**Name of Journal:** *World Journal of Gastrointestinal Oncology*

**Manuscript NO:** 87150

**Manuscript Type:** ORIGINAL ARTICLE

***Basic Study***

**Association between** **heat shock factor protein 4 methylation and colorectal cancer risk and potential molecular mechanisms: A bioinformatics study**

Zhang WJ *et al*. Analysis of *HSF4* methylation in CRC

Wen-Jing Zhang, Ke-Lin Yue, Jing-Zhai Wang, Yu Zhang

**Wen-Jing Zhang,** Department of Medical Oncology, The First People’s Hospital of Yunnan Province, Affiliated Hospital of Kunming University of Science and Technology, Kunming 650032, Yunnan Province, China

**Ke-Lin Yue, Jing-Zhai Wang, Yu Zhang,** Department of Gastroenterology, The First People’s Hospital of Yunnan Province, Affiliated Hospital of Kunming University of Science and Technology, Kunming 650032, Yunnan Province, China

**Author contributions:** Zhang WJ and Zhang Y conceived and designed the experiments; Zhang WJ, Yue KL, and Wang JZ analyzed the data; Zhang Y contributed to the data curation; Zhang WJ wrote-original draft preparation; Yue KL, Wang JZ, and Zhang Y participated in the writing-review and editing.

**Supported by** National Natural Science Foundation of China, No. 82260601; Joint Foundation of Kunming Medical University and Yunnan Provincial Science and Technology Department, No. 202201AY070001-256; Grant for Clinical Medical Center of Yunnan Provincial Health Commission, No. 2021LCZXXF-XH03; and Young Academic Talents Cultivation Foundation of Yunnan Province, No. 202205AC160070.

**Corresponding author: Yu Zhang, MD, PhD, Associate Professor,** Department of Gastroenterology, The First People’s Hospital of Yunnan Province, Affiliated Hospital of Kunming University of Science and Technology, No. 157 Jinbi Road, Kunming 650032, Yunnan Province, China. yuzhang320@sina.com

**Received:** July 28, 2023

**Revised:** October 16, 2023

**Accepted:** November 17, 2023

**Published online:**

**Abstract**

BACKGROUND

We previously demonstrated that heat shock factor protein 4 (HSF4) facilitates colorectal cancer (CRC) progression. DNA methylation, a major modifier of gene expression and stability, is involved in CRC development and outcome.

AIM

To investigate the correlation between *HSF4* methylation and CRC risk, and to uncover the underlying molecular mechanisms.

METHODS

Differences in β values of *HSF4* methylation loci in multiple malignancies and their correlation with *HSF4* mRNA expression were analyzed based on Shiny Methylation Analysis Resource Tool. *HSF4* methylation-related genes were identified by LinkedOmics in CRC, and Gene Ontology and Kyoto Encyclopedia of Genes and Genomes enrichment analyses were performed. Protein-protein interaction network of *HSF4* methylation-related genes was constructed by String database and MCODE algorithm.

RESULTS

A total of 19 CpG methylation loci were identified in *HSF4*, and their β values were significantly increased in CRC tissues and exhibited a positive correlation with *HSF4* mRNA expression. Unfortunately, the prognostic and diagnostic performance of these CpG loci in CRC patients was mediocre. In CRC, there were 1694 *HSF4* methylation-related genes; 1468 of which displayed positive and 226 negative associations, and they were involved in regulating phenotypes such as immune, inflammatory, and metabolic reprogramming. *EGFR*, *RELA*, *STAT3*, *FCGR3A*, *POLR2K*, and *AXIN1* are hub genes among the *HSF4* methylation-related genes.

CONCLUSION

HSF4 is highly methylated in CRC, but there is no significant correlation between it and the prognosis and diagnosis of CRC. *HSF4* methylation may serve as one of the ways in which HSF4 mediates the CRC process.

**Key Words:** Colorectal cancer; DNA methylation; Prognosis; Diagnosis; Bioinformatics; Heat shock factor protein 4

Zhang WJ, Yue KL, Wang JZ, Zhang Y. Association between heat shock factor protein 4 methylation and colorectal cancer risk and potential molecular mechanisms: A bioinformatics study. *World J Gastrointest Oncol* 2023; In press

**Core Tip:** Colorectal cancer (CRC) is a common malignant tumor of the gastrointestinal tract with clinical manifestations of diarrhea, constipation, and abdominal pain. We previously demonstrated that heat shock factor protein 4 (HSF4) accelerates the malignant biological behavior of CRC cells *in vivo* and in vitro. This study reveals that HSF4 is highly methylated and associated with *HSF4* overexpression in CRC. Although the diagnostic and prognostic value of *HSF4* methylation is poor, it may be involved in the process of CRC by mediating the expression of *HSF4* or related genes. Combined with the finding of our previous study, the present study suggests that the high expression of *HSF4* mRNA and protein and its oncogenic effects are likely to be associated with *HSF4* methylation.

