

Systematic Expression Analysis of Zinc-finger Transcription Factor Family Reveals the Importance of Snai1 in Colorectal Carcinogenesis

Jianxin Li

Affiliated Hospital of Southwest Medical University <https://orcid.org/0000-0001-9863-6493>

Ting Han

The Affiliated Hospital of Southwest Medical University

Xin Wang

The Affiliated Hospital of Southwest Medical University

Yinchun Wang

The Affiliated Hospital of Southwest Medical University

Qingqiang Yang (✉ yangqingqiang121@163.com)

Department of Gastrointestinal Surgery, The Affiliated Hospital of Southwest Medical University, 25 Taiping Street, Luzhou, Sichuan 646000, P.R. China

Research

Keywords: SNAI1, colorectal cancer, EMT, BGN, bioinformatic analysis

Posted Date: November 20th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-111313/v1>

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

[Read Full License](#)

Abstract

Background: Epithelial-mesenchymal transition (EMT) has long been recognized as one of the most important processes involved in cancer cells metastasis. SNAI1, a member of zinc-finger transcriptional factor family, has been identified as an EMT inducer, but its role in human colorectal cancer (CRC) remains largely unclear.

Methods: In the present study, we made a synthesis of the expression of zinc-finger transcription factor family and investigate the prognostic value of SNAI1 in human CRC by using a series of bioinformatic tools.

Results: SNAI1 is the only member of zinc-finger transcription factor family that frequently overexpressed in CRC. In addition, overexpression of SNAI1 was remarkably associated with high tumor stage, N classification and M classification in patients with CRC. Patients with high SNAI1 expression had worse prognosis than those with low SNAI1 expression. Moreover, cox regression analysis showed that SNAI1 was an independent risk factor for overall survival in CRC patients. KEGG pathway analysis showed that SNAI1 enriched in pathways of Focal adhesion, PI3K-Akt signaling pathway and ECM-receptor interaction. GO analysis revealed that SNAI1 participated in processes of extracellular matrix organization, cell adhesion and collagen catabolic process. Eventually, co-expression analysis revealed that overexpression of SNAI1 was significantly correlated with Biglycan (BGN), indicating that they may cooperate in regulating the process of EMT, thereby inducing cancer cells metastasis.

Conclusions: Our results demonstrated that SNAI1 is overexpressed in CRC and acts as an independent molecular marker for prognosis of CRC patients. SNAI1 is a potential biomarker for diagnosis, targeted therapy and prognostic evaluation of CRC.

Introduction

Malignancies are expected to be the leading cause of death and the major obstacle to increasing life expectancy worldwide in the 21st century. There are approximately 18.1 million newly diagnosed cancer cases and 9.6 million cancer deaths worldwide in 2018 [1]. Colorectal cancer (CRC) is one of the most frequently diagnosed malignancy of digestive system, and remains the third most common cause of cancer related death among both men and women [1][2]. The overall 5-year relative survival rate for CRC patients is about 65% through comprehensive treatment based on operation. For the patients diagnosed with localized stage, 5-year relative survival rate is about 90%, but the 5-year survival declines to 12% for patients diagnosed with advanced stage. Unfortunately, most of the patients are at advanced stage at the time of diagnosis [3]. Metastasis and recurrence are the leading causes of death in patients with CRC, approximately 60% of CRC patients have local or distant metastasis prior to receiving treatment, which leads to a significantly worse prognosis and a higher mortality for patients with CRC [4]. The molecular mechanisms underlying the carcinogenesis and metastasis of CRC remain largely unclear, but the evidence available suggested that tumor metastasis, in addition to carcinogenesis, is also a significant

factor which leads to the poor prognosis of patients with CRC. However, the underlying molecular mechanisms are elusive and are worthy of further research.

Epithelial-mesenchymal transition (EMT), an important step by which tumor epithelial cells gradually lose their epithelial features and acquire a mesenchymal features, has long been recognized as one of the most important steps in initiating cancer cell metastasis and is closely correlated with the ability of tumor invasion and metastasis [5]. Under normal conditions, EMT is integral in the process of embryogenesis, wound healing and stem cell behavior, while under pathological conditions, EMT contributes to fibrosis and cancer progression [6]. Tumor cells undergoing EMT are characterized by loss of apical-basal polarity, reduced intercellular contacts and adhesion to extracellular matrix, and enhanced invasive capacity [7]. In addition, changes in molecular phenotypes, such as downregulation of epithelial markers (e.g., E-cadherin and β -catenin) and upregulation of mesenchymal markers (e.g., N-cadherin and Vimentin), are also a feature of these tumor cells [7][8]. Among them, downregulation of E-cadherin is recognized as a significant hallmark of EMT [9]. E-cadherin, encoded by tumor suppressor gene CDH1, is a transmembrane glycoprotein which functions as a major contributor to maintaining tight cell-cell contacts [10]. The loss of E-cadherin is thought to be the leading cause of cancer progression because the majority of solid malignancies are carcinomas that arise from epithelial tissue [11]. Hence, exploring the molecular mechanisms underlying the dysregulation of E-cadherin will help to improve our understanding of EMT process, and provide new theoretical basis for the clinical treatment of malignant tumor patients.

The zinc finger molecule snail family transcriptional repressor 1 (SNAI1), a member of zinc-finger transcription factor family, is an important transcriptional factor which functions as a transcriptional repressor [12]. SNAI1 can specifically binds to the promoter of CDH1 gene through its zinc finger domain and inhibits the gene expression level of CDH1, leading to the loss of E-cadherin expression, and subsequently promotes metastasis of cancer cells [13]. Additionally, SNAI1 has been reported to directly or indirectly inhibit the activity of p53, one of the most important anti-oncogene during carcinogenesis [14]. Previous studies have reported that high expression of SNAI1 was significantly correlated with tumor metastasis, recurrence and worse prognosis in multiple types of malignancies, including CRC [15], lung cancer [16], breast cancer [17], and ovarian cancer [18]. However, the role of SNAI1 in the carcinogenesis and metastasis of CRC cells remains largely unclear and deserves more attention. In the present study, we systematically analyzed the expression level, prognostic value, biological function and mutation of SNAI1 to access its importance in human CRC by using a series of integrated bioinformatics approaches.

Materials And Methods

2.1 Pan-cancer analysis.

The Oncomine database (<https://www.oncomine.org>) is a web-based data mining platform which is designed to collect, standardize, analyze, and deliver cancer transcriptome data to the biomedical research community [19]. We used the Oncomine database to analyze the transcriptome data of zinc-

finger transcription factor family in multiple types of cancer tissues and respective normal tissues, the cut-off conditions were set to P -value $< 1E-4$ and fold change > 2 . Subsequently, the mRNA expression levels of zinc-finger transcription factor family in multiple types of cancer were analyzed by using the UALCAN database (<http://ualcan.path.uab.edu>), an online resource for exploring gene expression variations and survival associations in 31 cancer types [20]. Ultimately, the prognostic values of zinc-finger transcription factor family in patients with CRC were evaluated via the Gene Expression Profiling Interactive Analysis (GEPIA, <http://gepia.cancer-pku.cn>), a web-based tool to deliver fast and customizable functionalities based on The Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx) projects [21]. The threshold of statistical significance was set to logrank P -value < 0.05 .

2.2 Data collection from TCGA database.

To further explore the expression level and prognostic value of SNAI1 in human CRC, the level 3 transcriptome profiles (including 568 cancer samples and 44 normal samples) and clinical information (including 548 cases) of patients with CRC were downloaded from the TCGA database (<https://portal.gdc.cancer.gov>). Relationships between SNAI1 expression level and clinical parameters of CRC patients were analyzed and visualized by using R software (v 3.6.2, <http://www.r-project.org/>). Subsequently, we used Cox proportional-hazards model to evaluate the prognostic value of SNAI1 in patients with CRC. The cut-off condition was set to adjusted P -value < 0.05 .

2.3 Mutations of SNAI1 analysis.

The cBio Cancer Genomics Portal database (cBioPortal, <http://www.cbioportal.org>) is an online platform for interactive exploration of multidimensional cancer genomics datasets, which providing access to data from over 5,000 tumor samples from 20 cancer studies [22]. The Catalogue of Somatic Mutations in Cancer database (COSMIC, <http://cancer.sanger.ac.uk>) is an expert-curated database for exploring coding mutations across all types of human cancer disease [23][24]. We used these two databases to investigate the coding mutations in SNAI1 in CRC.

