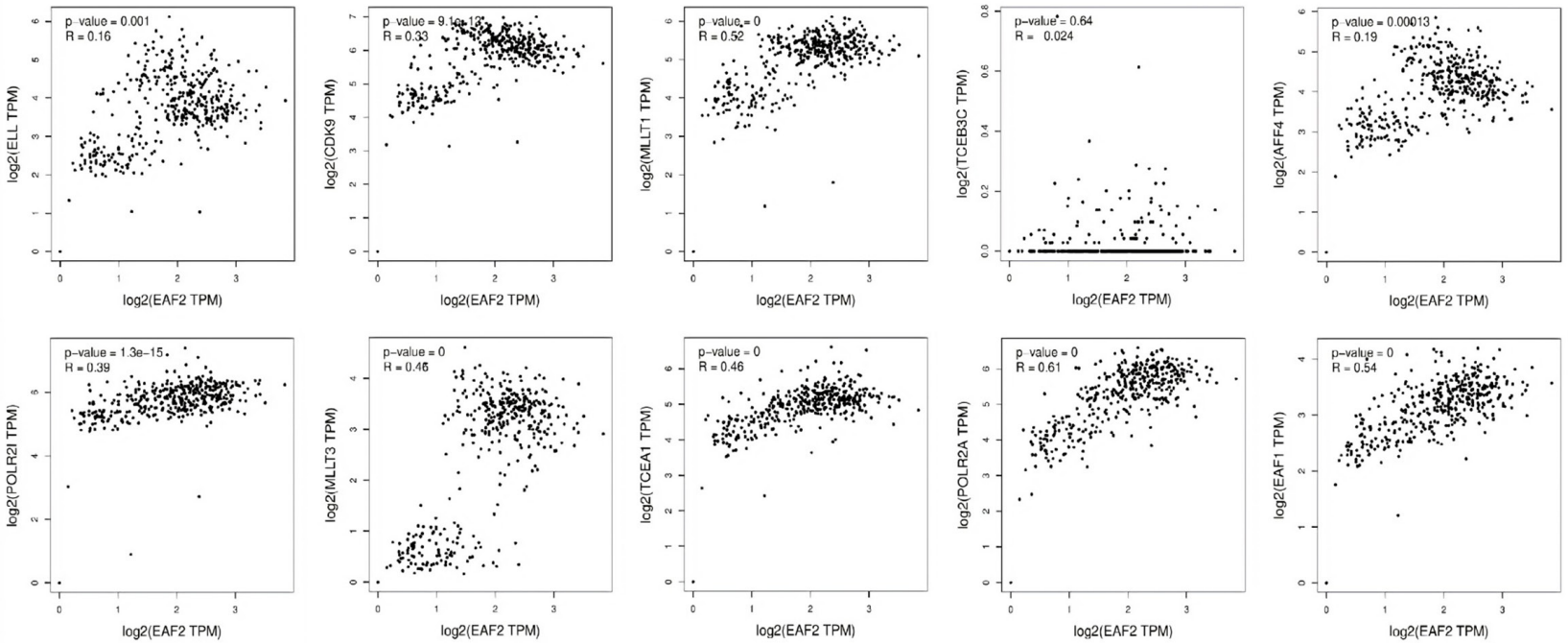


Figure 4. Correlations among EAF2 and its associated other genes' expression via GEPIA database.

Table 2. EAF2 expression regulatory drugs extracted from the CTD database

Sr. no	Name of the drug	Effect on the expression	Target gene
1	Valproic acid	Increase expression	EAF2
2	Vorinostat	Increase expression	
3	Cyclosporine	Increase expression	
4	Methionine	Decrease expression	
5	Tetrachlorodibenzodioxine	Decrease expression	
6	Vanadates	Decrease expression	
7	Tretinoin	Decrease expression	
8	Choline	Decrease expression	
9	Butylparaben	Decrease expression	
10	Bisphenol A	Decrease expression	

tumor purity in LIHC (Rho = -0.066, *p*-value = 2.23e-01) and LUSC (Rho = -0.124, *p*-value = 6.77e-03) (Supplementary Figure 7). Moreover, we have observed positive correlations between EAF2 expression and immune cell infiltration of CD4+ T cells (Rho = 0.035, *p*-value = 5.18e-01), CD8+ T cells (Rho = 0.269, *p*-value = 3.8e-07), B cells (Rho = 0.259, *p*-value = 1.06e-06), macrophages (Rho = 0.288, *p*-value = 5.19e-08), and neutrophils (Rho = 0.274, *p*-value = 2.22e-07) in LIHC while negative correlations between EAF2 expression and immune cells infiltration of CD4+ T cells (Rho = -0.082, *p*-value = 7.25e-02), CD8+ T cells (Rho = -0.087, *p*-value = 5.69e-02), B cells (Rho = -0.3, *p*-value = 2.18e-11), macrophages (Rho = -0.039, *p*-value = 3.99e-02), and neutrophils (Rho = -0.117, *p*-value = 1.07e-02) in LUSC (Supplementary Figure 7).

Screening of EAF2-associated therapeutic drugs

A gene-drug interaction of EAF2 was built using Cytoscape and explored in CTD databases to identify potential therapeutic agents for treating EAF2 in LIHC and LUSC. In total, 10 of the most verified unique chemicals by different previously reported studies were obtained, including Valporic acid, Vorinostat, Methionine, Tetrachlorodibenzodioxin, Vanadates, Tretinioin, Cyclosporine, Choline, Butylparaben, and bisphenol A (**Table 2**). Out of these 10 noted chemicals, 3 (Valporic acid, Vorinostat, and Cyclosporine) are reported to enhance the EAF2 expression (**Table 2**), while 7 (Methionine, Tetrachlorodibenzodioxin, Vanadates, Tretinioin, Choline, Butylparaben, and bisphenol A) are capable of lowering EAF2 expression and thus can be exploited as treatment options against overexpressed EAF2 in LIHC and LUSC.

Experimental in vitro validation of the EAF2 expression and methylation status

In the current study, by performing RNA-seq and targeted bisulfite-seq analyses of 4 lung cancer cell lines, including A549, ABC-1, EBC-1, LK-2, and one normal control lung cell line (MRC-9), the expression and methylation levels of the EAF2 gene were validated. The expression level of this gene was validated using FPKM, while the methylation level was validated using beta values. Both FPKM and beta are quantitative values with widespread use in RNA-seq and bisulfite-seq analyses. As shown in **Figure 5A**, it was noticed that the EAF2 gene was expressed in both normal and lung cancer cell lines, and FPKM values of EAF2 were notably higher in lung cancer cell lines (A549, ABC-1, EBC-1, LK-2) as compared to normal cell line (MRC-9) (**Figure 5A**). Moreover, the beta values of EAF2 were higher in the normal (MRC-9) cell line while lower in the lung cancer cell line (A549, ABC-1, EBC-1, LK-2) (**Figure 5B**).

Discussion

Cancer is the outcome of genomic and epigenomic alterations in normal cells [35]. Abnormal gene expression is known as the hallmark of cancer [36]. In line with the 2020 cancer states, the death rate due to this disease is still very high around the globe [3]. Hence, there is an urgent need to explore cancer biology to identify some sensitive diagnostic and prognostic biomarkers for the better management of the disease.