**INTRODUCTION**

Colorectal cancer (CRC) is a malignant tumor of the digestive tract that occurs in the rectum, cecum, and entire colon, with symptoms such as abdominal pain, difficulty passing stool, constipation, or diarrhea[1]. According to the latest statistics from the World Health Organization[2], CRC has become the third most common malignancy worldwide after lung cancer and breast cancer, with about 200000 new cases occurring worldwide each year, of which, 916000 die from CRC. According to the American Society of Clinical Oncology[3], the 5-year survival rate for patients with CRC is approximately 65%. Nevertheless, most CRC patients have already developed distal metastases by the time they receive a definitive diagnosis, which leads to a shrinking 5-year survival rate to 14%[3]. Consequently, the search for new biomarkers will facilitate the timely diagnosis of CRC and provide new insights into the mechanisms of CRC occurrence and development.

DNA methylation is a process of chemical modification of DNA that affects biological processes such as gene expression, cell differentiation and development[4-6]. In epigenetics, DNA methylation is an important marker of cellular genetic information and is widely applied in cancer prediction and diagnosis[7,8]. For CRC, the United States Food and Drug Administration currently approves SEPT9 (blood samples) and a combination of *NDRG4* and *BMP3* (stool samples) as commercially available biomarkers related to methylation[9]. In addition, *APC*, *SFRP1*, *SFRP2*, *SDC2*, *MGMT*, *VIM* and *NDRG4* are methylation-related candidate markers of CRC[10]. Mechanistically, DNA methylation can inhibit gene transcription or activate gene expression, thereby affecting protein synthesis to mediate the cancer process. For instance, teashirt zinc finger homeobox 3 (TSHZ3) promoter methylation effectively suppresses TSHZ3 expression, which facilitates CRC growth and metastasis[11]. Heparanase 2 (HPSE2) is highly methylated in CRC and is associated with poor patient prognosis, and high methylation of *HPSE2* reduces *HPSE2* expression, which inhibits the p53/p21 signaling cascade and facilitates proliferation of CRC cells *in vivo* and *in vitro*[12]. Heat shock response (HSR), an ancient cellular self-protective response, helps tumor cells to survive and proliferate smoothly under the stimulation of adverse microenvironment, oxidative stress and other stressors[13]. Heat shock factor protein 4 (HSF4), a member of the heat shock transcription factor family, plays an important role in HSR by preventing abnormal protein folding and aggregation to maintain intracellular protein homeostasis[13,14]. HSF4 has been identified as a cancer-promoting factor in lymphoma[15], breast cancer[16], and cervical cancer[17]. Our previous study demonstrated that *HSF4* is significantly upregulated in CRC, which predicts poor patient prognosis, and that it promotes CRC progression by enhancing the activity of c-MET and downstream ERK1/2 and AKT signaling pathways[18]. Nevertheless, whether DNA methylation is involved in HSF4-mediated CRC progression remains to be investigated.

This study investigated the correlation between *HSF4* methylation and *HSF4* expression, and its prognostic and diagnostic value in CRC, and aimed to identify the potential molecular mechanisms associated with *HSF4* methylation through bioinformatics analysis. The aim was to provide a theoretical basis and a novel perspective for *HSF4* as a methylation-related biomarker for future CRC diagnosis and treatment.

**MATERIALS AND METHODS**

***Differential analysis of HSF4 methylation and its prognostic and diagnostic value***

The Shiny Methylation Analysis Resource Tool (SMART) APP is an interactive and user-friendly web application for comprehensive analysis of DNA methylation in the The Cancer Genome Atlas (TCGA) project, with data from TCGA (https://portal.gdc.cancer.gov/)[19]. The level of methylation at each CpG loci of *HSF4* was assessed using the β value, which is the ratio of the methylation of the allele to the intensity of unmethylation, ranging from 0 to 1. In this study, we analyzed differences in the β values of 19 methylation probes associated with *HSF4* in 33 malignancies by SMART, including COAD, READ, BRCA, LAML, LGG, LIHC, BLCA, CESC, CHOL, KIRP, SKCM, LUAD, ACC, DLBC, KIRC, PCPG, OV, ESCA GBM, STAD, UCEC, UCS, HNSC, TGCT, THCA, THYM, KICH, PRAD, SARC, LUSC, MESO, PAAD, and UVM. Wilcoxon rank sum test was performed for difference analysis of β values, and data was adjusted using the Benjamini-Hochberg method. In addition, β values of 19 *HSF4*-related methylation probes were analyzed differentially in COAD stages and their correlation with HSF4 mRNA expression based on SMART. The differential analysis of β values in COAD stages was performed based on ANOVA, and the correlation between β values of each probe and *HSF4* mRNA expression was performed based on Pearson. The COAD dataset in SMART was extracted and *HSF4* methylation in prognostic and diagnostic value of COAD was assessed by survival (https://cran.r-project.org/web/packages/survival/index.html) and pROC[20]/timeROC[21] R packages, respectively. Kaplan-Meier survival curves, Receiver operating characteristic (ROC) curves and time-dependent ROC curves were visualized with the ggplot2 R package[22]. Patient information is shown in Table 1.