2.4 Validation using the Human Protein Atlas database.

The Human Protein Atlas (HPA, <https://www.proteinatlas.org/>) is an open platform which provides the information of all the human proteins in cells, tissues and organs by using integration of various omics technologies, including antibody-based imaging, mass spectrometry-based proteomics, transcriptomics and systems biology. We used the immunohistochemistry data from the HPA database to compare the differential protein expression of SNAI1 between CRC tissues and normal tissues.

2.5 Functional and KEGG pathway enrichment analysis.

To explore the biology function of SNAI1, the genes positively co-expressed with SNAI1 in human CRC were identified by using GEPIA and cBioPortal databases, respectively. Subsequently, the genes that appeared in the top 200 results of both GEPIA and cBioPortal databases were identified as SNAI1 co-

expressed genes by using the intersect function of Venn Diagram (<http://bioinformatics.psb.ugent.be/webtools/Venn/>). Ultimately, Gene Ontology (GO) analysis including biological process (BP), cellular component (CC), and molecular function (MF) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of SNAI1 co-expressed genes were performed by using the Database for Annotation, Visualization and Integrated Discovery database (DAVID, <https://david.ncifcrf.gov/>), a web-based bioinformatics resource which provides functional interpretation of large lists of genes derived from genomic studies [25][26]. Only the enrichment terms with P -value < 0.05 were considered statistical significance.

2.6 Gene co-expression analysis.

In the present study, a protein-protein interactions (PPI) network of SNAI1 co-expressed genes was generated by using the Search Tool for the Retrieval of Interacting Genes database (STRING, <http://string-db.org>) [27], and was visualized via Cytoscape software (v 3.7.2, <http://www.cytoscape.org/>) [28]. Subsequently, to extract core genes which significantly co-expressed with SNAI1, the PPI network was analyzed by using the CytoHubba plugin in Cytoscape software, and the top 10 genes were identified by using the Degree, Closeness and Betweenness algorithm, respectively. The genes that simultaneously appeared in all of three datasets were identified as core genes by using the intersect function of Venn Diagram, and the expression heat map of core genes was verified by UCSC Cancer Genomics Browser (<https://genome-cancer.ucsc.edu/>) [29]. We used GEPIA database to evaluate the prognostic values of core co-expressed genes in human CRC and the correlations between SNAI1 and core co-expressed genes. Eventually, relationships between the expression levels of core co-expressed genes and the clinical parameters of CRC patients were analyzed and visualized by using the R software. A P -value < 0.05 was considered statistically significant.

2.7 Statistical analysis

Most of the statistical analyses were performed by using the bioinformatic platforms mentioned above, and the rest of statistical analyses were performed via the R software (v 3.6.2). Significance of mRNA expression difference in Oncomine and UALCAN databases were estimated by Student's t-test. Survival analysis in GEPIA database was performed by Kaplan–Meier method and estimated by Log-rank test. The relationships between mRNA expression level and clinical parameters were evaluated by Wilcoxon test. Clinical parameters related to overall survival in patients with CRC were analyzed by using Cox regression. Fisher's exact test was applied to evaluate the significant GO terms and KEGG enrichment pathways. The correlations of gene expression were evaluated by Spearman correlation coefficient. P -value < 0.05 was considered to be statistically significant.

Results

3.1 Aberrantly overexpression of SNAI1 in human colorectal cancer.

The results of OncoPrint analysis showed that gene expression profiles of zinc-finger transcription factor family were dysregulated in multiple cancer types (Fig. 1A). In addition, expression profile of SNAI1 was obviously elevated in CRC tissues compared to normal tissues across multiple CRC datasets (Fig. 2). Subsequently, mRNA expression levels analyzed by UALCAN database further confirmed that SNAI1 was significantly overexpressed in CRC tissues compared with normal tissues (Fig. 1B). More importantly, survival analysis revealed that aberrantly expression of SNAI1, but not of SNAI2, SNAI3 and ZEB1, was significantly associated with poor overall survival in patients with CRC (Fig. 1C).

3.2 Associations between SNAI1 expression and clinical parameters in patients with colorectal cancer.

Clinical parameters of CRC patients, including survival status, survival time, age, gender, tumor stage, tumor, node and metastasis classification (TNM) were obtained from TCGA database. As shown in Fig. 3A-H, expression of SNAI1 was notably elevated in CRC tissues compared to normal tissues. In addition, SNAI1 expression was also significantly correlated with clinical stage ($P=0.027$), N classification ($P=0.003$) and M classification ($P=0.02$). However, no significant association was found between SNAI1 expression and other clinical variables including gender ($P=0.585$), age ($P=0.777$) and T classification ($P=0.092$). Subsequently, the prognostic value of SNAI1 in patients with CRC was evaluated by Cox proportional-hazards model. As shown in Fig. 4A&B, the univariate analysis revealed that overexpression of SNAI1 was significantly correlated with poor overall survival in human CRC [hazard ratio (HR): 1.133; 95%CI: 1.053–1.219; $P<0.001$]. Other clinical parameters, including age, stage, T, N, and M classification, are also closely related to poor survival. Multivariate analysis indicated that overexpression of SNAI1 remains as an independent prognostic factor for overall survival [HR: 1.110; 95%CI: 1.017–1.212; $P=0.020$]. Other clinical parameters remained as an independent prognostic factor for overall survival including age and T classification. The results above suggested that over expression of SNAI1 was significantly correlated with invasive tumor phenotype and worse prognosis in CRC patients.

3.3 Mutations of SNAI1 in colorectal cancer are rare.

We applied cBioPortal database to analysis the copy number variations (CNVs) and mutations of SNAI1 in human CRC. As shown in Fig. 5A&B, the frequency of genetic alterations is only 8%, most of which is gene amplification, and mutations of SNAI1 in human CRC is very low, only 0.8%. Subsequently, we used COSMIC database to evaluate the mutations of SNAI1 in human CRC. As shown in Fig. 5C&D, the types of mutations in SNAI1 include nonsense mutations, missense mutations, and synonymous mutations, the major type of which is missense mutations, up to 53.21%. The types of nucleotide alterations include C > T, G > A, C > A, C > G and G > C mutations, the largest proportion of which is C > T, up to 38.81%.

3.4 Protein expression of SNAI1 in cancer tissues compared to normal tissues.

We used the immunohistochemistry data from the HPA database to further examine the protein expression of SNAI1 in human CRC. As shown in Fig. 6A-D, protein expression of SNAI1 is mainly

localized in the nucleus. More importantly, the staining of SNAI1 was significantly higher in CRC tissues in relative to normal tissues, confirming our findings at the mRNA expression level.

3.5 KEGG and GO enrichment analysis revealing functional role of SNAI1 in colorectal carcinogenesis.

The top 200 co-expressed genes of SNAI1 in human CRC were obtained by using GEPIA and cBioPortal database, respectively (Fig. 7A&B). A total of 84 common co-expressed genes were identified through the intersect function of Venn Diagram (Fig. 7C). Subsequently, we applied DAVID online database to analysis the biological function of these co-expressed genes. KEGG pathway enrichment analysis revealed that the co-expressed genes of SNAI1 were mainly enriched in Focal adhesion, PI3K-Akt signaling pathway, ECM-receptor interaction, and MicroRNAs in cancer (Fig. 7D). GO function analysis showed that the co-expressed genes were mainly enriched in extracellular matrix organization, cell adhesion, collagen catabolic process, and collagen fibril organization in BP group. In CC group, co-expressed genes were mainly enriched in extracellular space, extracellular exosome, extracellular region, and extracellular matrix. In MF group, co-expressed genes were mainly enriched in protein binding, platelet-derived growth factor, and collagen binding. The top 10 GO terms according to *P*-value in each group are shown in Fig. 7E.

3.6 BGN was positively co-expressed with SNAI1 in human colorectal cancer.

A PPI network consisted of co-expressed genes were constructed by using the STRING online database, and the network was subsequently analyzed by using Cytoscape software (Fig. 8A). The top 10 genes of the PPI network were identified according to different algorithm generated by CytoHubba plug-in, and 5 genes (COL1A1, MMP2, BGN, LOX, and PDGFRB) were identified as core genes (Fig. 8B&C). Hierarchical clustering of the core co-expressed genes was performed through the use of UCSC Cancer Genomics Browser, revealing that these core genes were differentially expressed in most CRC samples (Fig. 8D). Ultimately, prognostic values of these core genes and the correlations between SNAI1 and core genes were evaluated by GEPIA database. Co-expression analysis demonstrated that all of these core genes were positively co-expressed with SNAI1 (Fig. 8E). However, only one of them (BGN) was significantly correlated with both poorer overall survival (HR = 2, *P* = 0.0017) and disease free survival (HR = 2, *P* = 0.0017) in patients with CRC (Fig. 8F). Therefore, BGN was identified as a hub gene which positively co-expressed with SNAI1 in human CRC.