EAF2 blocks the transcriptional function of hypoxia-induced factor 1α by disrupting its association with the co-activator CBP/p300 [37]. In *Liu et al.* study, EAF2 was found to suppress both TGF-β-induced G1 cell cycle arrest and TGF-β-induced cell migration by directly interacting with Smad3 [38]. Additionally, EAF2 controls the DNA repair process in prostate

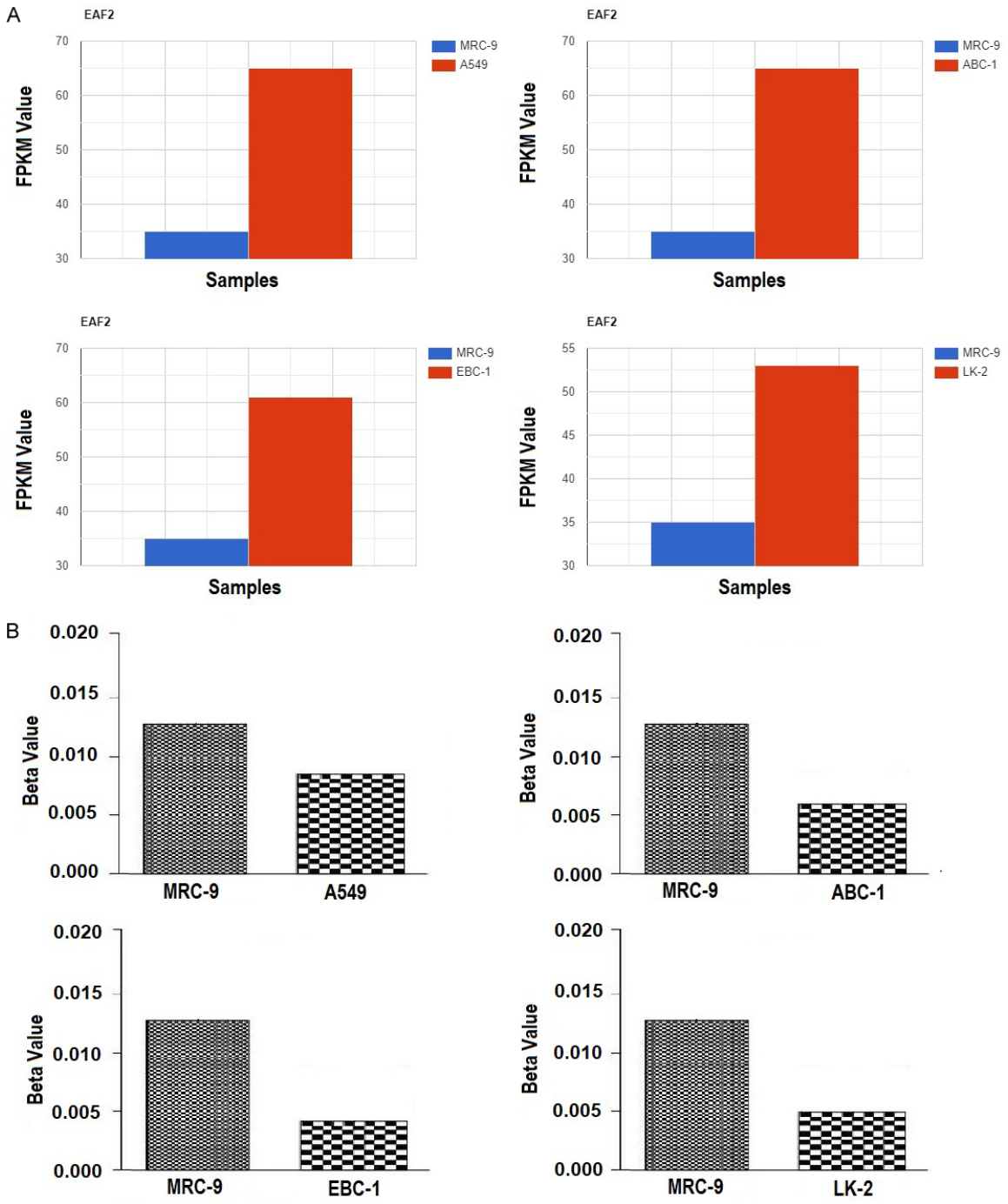


Figure 5. Validating EAF2 expression and methylation status using (MRC-9) and (A549, ABC-1, EBC-1, LK-2) cell lines via RNA-seq and targeted bisulfite-seq analyses. (A) FPKM values based expression plots of EAF2, and (B) Beta values based methylation plots of EAF2.

cancer via the Ku70/Ku80 complex to influence the radiation sensitization of androgen deprivation therapy [15]. In glioblastoma, the EAF2-HIF1α axis is associated with tumorigenesis and activation of glycolysis via EZH2 regulation [12]. To the best of our knowledge, until now, the oncogenic role of EAF2 has not been reported in other cancer types.

The findings of the present study showed that mRNA expression of EAF2 was up-regulated in the majority of targeted human cancers but

only associated with the decreased OS, RFS durations, and advanced metastasis in LIHC and LUSC. Earlier, it is known that aberrant expression of BECN1, LAMP2, and PINK1 genes in colorectal cancer is potentially regulated by CpG islands in the promoter region [39]. Therefore, to further identify the possible causes of EAF2 overexpression, we also performed a correlation analysis of EAF2 overexpression with its promoter methylation level, genetic mutations, and CNVs in both LIHC and LUSC patients. Results revealed significant negative correlations between EAF2 expression and its promoter methylation levels in LIHC and LUSC, therefore, this scenario of EAF2 promoter methylation highlighted the significant role of promoter hypomethylation in the up-regulation of EAF2 in LIHC and LUSC. Moreover, the EAF2 gene was found to be enriched in truncated mutations and deep amplification abnormality in small proportions of the LUSC and LIHC patients, respectively. Hence, we speculated that genetic mutations and CNVs participate insignificantly in expression regulation of EAF2 in these cancers. Furthermore, it was also observed that mutations in the EAF2 gene could change amino acids at different sites of the encoded protein.

Earlier, different expression-based biomarkers of LIHC have been reported in the medical literature, including AFP, GPC3, FCN3, PRC1, CLEC1B, AFP, GPC3, and CK19 [40, 41]. However, to the best of our knowledge, none of these or any other biomarker has been shortlisted for LIHC patients exhibiting different clinicopathological features. In the present study, we have revealed the significant up-regulation of EAF2 expression in LIHC patients exhibiting different clinicopathological variables, including different cancer stage, race, gender, and age. We have also shown that EAF2 overexpression is significantly associated with decreased OS, RFS durations, and metastasis in LIHC patients.

