

ARTICLE

A prognostic and immune infiltration analysis of ZEB2 in pan-cancer

Yihuang ZHAN , Hui REN , Chen WANG , Miao YAN , Zhiwei YU ,^{*} and Mingzhe LI^{*}

Digestive medicine center, The Seventh Affiliated Hospital, Sun Yat-sen University, No.628 Zhenyuan Road, Guangming district, Shenzhen 518017, China

^{*}Corresponding author. Email: ZhiweiYU,0697@hrbmu.edu.cn;<https://orcid.org/0000-0002-0737-3307>; MingzheLI,jackmarx@163.com;<https://orcid.org/0000-0003-1696-6064>

(Received 20 April 2023; revised 18 May 2023; accepted 5 June 2023; first published online 30 June 2023)

Abstract

Objective: ZEB2 is known to be participated in several cancers, yet there is no comprehensive analysis about its function in pan-cancer. This study aimed to investigate the role ZEB2 plays in pan-cancer. **Methods:** The data was downloaded from the University of California Santa Cruz (UCSC) Xena and the Cancer Genome Atlas (TCGA). The mRNA expression status of ZEB2 was studied in the TCGA_GTEX samples, TCGA samples and paired samples in TCGA, respectively, and protein expression status was obtained from the University of Alabama at Birmingham Cancer data analysis Portal (UALCAN). Kaplan-Meier analysis was applied to 33 kinds of tumors in TCGA, then among the cancers that ZEB2 can affect prognosis, clinical correlation analysis and univariate and multivariate Cox regression analysis were performed. Furthermore, to confirm the prognostic value of ZEB2 in cancers, nomogram models were constructed on lung adenocarcinoma (LUAD) and stomach adenocarcinoma (STAD). The relationship between ZEB2 mRNA expression and immune cell infiltrates was determined by tumor immune estimation resource (TIMER2.0). Protein-protein interaction (PPI) networks were constructed by the STRING. Functional enrichment analysis was also performed using the “ClusterProfiler” package. **Results:** ZEB2 was expressed in most tumors, either upregulated or downregulated, and correlated with clinical characteristics in a part of tumors. The expression level of ZEB2 was also found to impact the prognosis of tumors. Additionally, ZEB2 expression was significantly related to immune cells infiltration level in tumor microenvironment. **Conclusion:** we propose that ZEB2 has the potential to be a prognostic biomarker and a promising target for cancer immunotherapy.

Keywords: ZEB2; pan-cancer; nomogram model; immune infiltration

Introduction

According to WHO, cancer has become the number one cause of death worldwide. It can occur in any part of the body and can affect people of all ages, genders, and ethnicities, not only jeopardizing the lives of patients, but also diminishing their quality of life and having an indelible influence on the whole family. Cancer is expensive to treat, demanding large medical expenditures. Patients often require long-term therapy and care and are often absent from their social roles, which is also a loss of productivity for society. Such excessive consumption puts pressure on the medical and insurance systems that it is urgent to enhance early diagnosis and improve cure rate and survival rate. Although significant progress has been made in the diagnosis and treatment of cancer in recent years, it remains a major public health challenge.

One of the biggest challenges in cancer diagnosis is the lack of specific and sensitive biomarkers that can be used for early detection, prognosis, and treatment guidance. Fortunately, advances in high-throughput technologies and bioinformatics have enabled the mining of large-scale omics data to identify pan-cancer biomarkers. Pan-cancer refers to the analysis of multiple types of cancer using a common set of criteria, with the goal of identifying common features across different cancer types. Mining biomarkers that can sensitively identify pan-cancer is of great importance for early screening of tumors.

Zinc finger E-box binding homeobox 2 (ZEB2) is located in the nucleus and functions as a DNA-binding transcriptional repressor that interacts with activated SMADs. The protein it encodes is a member of the Zfh1 family of 2-handed zinc finger/homeodomain proteins. It is involved in several signaling pathways that contribute to tumor formation, development, and metastasis, such as epithelial-mesenchymal transition (EMT)(1), Wnt/-catenin signaling pathway(2), transforming growth factorbeta (TGF-) signaling pathway(3), and so on. Several studies have identified ZEB2 as a potential diagnostic biomarker for various types of cancer. For example, ZEB2 has been shown to be upregulated in breast cancer and may serve as a prognostic marker for poor survival(4). In lung cancer, ZEB2 expression has been linked to tumor grade and stage, and may be a useful biomarker for predicting patient outcomes(5).

In this project, the results of differential expression analysis showed that ZEB2 was differentially expressed in many tumors, and the results of survival analysis suggested that ZEB2 may have an impact on the prognosis of some tumors. Moreover, the expression of ZEB2 correlates with immune cells, and alterations in the immune microenvironment can also affect tumor prognosis. The results of the enrichment analysis also suggest that ZEB2 is involved in tumorigenesis. In summary, ZEB2 is widely involved in the development of pan-cancer through multiple pathways, which may provide new biomarkers for early screening and targets for tumor treatment.

Materials and methods

Data collection and processing

The expression profile data and clinical data of 33 tumors were downloaded from the Cancer Genome Atlas (TCGA, <https://portal.gdc.cancer.gov/>). The TCGA_GTEx dataset included TCGA samples and normal samples was collected from the University of California Santa

Cruz (UCSC) Xena (<https://xenabrowser.net/datapages/>). The 100 most relevant genes with ZEB2 were obtained from the TCGA datasets using the Gene Expression Profiling Interactive Analysis database (GEPIA2, <http://gepia2.cancer-pku.cn/index>).