3.7 Associations between BGN expression and clinical parameters in colorectal cancer patients.

Based on information of clinical samples obtained from TCGA database, the relationships between BGN expression pattern and clinical parameters of CRC patients were analyzed and visualized by using R software. As shown in Fig. 9A-H, expression of BGN was significantly elevated in CRC tissues compared to normal tissues. In addition, BGN expression was also remarkably associated with clinical stage (*P* = 0.042), T classification (*P* = 0.007) and N classification (*P* = 0.001). However, no significant association was found between BGN expression and other clinical variables including M classification (*P* = 0.28), gender (*P* = 0.443), and age (*P* = 0.092).

Discussion

The development, invasion and metastasis of tumor is a multistep and complex process. Colorectal cancer is one of the most common malignant tumors of the digestive system, with more than 1.8 million newly diagnosed cases and 800,000 cancer related deaths around the world in 2018 [1]. In addition to tumorigenesis, tumor metastasis seems to be another important cause of cancer related death in patients with CRC [30]. Despite the treatment of cancer has made great progress in recent years, the underlying molecular mechanisms of tumorigenesis and metastasis are complex and worthy of further study. Therefore, to explore the mechanisms of tumor metastasis will help to improve the clinical outcome of patients with CRC.

SNAI1 is an important transcription inhibitory factor, which has been described as the main molecule responsible for regulating the EMT process in multiple cancer types [31][32][33]. Blocking the Snail signaling pathways may partially or completely prevent the process of tumorigenesis, invasion and metastasis. However, the clinical implications of SNAI1 expression as a biomarker in human CRC have not been well understand. Therefore, we applied multiple online databases and R software in the present study to systematically analyze the role of SNAI1 in human CRC. Our study revealed that SNAI1 was expressed at significantly higher levels in CRC samples compared to normal samples, and overexpression of SNAI1 was significantly correlated with advanced tumor stage, lymph node metastasis, and distant metastasis. In agreement with the results of our study, previous studies have reported that the expression level of SNAI1 in colorectal cancer tissues and cell lines was significantly higher than that in normal control [34]. High expression level of SNAI1 was significantly correlated with lymph node metastasis, and SNAI1 was an independent predictor of lymph node metastasis in human CRC [35]. Enforced expression of SNAI1 significantly enhanced the invasive and metastatic ability of CRC cells in vitro and vivo [36]. Interestingly, studies have shown that SNAI1 expression can mediate chemoresistance and radiation resistance in human CRC cells [37][38]. Furthermore, a growing number of studies have reported that high SNAI1 expression was associated with poor overall survival in patients with CRC, and SNAI1 serve as a biomarker for poor prognosis in human CRC [39][40]. Consistent with those reports, our study found that patients with high SNAI1 expression caused decreased overall survival rate compared to those with low SNAI1 expression. Additionally, univariate and multivariate Cox analyses revealed that overexpression of SNAI1 was a potential independent biomarker for poor prognosis in human CRC. Thus, therapeutic targeting of SNAI1 may be a novel strategy to improve outcomes for CRC patients.

To further explore the molecular mechanisms underlying the role of SNAI1 in human CRC, the biological classification of co-expressed genes of SNAI1 was evaluated by DAVID database. Pathway enrichment analysis revealed that co-expressed genes are mainly enriched in Focal adhesion pathway, ECM-receptor interaction, and PI3K-Akt signaling pathway. Focal adhesion pathway has been reported to be responsible for cancer metastasis by weakening the cell–cell and cell–extracellular matrix adhesions [41]. Blocking the signaling cascades of Focal adhesion pathway may be a novel strategy to enhance therapeutic efficacy in colorectal cancer therapy [42]. ECM is the major components of the tumor microenvironment, which affects the biological behaviors of tumor cells [43]. Components of the ECM may be potential

therapeutic targets for reducing physical barriers to systemic treatments in patients with mCRC who receive anti-VEGF therapy [44]. Activation of the PI3K-Akt signaling pathway has been reported to induce the SNAI1-mediated EMT process, thereby facilitating the metastasis of CRC cells in vitro and in vivo [45]. Blocking the PI3K-Akt signaling pathway by using the AKT inhibitor can suppress the SNAI1-induced migration and invasion of CRC cells [46]. Together with our results, these data above indicated that the activation of PI3K-Akt signaling pathway is closely involved in the EMT process mediated by SNAI1, targeting PI3K-Akt signaling pathway via SNAI1 inhibition may be a potential strategy for CRC clinical therapy. Inevitably, there is a limitation of our study that all data and results were based on public databases and online bioinformatic tools, further experimental validation will be needed to verify the biological significance of SNAI1 expression in human CRC in the future.

Co-expression analysis indicated that the expression of BGN was significantly correlated with SNAI1. In line with our study, previous studies have shown that overexpression of BGN was significantly correlated with advanced tumor stage and poor clinical outcome in multiple cancer entities, including CRC [47], gastric cancer [48], and prostate cancer [49]. There are also increasing number of studies reported that BGN promotes the EMT process, thereby enhancing the invasive and metastatic capabilities of cancer cells [50][51]. In addition, BGN has been identified as a potential EMT biomarker, which participates in the integrated TGF β /Snail pathway in CRC cells [52]. These results above suggested that there is a potential link between SNAI1 and BGN expression during the EMT process, targeting these two collaborative genes simultaneously may be an efficient approach for CRC therapy.

Conclusions

In summary, our study performed a systematic analysis to calculate the expression level, prognostic value, biological function and mutation of SNAI1 in human CRC. We confirmed that overexpression of SNAI1 and its partner, BGN, are involved in the process of EMT in human CRC. We speculate that SNAI1 could be a potential biomarker for the early diagnosis, prognostic evaluation, and clinical treatment of CRC.

List Of Abbreviations

BGN, Biglycan; BP, biological process; CC, cellular component; COSMIC, Catalogue of Somatic Mutations in Cancer database; CRC, Colorectal cancer; EMT, Epithelial-mesenchymal transition; GEPIA, Gene Expression Profiling Interactive Analysis; GO, Gene Ontology; GTEx, Genotype-Tissue Expression; SNAI1, Zinc finger molecule snail family transcriptional repressor 1; TCGA, The Cancer Genome Atlas; cBioPortal, cBio Cancer Genomics Portal; HPA, Human Protein Atlas; MF, molecular function; KEGG, Kyoto Encyclopedia of Genes and Genomes; DAVID, Database for Annotation, Visualization and Integrated Discovery; PPI, protein-protein interactions; STRING, Search Tool for the Retrieval of Interacting Genes;

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and/or analyzed during the current study are available in The Cancer Genome Atlas (TCGA) repository (<https://www.cancer.gov/tcga>).

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

No funding was received.

Authors' contributions

JL, TH and QY conceived and designed the study. JL, TH, YW and XW performed the bioinformatics analysis and interpretation of the data. JL drafted the manuscript. QY agreed to be responsible for all aspects of the work to ensure that issues of accuracy or completeness of the study were properly investigated and addressed. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68:394-424.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* 2020;70:7-30.
3. Miller KD, Nogueira L, Mariotto AB, et al. Cancer treatment and survivorship statistics, 2019. *CA Cancer J Clin.* 2019;69:363-385.
4. Siegel RL, Miller KD, Fedewa SA, et al. Colorectal cancer statistics, 2017. *CA Cancer J Clin.* 2017;67:177-193.

