Abnormal Expression of EPB41L1 in Colon Adenocarcinoma and Effect on Prognosis

jiaojiao Liu
Zhengzhou University https://orcid.org/0000-0002-8149-5281
Taotao Liang
Zhengzhou University
Cong Ding
Zhengzhou University
Zhenyu Ji
Zhengzhou University
Ting Wang
Zhengzhou University
Xin Liu (liux@zzu.edu.cn)
Zhengzhou University
Qiaozhen Kang
Zhengzhou University

Primary research

Keywords: EPB41L1, Colon adenocarcinoma, Prognosis, Bioinformatics analysis, Biomarker.

Posted Date: April 12th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-358130/v1

Abstract

Background: Colon adenocarcinoma (COAD) is the most common pathological type of colorectal cancer (CRC) and further study of the molecular mechanism will help to improve the quality of life of patients with COAD. Erythrocyte membrane protein band 4.1 like 1(EPB41L1), a gene encoding protein 4.1N, has been reported to be closely associated with tumorigenesis and progression. However, the role of EPB41L1 in COAD is largely unknown and remains to be fully studied.

Methods: In this study, we analyzed the mRNA expression of EPB41L1 in COAD through the Oncomine and GEO databases. Then, the relative expressions of EPB41L1 across the sub-groups of COAD were performed by the UALCAN data portal. Next, we investigated the prognostic value of EPB41L1 in COAD patients by using the UALCAN and HPA online databases. Further, the mutation of EPB41L1 in COAD was analyzed by c-Bioportal. The co-expression genes of EPB41L1 in COAD were displayed from the LinkedOmics database, and function enrichment analysis was analyzed by DAVID. The co-expression gene network was constructed through the STRING database, and the MCODE plug-in of which was used to build the gene modules, both of them were visualized by Cytoscape software. Finally, the pathway enrichment of the top modular genes in the co-expression gene network was analyzed by DAVID.

Results: The results revealed that EPB41L1 is significantly upregulated with the development of COAD, leading to a poor prognosis. There are mutations in the EPB41L1 gene, but it has no significant effect on the prognosis of COAD. Moreover, the genes correlated with EPB41L1 in COAD, identified in the most highly connected sub-network, were enriched in the cell cycle.

Conclusions: In summary, the above results suggest that EPB41L1 was significantly upregulated in the COAD tissues and the high expression level of EPB41L1 predicts a poor prognosis of COAD patients. Therefore, we suggest that EPB41L1 can be a potential candidate biomarker for the diagnosis and prognosis of COAD.

Introduction

Colon adenocarcinoma (COAD) is one of the most common cancer types and also a major cause of cancerous death throughout the world[1]. COAD is a kind of malignant epithelial neoplasm with adenoid differentiation of colonic mucosa originating from epithelial dysplasia followed by malignant infiltration and growth [2, 3]. Many factors such as environmental factors, poor diet, and genetic factors significantly contribute to the onset of COAD[4, 5]. Cell cycle dysregulation and abnormal metabolism are the important characteristics of COAD[1, 6]. Although surgical techniques and chemotherapy have made progress in recent decades, the survival rate of COAD does not seem to be substantially improved, and many patients die of recurrence events, leading to poor prognosis [7]. Effective and reliable diagnostic and prognostic biomarkers can provide useful prognostic information and help to guide treatment through identifying "high-risk" COAD patients[8, 9]. Therefore, there is an urgent need to identify reliable prognostic biomarkers in COAD to guide effective targeted treatment for clinical benefits.

Members of the protein 4.1 superfamily are cytoskeletal proteins that play a key role in maintaining normal cell morphology and regulating cell division, cell proliferation, and intercellular signaling[10]. It includes Protein 4.1R, Protein 4.1G, Protein 4.1N, and Protein 4.1B. There are three conserved domains in all members of this family: an N-terminal FERM domain, a spectrin-actin-binding domain (SABD), and a C-terminal domain (CTD)[11]. Just like other protein 4.1 family members, protein 4.1N serves as critical components of the membrane skeleton and a bridge connecting actin-networks and the transmembrane proteins[12]. Protein 4.1N has been reported to be associated with tumorigenesis and progression such as breast and ovarian cancer[13, 14], and its abnormal expression affects tumor cell movement, proliferation, and metastasis[10]. However, the biological function of 4.1N coding gene EPB41L1 in COAD is largely unknown and remains to be fully studied. Our research aims to explore the biological roles of EPB41L1 in COAD.

In the present work, we compared the EPB41L1 expression patterns between COAD and normal colon tissues using public databases, assessed the influences of EPB41L1 expression on COAD prognostic values, and predicted co-expression genes of EPB41L1 in COAD. Further, we constructed the Protein–Protein Interaction (PPI) network and performed functional enrichment of EPB41L1-related genes. The present study was the first to report the value of EPB41L1 in predicting the survival rate of COAD patients and its potential mechanism in regulating the malignant phenotype of tumors.

Materials And Methods

mRNA expression analysis

Oncomine (http://www.oncomine.org) is a large cancer microarray database and web-based data-mining platform which covers thousands of samples for gene differential expression analysis[15]. The mRNA expression of EPB41L1 was compared between colorectal cancer tissues and normal tissues by Oncomine database, as well as the DNA copy number variations (CNVs). This analysis based on a series of researches about COAD, including TCGA Colorectal 2, Kurashina Colon, and Ki Colon[16, 17]. We also performed a meta-analysis on related colorectal cancer studies to further confirm EPB41L1 expression levels. The student's t-test was performed to assess whether EPB41L1 expressed higher in cancer tissue than in normal tissues. The threshold of p-value was set to 1E-4 and the threshold of fold change was set to 2 as conditions for screening analyses.

Meanwhile, two microarray datasets GSE41328 and GSE81558, which include 10 and 23 colorectal cancer patients, respectively, were analyzed by GEO2R software for external validation[18]. The normalized expression matrix of microarray data could be directly downloaded from the dataset. The probes were annotated by using the corresponding annotation files from the dataset as well.