Several expression-based indicators are now being investigated to distinguish LUSC patients from healthy individuals. For example, according to research by *Li et al.* [42], BIRC5 was significantly overexpressed in LUSC tissues compared to normal tissues, making it a unique target for anti-LUSC therapy [42]. Similarly, GAPDH has also been revealed to play a key role in the regulation of glycolysis within LUSC cells. GAPDH protein can act as a new target for anti-LUSC therapy since it reduces ATP in cancer cells by inhibiting glycolysis, which kills cancer cells [43, 44]. However, to the best of our knowledge, none of the previously described LUSC biomarkers have been generalized in LUSC patients with various clinicopathological characteristics. However, in the present study, we revealed the significant up-regulation of EAF2 expression in LUSC patients with various clinicopathological variables including different cancer stage, race, gender, and age as compared to the normal controls. We have also shown that EAF2 overexpression is significantly associated with decreased OS, RFS durations, and metastasis in LUSC patients.

Although single gene-based indicators have shown success in predicting the diagnostic and prognostic outcomes of cancer patients. However, multi-gene-based diagnostic and prognostic systems have also gained massive attraction in recent years [45]. Moreover, still, there is no appropriate therapy is known for LIHC and LUSC patients harboring different kinds of mutations in different oncogenic genes [46]. That is why, in this study, we used muTarget to explore different mutant genes that may alter EAF2 expression. In view of the analysis results, we have identified 5 top mutant genes in each LIHC and LUSC, respectively, for EAF2, including TP53, COBLL1, CSMD3, SLC6A11, and TRIM66 in LIHC, and RREB1, AADACL2, CD244, CD1B, and ZFHX3 in LUSC. By linking the explored mutant genes to EAF2 expression, it will be more effective to identify potential multi-gene-based therapies for LIHC and LUSC.

Cancer treatment is a very tough task in the medical field [47]. However, one of the most successful cancer treatment methods is cancer immunotherapy, which has shown some promising results in the field of anticancer research [48]. The use of different drugs, including cytotoxic T lymphocyte antigen 4 (CTLA4), programmed cell death 1 (PD-1), and programmed cell death 1 ligand 1 (PD-L1) inhibitors, is considered as the core of immunotherapy [49]. However, due to drug resistance and tumor immune escape, immunotherapy is facing difficulties in achieving the expected higher outcomes [50]. Therefore, we speculate that understanding the tumor im-

mune microenvironment (TIME) of LIHC and LUSC may improve our understanding of immunotherapy and help clinicians to obtain better clinical results.

Stromal cells, immune cells, and tumor purity are recognized as the essential elements of the tumor microenvironment and have been previously documented to alter the response to immunotherapy for enhancing tumor growth [51, 52]. In this study, EAF2 expression has been revealed to be negatively correlated with tumor purity in both LIHC and LUSC samples, suggesting that EAF2 is expressed highly in stromal cells than in epithelial cells and may participate in LIHC and LUSC development by operating stromal cells. However, further research is needed to clarify the stromal cell-dependent tumor operating mechanism of EAF2. Moreover, we also observed significant positive correlations between CD4+ T cells, CD8+ T cells, B cells, macrophages, neutrophils, and EAF2 expression in LIHC, and significant negative correlations between the same variables in LUSC. Taken together these results, we also speculated that EAF2 may also exert its tumorigenic effect by activating immune cells in LIHC while suppressing immune cells in LUSC. In a nutshell, the found connections may aid clinicians in developing more accurate immunotherapies by providing them with a greater understanding of the LIHC and LUSC tumor microenvironment landscape.

In the current study, the interaction network of EAF2-associated genes was constructed and visualized. In total, 10 EAF2 associate genes were identified. Pathway enrichment analysis of EAF2-associated genes revealed their involvement in four diverse signaling pathways, including “Transcriptional misregulation in cancer”, “RNA polymerase”, and “Pyrimidine metabolism”. These findings are consistent with the results of earlier studies, where the role of identified pathways is well established in the development and prognosis of cancer [53]. Moreover, correlation analyses among EAF2 and its associated genes expression have revealed strong positive correlations, which further validated our findings regarding KEGG analysis. Finally, we have also explored a few potential therapeutic drugs, including Methionine, Tetrachlorodibenzodioxin, Vanadates, Tretinioin, Choline, Butylparaben, and bisphenol A that can be used against overexpressed EAF2 in the treatment of LIHC and LUSC. However, the identification of these drugs is based on the previous reported limited number of studies. Therefore, further extensive work is needed before clinical application.

Despite the many merits of this study such as utilization of the large cancer cohorts, it is inevitable that there are limitations. Our findings were largely derived from bioinformatics analyses. Therefore, it would be valuable to validate the efficacy of EAF2 by examining additional in-house clinical samples of LIHC and LUSC. Furthermore, gaining a more comprehensive understanding of the potential role of EAF2 as a diagnostic and prognostic tool in these cancers could lead to the development of innovative diagnostic tools for early cancer detection.

Conclusion

In this comprehensive study, we utilized various online bioinformatics platforms, web tools, and cell line-based experiments to systematically examine the impact of EAF2 on cancer development and assess its potential for diagnosis and prognosis. Our multi-omics analysis demonstrated that EAF2 was up-regulated and linked to reduced OS in LIHC and LUSC, revealing the significance of EAF2-related pathways in the progression of these cancers. The results of this study suggest that EAF2 could be a promising therapeutic target for LIHC and LUSC. However, further molecular studies involving large cohorts of LIHC and LUSC are necessary to confirm the role of EAF2 in these cancers.

Disclosure of conflict of interest

None.

Address correspondence to: Chao Peng, Gansu University of Chinese Medicine, Lanzhou, Gansu, China. E-mail: yxgszyy@163.com

References

[1] Sarkar S, Horn G, Moulton K, Oza A, Byler S, Kokolus S and Longacre M. Cancer development, progression, and therapy: an epigenetic overview. Int J Mol Sci 2013; 14: 21087-21113.