Expression analysis of ZEB2

In TCGA samples, TCGA_GTEX samples, and TCGA paired samples, the mRNA expression of ZEB2 in normal tissues and tumor tissues was compared. Additionally, protein expression has been studied using UALCAN, an online tool for analyzing openly accessible cancer data.

Prognosis analyses

To determine the relationship between ZEB2 expression and clinical outcomes such as overall survival (OS), progression-free interval (PFI), and disease-specific survival (DSS) in pan-cancer in TCGA, Kaplan-Meier analysis with log-rank test was used. The survival curves with $p < 0.05$ were then displayed. In cancers where ZEB2 might influence prognosis, the receiver operating characteristic curve (ROC) was also drawn.

Correlation analysis between ZEB2 expression and clinical features

In cancers where ZEB2 can affect prognosis, the relationship between ZEB2 expression and an array of significant clinical characteristics, including T stage, N stage, pathologic stage, and so on, was examined.

Establishment and evaluation of the nomogram models

In cancers where ZEB2 can affect OS, univariate and multivariate Cox regression analyses for OS were conducted. Tumors with $p < 0.05$ and sample sizes greater than 500 were selected to construct a nomogram model, an efficient and practical method for predicting the OS in specific patients. The calibration curves were used to evaluate the nomograms' prognostication accuracy at 1, 3, and 5 years.

Immune infiltration analysis

The Tumor Immune Estimation Resource 2.0 (TIMER2.0, <http://timer.cistrome.org/>) was used to analyze the relationships between ZEB2 expression and several immune cell types, such as macrophages, B cells, Tregs, T cells CD4+, T cells CD8+, DC, and NK, across all of the tumors in the TCGA.

Functional enrichment analysis and protein-protein interaction network analysis

From the GEPIA2 database, the 100 genes that are most similar to ZEB2 in terms of expression pattern were identified. To further examine the potential roles of ZEB2, gene ontology (GO) analysis, which mostly included biological pathways (BP), cellular components (CC), and molecular functions (MF), and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, were carried out based on the ZEB2-related genes. The Search Tool for the Retrieval of Interacting Genes (STRING) database (<https://cn.string-db.org>) was also used to generate a PPI network using the 100 ZEB2-related genes, with 0.4 chosen as the minimal interaction threshold.

Statistical analysis

The Wilcoxon rank-sum test was used to compare differences between two groups, whereas the Spearman rank test was used to determine correlations between two groups. The factors influencing the prognosis were examined using univariate and multivariate Cox proportional hazard regression. The survival analysis method employed was Kaplan-Meier analysis with a log-rank test. Statistical analysis was executed by R (version 4.2.1), and $p < 0.05$ was regarded as statistically significant (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$).

Results

The expression of ZEB2 in pan-cancer

In order to clarify the expression of ZEB2 in pan-cancer, the uniformed TCGA_GTEEx data collected from the UCSC were analyzed. The results showed that ZEB2 was differentially expressed in most tumors, most with low expression such as BRCA, COAD, STAD, LUAD, LUSC and so on, and some with high expression (Figure 1A), which is roughly consistent with the results from TCGA database. Additionally, the expression of ZEB2 in 20 kinds of tumors with paired samples in TCGA was also analyzed (Figure 1C), which kept consistence with the results from TCGA.

We used UALCAN to analyze CPTAC in order to give a comprehensive characterization of ZEB2 expression. The findings demonstrated that ZEB2 protein expression was lower in LUAD, LUSC, and BRCA than in normal tissues (Figure 2), which was consistent with mRNA expression in these cancers from the TCGA.

The correlation between mRNA expression of ZEB2 and prognosis in pan-cancer

Kaplan-Meier survival analysis was carried out to examine the relationship between the expression of ZEB2 and clinical outcomes in tumors with significantly different ZEB2 expression levels to better understand the prognostic evaluation value of ZEB2 in pan-cancer. First, we looked into the relationship between OS and ZEB2 expression. The results indicated that high ZEB2 expression was protective in LUAD and GBMLGG whereas it was associated with poor prognosis in STAD, MESO, and LUSC (Figure 3). Meanwhile, we explored the link between the expression of ZEB2 and PFI, and it was still a protective factor in GBMLGG, maintaining consistent with the result in OS, while in ACC it exhibited a worse prognosis (Figure 4). Similarly, we explored the relationship between ZEB2 expression and DSS and the results demonstrated that upregulation of ZEB2 was associated with a better prognosis in both GBMLGG and LUAD (Figure 5). ROC curves were presented in supplementary figure.

The relationships between ZEB2 mRNA level and clinicopathological characteristics

We then examined the correlation between clinicopathological data and ZEB2 expression using the TCGA database in order to comprehensively clarify which aspect of the tumor process ZEB2 expression was most associated with. The results showed that among the tumors known to be prognostically relevant in the aforementioned analysis, ZEB2 expression was only connected with tumor size and lymph node metastasis in STAD (Figure 6B,E), and the relevance between other of them and ZEB2 expression was not statistically significant.

In contrast, in some prognostically irrelevant tumors, such as SARC, ZEB2 expression was connected with gender (Figure 6A), while in THCA, ESCA, and COADREAD, for example, ZEB2 expression was significantly correlated with tumor size (Figure 6D,F,G,H). In BLCA, ZEB2 was found to be connected with pathologic stage (Figure 6I).

Nomogram models in LUAD and STAD

We conducted a univariate and multivariate Cox regression analysis of ZEB2 expression and clinical features in LUAD and STAD respectively in order to investigate the potential of ZEB2 in predicting the prognosis of specific cancers (Supplementary Tables 1–2). Based on which relevant factors were selected to construct nomogram models (Figure 7A–B), and the accuracy of the models was evaluated using calibration curves (Figure 7C–D). The results showed that the two nomogram models were reliable in predicting patient prognosis.