Expression of EPB41L1 in various COAD sub-groups

UALCAN (http://ualcan.path.uab.edu) is a comprehensive and interactive web resource that performs to in-depth analyses of TCGA gene expression data by using TCGA level 3 RNA-seq and clinical data. It

allows users to analyze the relative expression of a certain gene across tumor and normal samples, and in various tumor sub-groups based on individual cancer stages, tumor grade, race, body weight and other clinicopathologic features[19].

Prognostic analysis of Differentially Expressed EPB41L1 Gene in COAD patients.

The association of differential expression status of EPB41L1 gene with patient survival was examined using TCGA database information via The Human Protein Atlas (https://www.proteinatlas.org/)[20]. Kaplan–Meier (KM) survival plot were drawn for prognostic values of EPB41L1 gene that had significant up-or downregulation in COAD patients at the mRNA level. Meanwhile, the prognostic analysis by Kaplan–Meier survival curve was obtained by the UALCAN database for validation. Kaplan meier plot showing effect of gene expression on patient survival. Significance of survival impact is measured by log rank test.

EPB41L1 mutation in COAD

The cBioPortal for Cancer Genomics (http://cbioportal .org) is a Web resource that can be used to explore and analyze multidimensional cancer genomics datasets[21, 22]. We used c-BioPortal to analyze EPB41L1 mutation in the CPTAC-2 Prospective whole Exome Sequencing of 110 COAD tumor samples[23]. And the OncoPrint is a graphical summary that can display genetic alterations across a set of tumor samples in EPB41L1.

Co-expression gene prediction Correlated with EPB41L1 in COAD

The LinkedOmics database (http://www.linkedomics.org) is a publicly available portal that contains multi-omics and clinical data for 32 cancer types from the TCGA project and allows users to analyze these data comprehensively[24]. The differentially expressed genes related to EPB41L1 were screened from the TCGA COAD cohort (n = 379) through the LinkFinder analytical module in the LinkedOmics database. And the correlation of results was tested by the Spearman correlation coefficient.

GO and KEGG Pathway Enrichment Analysis

The differentially expressed genes related to EPB41L1 from the LinkedOmics database were annotated by the Database for Annotation, Visualization and Integrated Discovery (DAVID) version 6.8 (https://david.ncifcrf.gov/), an bioinformatics resource provides a functional annotation tools for investigators to understand biological meaning from large genes lists[25, 26]. Gene Ontology (GO) [containing cellular component (CC), biological process (BP), and molecular function (MF)] and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of co-expression genes and the genes in the most highly connected module from MCODE plugin in were also performed by the DAVID tool. The results were evaluated significantly at $P < 0.05$ statistical level verified by Fisher's exact test.

Establishment of interactive network and modules

The interaction between proteins could be sought through the STRING database that is available online at https://string-db.org/, we screened out co-expressed genes with interaction scores greater than 0.4 to establish a protein–protein interactive (PPI) network[27]. Then the PPI network was visualized by Cytoscape analysis software with version 3.7.2[28], in which we could find the densely interconnected protein-interactive regions, and cluster them into hub gene modules with the degree cut-off, haircut on, k-core, node score cut-off, and max depth set as 2, 0.1, 2, 0.2, and 100 through the Molecular Complex Detection (MCODE) version 1.6.1 plugin in Cytoscape to prepare for the next analysis[29].

Statistical analysis

T test was used for differential expression analysis, Log-rank test was used to indicate statistical significance of survival correlation between groups, the differences were considered significant when P 0.05.

Results

Expression of EPB41L1 in Human COAD

We analyzed the transcription level of EPB41L1 in COAD tissues from a series of studies linked to the Oncomine database and found that the mRNA expression of EPB41L1 in COAD tissues was obviously higher compared to normal tissues (P 0.0001), as well as CNVs. As shown in Figures.1a–c, both CNVs and the mRNA expression of EPB41L1 was among the top 5%, although the differences were not more than twofold between COAD tissues and normal tissues. Furthermore, we conducted a pooled meta-analysis including all the 2093 samples of 34 researches from 12 different datasets in Oncomine. The pooled meta-analysis demonstrates that EPB41L1 significantly up-regulates in colorectal cancer (p = 0.009) (Figure. 1e). GEO datasets GSE41328 as test dataset confirmed it (Figs. 1d). Thus, we found a difference in the expression of EPB41L1 between COAD and normal tissues by analyzing data from the Oncomine and GEO databases.

EPB41L1 expression in sub-groups of human COAD

To further prove the specificity of EPB41L1 in COAD, we integrated various clinic COAD patient data by using the TCGA database, for example, pathological stages, lymph nodal metastasis status, patient gender, body weight, patient age or COAD subtype, and to compare the transcription levels of EPB41L1 in each group. The result showed that the pathological tumor stages 3 and 4 COAD patients, compared with normal subjects, maintained higher transcription levels of EPB41L1 (Fig. 2a). Figure 2b showed that EPB41L1 transcriptional expression is increased in COAD patients accompanied by more axillary lymph node metastasis. There were significant differences between male/female COAD patients and normal subjects, and the transcription level of EPB41L1 was higher in male COAD patients (Fig. 2c). The results showed that the expression level of EPB41L1 increased with the increase of COAD patients body weight (Fig. 2d). Figure 2e showed that EPB41L1 transcriptional expression of COAD patients aged 40 to 80 was

higher than that of the normal group. The expression of EPB41L1 in adenocarcinoma subtype was significantly higher than normal group (Fig. 2f).

High expression of EPB41L1 predicts poor prognosis of COAD Patients

We then studied the prognostic value of EPB41L1 in COAD patients by UALCAN and HPA databases. As shown in the KM curve, there was a close relationship between the expression of EPB41L1 and the survival of COAD patients that the high expression of EPB41L1 with a significantly worse prognosis ($P < 0.05$; Figs. 3a-b). These results suggested that EPB41L1 gene may be considered as prognostic biomarker for patients with COAD.