***Identification of HSF4 methylation-related genes and their enrichment analysis***

LinkedOmics is a publicly available portal that includes three analysis modules, LinkFinder, LinkInterpreter and LinkCompare, to support users in performing multi-omics analysis in cancer, with data from TCGA (https://www.cancer.gov/ccg/ research/genome-sequencing/tcga) and Clinical Proteomic Tumor Analysis Consortium (https://proteomics.cancer.gov/programs/cptac)[23]. In this study, genes associated with *HSF4* methylation were identified at COAD through LinkedOmics. *HSF4* methylation-associated genes were identified by Spearman and subjected to correction by the Benjamini-Hochberg method. Finally, *HSF4* methylation-related genes were displayed by volcano plot and heatmap. Enrichment analysis of *HSF4* methylation-related genes was performed by hypergeometric distribution algorithm based on Gene Ontology (GO)[24] and Kyoto Encyclopedia of Genes and Genomes (KEGG)[25] databases, and presented by bubble and histogram plots. The above results was visualized with the ggplot2 R package[22].

***Protein-protein interaction network construction for HSF4 methylation-associated genes***

The protein-protein interaction (PPI) network construction for *HSF4* methylation-related genes was based on the String database[26], CytoScape software[27] and the MCODE plugin[28]. Briefly, *HSF4* methylation-related genes obtained from LinkedOmics were extracted, and the interactions of these genes were predicted from the String database. The minimum required interaction score of the String database was set to highest confidence. The interactions were imported into CytoScape software (version:3.8.2) for visualization and clustering analysis of the PPI network was performed by the MCODE algorithm. The parameters of MCODE are degree cutoff = 2, node density cutoff = 0.1, node score cutoff = 0.2, K-core = 2, Max depth = 100.

**RESULTS**

***Identification of HSF4 methylation levels***

HSF4 is located on chromosome 16 with 19 CpG loci, with 14 on CpG island, three on N Shore and two on S Shore (Figure 1). Differential analysis revealed that β values of *HSF4* CpG-aggregation methylation were significantly enhanced in most malignancies, including COAD, and READ (Figure 1B). Similarly, the β values of each CpG site were significantly higher in most malignant tumors than in the corresponding paracancerous tissues (Supplementary Figure 1). It is notable that all CpG loci of *HSF4* had significantly elevated β values in these malignancies only in COAD (Supplementary Figure 1 and Figure 2). In READ, only two probes, cg07188665 and cg09567485, exhibited no significant difference in β values. We analyzed the methylation levels of HSF4 CpG loci in different tumor stages. The β values of cg06277900, cg03811260, cg04580872, cg06621126, cg03887094 and cg09567485 probes were significantly different at various stages of COAD (Supplementary Figure 2). Therefore, we further explored the correlation between *HSF4* methylation and *HSF4* expression. In COAD, the β values of the probes displayed a significant positive correlation with *HSF4* expression, except for cg07188665 (Figure 3). Combined with our previous findings, we believe that HSF4 promotes the CRC process at least through DNA methylation.

***HSF4 methylation correlates poorly with CRC prognosis and diagnosis***

In view of the differences in *HSF4* methylation in CRC, we further analyzed the prognostic and diagnostic value of *HSF4* methylation. Kaplan-Meier curves indicated no significant difference in survival among COAD patients with high and low methylation levels for each CpG loci (Figure 4). Nevertheless, most patients with hypermethylated CpG loci had better prognosis. The ROC curve revealed that the area under the curve (AUC) of each CpG loci ranged from 0.498 to 0.574 in COAD patients, suggesting the mediocre diagnostic value of *HSF4* methylation in COAD patients (Figure 5A). The time-dependent ROC curves suggested that the AUC of each CpG loci was greater with increasing time (Figure 5B). The above results indicated that the performance of *HSF4* methylation as a prognostic and diagnostic biomarker in CRC was ordinary, which may have been caused by relatively low accumulation of single genes.

***Identification of HSF4 methylation-related genes in CRC and their functional enrichment analysis***

We analyzed the genes associated with *HSF4* methylation in CRC by LinkedOmics. The expression of 1468 genes was positively correlated with *HSF4* methylation levels, and expression of 226 genes was negatively correlated with *HSF4* methylation levels in the COAD cohort (Figure 6A). The heatmap illustrated the top 50 genes with absolute correlation coefficients (Figures 6B and C). To further understand the functions and pathways involved in these genes, we performed GO and KEGG enrichment analysis. GO identified that the proteins encoded by these genes were mainly extracellular matrix, and associated with processes such as positive regulation of mitogen-activated protein kinase cascade, tumor necrosis factor superfamily cytokine production, neutrophil mediated cytotoxicity, and chemokine activity (Figure 6D). KEGG enrichment revealed that *HSF4* methylation-related genes were involved in pathways including chemokine signaling pathway, calcium signaling pathway, glycosphingolipid biosynthesis - lacto and neolacto series, inflammatory bowel disease and inflammatory bowel disease (Figure 6E). It is suggested that *HSF4* methylation mediates the phenotypic involvement of immune, inflammatory, and metabolic reprogramming in the CRC process.