5. Cao H, Xu E, Liu H, Wan L, Lai M. Epithelial-mesenchymal transition in colorectal cancer metastasis: A system review. *Pathol Res Pract*. 2015;211:557-569.
6. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol*. 2014;15:178-196.
7. Nieto MA, Huang RY, Jackson RA, Thiery JP. EMT: 2016. *Cell*. 2016;166:21-45.
8. Dongre A, Weinberg RA. New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat Rev Mol Cell Biol*. 2019;20:69-84.
9. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *J Clin Invest*. 2009;119:1420-1428.
10. Canel M, Serrels A, Frame MC, Brunton VG. E-cadherin-integrin crosstalk in cancer invasion and metastasis. *J Cell Sci*. 2013;126:393-401.
11. Wong SHM, Fang CM, Chuah LH, Leong CO, Ngai SC. E-cadherin: Its dysregulation in carcinogenesis and clinical implications. *Crit Rev Oncol Hematol*. 2018;121:11-22.
12. Nieto MA. The ins and outs of the epithelial to mesenchymal transition in health and disease. *Annu Rev Cell Dev Biol*. 2011;27:347-376.
13. Sánchez-Tilló E, Liu Y, de Barrios O, et al. EMT-activating transcription factors in cancer: beyond EMT and tumor invasiveness. *Cell Mol Life Sci*. 2012;69:3429-3456.
14. Ni T, Li XY, Lu N, et al. Snail1-dependent p53 repression regulates expansion and activity of tumour-initiating cells in breast cancer. *Nat Cell Biol*. 2016;18:1221-1232.
15. Wang S, Yan S, Zhu S, et al. FOXF1 Induces Epithelial-Mesenchymal Transition in Colorectal Cancer Metastasis by Transcriptionally Activating SNAI1. *Neoplasia*. 2018;20:996-1007.
16. Wang X, Liu R, Zhu W, et al. UDP-glucose accelerates SNAI1 mRNA decay and impairs lung cancer metastasis. *Nature*. 2019;571:127-131.
17. Jia Z, Wang M, Li S, et al. U-box ubiquitin ligase PPIL2 suppresses breast cancer invasion and metastasis by altering cell morphology and promoting SNAI1 ubiquitination and degradation. *Cell Death Dis*. 2018;9:63.
18. Lu ZY, Dong R, Li D, et al. SNAI1 overexpression induces stemness and promotes ovarian cancer cell invasion and metastasis. *Oncol Rep*. 2012;27:1587-1591.
19. Rhodes DR, Kalyana-Sundaram S, Mahavisno V, et al. OncoPrint 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia*. 2007;9:166-180.
20. Chandrashekar DS, Bashel B, Balasubramanya SAH, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia*. 2017;19:649–658.
21. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017;45:W98–W102.
22. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012;2:401-404.

23. Forbes SA, Beare D, Boutselakis H, et al. COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res.* 2017;45:D777-D783.
24. Forbes SA, Beare D, Bindal N, et al. COSMIC: High-Resolution Cancer Genetics Using the Catalogue of Somatic Mutations in Cancer. *Curr Protoc Hum Genet.* 2016;91:10.11.1-10.11.37.
25. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4:44-57.
26. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2009;37:1-13.
27. Franceschini A, Szklarczyk D, Frankild S, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 2013;41(Database issue):D808-D815.
28. Kohl M, Wiese S, Warscheid B. Cytoscape: software for visualization and analysis of biological networks. *Methods Mol Biol.* 2011;696:291-303.
29. Haeussler M, Zweig AS, Tyner C, et al. The UCSC Genome Browser database: 2019 update. *Nucleic Acids Res.* 2019;47:D853-D858.
30. Tang L, Zhao P, Kong D. Muscleblind-like 1 destabilizes Snail mRNA and suppresses the metastasis of colorectal cancer cells via the Snail/E-cadherin axis. *Int J Oncol.* 2019;54:955-965.
31. Hemavathy K, Ashraf SI, Ip YT. Snail/slug family of repressors: slowly going into the fast lane of development and cancer. *Gene.* 2000;257:1-12.
32. Deep G, Jain AK, Ramteke A, et al. SNAI1 is critical for the aggressiveness of prostate cancer cells with low E-cadherin. *Mol Cancer.* 2014;13:37.
33. Ye T, Xu J, Du L, Mo W, Liang Y, Xia J. Downregulation of UBAP2L Inhibits the Epithelial-Mesenchymal Transition via SNAI1 Regulation in Hepatocellular Carcinoma Cells. *Cell Physiol Biochem.* 2017;41:1584-1595.
34. Zheng P, Meng HM, Gao WZ, et al. Snail as a key regulator of PRL-3 gene in colorectal cancer. *Cancer Biol Ther.* 2011;12:742-749.
35. Fan XJ, Wan XB, Yang ZL, et al. Snail promotes lymph node metastasis and Twist enhances tumor deposit formation through epithelial-mesenchymal transition in colorectal cancer. *Hum Pathol.* 2013;44:173-180.
36. Fan F, Samuel S, Evans KW, et al. Overexpression of snail induces epithelial-mesenchymal transition and a cancer stem cell-like phenotype in human colorectal cancer cells. *Cancer Med.* 2012;1:5-16.
37. Hoshino H, Miyoshi N, Nagai K, et al. Epithelial-mesenchymal transition with expression of SNAI1-induced chemoresistance in colorectal cancer. *Biochem Biophys Res Commun.* 2009;390:1061-1065.
38. Zhu Y, Wang C, Becker SA, et al. miR-145 Antagonizes SNAI1-Mediated Stemness and Radiation Resistance in Colorectal Cancer. *Mol Ther.* 2018;26:744-754.
39. Kim YH, Kim G, Kwon CI, Kim JW, Park PW, Hahm KB. TWIST1 and SNAI1 as markers of poor prognosis in human colorectal cancer are associated with the expression of ALDH1 and TGF- β 1. *Oncol Rep.* 2014;31:1380-1388.

40. Kwon CH, Park HJ, Choi JH, et al. Snail and serpinA1 promote tumor progression and predict prognosis in colorectal cancer. *Oncotarget*. 2015;6:20312-20326.
41. Béraud-Dufour S, Devader C, Massa F, Roulot M, Coppola T, Mazella J. Focal Adhesion Kinase-Dependent Role of the Soluble Form of Neurotensin Receptor-3/Sortilin in Colorectal Cancer Cell Dissociation. *Int J Mol Sci*. 2016;17:1860.
42. Jeong KY. Inhibiting focal adhesion kinase: A potential target for enhancing therapeutic efficacy in colorectal cancer therapy. *World J Gastrointest Oncol*. 2018;10:290-292.
43. Fang M, Yuan J, Peng C, Li Y. Collagen as a double-edged sword in tumor progression. *Tumour Biol*. 2014;35:2871-2882.
44. Rahbari NN, Kedrin D, Incio J, et al. Anti-VEGF therapy induces ECM remodeling and mechanical barriers to therapy in colorectal cancer liver metastases. *Sci Transl Med*. 2016;8:360ra135.
45. Zhao GX, Xu YY, Weng SQ, et al. CAPS1 promotes colorectal cancer metastasis via Snail mediated epithelial mesenchymal transformation. *Oncogene*. 2019;38:4574-4589.
46. Wang F, Ruan XJ, Zhang HY. BDE-99 (2,2',4,4',5-pentabromodiphenyl ether) triggers epithelial-mesenchymal transition in colorectal cancer cells via PI3K/Akt/Snail signaling pathway. *Tumori*. 2015;101:238-245.
47. Gu X, Ma Y, Xiao J, et al. Up-regulated biglycan expression correlates with the malignancy in human colorectal cancers. *Clin Exp Med*. 2012;12:195-199.
48. Wang B, Li GX, Zhang SG, et al. Biglycan expression correlates with aggressiveness and poor prognosis of gastric cancer. *Exp Biol Med (Maywood)*. 2011;236:1247-1253.
49. Jacobsen F, Kraft J, Schroeder C, et al. Up-regulation of Biglycan is Associated with Poor Prognosis and PTEN Deletion in Patients with Prostate Cancer. *Neoplasia*. 2017;19:707-715.
50. Schulz GB, Grimm T, Sers C, et al. Prognostic value and association with epithelial-mesenchymal transition and molecular subtypes of the proteoglycan biglycan in advanced bladder cancer. *Urol Oncol*. 2019;37:530.e9-530.e18.
51. Thakur AK, Nigri J, Lac S, et al. TAp73 loss favors Smad-independent TGF- β signaling that drives EMT in pancreatic ductal adenocarcinoma. *Cell Death Differ*. 2016;23:1358-1370.
52. Li H, Zhong A, Li S, et al. The integrated pathway of TGF β /Snail with TNF α /NF κ B may facilitate the tumor-stroma interaction in the EMT process and colorectal cancer prognosis. *Sci Rep*. 2017;7:4915.