The mutation of EPB41L1 in COAD and its prognosis

The occurrence and prognosis of some tumors are related to gene alterations. Here, we evaluated the frequency of EPB41L1 mutations in 110 sequencing data of COAD patients in the CPTAC-2 Prospective database of cBioPortal. The OncoPrint showed that EPB41L1 is altered in 5 of queried samples, mainly accounted by amplification in 1 case and missense mutation in 4 cases (Fig. 4a). Figure 4b showed that the percentage of samples with a somatic mutation in EPB41L1 is 3.8% and the mutation type all is missense. The mutation sites occurred in FERM_C (D297N), FA(R419C), SAB(S523I) and upstream of FERM_N fragment(G6S). The overall patient survival status and overall disease free status curves analysis of the prognostic value between the EPB41L1 altered group and the unaltered group showed no significance ($P = 0.173$ and $P = 0.573$) (Fig. 4c-d). This result suggested that the poor prognosis of COAD patients is not caused by EPB41L1 mutation.

Coexpression genes correlated with EPB41L1 in COAD

We speculated that the role of EPB41L1 in COAD might be closely related to the function of its neighbor genes in COAD. We used the LinkedOmics database to analyze the co-expressed genes of EPB41L1 in 379 COAD cases. As shown in Fig. 5a, there were 2674 genes represented by dark red dots, having an obviously positive connection with EPB41L1. Conversely, there were 1882 genes, represented by dark green dots, having a notably negative correlation with EPB41L1 (false discovery rate [FDR] < 0.001). Fifty significant gene sets that they were positively and negatively correlated with EPB41L1 were shown in the heat map (Fig. 5b, c).

GO and KEGG analysis of EPB41L1-related co-expressed genes in COAD

The outcomes of GO analysis carried out by DAVID indicated that differentially expressed genes correlated with EPB41L1 in COAD were mainly located in nucleus, spindle microtubule and cytoplasm, where they primarily participated in cell division, mitotic nuclear division, sister chromatid cohesion,

protein ubiquitination involved in ubiquitin-dependent protein catabolic process, cellular response to DNA damage stimulus, DNA replication, chromosome segregation and mitotic spindle organization. They acted as structural constituents in protein binding, ubiquitin protein ligase activity, ligase activity, microtubule binding and proton-transporting ATPase activity, rotational mechanism (Fig. 6a-c). The functions of these EPB41L1-related genes were principally enriched in ubiquitin mediated proteolysis, cell cycle, amino sugar and nucleotide sugar metabolism and fructose and mannose metabolism through the KEGG pathway analysis (Fig. 6d).

Construction of coexpression gene protein–protein interaction (PPI) network

The co-expressed genes were built into a PPI network by using the STRING database (Fig. 7a). Biological network consists of several functional modules in which these sub-units of a molecular complex generally function towards the same biological goal. We performed the module analysis of the PPI network using MCODE version 1.6.1 plugin in Cytoscape. The most highly connected sub-network (cluster rank 1; Score 58.299) highlighted in yellow was obtained from the complex (Fig. 7b), consisting of 68 nodes and 1953 interactions. The KEGG pathway analysis of 68 proteins was further performed by the DAVID database. As the result indicated in Fig. 7c, for the genes identified in the most highly connected sub-network, a total of 6 KEGG pathways were enriched. The top pathway was Cell cycle in which CCNA2 (degree, 106), CDC20 (degree, 113), CCNB2 (degree, 91), CDC45 (degree, 80), PTTG1 (degree, 70), ORC1 (degree, 60), BUB1B (degree, 99), CDC7 (degree, 53), MCM5 (degree, 70), MCM6 (degree, 68), BUB1 (degree, 91), MAD2L1 (degree, 93) were enriched.

Discussion

COAD is the most common diagnosed malignancy among many cancers[1]. And the clinical diagnosis of COAD at an early stage is a severe challenge. While improvements in early detection through screening methods with endoscopy and guaiac fecal occult blood test have decreased the incidence of death from COAD, patients continue to present with advanced stage disease and the accuracy is highly variable[30]. Therefore, it is very urgent to find an effective and reliable biomarker to improve the accuracy of the clinical diagnosis of COAD. With the development of genomics and proteomics, high-throughput technology has been applied to explore biomarker related to the diagnosis and prognosis of cancer patients. An increasing number of prognostic biomarkers have been found in colon cancer, and these biomarkers are helpful to determine the prognosis and provide a theoretical basis for individualized treatment[9].

Protein 4.1N is a member of the 4.1 protein family that is involved in cellular processes such as cell proliferation and migration[31]. Increasing studies have shown that 4.1N coding gene EPB41L1 plays an important role in a variety of cancers. Protein 4.1N was reported to a negative regulator of cell metastasis in breast cancer and involved in the process of cell adhesion, migration and invasion of breast cancer cells[13]. Recent study found that 4.1N protein expression level was significantly decreased during

malignant transformation of epithelial ovarian tumors and 4.1N might be a potential marker for target therapy in epithelial ovarian cancer[14]. 4.1N has been reported to suppress cell proliferation and migration in non-small cell lung cancer cell lines[31]. EPB41L1 has been reported to be down-expressed in KIRC tissues, resulting a poor prognosis and it can be an effective biomarker for the diagnosis of KIRC[32]. Therefore, EPB41L1 is a potential biomarker for the diagnosis and prognosis of various cancer types. However, studies of the the biological roles of EPB41L1 in COAD have not been reported, and its relationship with the prognosis of COAD patients and the potential mechanism that EPB41L1 affects the prognosis of COAD patients remain unknown. Our work aims to explore the biological roles of EPB41L1 in COAD.

The present study first analyzed the expression level and prognostic value of EPB41L1 in COAD with the databases. Here, we screened out available data related to COAD from the public databases to study the function of EPB41L1 on the oncoming, progression, and prognosis of COAD. The result revealed that EPB41L1 is significantly upregulated with the emergence and development of COAD compared with normal tissues from the Oncomine and GEO databases. Furthermore, COAD patients with a high EPB41L1 level in their tumor tissues usually had a poor prognosis compared with that of patients with medium/low levels. This suggested that EPB41L1 could be identified as a potential biomarker of COAD diagnosis. We also queried the mutation of EPB41L1 in COAD through the cBioportal database. Only 5 cases were found in the cBioportal database, indicated that the mutation frequency of EPB41L1 in COAD was very low. Comparing EPB41L1 altered group with the unaltered group, it was found that there was no significant difference in the prognosis between them, which suggested that the poor prognosis caused by EPB41L1 high expression was not due to its mutation. However, this result may be due to the small sample size from database. If there are more mutation samples in the future, more interesting results can be obtained.