Correlation of ZEB2 expression and tumor immune environment

To investigate the relationship between ZEB2 and tumor immune microenvironment, we conducted a correlation analysis of its expression with an array of immune cells and heatmaps were generated. Our findings suggested that ZEB2 expression was closely related to the infiltration of immune cells in the tumor microenvironment, with a significant correlation with immune cells such as macrophages, B cells, Tregs, T cells CD4+, T cells CD8+, DCs, and NK cells (Figure 8).

PPI networks and functional annotations

With the 100 ZEB2-related genes derived from TCGA database with GEPIA2, a PPI network was generated on the STRING website and shown in Figure 9A. Further, we conducted GO and KEGG analyses. GO analysis indicated that ZEB2-related genes might participate in biosynthetic processes like establishment of organelle localization, TOR signaling, phosphatidylinositol phosphate and phosphatidylinositol-3-phosphate biosynthetic process, as long as in cell components like distal axon, centriole, DNA repair complex and inflammasome complex, and in molecular functions such as protein serine/threonine kinase activity, protein serine kinase activity, GTPase activator activity and phosphatidylinositol kinase activity (Figure 9B). KEGG enrichment analysis did not produce a result, which may be related to the low transformation rate of the relevant genes.

Discussion

The diagnosis and treatment of cancer poses a formidable challenge to the global medical community. Despite years of extensive research and exploration, the mortality rate associated with cancer remains high. Presently, biomarkers such as carcinoembryonic antigen (CEA) and carbohydrate antigen 199 (CA199) have been identified to aid in the early diagnosis and prognostication of certain cancers. However, these achievements represent a mere fraction of the overall challenge. Furthermore, these biomarkers often exhibit association with only limited types of cancers, whereas the actual range of cancer subtypes is so vast that the currently discovered tumor markers fail to meet the demand for comprehensive early cancer screening.

The invasive growth of tumors is closely related to the adhesive properties of tumor cells, cell polarity, and the tumor microenvironment. Studies have found that the ZEB gene family is expressed in various tumors and may promote tumor invasion and metastasis. As an important member of the ZEB gene family, ZEB2 can directly downregulate the expression of cell adhesion-related coding genes, including adhesive connections, tight junctions, desmosomes, gap junctions, and also has the potential to promote tumor invasion and metastasis(6). However, there is no analysis of what role ZEB2 plays in pan-cancer.

This study demonstrated differential expression of the transcription factor ZEB2 in various types of tumors, which was correlated with clinical characteristics in different tumors. The expression level of ZEB2 was also found to impact the prognosis of tumors. The predictive nomogram models constructed based on these analyses exhibited good reliability. Furthermore, evidence was presented to support the role of ZEB2 in immune infiltration of the tumor microenvironment, providing great insights into the development and progression of cancer.

Numerous studies have been performed already, revealing that ZEB2 expression is upregulated in tumors and plays a role in promoting tumor progression. According to the study of Francescangeli *et al.*, ZEB2 overexpression plays a crucial role in increased chemoresistance, and is associated with worse relapse-free survival in human colorectal cancer(7). Another research on triple negative breast cancer suggests that ZEB2 expression is increased in breast cancer with poor prognosis(8). There are also articles suggest that ZEB2 is upregulated when it is connected to poor prognosis in gastric cancer as it promotes tumor invasion and metastasis(9), induces epithelial-mesenchymal transition (EMT), and correlates with advanced disease stage, lymph node metastasis(10). However, these results partly contradict our results from analysis of the TCGA database, where we observed that ZEB2 expression is usually downregulated in tumors, and according to the results of OS that upregulation of ZEB2 expression in tumors is normally associated with a better prognosis. This discrepancy may be due to the fact that the samples in the TCGA database are mainly from tumor tissues and fewer normal control samples, which may have an impact on the analytical results of the data. In addition, the heterogeneity of the tumors themselves is also a factor that cannot be ignored. The most important point is that tumorigenesis and progression are not determined by a single gene, but by a combination of factors. To increase the credibility of the analysis results, further validation in conjunction with clinical databases is essential.

Although the results of differential expression between tumors and non-tumors were not as expected, the results of differential expression on the subgroups of clinical characteristics were more favorable for comparison. In terms of TNM stage, ZEB2 showed a correlation between upregulated expression and clinical progression in different cancers. The prediction model constructed by combining clinical characteristics and prognostic analysis was also credible.

Nowadays, an increasing number of studies have concluded that the level of immune cell infiltration in the tumor microenvironment has a great impact on tumor progression and prognosis. Most of them believes that various factors can contribute to the complex relationship between immune cell infiltration and tumor prognosis, such as the type and location of immune cells, the tumor microenvironment, and the immune suppression

mechanisms employed by tumors(11). In general, higher levels of immune cell infiltration are associated with better prognosis in some types of cancer, particularly those that are responsive to immunotherapy(12). Our findings also suggest that ZEB2 expression is closely related to the infiltration of immune cells such as macrophages, B cells, Tregs, T cells CD4+, T cells CD8+, DCs, and NK cells. However, the relationship between immune cell infiltration and prognosis can vary depending on the type and stage of cancer, as well as other individual factors such as patient age and overall health(13). To explore the specific mechanism and role of ZEB2 in immune infiltration, further analysis of its levels in different tumors and validation with experimental and clinical data such as immunohistochemistry are still needed. The results obtained from the mining and clustering analysis for ZEB2-associated genes can also serve as a hint for the next step of the study, highlighting the complex mechanisms of ZEB2 action in tumors and the need for further studies to fully understand its impact on tumor prognosis.