***PPI network of HSF4 methylation-associated genes in CRC***

To identify the hub genes in *HSF4* methylation-related genes, we constructed a relevant PPI network based on the String database and the MCODE algorithm. The PPI network constructed for *HSF4* methylation positively correlated genes contained 422 nodes and 702 edges, and 22 clustering networks were obtained (Figure 7A). The top 20 genes in this network with the highest number of edges are displayed in Figure 7B, where *EGFR*, *RELA*, *STAT3*, *ESR1*, and *F2* had the highest number of edges. The top 10 interworking networks with clustering scores are illustrated in Figure 7C. The network consisting of *NUP98*, *SUMO3*, *IPO8*, and *HSPA6* had the highest clustering score, which contained 11 nodes and 35 edges (Figure 7C). In the same way, the network constructed for negatively associated genes contained 110 genes, 122 interactions and five clusters (Figure 8A). The edge numbers TOP5 of *FCGR3A*, *POLR2K*, *AXIN1*, *CCL2* and COPS5 had eight, seven, six, five and five edges, respectively (Figure 8B). The five clustering networks composed of genes and interactions are shown in Figure 8C. It is suggested that these genes are involved in *HSF4* methylation mediation of the CRC process.

**DISCUSSION**

CRC is the second leading cause of cancer-related deaths worldwide. CRC is the outcome of progressive accumulation of a series of mutations and epigenetic changes in the rectum, cecum, and colon, leading to the development of colorectal adenoma and invasive adenocarcinoma. DNA methylation, one of the major epigenetic modifications, has been partially identified as a commercial diagnostic and prognostic biomarker for CRC. We have previously identified *HSF4* as an oncogenic gene in CRC[18]. Therefore, we tapped the diagnostic and prognostic value of *HSF4* and its possible molecular mechanisms in CRC. Unfortunately, *HSF4*, like most single-gene markers[29-32], has a mediocre diagnostic and prognostic value for its methylation levels in CRC. As in previous studies[29-32], this may be due to the small sample size analyzed or the insufficient accumulation of single gene methylation. Therefore, we analyzed the role of *HSF4* methylation in CRC at the molecular mechanism level. It is noteworthy that we identified 1694 genes associated with *HSF4* methylation, and their possible involvement in immune, inflammatory, and metabolic reprogramming. In addition, the constructed PPI network demonstrated that *EGFR*, *RELA*, *STAT3*, *ESR1*, *FCGR3A*, *AXIN1*, *CCL2*, and *COPS5* are hub genes among *HSF4* methylation-related genes.

Most of these hub genes have been demonstrated to be involved in the CRC process and have been applied as therapies for CRC. For instance, EGFR is a transmembrane receptor that plays a regulatory role in tumor cell function by binding to EGFs, promoting cell proliferation, differentiation, and survival[33]. Currently, monoclonal antibodies against EGFR, such as cetuximab or panitumumab, are utilized in the clinical treatment of patients with metastatic CRC[34,35]. RELA, also known as p65 or nuclear factor (NF)-κB p65, is known to be a key transcription factor in tumors, and it mediates immune and inflammatory responses to facilitate cancer cell survival and metastasis, which leads to it being a key target in tumor therapy[36,37]. Similarly, FCGR3A belongs to the Fc γ receptor family, which is mainly expressed on the surface of natural killer cells, monocytes, and macrophages and plays an important role in antibody-mediated immune responses[38]. Polymorphisms in *FCGR3A* are associated with progression-free survival in patients with metastatic CRC treated with cetuximab[39,40]. COPS5, also known as CSN5 or JAB1, is one of the constituent proteins of the COP9 signalosome, is a nuclear-plasmid transmembrane protein with multiple functions, and is involved in the regulation of various cellular processes such as cell proliferation, differentiation, apoptosis, and DNA replication and repair[41]. It has been demonstrated that COPS5 plays a role as a pro-cancer factor in CRC by regulating Wnt and PI14K/AKT pathways[42-44]. Some of these hub genes have also been proven to be related to HSF family proteins. For example, HSR-induced activation of HSP1 is regulated by the NF-κB, which activates transcription of HSPA1A[45]. In turn, HSP1 inhibits the activation of NF-κB pathway[46,47]. Stephanou and Latchman[48] showed that the activation of *STAT3* alone facilitates the mobilization of the HSP promoter. Nevertheless, whether *HSF4* methylation-mediated alterations in HSF4 expression are crosstalk with these hub genes in CRC remains to be further investigated.