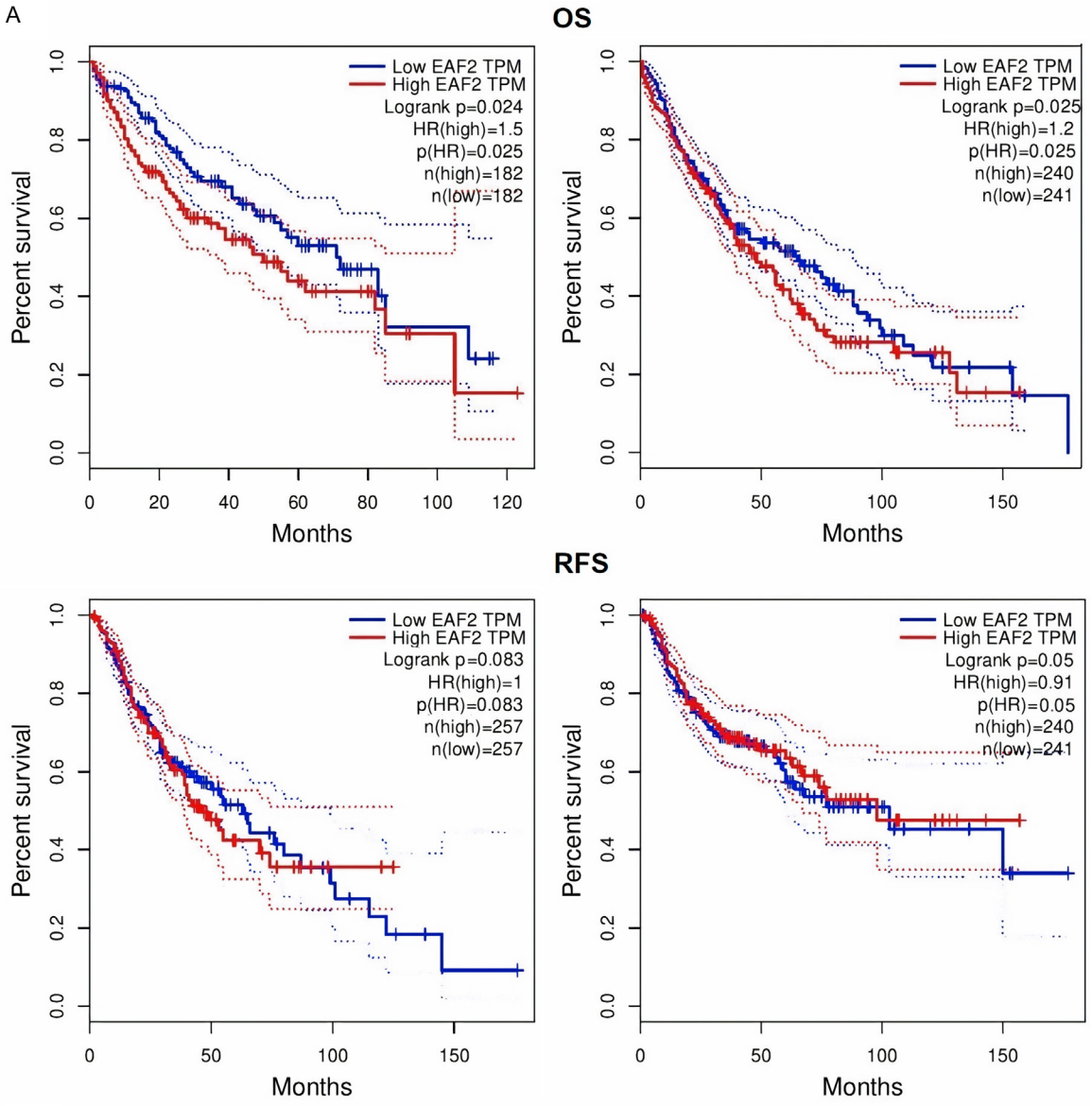
[2] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global can-

cer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021; 71: 209-249.
[3] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2020. CA Cancer J Clin 2020; 70: 7-30.
[4] Sarwar MR and Saqib A. Cancer prevalence, incidence and mortality rates in Pakistan in 2012. Cogent Medicine 2017; 4: 1288773.
[5] Akram M, Iqbal M, Daniyal M and Khan AU. Awareness and current knowledge of breast cancer. Biol Res 2017; 50: 33.
[6] Ferber MJ, Montoya DP, Yu C, Aderca I, McGee A, Thorland EC, Nagorney DM, Gostout BS, Burgart LJ, Boix L, Bruix J, McMahon BJ, Cheung TH, Chung TK, Wong YF, Smith DI and Roberts LR. Integrations of the hepatitis B virus (HBV) and human papillomavirus (HPV) into the human telomerase reverse transcriptase (hTERT) gene in liver and cervical cancers. Oncogene 2003; 22: 3813-3820.
[7] Hameed Y, Usman M and Ahmad M. Does mouse mammary tumor-like virus cause human breast cancer? Applying Bradford Hill criteria postulates. Bull Natl Res Cent 2020; 44: 183.
[8] Hashibe M, Brennan P, Chuang SC, Boccia S, Castellsague X, Chen C, Curado MP, Dal Maso L, Daudt AW, Fabianova E, Fernandez L, Wünsch-Filho V, Franceschi S, Hayes RB, Herrero R, Kelsey K, Koifman S, La Vecchia C, Lazarus P, Levi F, Lence JJ, Mates D, Matos E, Menezes A, McClean MD, Muscat J, Eluf-Neto J, Olshan AF, Purdue M, Rudnai P, Schwartz SM, Smith E, Sturgis EM, Szeszenia-Dabrowska N, Talamini R, Wei Q, Winn DM, Shangina O, Pilarska A, Zhang ZF, Ferro G, Berthiller J and Boffetta P. Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. Cancer Epidemiol Biomarkers Prev 2009; 18: 541-550.
[9] Ascierto PA, Minor D, Ribas A, Lebbe C, O'Hagan A, Arya N, Guckert M, Schadendorf D, Kefford RF, Grob JJ, Hamid O, Amaravadi R, Simeone E, Wilhelm T, Kim KB, Long GV, Martin AM, Mazumdar J, Goodman VL and Trefzer U. Phase II trial (BREAK-2) of the BRAF inhibitor dabrafenib (GSK2118436) in patients with metastatic melanoma. J Clin Oncol 2013; 31: 3205-3211.
[10] Guo J, Qin S, Liang J, Lin T, Si L, Chen X, Chi Z, Cui C, Du N, Fan Y, Gu K, Li F, Li J, Li Y, Liang H, Liu J, Lu M, Lu A, Nan K, Niu X, Pan H, Ren G, Ren X, Shu Y, Song X, Tao M, Wang B, Wei W, Wu D, Wu L, Wu A, Xu X, Zhang J, Zhang X, Zhang Y and Zhu H; written; Chinese Society of Clinical Oncology (CSCO) Melanoma Panel. Chinese guidelines on the diagnosis and treatment of melanoma (2015 edition). Ann Transl Med 2015; 3: 322.
[11] Li M, Wu X, Zhuang F, Jiang S, Jiang M and Liu YH. Expression of murine ELL-associated factor 2 (Eaf2) is developmentally regulated. Dev Dyn 2003; 228: 273-280.
[12] Ai J, Pascal LE, Wei L, Zang Y, Zhou Y, Yu X, Gong Y, Nakajima S, Nelson JB, Levine AS, Lan L and Wang Z. EAF2 regulates DNA repair through Ku70/Ku80 in the prostate. Oncogene 2017; 36: 2054-2065.
[13] Guo W, Keener AL, Jing Y, Cai L, Ai J, Zhang J, Fisher AL, Fu G and Wang Z. FOXA1 modulates EAF2 regulation of AR transcriptional activity, cell proliferation, and migration in prostate cancer cells. Prostate 2015; 75: 976-987.
[14] Jiang F, Ai J, Xiao W and Wang Z. FB1, an E2A fusion partner in childhood leukemia, interacts with U19/EAF2 and inhibits its transcriptional activity. Cancer Lett 2007; 253: 265-272.
[15] Yang T, Jing Y, Dong J, Yu X, Zhong M, Pascal LE, Wang D, Zhang Z, Qiao B and Wang Z. Regulation of ELL2 stability and polyubiquitination by EAF2 in prostate cancer cells. Prostate 2018; 78: 1201-1212.
[16] Jo YS, Kim SS, Kim MS, Yoo NJ and Lee SH. Candidate tumor suppressor gene EAF2 is mutated in colorectal and gastric cancers. Pathol Oncol Res 2019; 25: 823-824.
[17] Feng ML, Wu C, Zhang HJ, Zhou H, Jiao TW, Liu MY and Sun MJ. Overexpression of ELL-associated factor 2 suppresses invasion, migration, and angiogenesis in colorectal cancer. World J Gastrointest Oncol 2022; 14: 1949-1967.
[18] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK and Varambally S. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 2017; 19: 649-658.
[19] Koch A, De Meyer T, Jeschke J and Van Criekinge W. MEXPRESS: visualizing expression, DNA methylation and clinical TCGA data. BMC Genomics 2015; 16: 636.
[20] Park SJ, Yoon BH, Kim SK and Kim SY. GENT2: an updated gene expression database for normal and tumor tissues. BMC Med Genomics 2019; 12 Suppl 5: 101.
[21] Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B and Liu XS. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. Cancer Res 2017; 77: e108-e110.
[22] Tang Z, Li C, Kang B, Gao G, Li C and Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017; 45: W98-W102.