Although there are many studies on ZEB2, they are mainly focused on single cancer types. The aim of this study is to use a large public database to do a pooled analysis of the role of ZEB2 in pan-cancer to find new possible indicators for cancer screening. The results of this study also suggest that ZEB2 is involved in multiple tumor processes and is significantly associated with tumor prognosis and with the level of immune infiltration in the tumor microenvironment, and has the potential to become a tumor biomarker. However, the limitations of this study are obvious, as the results of the analysis are not fully consistent with the actual results due to the complexity of the data in the database and tumor heterogeneity, and further wet experiments are still needed to verify the results. We need to explore in greater depth the molecular mechanisms by which ZEB2 interacts with immune cells and how it is regulated in different tumor types and settings to better explain the relationship between ZEB2 expression and prognosis.

Data availability statement

The TCGA database and the UCSC database both contain the data that were used in this study's analysis of publicly accessible datasets.

ACKNOWLEDGMENT

None.

STATEMENT

There is no conflict of interest in this article.

Funding

This study was funded by the Free Exploration Foundation of Science, Technology and Innovation Commission of Shenzhen Municipality (No. JCYJ20180307151238174); Shenzhen Fundamental Research Program (JCYJ20190809142807444).

Authors

Yihuang ZHAN and Hui REN contributed equally to this work.

References

- [1] Yuan J h ang, Yang F, Wang F, Ma J zhao, Guo Y jun, Tao Q fei, et al. A long noncoding RNA activated by TGF- promotes the invasion-metastasis cascade in hepatocellular carcinoma. *Cancer Cell*. 2014 May 12;25(5):666–81.
- [2] Ang HL, Mohan CD, Shanmugam MK, Leong HC, Makvandi P, Rangappa KS, et al. Mechanism of epithelial-mesenchymal transition in cancer and its regulation by natural compounds. *Med Res Rev*. 2023 Jul;43(4):1141–200.
- [3] Wang Y, Guo Y, Hu Y, Sun Y, Xu D. Endosulfan triggers epithelial-mesenchymal transition via PTP4A3-mediated TGF- signaling pathway in prostate cancer cells. *Sci Total Environ*. 2020 Aug 20;731:139234.
- [4] Si W, Huang W, Zheng Y, Yang Y, Liu X, Shan L, et al. Dysfunction of the Reciprocal Feedback Loop between GATA3- and ZEB2-Nucleated Repression Programs Contributes to Breast Cancer Metastasis. *Cancer Cell*. 2015 Jun 8;27(6):822–36.
- [5] Wu DM, Zhang T, Liu YB, Deng SH, Han R, Liu T, et al. The PAX6-ZEB2 axis promotes metastasis and cisplatin resistance in non-small cell lung cancer through PI3K/AKT signaling. *Cell Death Dis*. 2019 Apr 25;10(5):349.
- [6] Slowicka K, Petta I, Blancke G, Hoste E, Dumas E, Sze M, et al. Zeb2 drives invasive and microbiota-dependent colon carcinoma. *Nat Cancer*. 2020 Jun;1(6):620–34.
- [7] Francescangeli F, Contavalli P, De Angelis ML, Careccia S, Signore M, Haas TL, et al. A pre-existing population of ZEB2+ quiescent cells with stemness and mesenchymal features dictate chemoresistance in colorectal cancer. *J Exp Clin Cancer Res CR*. 2020 Jan 8;39(1):2.
- [8] di Gennaro A, Damiano V, Brisotto G, Armellin M, Perin T, Zucchetto A, et al. A p53/miR-30a/ZEB2 axis controls triple negative breast cancer aggressiveness. *Cell Death Differ*. 2018 Dec;25(12):2165–80.
- [9] Gong P, Qiao F, Wu H, Cui H, Li Y, Zheng Y, et al. LncRNA UCA1 promotes tumor metastasis by inducing miR-203/ZEB2 axis in gastric cancer. *Cell Death Dis*. 2018 Nov 21;9(12):1158.
- [10] Li H, Xu L, Li C, Zhao L, Ma Y, Zheng H, et al. Ubiquitin ligase Cbl-b represses IGF-I-induced epithelial mesenchymal transition via ZEB2 and microRNA-200c regulation in gastric cancer cells. *Mol Cancer*. 2014 Jun 2;13:136.
- [11] Bruni D, Angell HK, Galon J. The immune contexture and Immunoscore in cancer prognosis and therapeutic efficacy. *Nat Rev Cancer*. 2020 Nov;20(11):662–80.

- [12] Fridman WH, Pagès F, Sautès-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer*. 2012 Mar 15;12(4):298–306.
- [13] Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pagès C, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science*. 2006 Sep 29;313(5795):1960–4.