Unfortunately, there were some limitations to this study. For instance, the sample size was small, and a larger sample might reveal satisfactory diagnostic and prognostic values[49,50]. Exploring the correlation between *HSF4* methylation and CRC subtypes or *HSF4*-related gene methylation combinations may improve the value of *HSF4* in CRC[9,49]. It is essential to verify *HSF4* methylation in CRC tissues by methylation sequencing or microarrays, which could support the findings of this study[51]. Although we demonstrated that *HSF4* methylation levels exhibited a positive correlation with *HSF4* mRNA expression, *in vivo* and *in vitro* experiments are lacking for validation. In the same way, the molecular mechanisms associated with *HSF4* methylation remain to be explored *in vivo* and *in vitro*. The tumor immune microenvironment (TME) consists of immune cells, blood vessels, and extracellular matrix, and has a dual role in the growth and metastasis of tumor cells[52,53]. HSF family proteins, especially HSF1, have been demonstrated to mediate tumor cell associated immune responses in the TME[54,55]. This predicts that the function of HSF4 in CRC-associated TME is also worthy of investigation.

**CONCLUSION**

In conclusion, this study reveals that HSF4 is highly methylated in CRC and is associated with *HSF4* overexpression. Although *HSF4* methylation has poor diagnostic and prognostic value, it may be involved in the CRC process by mediating expression of *HSF4* or related genes with potential mechanisms. Combined with our previously described findings[18], the present study believes that high expression of *HSF4* mRNA and protein and its oncogenic effects are most probably due to *HSF4* methylation (Figure 9). Specific mechanisms need to be confirmed by more *in vivo* and *in vitro* experiments.

**ARTICLE HIGHLIGHTS**

***Research background***

DNA methylation is involved in the regulation of gene expression and has been implicated in development and outcome of colorectal cancer (CRC).

***Research motivation***

We previously demonstrated that heat shock factor protein 4 (HSF4) expression is abnormally high, and contributes to the malignant biological behavior of CRC *in vivo* and *in vitro*. However, the correlation of *HSF4* methylation with *HSF4* expression and prognosis of CRC patients, and other potential molecular mechanisms need to be further investigated.

***Research objectives***

The present study was proposed to investigate the correlation between *HSF4* methylation and CRC risk, and to uncover the underlying molecular mechanisms.

***Research methods***

Identification of *HSF4* methylation sites, and analysis of the differences in β values of *HSF4* methylation sites and their correlation with *HSF4* mRNA expression were performed using Shiny Methylation Analysis Resource Tool Web. The genes associated with *HSF4* methylation were identified by LinkedOmics Web for CRC, and Gene Ontology and Kyoto Encyclopedia of Genes and Genomes enrichment analyses were performed to reveal the functions and signaling that these associated genes may be involved in. The String database and MCODE algorithm were performed to construct protein-protein interaction (PPI) networks of *HSF4* methylation-related genes.

***Research results***

The *HSF4* gene had 19 CpG methylation sites, and their β-values were significantly higher in CRC tissues, positively correlating with *HSF4* mRNA expression. The β value of the *HSF4* methylation site was not associated with the prognosis of CRC patients. Notably, there are 1694 genes in CRC tissues whose expression is associated with *HSF4* methylation and which are involved in immune, inflammatory, and metabolic reprogramming. EGFR, STAT3 and AXIN1 are hub genes in the PPI network constructed by these *HSF4* methylation-related genes.

***Research conclusions***

The *HSF4* gene is highly methylated in CRC, and is associated with the overexpression of *HSF4* mRNA. *HSF4* methylation may be involved in the process of CRC by mediating the expression of *HSF4* or related genes.

***Research perspectives***

The finding will provide a theoretical basis and a new perspective on *HSF4* as a methylation-related biomarker for future CRC diagnosis and treatment.

**REFERENCES**

1 **Dekker E**, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. *Lancet* 2019; **394**: 1467-1480 [PMID: 31631858 DOI: 10.1016/S0140-6736(19)32319-0]

2 **Singh D**, Vignat J, Lorenzoni V, Eslahi M, Ginsburg O, Lauby-Secretan B, Arbyn M, Basu P, Bray F, Vaccarella S. Global estimates of incidence and mortality of cervical cancer in 2020: a baseline analysis of the WHO Global Cervical Cancer Elimination Initiative. *Lancet Glob Health* 2023; **11**: e197-e206 [PMID: 36528031 DOI: 10.1016/S2214-109X(22)00501-0]

3 **Morris VK**, Kennedy EB, Baxter NN, Benson AB 3rd, Cercek A, Cho M, Ciombor KK, Cremolini C, Davis A, Deming DA, Fakih MG, Gholami S, Hong TS, Jaiyesimi I, Klute K, Lieu C, Sanoff H, Strickler JH, White S, Willis JA, Eng C. Treatment of Metastatic Colorectal Cancer: ASCO Guideline. *J Clin Oncol* 2023; **41**: 678-700 [PMID: 36252154 DOI: 10.1200/JCO.22.01690]

4 **Mattei AL**, Bailly N, Meissner A. DNA methylation: a historical perspective. *Trends Genet* 2022; **38**: 676-707 [PMID: 35504755 DOI: 10.1016/j.tig.2022.03.010]

5 **Law PP**, Holland ML. DNA methylation at the crossroads of gene and environment interactions. *Essays Biochem* 2019; **63**: 717-726 [PMID: 31782496 DOI: 10.1042/EBC20190031]