The co-expression genes related to EPB41L1 in COAD were analyzed through the Linkedomics database. The GO and KEGG analyses further indicated that EPB41L1 co-expression genes are mainly involved in regulating ubiquitin mediated proteolysis, cell cycle, amino sugar and nucleotide sugar metabolism, and fructose and mannose metabolism, which may be one of the reasons for affecting the prognosis of COAD patients. Cell cycle dysregulation is one of the characteristics of COAD. Abnormal cell cycle regulation has been a critical inducement of the imbalance between cell multiplication and apoptosis, also preventing cell damage[1]. At the same time, abnormal metabolism is another important characteristic of cancer. Cancer cells exhibit distinct metabolic phenotypes, which are essential for supporting high proliferative rates[6]. As we all know that high energetic requirements of tumors are mainly satisfied by enhanced glycolytic rate, whereas biosynthetic need to support the high proliferation rate is provided by enhanced lipogenesis and nucleotide synthesis[33]. The key metabolic pathways along with distinguishing metabolites have been the focus of many cancer studies[6]. Next, we performed the KEGG pathway analyses of the most highly connected sub-network indicated that the genes were enriched in cell cycle, which may be the major reason for its influence on the prognosis of COAD patients. However, the present study provides little direct empirical evidence and is limited.

Conclusion

This study integrated public data and found that EPB41L1 was significantly upregulated in COAD tissues, and a high expression level of EPB41L1was predictive of a poor prognosis. In addition, the KEGG pathway revealed that EPB41L1 co-expression genes are mainly involved in regulating cell cycle and metabolism. To the best of our knowledge, the present study was the first to report on the role of EPB41L1 in the prognosis of COAD and provided novel insight EPB41L1 affects the diagnosis and prognosis of COAD patients. But whether EPB41L1 becomes a potential candidate biomarker for early diagnosis still needs to be verified with a lot of research.

Abbreviations

COAD: Colon adenocarcinoma; EPB41L1: Erythrocyte membrane protein band 4.1 like 1; CRC: Colorectal cancer; CNVs: copy number variations; KM: Kaplan–Meier; TCGA: The Cancer Genome Atlas; GEO: Gene Expression Omnibus; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; PPI: Protein–protein interaction.

Declarations

Acknowledgements

The data of this study is from Oncomine, GEPIA, UALCAN, cBioportal, HPA, GEO and LinkedOmics databases. We are grateful to them for providing the data source for our research.

Authors' contributions

QK and XL designed and conceived the research. JL did the bioinformatics analysis and wrote the manuscript. ZJ and TW revised the article. TL collected the datasets. CD searched the literatures. All authors read and approved the final manuscript.

Funding

The present study was supported by the grants from the National Natural Science Foundation of China (NOs. 81870093 and 81901584); The Training Project for Young Outstanding Teachers of University in Henan Province (2019GGJS016 and 2018GGJS009); Foundation for the Cultivation of Basic Research for the Youth Scholars of Zhengzhou University (JC202043031).

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The datasets used during the current study are available from http://www.oncomine.org, http://ualcan.path.uab.edu, https://www.proteinatlas.org/, http://cbioportal.org, and http://www.linkedomics.org.

Competing interests

The authors declare that they have no competing interests.

References

1. Yang M, Hu C, Cao Y, Liang W, Yang X, Xiao T: **Ursolic Acid Regulates Cell Cycle and Proliferation in Colon Adenocarcinoma by Suppressing Cyclin B1**. *FRONT PHARMACOL* 2021, **11**:622212.
2. Liu T, Li C, Jin L, Li C, Wang L. The Prognostic Value of m6A RNA Methylation Regulators in Colon Adenocarcinoma. Medical science monitor: international medical journal of experimental clinical research. 2019;25:9435–45.
3. Yang C, Zhang Y, Xu X, Li W. Molecular subtypes based on DNA methylation predict prognosis in colon adenocarcinoma patients. Aging. 2019;11(24):11880–92.
4. Ye Y, Gu B, Wang Y, Shen S, Huang W. E2F1-mediated MNX1-AS1-miR-218-5p-SEC61A1 feedback loop contributes to the progression of colon adenocarcinoma. J CELL BIOCHEM. 2019;120(4):6145–53.
5. Pak H, Maghsoudi LH, Soltanian A, Gholami F. Surgical complications in colorectal cancer patients. Annals of Medicine Surgery. 2020;55:13–8.
6. Zhang H, Qiao L, Li X, Wan Y, Yang L, Wang H. Tissue metabolic profiling of lymph node metastasis of colorectal cancer assessed by 1H NMR. ONCOL REP. 2016;36(6):3436–48.
7. Wang X, Zhang L, Li H, Sun W, Zhang H, Lai M. THBS2 is a Potential Prognostic Biomarker in Colorectal Cancer. SCI REP-UK. 2016;6:33366.
8. Koncina E, Haan S, Rauh S, Letellier E. **Prognostic and Predictive Molecular Biomarkers for Colorectal Cancer: Updates and Challenges**. *CANCERS* 2020, 12(2).
9. Xu Y, Zhao J, Dai X, Xie Y, Dong M. High expression of CDH3 predicts a good prognosis for colon adenocarcinoma patients. EXP THER MED. 2019;18(1):841–7.
10. Feng G, Guo K, Yan Q, Ye Y, Shen M, Ruan S, Qiu S. Expression of Protein 4.1 Family in Breast Cancer: Database Mining for 4.1 Family Members in Malignancies. MED SCI MONITOR. 2019;25:3374–89.
11. Sun C, Robb VA, Gutmann DH. Protein 4.1 tumor suppressors: getting a FERM grip on growth regulation. J CELL SCI. 2002;115(Pt 21):3991–4000.
12. Wang Z, Ma B, Li H, Xiao X, Zhou W, Liu F, Zhang B, Zhu M, Yang Q, Zeng Y, et al. Protein 4.1N acts as a potential tumor suppressor linking PP1 to JNK-c-Jun pathway regulation in NSCLC. Oncotarget. 2016;7(1):509–23.