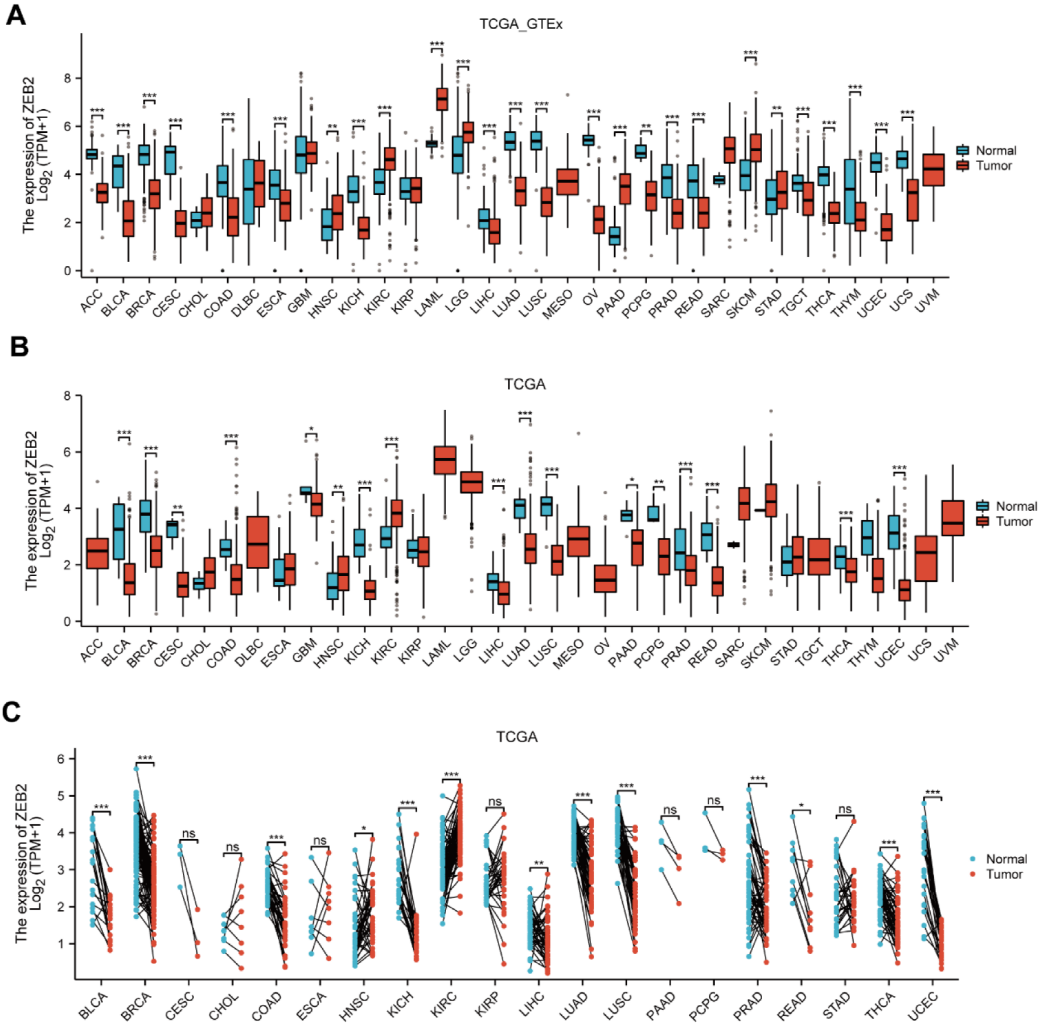


Figure 1. The mRNA expression of ZEB2 in pan-cancer. (A) The mRNA expression of ZEB2 in 33 tumors in TCGA_GTEX samples. (B) The mRNA expression of ZEB2 in 33 tumors in TCGA database. (C) Expression of ZEB2 in paired samples of 20 tumors in TCGA database.

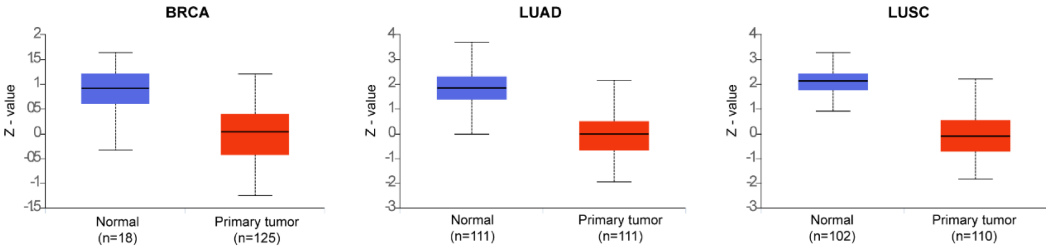


Figure 2. The protein expression of ZEB2 extracted from UALCAN.

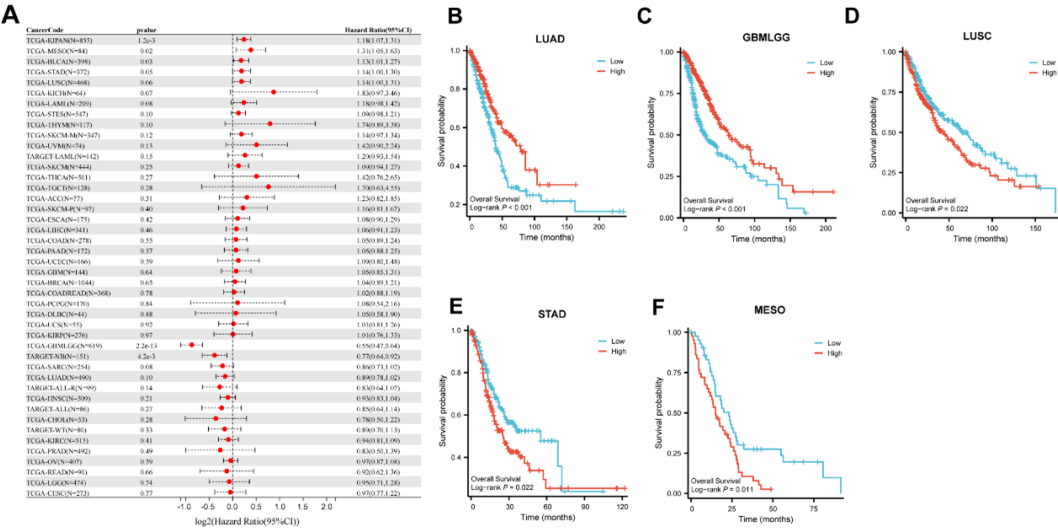


Figure 3. The association between ZEB2 expression and OS in pan-cancer. (A) The effects of ZEB2 expression on OS in pan-cancer were exhibited by a forest plot. (B-F) Effects of ZEB2 expression on OS in LUAD, GBMLGG, LUSC, STAD and MESO, respectively.