6 **Locke WJ**, Guanzon D, Ma C, Liew YJ, Duesing KR, Fung KYC, Ross JP. DNA Methylation Cancer Biomarkers: Translation to the Clinic. *Front Genet* 2019; **10**: 1150 [PMID: 31803237 DOI: 10.3389/fgene.2019.01150]

7 **Saghafinia S**, Mina M, Riggi N, Hanahan D, Ciriello G. Pan-Cancer Landscape of Aberrant DNA Methylation across Human Tumors. *Cell Rep* 2018; **25**: 1066-1080.e8 [PMID: 30355485 DOI: 10.1016/j.celrep.2018.09.082]

8 **Nishiyama A**, Nakanishi M. Navigating the DNA methylation landscape of cancer. *Trends Genet* 2021; **37**: 1012-1027 [PMID: 34120771 DOI: 10.1016/j.tig.2021.05.002]

9 **Jung G**, Hernández-Illán E, Moreira L, Balaguer F, Goel A. Epigenetics of colorectal cancer: biomarker and therapeutic potential. *Nat Rev Gastroenterol Hepatol* 2020; **17**: 111-130 [PMID: 31900466 DOI: 10.1038/s41575-019-0230-y]

10 **Müller D**, Győrffy B. DNA methylation-based diagnostic, prognostic, and predictive biomarkers in colorectal cancer. *Biochim Biophys Acta Rev Cancer* 2022; **1877**: 188722 [PMID: 35307512 DOI: 10.1016/j.bbcan.2022.188722]

11 **Zhou Y**, Wang S, Yin X, Gao G, Wang Q, Zhi Q, Han Y, Kuang Y. TSHZ3 functions as a tumor suppressor by DNA methylation in colorectal cancer. *Clin Res Hepatol Gastroenterol* 2021; **45**: 101725 [PMID: 34089916 DOI: 10.1016/j.clinre.2021.101725]

12 **Zhang H**, Xu C, Shi C, Zhang J, Qian T, Wang Z, Ma R, Wu J, Jiang F, Feng J. Hypermethylation of heparanase 2 promotes colorectal cancer proliferation and is associated with poor prognosis. *J Transl Med* 2021; **19**: 98 [PMID: 33663522 DOI: 10.1186/s12967-021-02770-0]

13 **Lang BJ**, Guerrero ME, Prince TL, Okusha Y, Bonorino C, Calderwood SK. The functions and regulation of heat shock proteins; key orchestrators of proteostasis and the heat shock response. *Arch Toxicol* 2021; **95**: 1943-1970 [PMID: 34003342 DOI: 10.1007/s00204-021-03070-8]

14 **Syafruddin SE**, Ling S, Low TY, Mohtar MA. More Than Meets the Eye: Revisiting the Roles of Heat Shock Factor 4 in Health and Diseases. *Biomolecules* 2021; **11** [PMID: 33807297 DOI: 10.3390/biom11040523]

15 **Jin X**, Eroglu B, Cho W, Yamaguchi Y, Moskophidis D, Mivechi NF. Inactivation of heat shock factor Hsf4 induces cellular senescence and suppresses tumorigenesis in vivo. *Mol Cancer Res* 2012; **10**: 523-534 [PMID: 22355043 DOI: 10.1158/1541-7786.MCR-11-0530]

16 **Chen R**, Liliental JE, Kowalski PE, Lu Q, Cohen SN. Regulation of transcription of hypoxia-inducible factor-1α (HIF-1α) by heat shock factors HSF2 and HSF4. *Oncogene* 2011; **30**: 2570-2580 [PMID: 21258402 DOI: 10.1038/onc.2010.623]

17 **Tu N**, Hu Y, Mivechi NF. Heat shock transcription factor (Hsf)-4b recruits Brg1 during the G1 phase of the cell cycle and regulates the expression of heat shock proteins. *J Cell Biochem* 2006; **98**: 1528-1542 [PMID: 16552721 DOI: 10.1002/jcb.20865]

18 **Zhang W**, Zhang X, Cheng P, Yue K, Tang M, Li Y, Guo Q, Zhang Y. HSF4 promotes tumor progression of colorectal cancer by transactivating c-MET. *Mol Cell Biochem* 2023; **478**: 1141-1150 [PMID: 36229759 DOI: 10.1007/s11010-022-04582-2]

19 **Li Y**, Ge D, Lu C. The SMART App: an interactive web application for comprehensive DNA methylation analysis and visualization. *Epigenetics Chromatin* 2019; **12**: 71 [PMID: 31805986 DOI: 10.1186/s13072-019-0316-3]

20 **Robin X**, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011; **12**: 77 [PMID: 21414208 DOI: 10.1186/1471-2105-12-77]

21 **Blanche P**, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med* 2013; **32**: 5381-5397 [PMID: 24027076 DOI: 10.1002/sim.5958]

22 **Ito K**, Murphy D. Application of ggplot2 to Pharmacometric Graphics. *CPT Pharmacometrics Syst Pharmacol* 2013; **2**: e79 [PMID: 24132163 DOI: 10.1038/psp.2013.56]