13. Ji Z, Shi X, Liu X, Shi Y, Zhou Q, Liu X, Li L, Ji X, Gao Y, Qi Y, et al. The membrane-cytoskeletal protein 4.1N is involved in the process of cell adhesion, migration and invasion of breast cancer cells. EXP THER MED. 2012;4(4):736–40.
14. Xi C, Ren C, Hu A, Lin J, Yao Q, Wang Y, Gao Z, An X, Liu C. Defective expression of Protein 4.1N is correlated to tumor progression, aggressive behaviors and chemotherapy resistance in epithelial ovarian cancer. GYNECOL ONCOL. 2013;131(3):764–71.
15. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pander A, Chinnaiyan AM. ONCOMINE: A Cancer Microarray Database and Integrated Data-Mining Platform. Neoplasia. 2004;6(1):1–6.
16. Kurashina K, Yamashita Y, Ueno T, Koinuma K, Ohashi J, Horie H, Miyakura Y, Hamada T, Haruta H, Hatanaka H, et al. Chromosome copy number analysis in screening for prognosis-related genomic regions in colorectal carcinoma. CANCER SCI. 2008;99(9):1835–40.
17. Ki DH, Jeung H, Park CH, Kang SH, Lee GY, Lee WS, Kim NK, Chung HC, Rha SY. Whole genome analysis for liver metastasis gene signatures in colorectal cancer. INT J CANCER. 2007;121(9):2005–12.
18. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. 2007;23(14):1846–7.
19. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK, Varambally S. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. Neoplasia. 2017;19(8):649–58.
20. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419.
21. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. SCI SIGNAL. 2013;6(269):l1.
22. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. CANCER DISCOV. 2012;2(5):401–4.
23. Vasaikar S, Huang C, Wang X, Petyuk VA, Savage SR, Wen B, Dou Y, Zhang Y, Shi Z, Arshad OA, et al. Proteogenomic Analysis of Human Colon Cancer Reveals New Therapeutic Opportunities. CELL. 2019;177(4):1035–49.
24. Vasaikar SV, Straub P, Wang J, Zhang B. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. NUCLEIC ACIDS RES. 2018;46(D1):D956–63.
25. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. NAT PROTOC. 2009;4(1):44–57.
26. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. NUCLEIC ACIDS RES. 2009;37(1):1–13.

27. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. NUCLEIC ACIDS RES. 2019;47(D1):D607–13.
28. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. GENOME RES. 2003;13(11):2498–504.
29. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC BIOINFORMATICS. 2003;4:2.
30. Turner MC, Becerra D, Sun Z, Watson J, Leung K, Migaly J, Mantyh CR, Blazer DG. The side of the primary tumor affects overall survival in colon adenocarcinoma: an analysis of the national cancer database. TECH COLOPROCTOL. 2019;23(6):537–44.
31. Yang Q, Zhu M, Wang Z, Li H, Zhou W, Xiao X, Zhang B, Hu W, Liu J. 4.1N is involved in a flotillin-1/β-catenin/Wnt pathway and suppresses cell proliferation and migration in non-small cell lung cancer cell lines. Tumour biology: the journal of the International Society for Oncodevelopmental Biology Medicine. 2016;37(9):12713–23.
32. Liang T, Sang S, Shao Q, Chen C, Deng Z, Wang T, Kang Q. Abnormal expression and prognostic significance of EPB41L1 in kidney renal clear cell carcinoma based on data mining. CANCER CELL INT. 2020;20:356.
33. Vizán P, Alcarraz-Vizán G, Díaz-Moralli S, Solovjeva ON, Frederiks WM, Cascante M. Modulation of pentose phosphate pathway during cell cycle progression in human colon adenocarcinoma cell line HT29. INT J CANCER. 2009;124(12):2789–96.

Figures

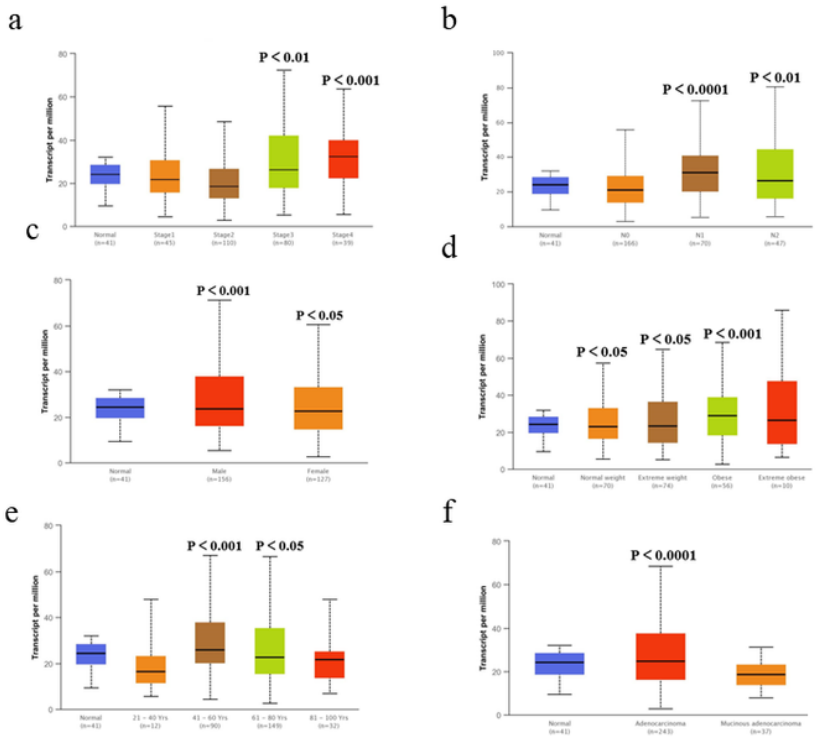


Figure 1

The expression of EPB41L1 in human COAD. Box plot showing EPB41L1 copy number in the TCGA Colorectal 2 and Kurashina Colon based on Oncomine analysis(a-b). Box plot showing associated p value of EPB41L1 mRNA level in the Ki Colon(c). EPB41L1 expression was analyzed in GEO datasets: GSE41328 (d). Meta-analysis of related studies in colorectal cancer demonstrated EPB41L1 were significantly upregulated in COAD tissues compared to normal tissues using the Oncomine (e).