Figures

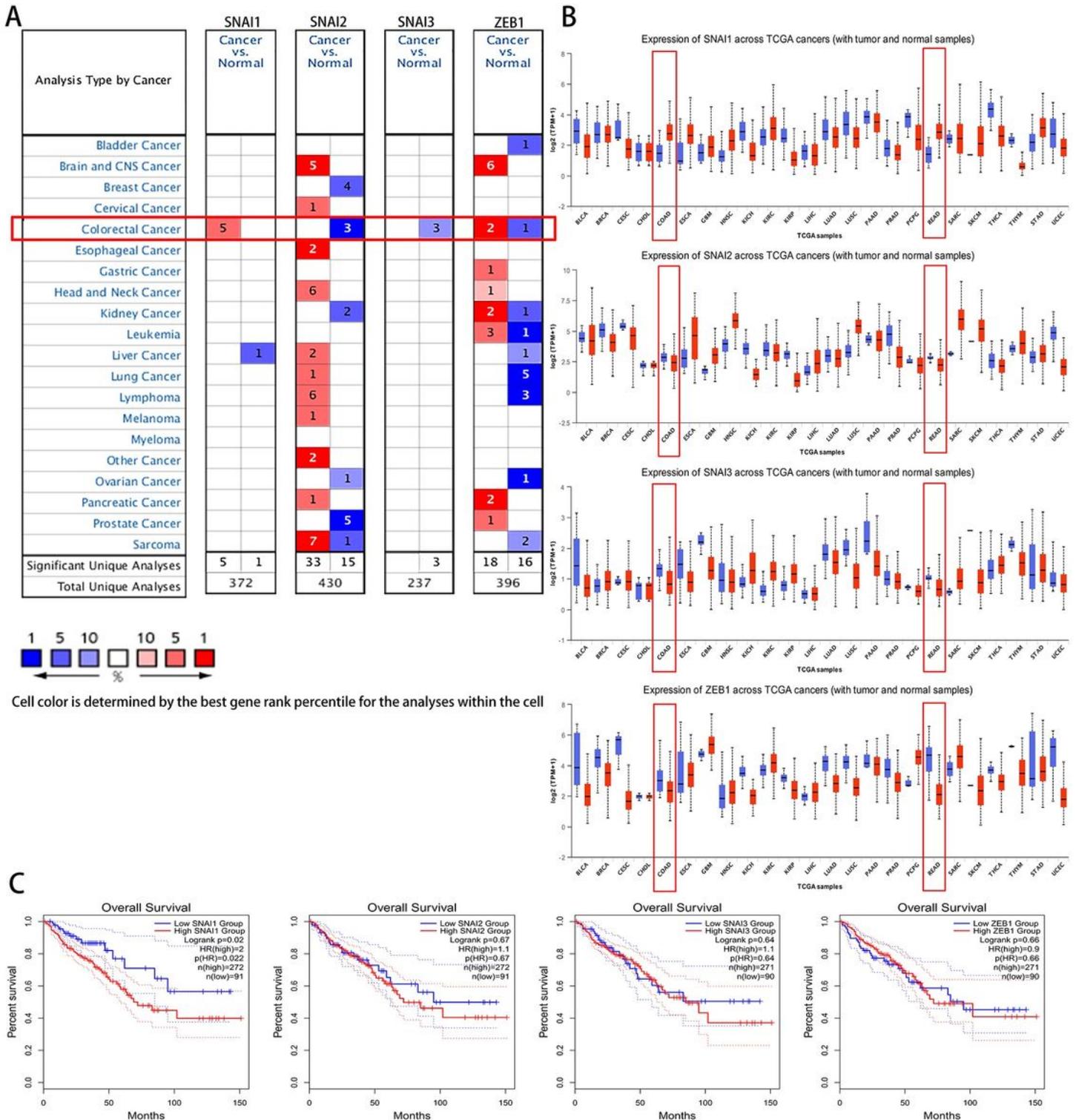
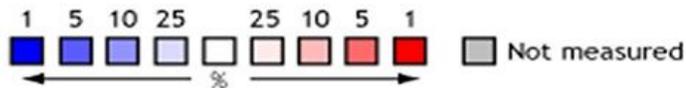
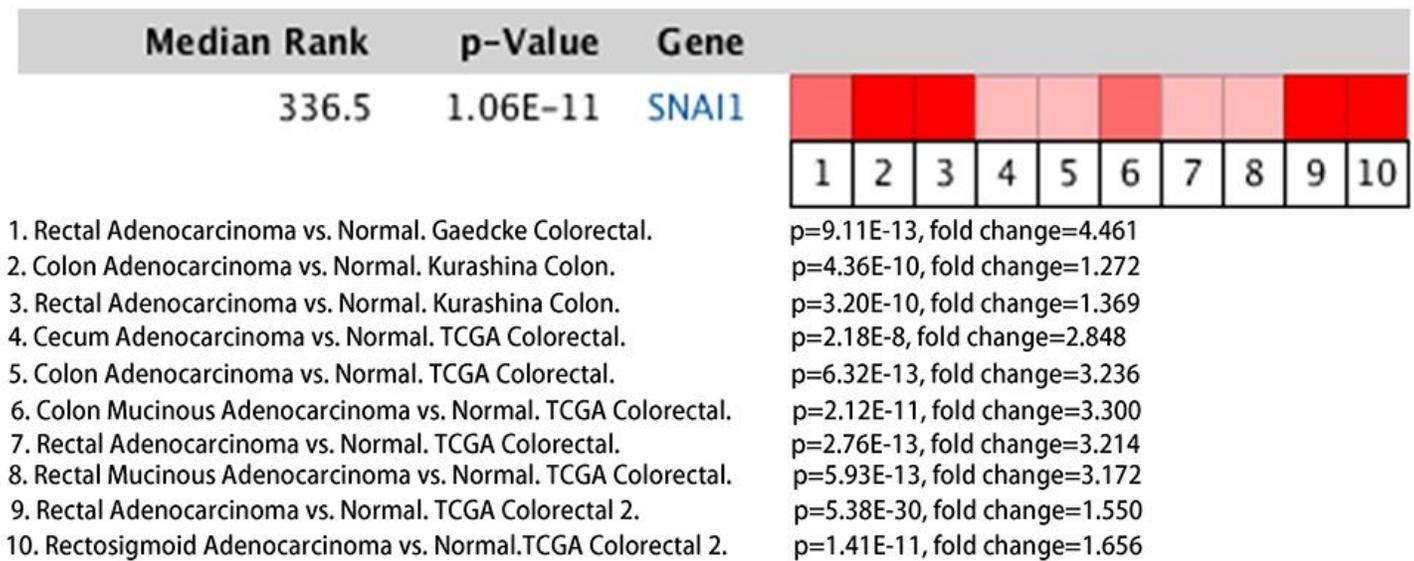


Figure 1

Identification the role of SNAI1 mRNA in human colorectal cancer. (A) The numbers of datasets with statistically significant mRNA overexpression (red) or downexpression (blue) of SNAI1, SNAI2, SNAI3 and ZEB1 in cancer samples compared with noncancerous samples. The threshold was defined with the following parameters: P-value of $1E-4$, fold change of 2 and gene ranking of 10%. (B) The box plot of zinc-finger transcription factor family gene expression in colorectal cancer samples (red) compared with

noncancerous samples (blue). (C) Prognostic values of zinc-finger transcription factor family gene expression in patients with colorectal cancer. HR, hazard ratio.

Comparison of SNAI1 Across 10 Analyses



The rank for a gene is the median rank for that gene across each of the analyses.
 The p-Value for a gene is its p-Value for the median-ranked analysis.

Figure 2

Heat map of SNAI1 mRNA expression in multiple CRC tissues vs. normal tissues.