[23] Maciejczyk A, Szelachowska J, Czapiga B, Matkowski R, Hałoń A, Györffy B and Surowiak P. Elevated BUBR1 expression is associated with poor survival in early breast cancer patients: 15-year follow-up analysis. J Histochem Cytochem 2013; 61: 330-339.
[24] Bartha Á and Győrffy B. TNMplot.com: a web tool for the comparison of gene expression in normal, tumor and metastatic tissues. Int J Mol Sci 2021; 22: 2622.
[25] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C and Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012; 2: 401-404.
[26] von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P and Snel B. STRING: a database of predicted functional associations between proteins. Nucleic Acids Res 2003; 31: 258-261.
[27] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498-2504.
[28] Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, Stephens R, Baseler MW, Lane HC and Lempicki RA. The DAVID gene functional classification tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol 2007; 8: R183.
[29] Nagy Á and Győrffy B. muTarget: a platform linking gene expression changes and mutation status in solid tumors. Int J Cancer 2021; 148: 502-511.
[30] Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, Li B and Liu XS. TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res 2020; 48: W509-W514.
[31] Mattingly CJ, Colby GT, Forrest JN and Boyer JL. The comparative toxicogenomics database (CTD). Environ Health Perspect 2003; 111: 793-795.
[32] Rio DC, Ares M Jr, Hannon GJ and Nilsen TW. Purification of RNA using TRIzol (TRI reagent). Cold Spring Harb Protoc 2010; 2010: pdb.prot5439.
[33] Ghatak S, Muthukumaran RB and Nachimuthu SK. A simple method of genomic DNA extraction from human samples for PCR-RFLP analysis. J Biomol Tech 2013; 24: 224-231.
[34] Ziai J, Gilbert HN, Foreman O, Eastham-Anderson J, Chu F, Huseni M and Kim JM. CD8+ T cell infiltration in breast and colon cancer: a histologic and statistical analysis. PLoS One 2018; 13: e0190158.
[35] Carter SL, Eklund AC, Kohane IS, Harris LN and Szallasi Z. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. Nat Genet 2006; 38: 1043-1048.
[36] Gutschner T and Diederichs S. The hallmarks of cancer: a long non-coding RNA point of view. RNA Biol 2012; 9: 703-719.
[37] Chen Z, Liu X, Mei Z, Wang Z and Xiao W. EAF2 suppresses hypoxia-induced factor 1α transcriptional activity by disrupting its interaction with coactivator CBP/p300. Mol Cell Biol 2014; 34: 1085-1099.
[38] Zhang Y, Alexander PB and Wang XF. TGF-β family signaling in the control of cell proliferation and survival. Cold Spring Harb Perspect Biol 2017; 9: a022145.
[39] Bednarczyk M, Fatyga E, Dzięgielewska-Gęsiak S, Waniczek D, Grabarek B, Zmarzły N, Janikowska G and Muc-Wierzgoń M. The expression patterns of BECN1, LAMP2, and PINK1 genes in colorectal cancer are potentially regulated by microrna and CpG islands: an in silico study. J Clin Med 2020; 9: 4020.
[40] Kaur H, Dhall A, Kumar R and Raghava GPS. Identification of platform-independent diagnostic biomarker panel for hepatocellular carcinoma using large-scale transcriptomics data. Front Genet 2020; 10: 1306.
[41] Ocker M. Biomarkers for hepatocellular carcinoma: what's new on the horizon? World J Gastroenterol 2018; 24: 3974-3979.
[42] Li S, Sun X, Miao S, Liu J and Jiao W. Differential protein-coding gene and long noncoding RNA expression in smoking-related lung squamous cell carcinoma. Thorac Cancer 2017; 8: 672-681.
[43] Ganapathy-Kanniappan S, Geschwind JF, Kunjithapatham R, Buijs M, Vossen JA, Tchernyshyov I, Cole RN, Syed LH, Rao PP, Ota S and Vali M. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is pyruvylated during 3-bromopyruvate mediated cancer cell death. Anticancer Res 2009; 29: 4909-4918.
[44] Xu RH, Pelicano H, Zhou Y, Carew JS, Feng L, Bhalla KN, Keating MJ and Huang P. Inhibition of glycolysis in cancer cells: a novel strategy to overcome drug resistance associated with mitochondrial respiratory defect and hypoxia. Cancer Res 2005; 65: 613-621.
[45] Győrffy B, Hatzis C, Sanft T, Hofstatter E, Aktas B and Pusztai L. Multigene prognostic tests in breast cancer: past, present, future. Breast Cancer Res 2015; 17: 11.
[46] Li J, Hu K, Huang J, Zhou L, Yan Y and Xu Z. Insights of fibroblast growth factor receptor 3 aberrations in pan-cancer and their roles in potential clinical treatment. Aging (Albany NY) 2021; 13: 16541-16566.

[47] Chakraborty S and Rahman T. The difficulties in cancer treatment. Ecancermedicalscience 2012; 6: ed16.
[48] Shanmugam R, Zhang F, Srinivasan H, Charles Richard JL, Liu KI, Zhang X, Woo CWA, Chua ZHM, Buschdorf JP, Meaney MJ and Tan MH. SRSF9 selectively represses ADAR2-mediated editing of brain-specific sites in primates. Nucleic Acids Res 2018; 46: 7379-7395.
[49] Somberg M, Li X, Johansson C, Orru B, Chang R, Rush M, Fay J, Ryan F and Schwartz S. Serine/arginine-rich protein 30c activates human papillomavirus type 16 L1 mRNA expression via a bimodal mechanism. J Gen Virol 2011; 92: 2411-2421.
[50] Restifo NP, Smyth MJ and Snyder A. Acquired resistance to immunotherapy and future challenges. Nat Rev Cancer 2016; 16: 121-126.
[51] Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, Barron DA, Zehir A, Jordan EJ, Omuro A, Kaley TJ, Kendall SM, Motzer RJ, Hakimi AA, Voss MH, Russo P, Rosenberg J, Iyer G, Bochner BH, Bajorin DF, Al-Ahmadie HA, Chaft JE, Rudin CM, Riely GJ, Baxi S, Ho AL, Wong RJ, Pfister DG, Wolchok JD, Barker CA, Gutin PH, Brennan CW, Tabar V, Mellinghoff IK, DeAngelis LM, Ariyan CE, Lee N, Tap WD, Gounder MM, D'Angelo SP, Saltz L, Stadler ZK, Scher HI, Baselga J, Razavi P, Klebanoff CA, Yaeger R, Segal NH, Ku GY, DeMatteo RP, Ladanyi M, Rizvi NA, Berger MF, Riaz N, Solit DB, Chan TA and Morris LGT. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat Genet 2019; 51: 202-206.
[52] Mouradov D, Domingo E, Gibbs P, Jorissen RN, Li S, Soo PY, Lipton L, Desai J, Danielsen HE, Oukrif D, Novelli M, Yau C, Holmes CC, Jones IT, McLaughlin S, Molloy P, Hawkins NJ, Ward R, Midgely R, Kerr D, Tomlinson IP and Sieber OM. Survival in stage II/III colorectal cancer is independently predicted by chromosomal and microsatellite instability, but not by specific driver mutations. Am J Gastroenterol 2013; 108: 1785-1793.
[53] Ramadan RA, Moghazy TF, Hafez R, Morsi H, Samir M and Shamesya M. Significance of expression of pyrimidine metabolizing genes in colon cancer. Arab J Gastroenterol 2020; 21: 189-193.