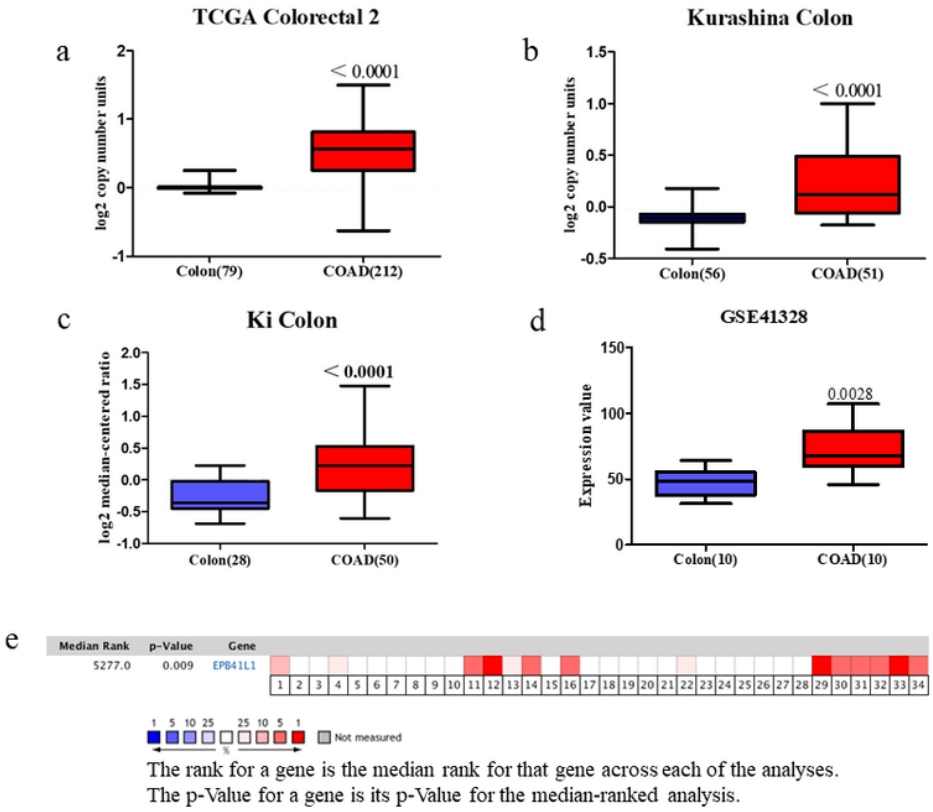


Figure 2

EPB41L1 expression in sub-groups of human COAD. EPB41L1 transcription in sub-groups of patients with COAD stratified based on gender, age, and other criteria (UALCAN). Expression of EPB41L1 (a) in normal individuals or COAD patients in stage 1, 2, 3 or 4; (b) in normal individuals or COAD patients with any lymph node metastasis status N0, N1 or N2; (c) in normal individuals of either gender or male or female COAD patients; (d) in normal individual of any weight or in COAD patients with normal weight, extreme weight, obese, or extreme obese; (e) in normal individuals of any age or COAD patients of age 21 to 40, 41 to 60, 61 to 80, or 81 to 100 years; and (f) in normal individual or adenocarcinoma subtype or mucinous adenocarcinoma subtype in COAD patients.

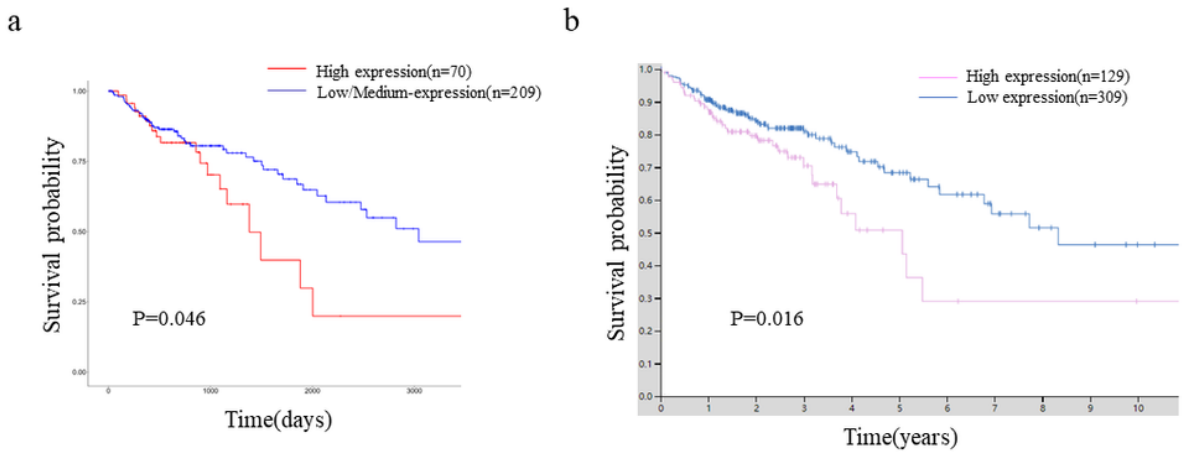


Figure 3

Prognostic analysis of EPB41L1 in COAD patients. Kaplan–Meier (KM) survival curves for survival in COAD patients from the UALCAN (a) and HPA(b) databases. P value is marked to assess the impact of EPB41L1 on the prognosis of COAD.