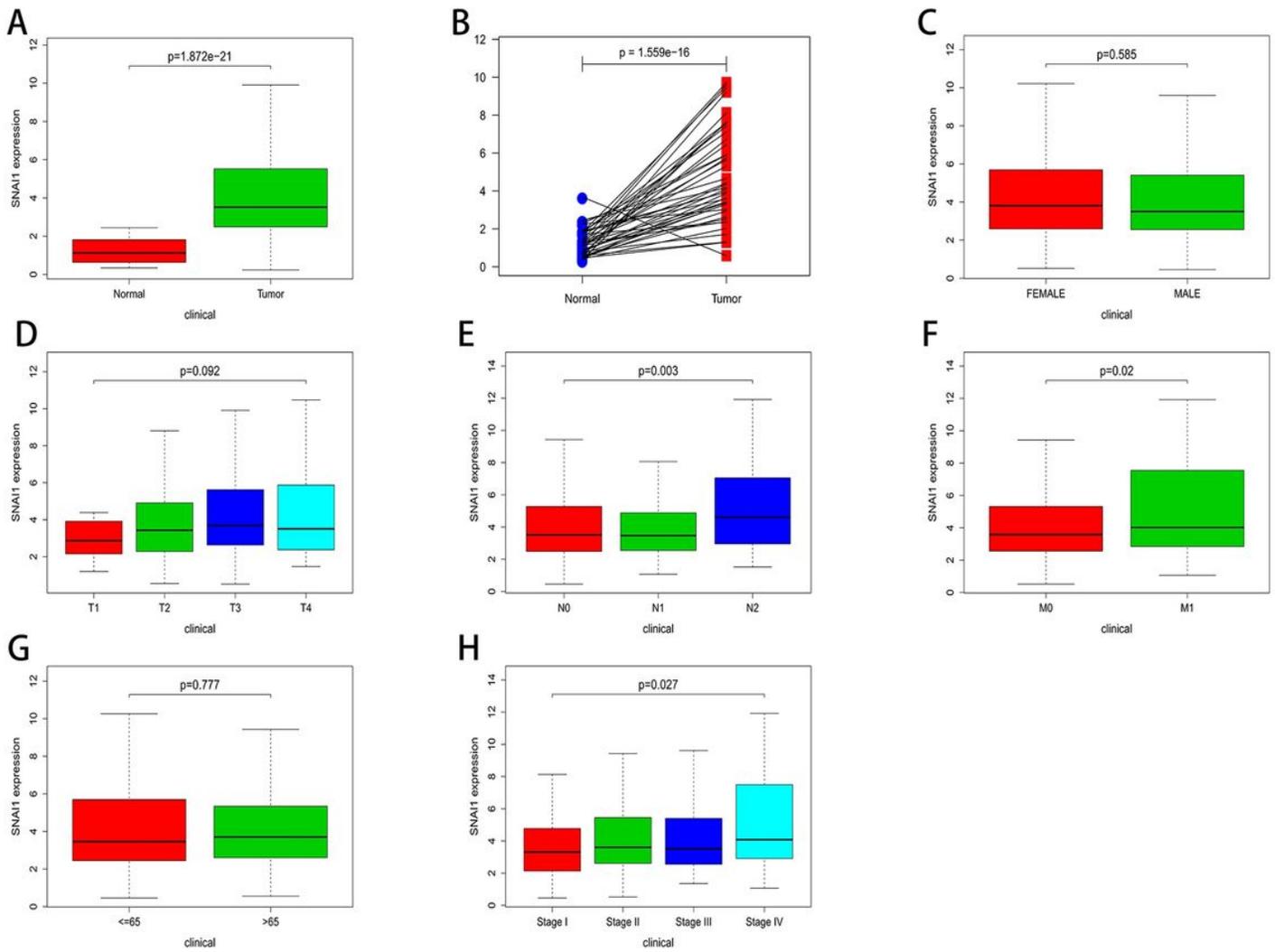


Figure 3

Associations of SNAI1 expression with clinical parameters in colorectal cancer. (A) Sample type (cancer vs. normal). (B) Sample type (cancer vs. paired normal). (C) Gender (female vs. male). (D) Tumor size (T classification). (E) Lymph node metastasis (N classification). (F) Distant metastases (M classification). (G) Age (<=65 vs. >65). (H) Clinical stage.

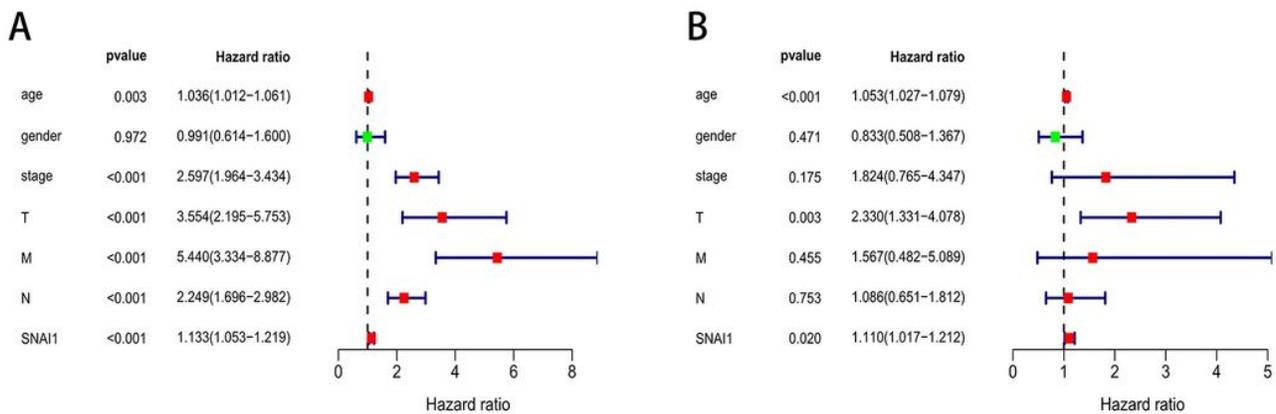


Figure 4

Prognostic value of SNAI1 in human CRC by using Cox proportional-hazards model. (A) Univariate analysis of the relationship between SNAI1 expression and overall survival among colorectal cancer patients. (B) Multivariate analysis of the relationship between SNAI1 expression and overall survival among colorectal cancer patients.

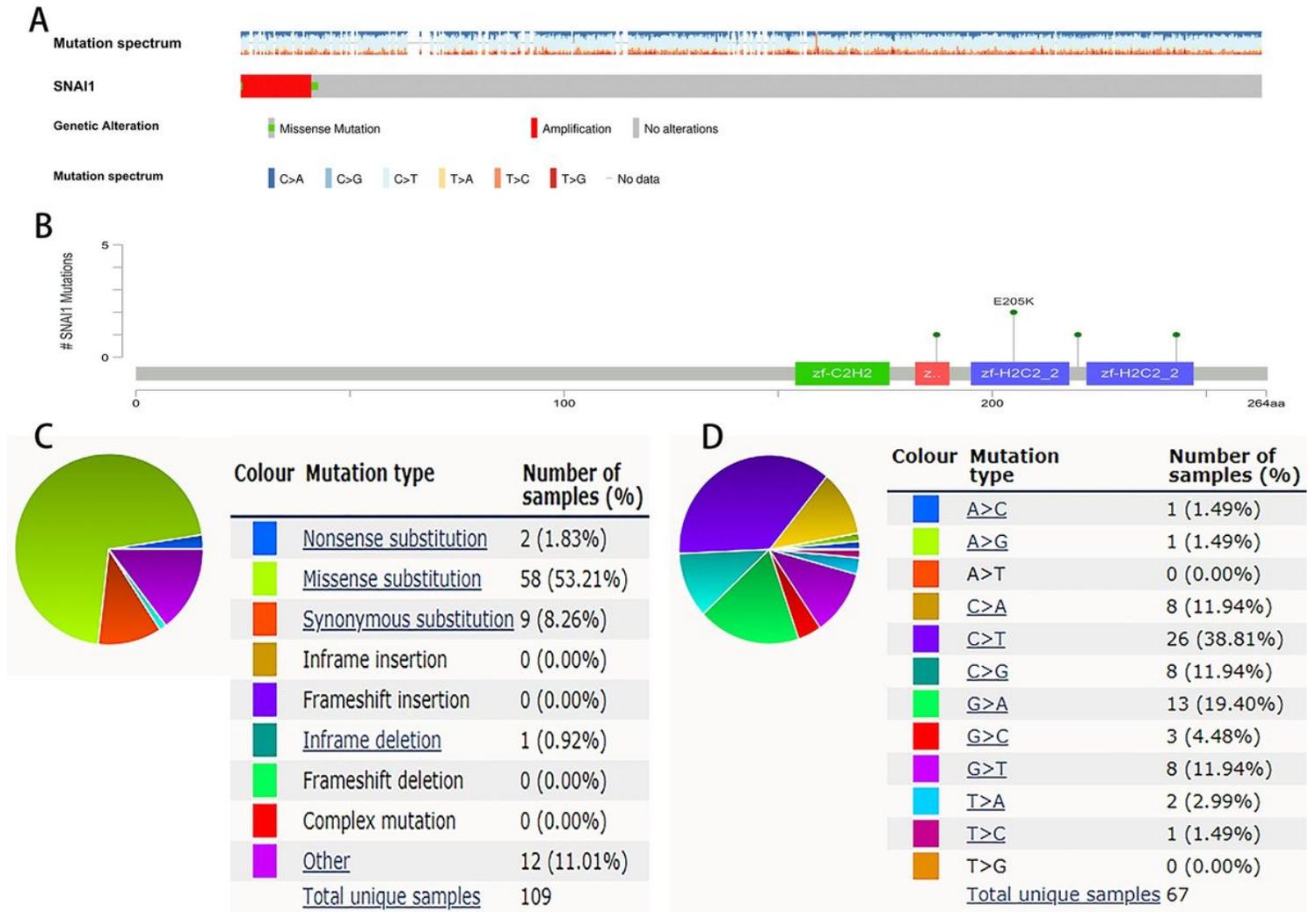


Figure 5

Mutations of SNAI1 in human colorectal cancer. (A & B) Schematic representation of SNAI1 mutations by using the cBioPortal database. (C & D) The percentages of mutation types of SNAI1 in human colorectal cancer were presented in a pie chart generated by using the COSMIC database.