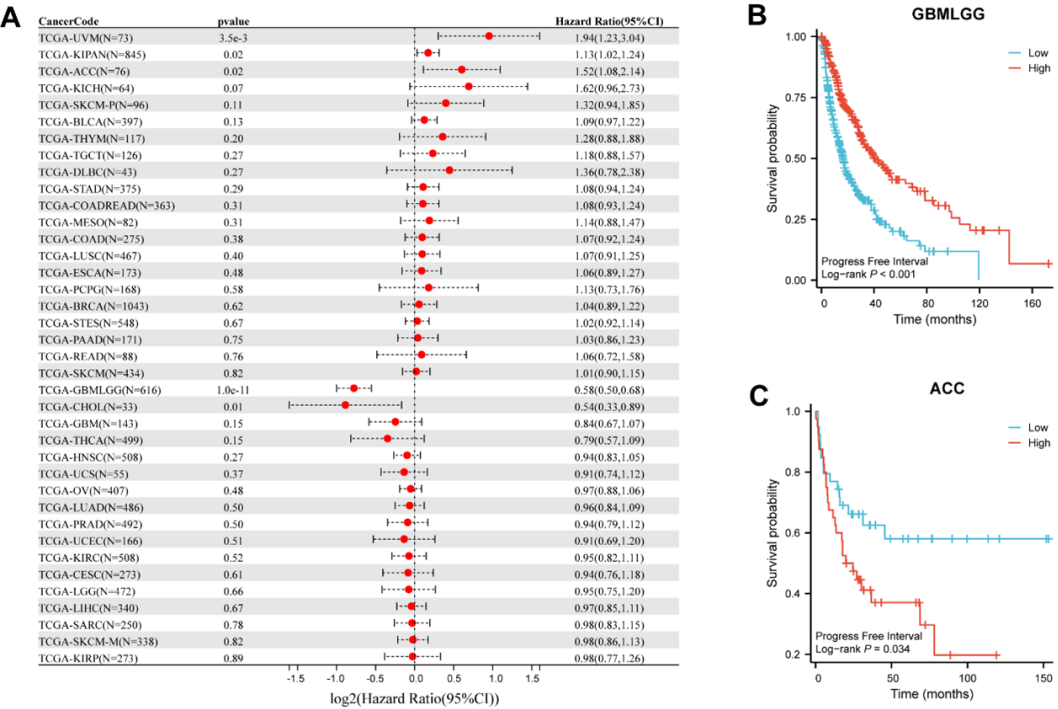


Figure 4. The association between ZEB2 expression and PFI in pan-cancer. (A) The associations of ZEB2 expression on PFI in various cancers were displayed by a forest diagram. (B–D) Survival curves of ZEB2 on PFI in GBMLGG and ACC, respectively.

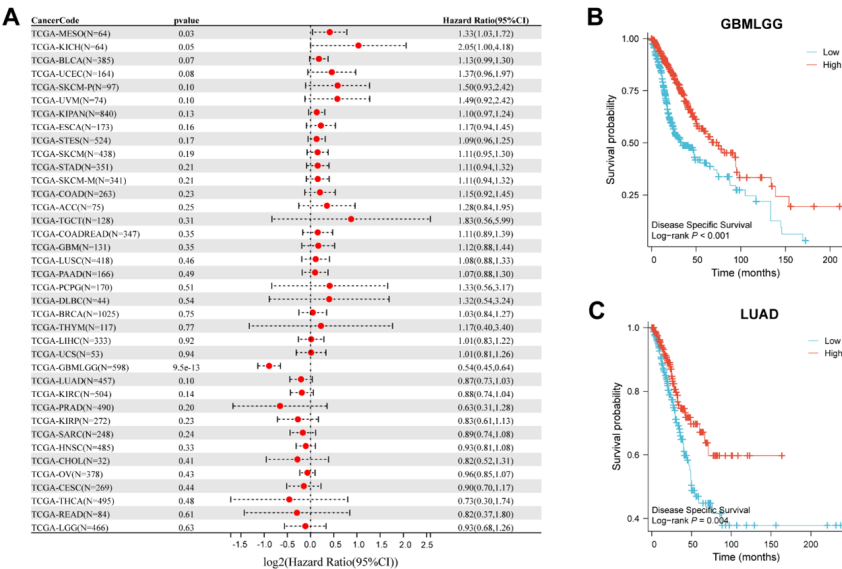


Figure 5. The association between ZEB2 expression and DSS in pan-cancer. (A) A forest map of the relationship of ZEB2 expression on DSS in TCGA. (B–C) Effects of ZEB2 expression on DSS in GBMLGG, and LUAD, respectively.