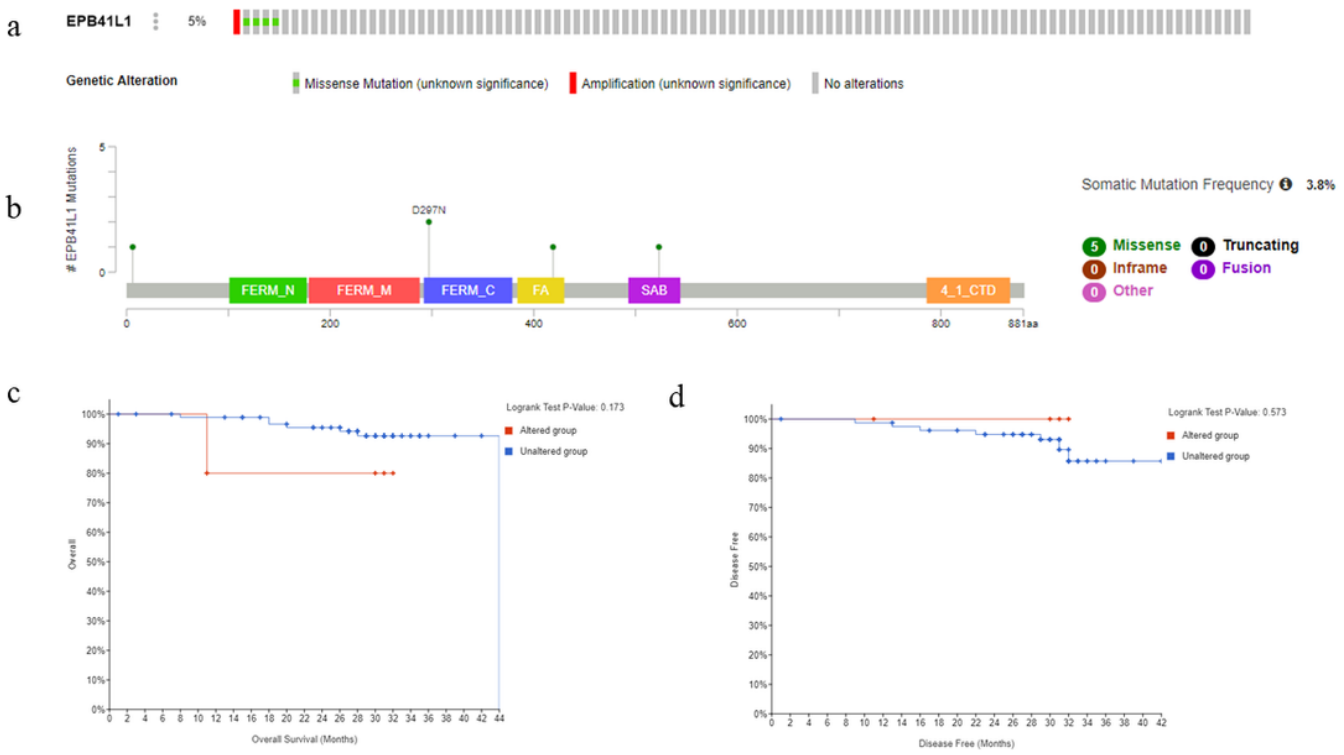


Figure 4

Gene alteration of EPB41L1 in COAD and its prognosis. (a) The OncoPrint schematic showed that gene alteration of EPB41L1 occurred in 5 (5%) of all 110 sequenced cases, including amplification in 1 case (0.90%), and missense mutation in 4 cases (3.63%); (b) the percentage of samples with a somatic mutation in EPB41L1 and the mutation sites; The overall patient survival status(c) and overall disease-free status(d) survival curves between EPB41L1 altered group and unaltered group by c-Bioportal.

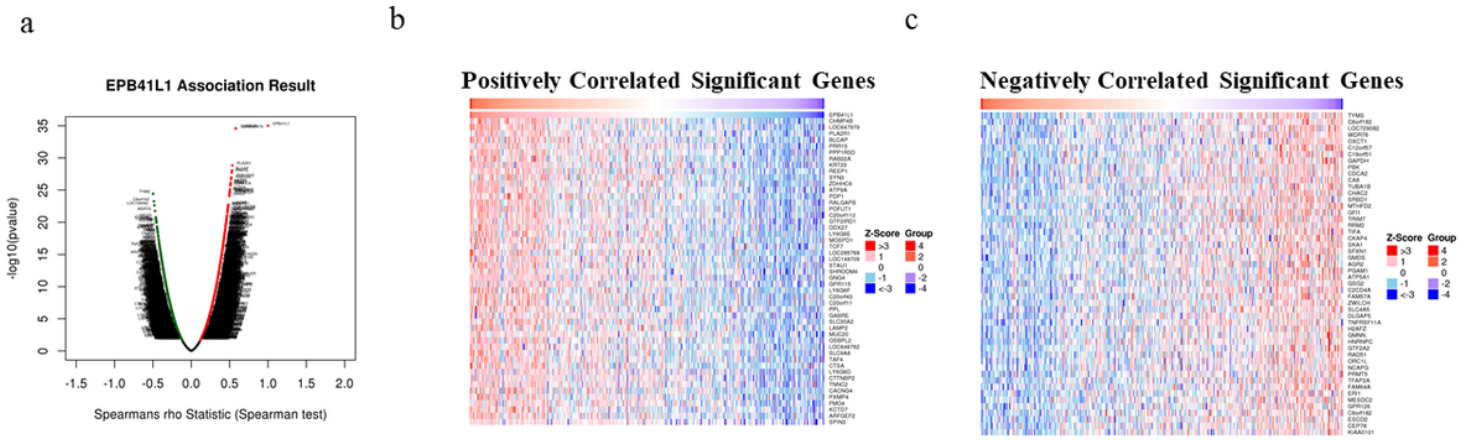


Figure 5

Co-expression genes of EPB41L1 in COAD (LinkedOmics). A Spearman test was used to analyze correlations between EPB41L1 and genes differentially expressed in COAD, red indicates positively correlated genes and green indicates negatively correlated genes (a). Heat maps show the top 50 significant genes positively and negatively correlated with EPB41L1 in COAD (b, c).