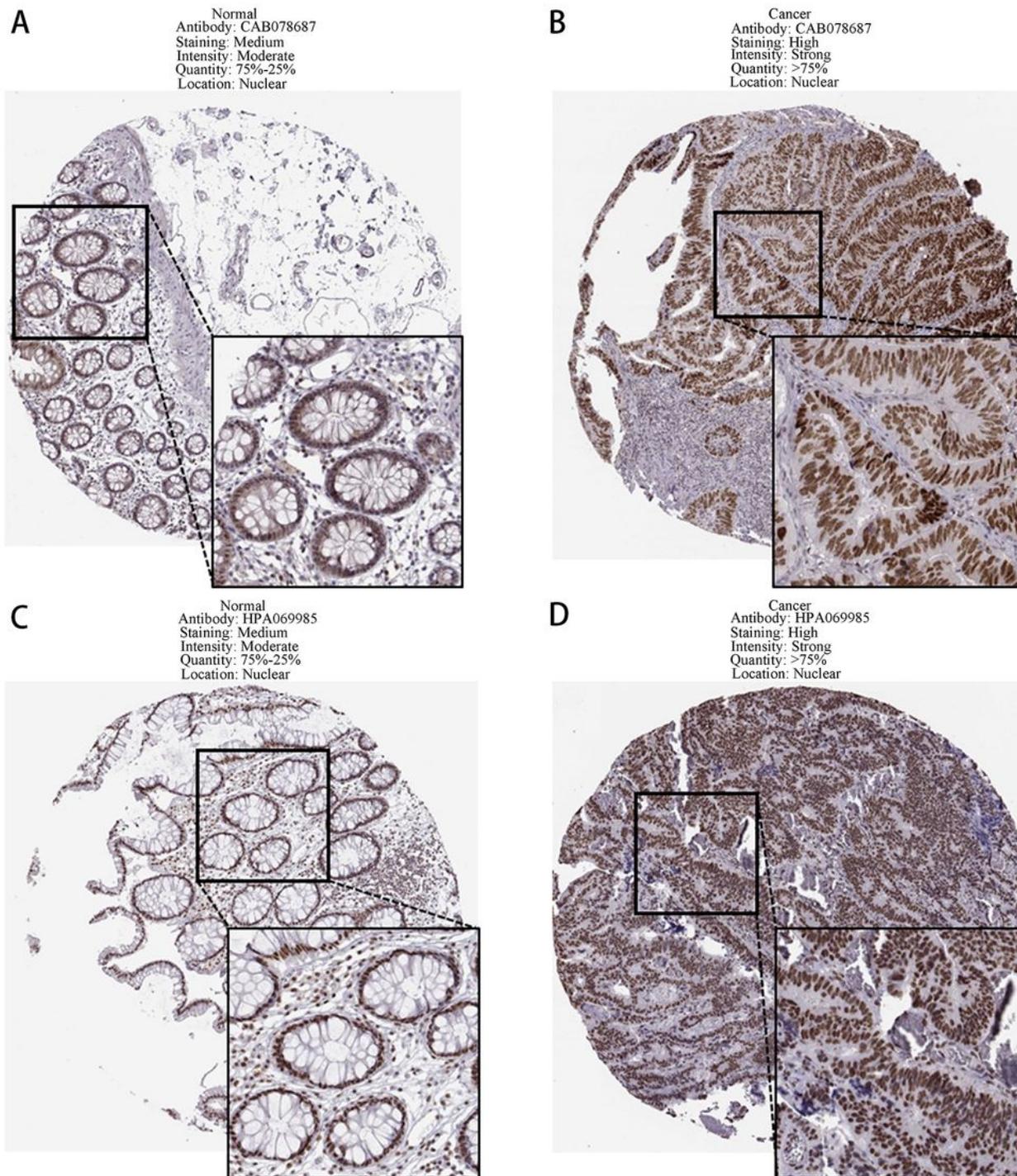


Figure 6

Validation of protein expression of SAN11 in colorectal cancer (B & D) and noncancerous (A & C) tissues by using different antibody through the Human Protein Atlas database.

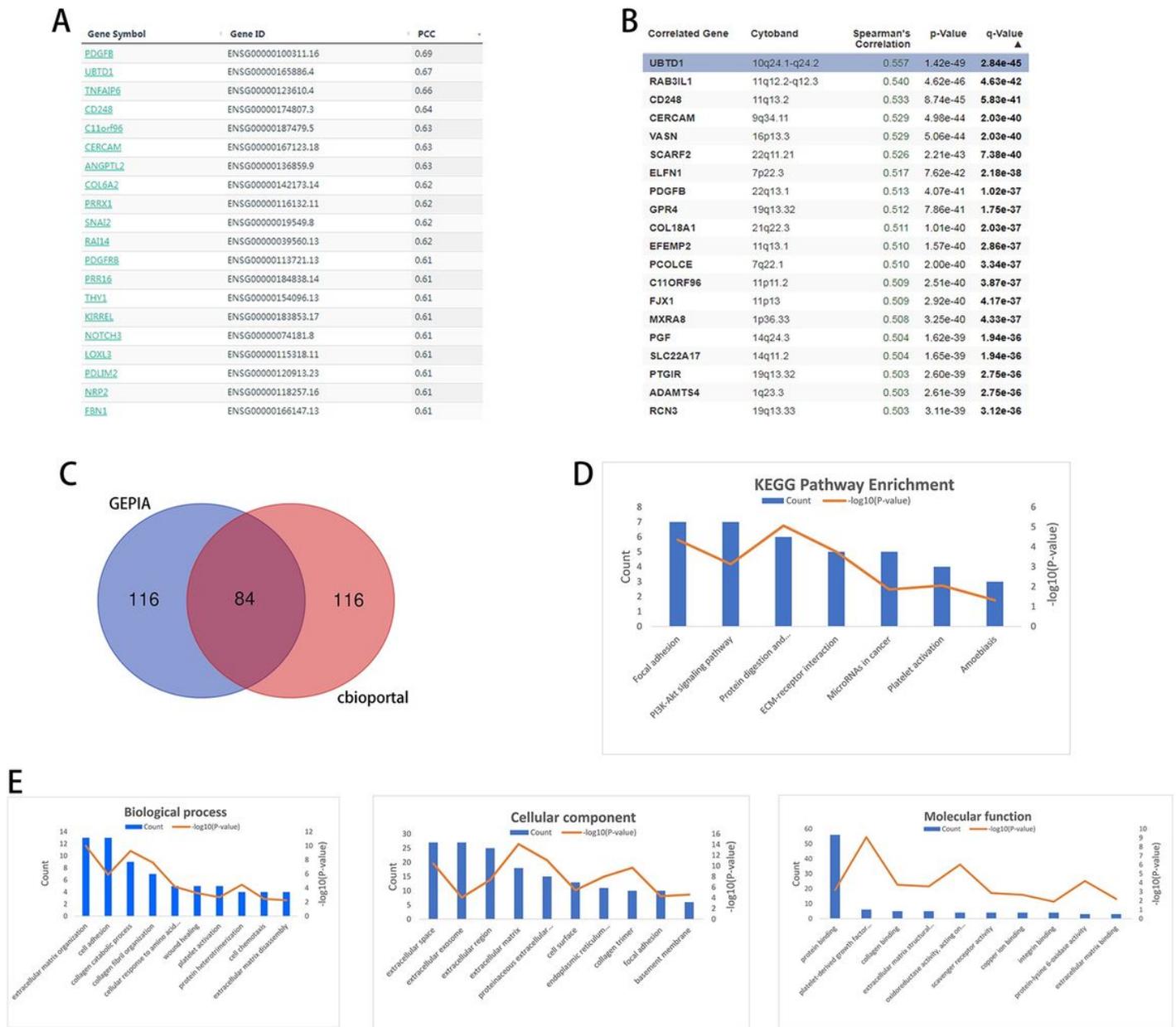


Figure 7

KEGG and GO enrichment analysis of genes which co-expressed with SNAI1 in colorectal cancer samples. (A) The top 200 genes in colorectal cancer positively correlated with SNAI1 transcript level based on the GEPIA database. (B) The top 200 genes in colorectal cancer positively correlated with SNAI1 transcript level based on the cBioPortal database. (C) Venn diagram indicates the intersection of top positively associated genes between the GEPIA database and the cBioPortal database. (D) KEGG enrichment analysis of co-expressed genes with SNAI1. (E) GO enrichment analysis of co-expressed genes with SNAI1.