23 **Vasaikar SV**, Straub P, Wang J, Zhang B. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res* 2018; **46**: D956-D963 [PMID: 29136207 DOI: 10.1093/nar/gkx1090]

24 **Mi H**, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res* 2019; **47**: D419-D426 [PMID: 30407594 DOI: 10.1093/nar/gky1038]

25 **Kanehisa M**, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res* 2021; **49**: D545-D551 [PMID: 33125081 DOI: 10.1093/nar/gkaa970]

26 **Szklarczyk D**, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S, Bork P, Jensen LJ, von Mering C. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* 2023; **51**: D638-D646 [PMID: 36370105 DOI: 10.1093/nar/gkac1000]

27 **Otasek D**, Morris JH, Bouças J, Pico AR, Demchak B. Cytoscape Automation: empowering workflow-based network analysis. *Genome Biol* 2019; **20**: 185 [PMID: 31477170 DOI: 10.1186/s13059-019-1758-4]

28 **Bader GD**, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 2003; **4**: 2 [PMID: 12525261 DOI: 10.1186/1471-2105-4-2]

29 **Shao C**, Dai W, Li H, Tang W, Jia S, Wu X, Luo Y. The relationship between RASSF1A gene promoter methylation and the susceptibility and prognosis of melanoma: A meta-analysis and bioinformatics. *PLoS One* 2017; **12**: e0171676 [PMID: 28207831 DOI: 10.1371/journal.pone.0171676]

30 **Laugsand EA**, Brenne SS, Skorpen F. DNA methylation markers detected in blood, stool, urine, and tissue in colorectal cancer: a systematic review of paired samples. *Int J Colorectal Dis* 2021; **36**: 239-251 [PMID: 33030559 DOI: 10.1007/s00384-020-03757-x]

31 **Vedeld HM**, Nesbakken A, Lothe RA, Lind GE. Re-assessing ZNF331 as a DNA methylation biomarker for colorectal cancer. *Clin Epigenetics* 2018; **10**: 70 [PMID: 29854011 DOI: 10.1186/s13148-018-0503-2]

32 **Gogna P**, King WD. The relationship between colorectal cancer risk factors and LINE-1 DNA methylation in healthy colon tissue. *Epigenomics* 2020; **12**: 1087-1093 [PMID: 32790479 DOI: 10.2217/epi-2019-0340]

33 **Sigismund S**, Avanzato D, Lanzetti L. Emerging functions of the EGFR in cancer. *Mol Oncol* 2018; **12**: 3-20 [PMID: 29124875 DOI: 10.1002/1878-0261.12155]

34 **Zhou J**, Ji Q, Li Q. Resistance to anti-EGFR therapies in metastatic colorectal cancer: underlying mechanisms and reversal strategies. *J Exp Clin Cancer Res* 2021; **40**: 328 [PMID: 34663410 DOI: 10.1186/s13046-021-02130-2]

35 **Piawah S**, Venook AP. Targeted therapy for colorectal cancer metastases: A review of current methods of molecularly targeted therapy and the use of tumor biomarkers in the treatment of metastatic colorectal cancer. *Cancer* 2019; **125**: 4139-4147 [PMID: 31433498 DOI: 10.1002/cncr.32163]

36 **Patel M**, Horgan PG, McMillan DC, Edwards J. NF-κB pathways in the development and progression of colorectal cancer. *Transl Res* 2018; **197**: 43-56 [PMID: 29550444 DOI: 10.1016/j.trsl.2018.02.002]

37 **Peng C**, Ouyang Y, Lu N, Li N. The NF-κB Signaling Pathway, the Microbiota, and Gastrointestinal Tumorigenesis: Recent Advances. *Front Immunol* 2020; **11**: 1387 [PMID: 32695120 DOI: 10.3389/fimmu.2020.01387]

38 **Wang TT**, Ravetch JV. Functional diversification of IgGs through Fc glycosylation. *J Clin Invest* 2019; **129**: 3492-3498 [PMID: 31478910 DOI: 10.1172/JCI130029]

39 **Zhang W**, Gordon M, Schultheis AM, Yang DY, Nagashima F, Azuma M, Chang HM, Borucka E, Lurje G, Sherrod AE, Iqbal S, Groshen S, Lenz HJ. FCGR2A and FCGR3A polymorphisms associated with clinical outcome of epidermal growth factor receptor expressing metastatic colorectal cancer patients treated with single-agent cetuximab. *J Clin Oncol* 2007; **25**: 3712-3718 [PMID: 17704420 DOI: 10.1200/jco.2006.08.8021]

40 **Pander J**, Gelderblom H, Antonini NF, Tol J, van Krieken JH, van der Straaten T, Punt CJ, Guchelaar HJ. Correlation of FCGR3A and EGFR germline polymorphisms with the efficacy of cetuximab in KRAS wild-type metastatic colorectal cancer. *Eur J Cancer* 2010; **46**: 1829-1834 [PMID: 20418097 DOI: 10.1016/j.ejca.2010.03.017]