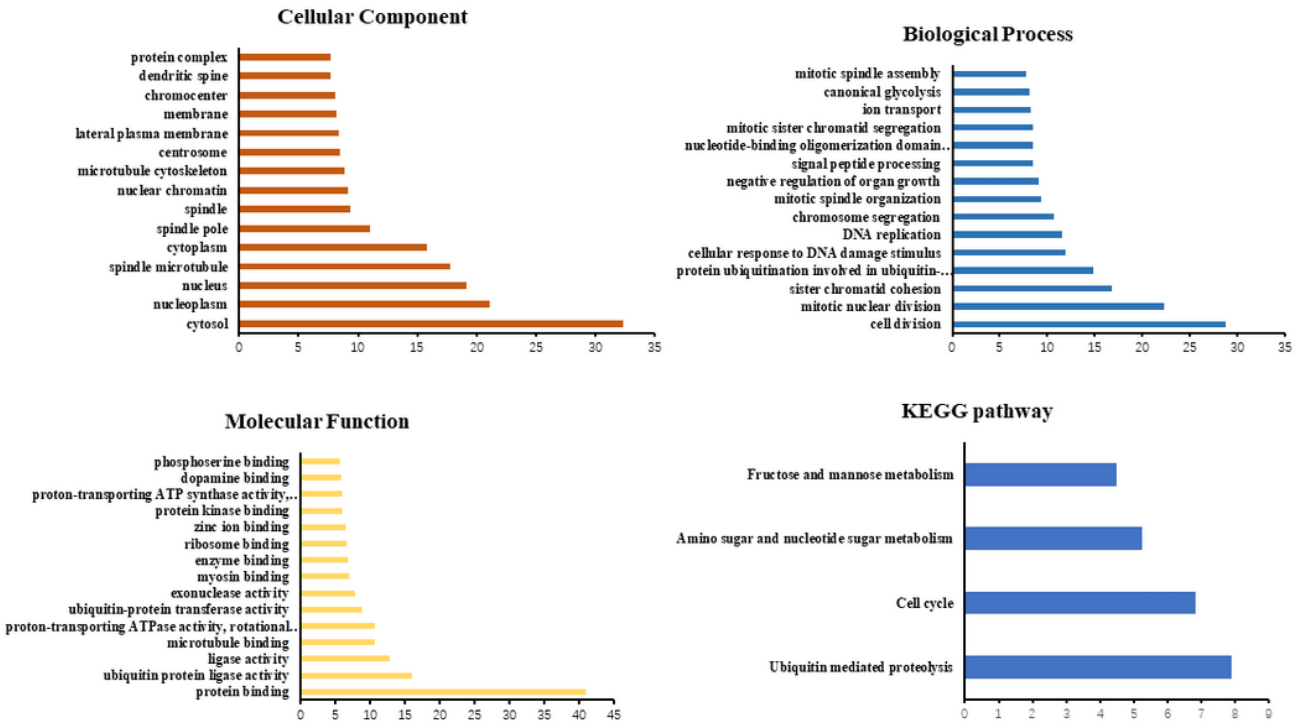
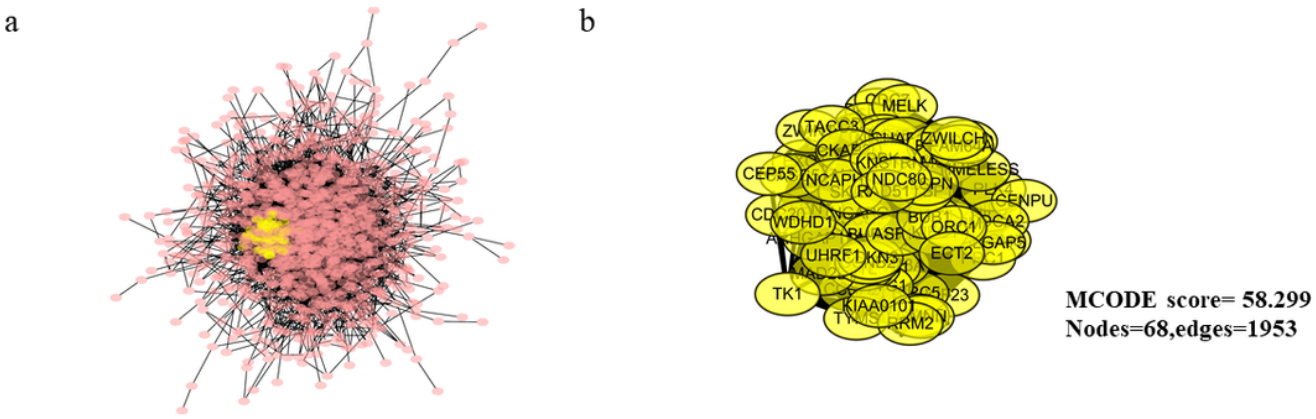


Figure 6

Significantly enriched Gene Ontology (GO) annotations and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of EPB41L1 in COAD. The significantly enriched GO annotations and KEGG pathways of EPB41L1 co-expression genes in COAD were analyzed using DAVID. (a) Cellular components. (b) Biological processes. (c) Molecular functions. (d) KEGG pathway analysis. The x-axis represents the logarithmized p-value, and the y-axis represents the term of GO and KEGG.



c

Category	KEGG Pathway	Court	P-value	FDR
KEGG_PATHWAY	Cell cycle	12	2.5E-14	7.0E-13
KEGG_PATHWAY	Oocyte meiosis	5	3.8E-4	5.3E-3
KEGG_PATHWAY	Progesterone-mediated oocyte maturation	4	2.5E-3	2.4E-2
KEGG_PATHWAY	DNA replication	3	5.8E-3	4.0E-2

Figure 7

PPI network of co-expression genes and the top hub gene module in PPI network. (a) PPI network; (b) The most highly connected module (score = 58.299). (c) KEGG analysis of 68 nodal genes by DAVID.