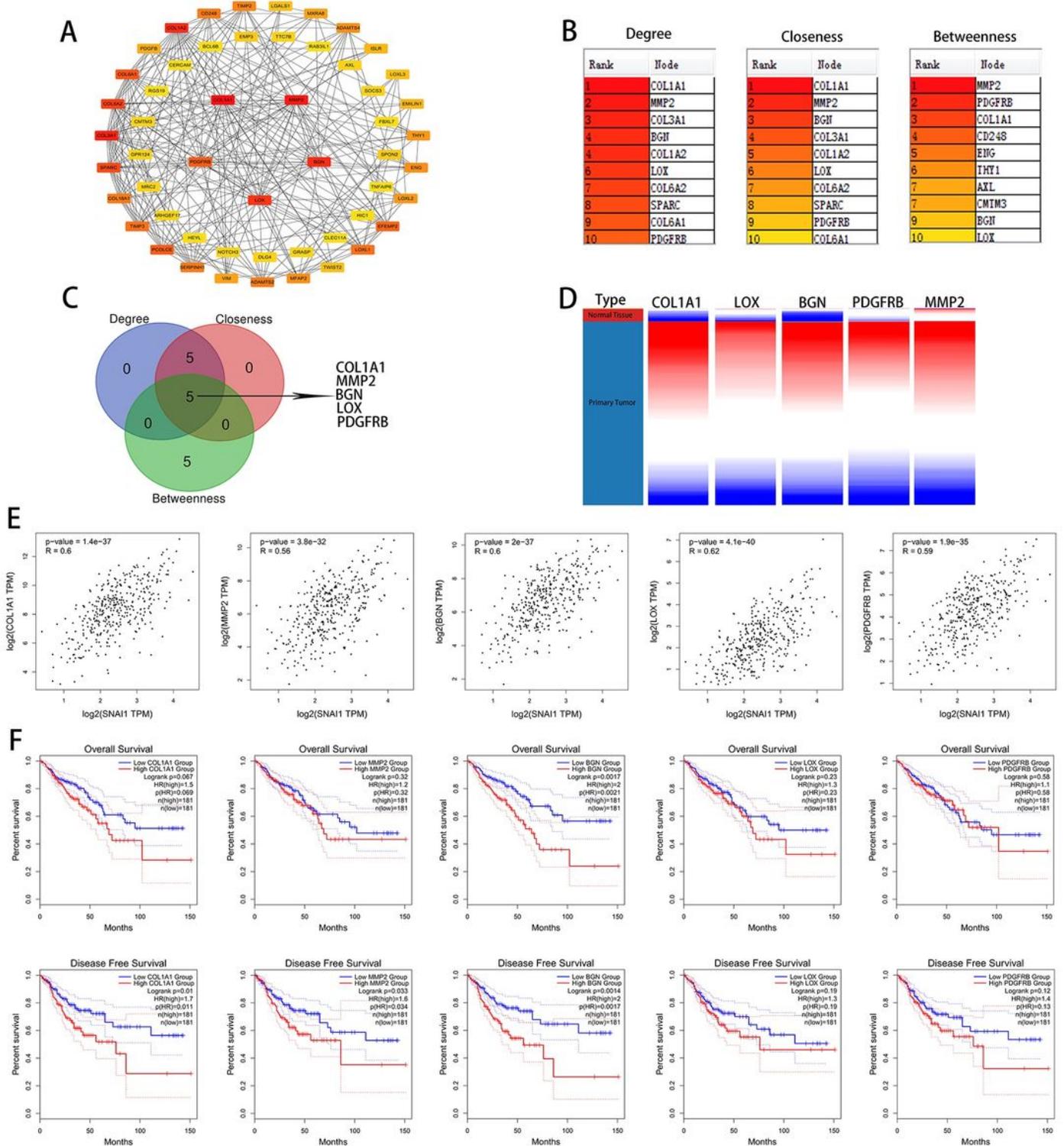


Figure 8

Construction of PPI network consisted of SNAI1 co-expressed genes and analysis of key genes. (A) PPI network of SNAI1 co-expressed genes generated from the STRING database. (B) The top 10 genes screened by using different algorithm of cytoHubba plugin in Cytoscape software. (C) Venn diagram represents the intersection of top genes among the different algorithm. (D) The hierarchical clustering of key genes was generated by using the UCSC database. (E) Correlation between SNAI1 and co-expressed

key genes determined by using the GEPIA database. (F) Overall Survival and Disease Free Survival analysis of key genes in patients with colorectal cancer by using the GEPIA database.

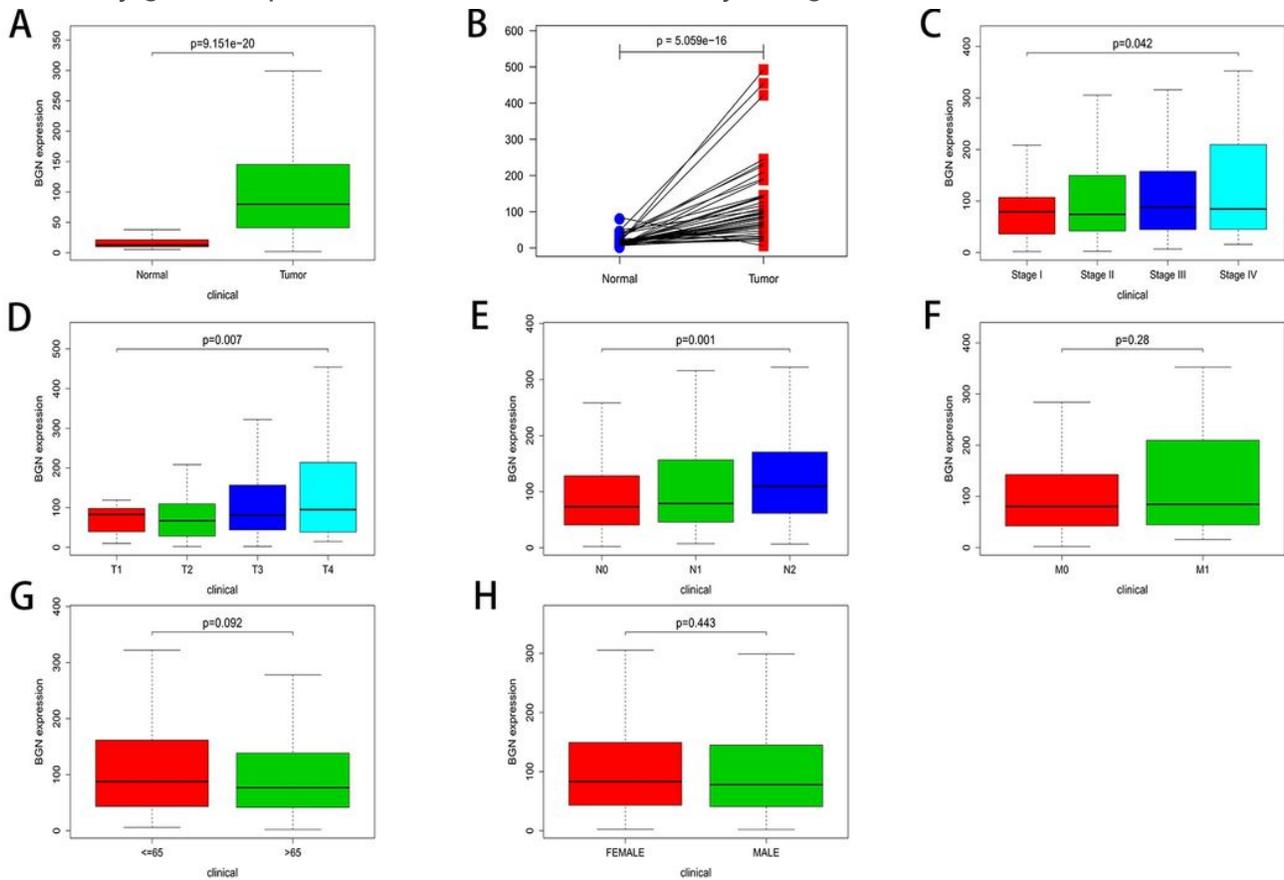


Figure 9

Associations of BGN expression with clinical parameters in colorectal cancer. (A) Sample type (cancer vs. normal). (B) Sample type (cancer vs. paired normal). (C) Clinical stage. (D) Tumor size (T classification). (E) Lymph node metastasis (N classification). (F) Distant metastases (M classification). (G) Age (≤ 65 vs. > 65). (H) Gender (female vs. male).