A



B

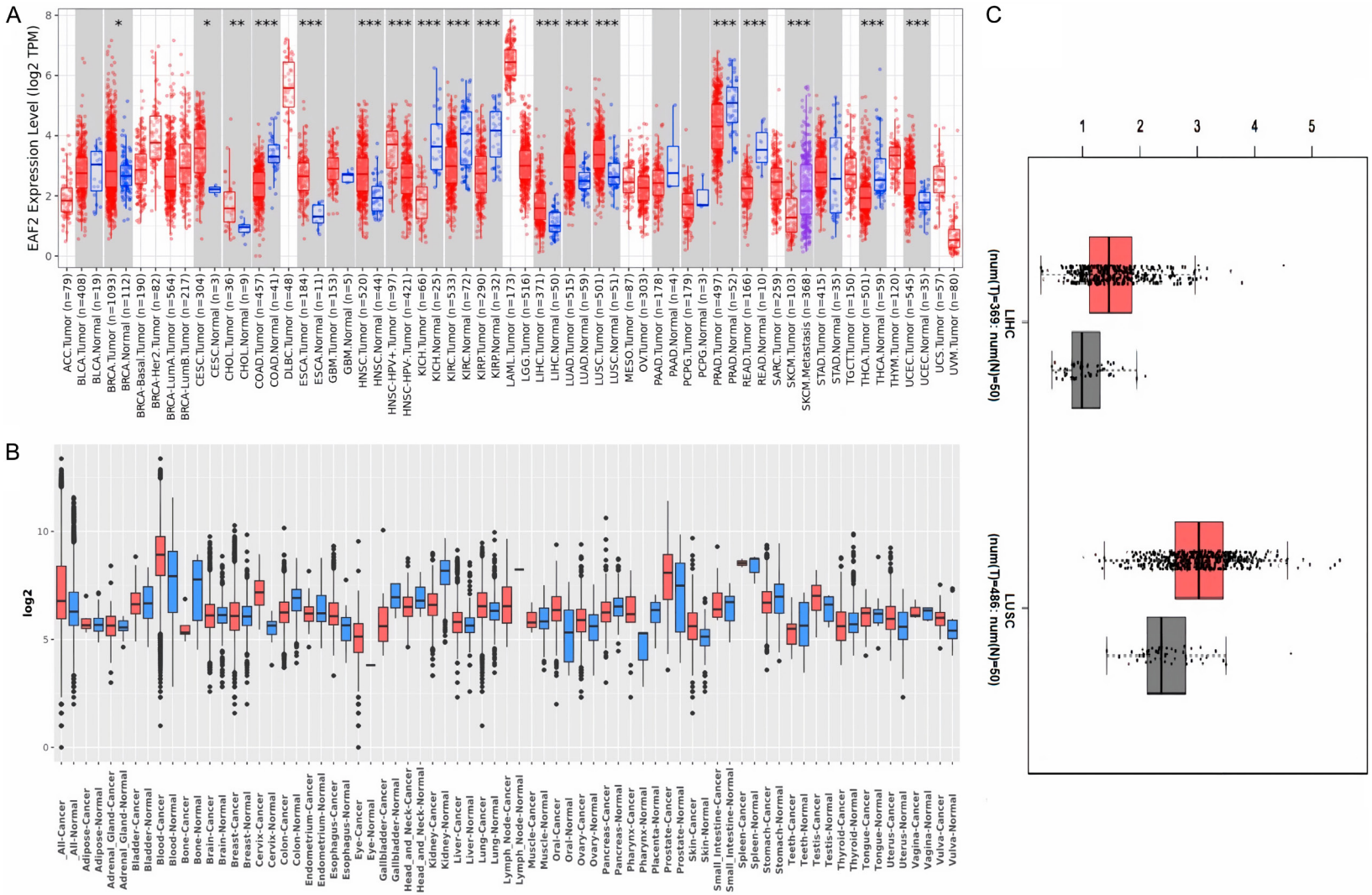
EAF2

Variables	Univariate analysis			Multivariate analysis			Dataset
	p value	HR	95% CI	p value	HR	95% CI	
Stage 1	0.0108	2.9174	0.5866 ~ 14.5086	0.3392	2.3195	0.4132~13.0214	GSE76427
Stage 2	0.0189	3.8310	1.2480 ~ 11.7600	0.0317	3.2472	1.1087~9.5101	
Stage 3	0.030	2.1428	0.3867 ~ 11.8741	0.0623	11.3302	0.8829~145.3997	
Male	0.0028	6.6350	1.9220 ~ 22.9000	0.0081	14.9754	2.0231~110.8509	
Female	0.0167	3.6756	0.3627 ~ 1.2586	0.6887	0.9383	0.6872~1.2811	