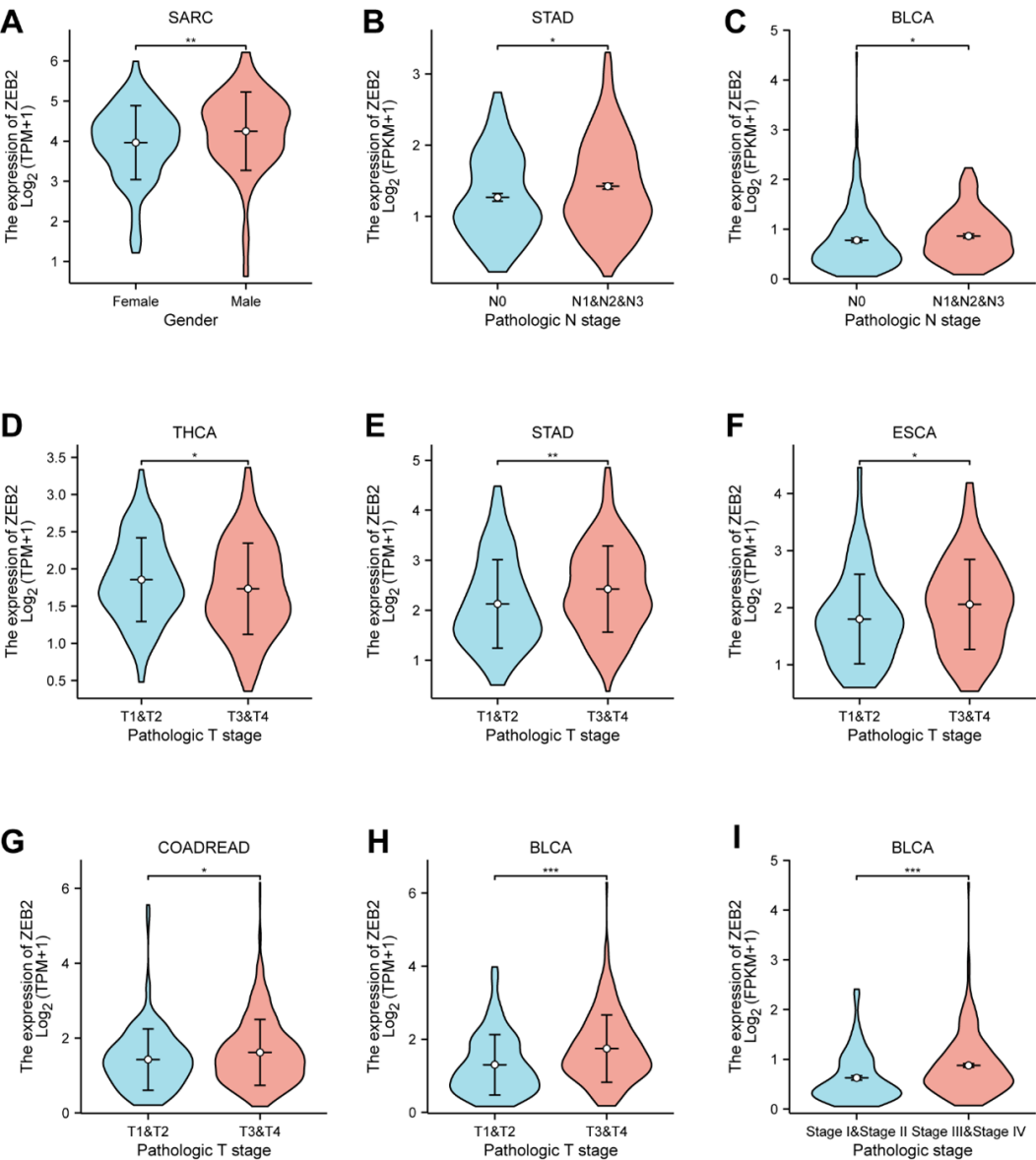


Figure 6. The correlation between ZEB2 expression and clinicopathological parameters. (A) ZEB2 expression was correlated with gender in SARC. (B-C) ZEB2 expression was correlated with N stage in STAD and BLCA. (D-H) ZEB2 expression was correlated with T stage in THCA, STAD, ESCA, COADREAD and BLCA. (I) ZEB2 expression was correlated with pathologic stage in BLCA.

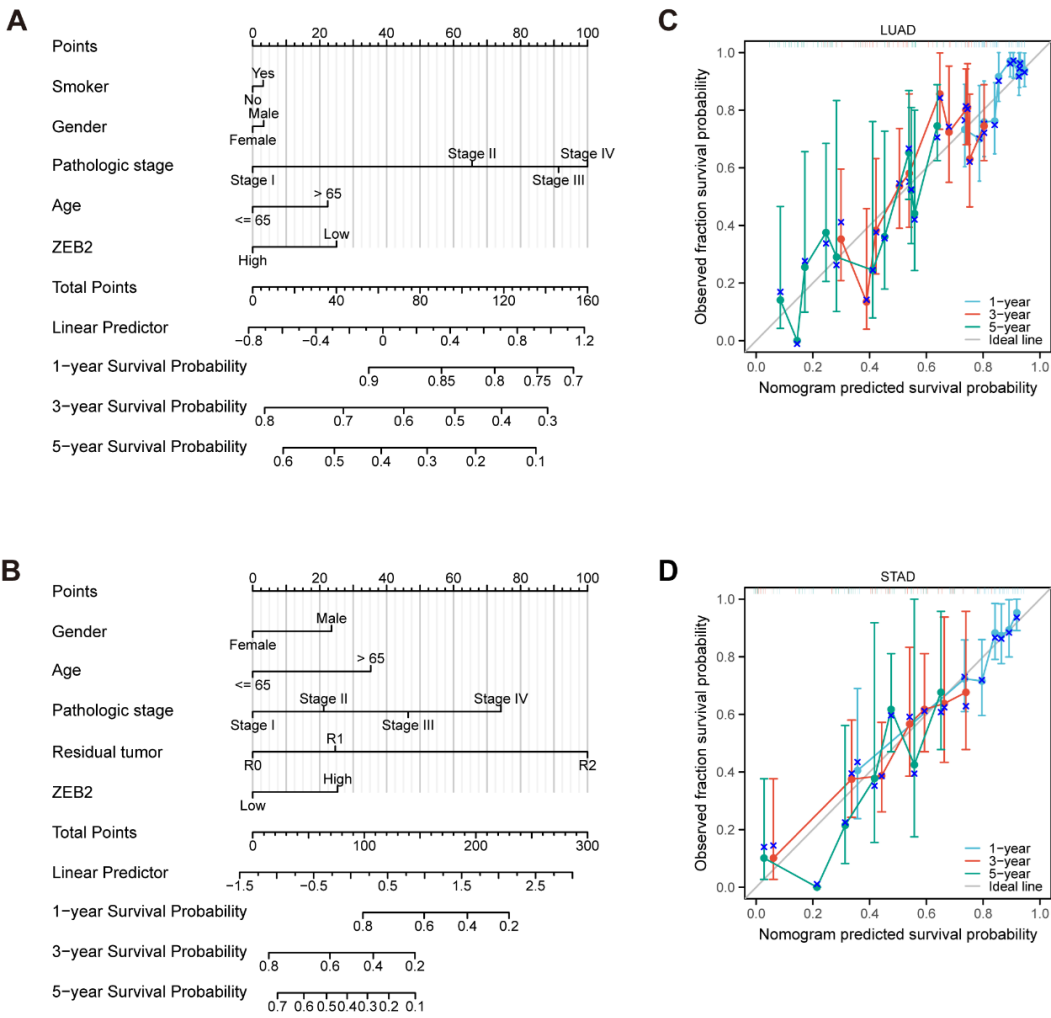


Figure 7. Nomogram models were established and evaluated in LUAD and STAD.

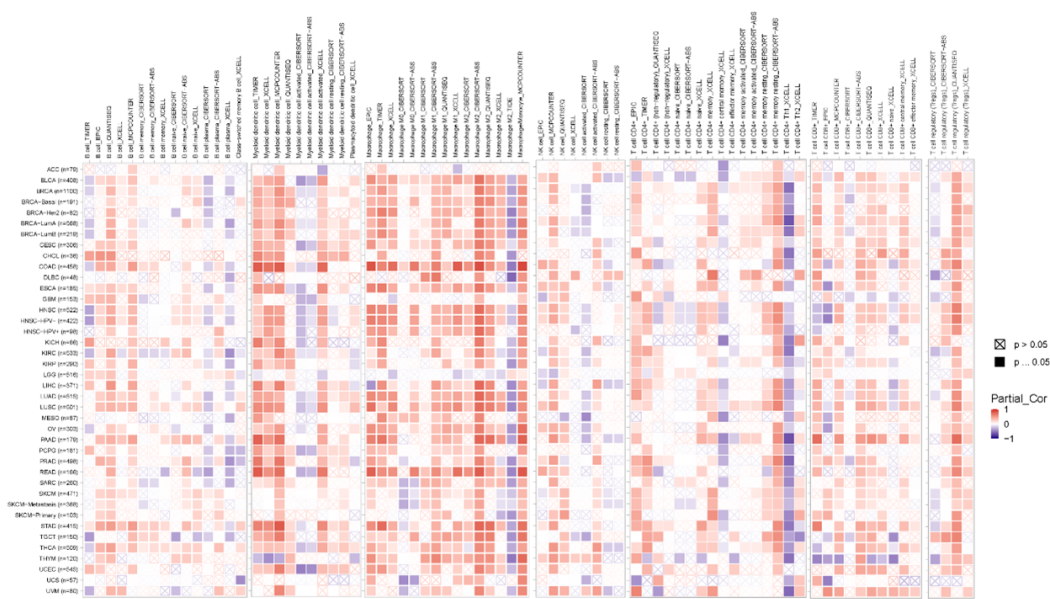


Figure 8. The correlation of ZEB2 expression and immune cell infiltration.

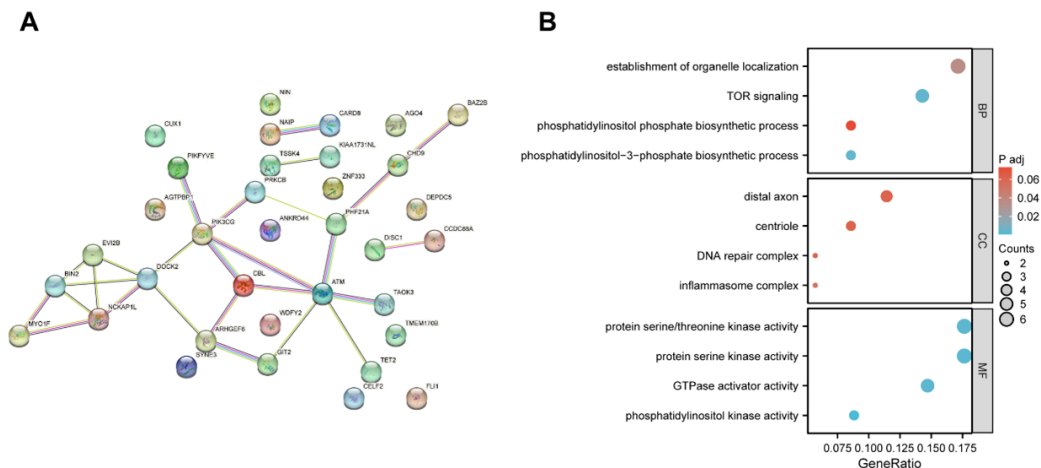


Figure 9. PPI networks and functional enrichment analyses. (A) A network of ZEB2 and its co-expression genes. (B) GO enrichment analysis based on 100 ZEB2-related genes, including BP, CC, and MF.