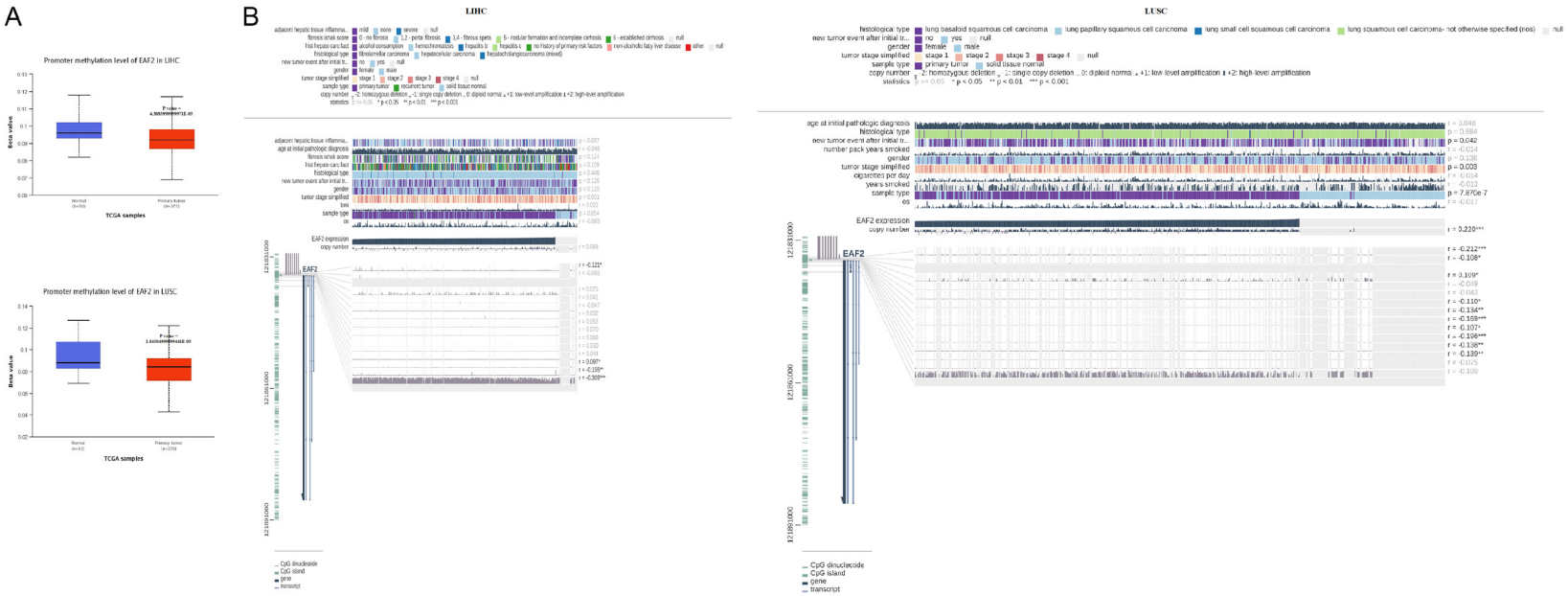
EAF2

Dataset	Variables	Univariate analysis			Multivariate analysis		
		p value	HR	95% CI	p value	HR	95% CI
GSE41271	Stage 1	0.0160	1.18	0.71~1.98	0.5175	0.90	0.67~1.23
	Stage 2	<0.0001	2.06	1.55~2.74	<0.0001	2.07	1.75~2.44
	Stage 3	0.0026	0.58	0.25~1.34	0.9036	1.03	0.66~1.59
	Male	0.0089	2.44	0.85~7.06	0.0040	2.47	1.34~4.57
	Female	0.0039	0.89	0.58~1.37	0.5894	0.89	0.58~1.36

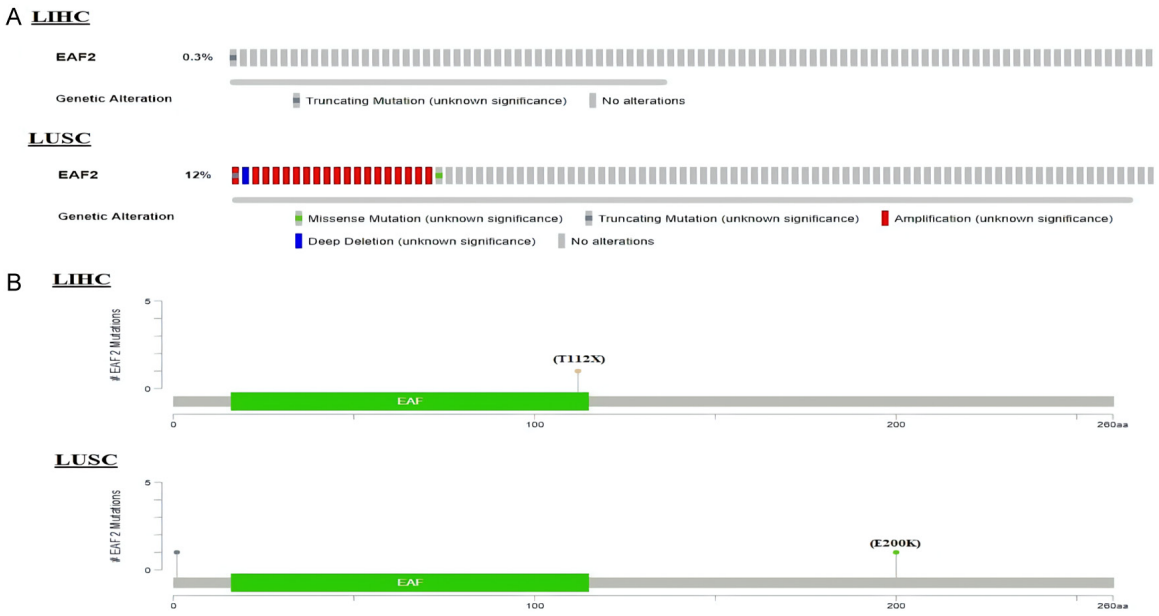
Supplementary Figure 1. Prognostic values via GEPIA and multivariate cox regression analysis of EAF2. (A) Prognostic values of EAF2 via GEPIA across LIHC and LUSC, (B) Multivariate cox regression analysis of EAF2 across LIHC (GSE76427) and LUSC (GSE41271) GEO datasets.



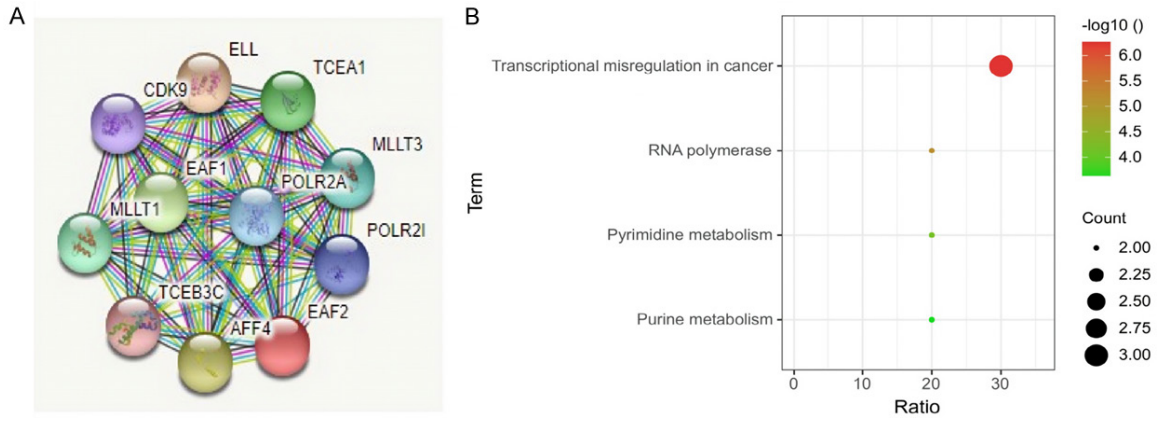
Supplementary Figure 2. Transcription expression level validation of EAF2 via TIMER, GENT2, and GEPIA. (A) TIMER, (B) GENT2, and (C) GEPIA.



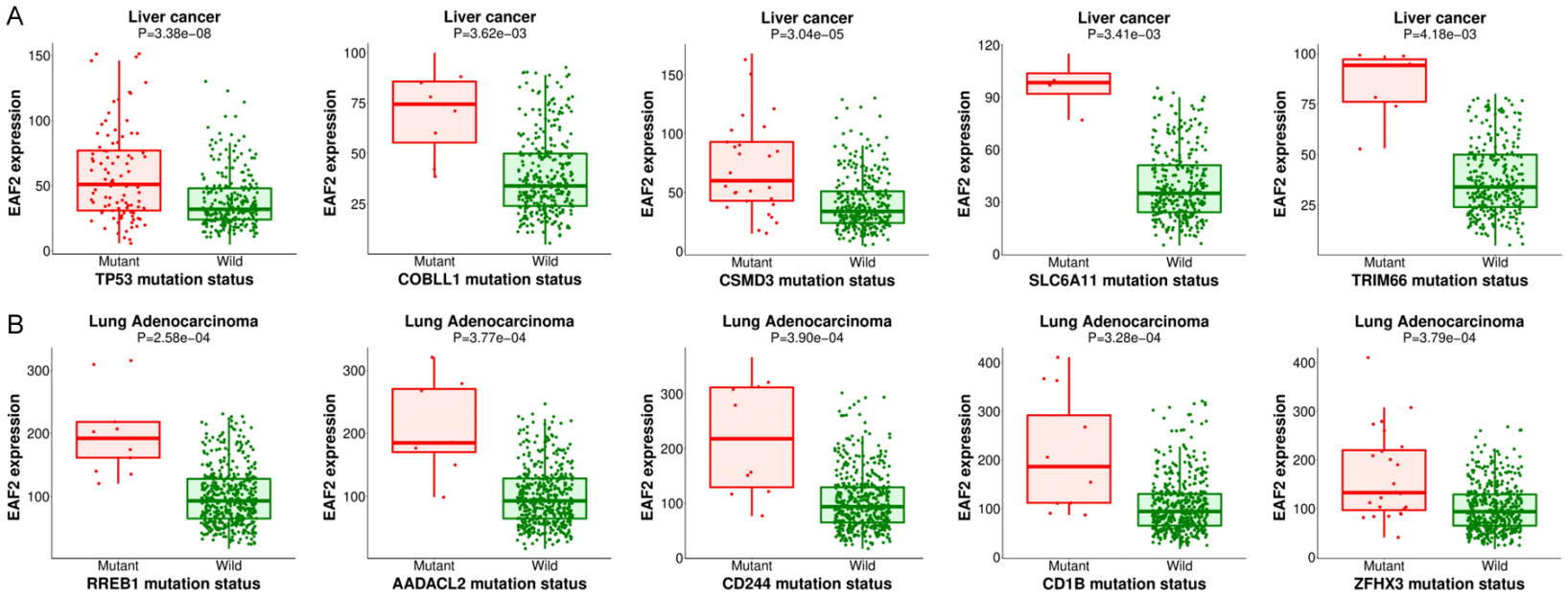
Supplementary Figure 3. EAF2 promoter methylation levels. (A) Promoter methylation level inquiry via UALCAN, and (B) Promoter methylation level validation via MEXPRESS.



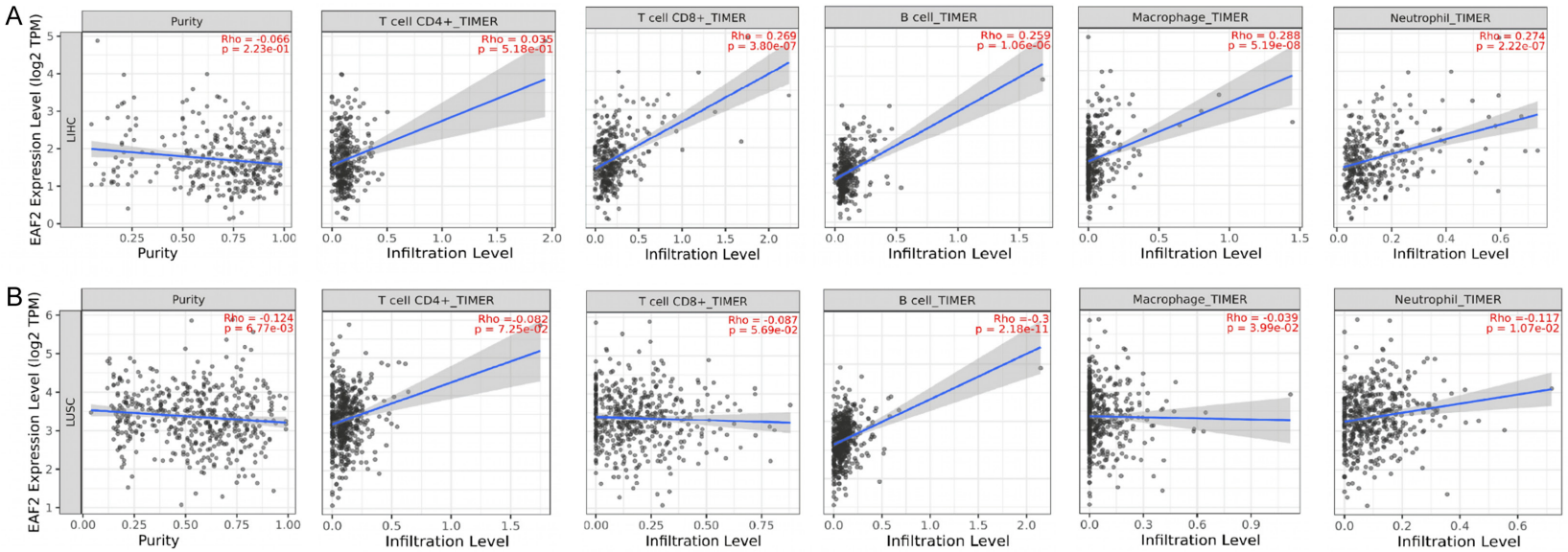
Supplementary Figure 4. Analysis of genetic mutations, CNVs alterations, and mutational hotspots identification of EAF2 in TCGA LIHC and LUSC datasets using cBioPortal. (A) The frequency of genetic alterations and CNVs (mutations, deep amplification, and deep deletion) in LIHC and LUSC samples, and (B) Protein architecture of EAF2 protein showing most frequently observed mutations in TCGA LIHC and LUSC samples. A green color region in protein architecture is showing the EAF domain of EAF2 protein.



Supplementary Figure 5. A PPI network and pathway of EAF2-interacting genes. (A) A PPI network of EAF2-interacting genes, and (B) A bubble chart showing KEGG terms.



Supplementary Figure 6. Correlations among mutant genes and EAF2 overexpression in LIHC and LUSC samples obtained via MuTarget. (A) A list of top 5 positively correlated genes with EAF2 overexpression in LIHC, and (B) A list of top 5 positively correlated genes with EAF2 overexpression in LUSC. Green color box whisker plots showing EAF2 expression in wild type samples while red color box whisker plots showing EAF2 expression in mutant samples for a specific gene.



Supplementary Figure 7. Correlation analysis between tumor purity, different immune cells infiltration level, and EAF2 expression in LIHC and LUSC. (A) A correlation analysis between CD4+ T cells, CD8+ T cells, B cells, macrophages, and neutrophils immune cells infiltration and EAF2 expression in LIHC, and (B) A correlation analysis between CD4+ T cells, CD8+ T cells, B cells, macrophages, and neutrophils immune cells infiltration and EAF2 expression in LUSC.