41 **Liu G**, Claret FX, Zhou F, Pan Y. Jab1/COPS5 as a Novel Biomarker for Diagnosis, Prognosis, Therapy Prediction and Therapeutic Tools for Human Cancer. *Front Pharmacol* 2018; **9**: 135 [PMID: 29535627 DOI: 10.3389/fphar.2018.00135]

42 **Zhou R**, Shao Z, Liu J, Zhan W, Gao Q, Pan Z, Wu L, Xu L, Ding Y, Zhao L. COPS5 and LASP1 synergistically interact to downregulate 14-3-3σ expression and promote colorectal cancer progression via activating PI3K/AKT pathway. *Int J Cancer* 2018; **142**: 1853-1864 [PMID: 29226323 DOI: 10.1002/ijc.31206]

43 **Jumpertz S**, Hennes T, Asare Y, Schütz AK, Bernhagen J. CSN5/JAB1 suppresses the WNT inhibitor DKK1 in colorectal cancer cells. *Cell Signal* 2017; **34**: 38-46 [PMID: 28229932 DOI: 10.1016/j.cellsig.2017.02.013]

44 **Schütz AK**, Hennes T, Jumpertz S, Fuchs S, Bernhagen J. Role of CSN5/JAB1 in Wnt/β-catenin activation in colorectal cancer cells. *FEBS Lett* 2012; **586**: 1645-1651 [PMID: 22668871 DOI: 10.1016/j.febslet.2012.04.037]

45 **Sasi BK**, Sonawane PJ, Gupta V, Sahu BS, Mahapatra NR. Coordinated transcriptional regulation of Hspa1a gene by multiple transcription factors: crucial roles for HSF-1, NF-Y, NF-κB, and CREB. *J Mol Biol* 2014; **426**: 116-135 [PMID: 24041570 DOI: 10.1016/j.jmb.2013.09.008]

46 **Shang L**, Wang L, Shi X, Wang N, Zhao L, Wang J, Liu C. HMGB1 was negatively regulated by HSF1 and mediated the TLR4/MyD88/NF-κB signal pathway in asthma. *Life Sci* 2020; **241**: 117120 [PMID: 31825792 DOI: 10.1016/j.lfs.2019.117120]

47 **Chen Y**, Currie RW. Small interfering RNA knocks down heat shock factor-1 (HSF-1) and exacerbates pro-inflammatory activation of NF-kappaB and AP-1 in vascular smooth muscle cells. *Cardiovasc Res* 2006; **69**: 66-75 [PMID: 16061216 DOI: 10.1016/j.cardiores.2005.07.004]

48 **Stephanou A**, Latchman DS. Transcriptional regulation of the heat shock protein genes by STAT family transcription factors. *Gene Expr* 1999; **7**: 311-319 [PMID: 10440232]

49 **Petit J**, Carroll G, Zhao J, Roper E, Pockney P, Scott RJ. Evaluation of epigenetic methylation biomarkers for the detection of colorectal cancer using droplet digital PCR. *Sci Rep* 2023; **13**: 8883 [PMID: 37264006 DOI: 10.1038/s41598-023-35631-5]

50 **Fang Q**, Yuan Z, Hu H, Zhang W, Wang G, Wang X. Genome-wide discovery of circulating cell-free DNA methylation biomarkers for colorectal cancer detection. *Clin Epigenetics* 2023; **15**: 119 [PMID: 37501075 DOI: 10.1186/s13148-023-01518-5]

51 **Ye Z**, Song G, Liang J, Yi S, Gao Y, Jiang H. Optimized screening of DNA methylation sites combined with gene expression analysis to identify diagnostic markers of colorectal cancer. *BMC Cancer* 2023; **23**: 617 [PMID: 37400791 DOI: 10.1186/s12885-023-10922-2]

52 **Lv B**, Wang Y, Ma D, Cheng W, Liu J, Yong T, Chen H, Wang C. Immunotherapy: Reshape the Tumor Immune Microenvironment. *Front Immunol* 2022; **13**: 844142 [PMID: 35874717 DOI: 10.3389/fimmu.2022.844142]

53 **El-Arabey AA**, Abdalla M, Abd-Allah AR. SnapShot: TP53 status and macrophages infiltration in TCGA-analyzed tumors. *Int Immunopharmacol* 2020; **86**: 106758 [PMID: 32663767 DOI: 10.1016/j.intimp.2020.106758]

54 **Shan Q**, Ma F, Wei J, Li H, Ma H, Sun P. Physiological Functions of Heat Shock Proteins. *Curr Protein Pept Sci* 2020; **21**: 751-760 [PMID: 31713482 DOI: 10.2174/1389203720666191111113726]

55 **Calderwood SK**, Hightower LE. Report on the VIIth International Symposium on Heat Shock Proteins in Biology & Medicine. *Cell Stress Chaperones* 2015; **20**: 213-216 [PMID: 25542250 DOI: 10.1007/s12192-014-0562-z]

**Footnotes**

**Institutional review board statement:** This study did not involve any animal and human experimentation.

**Conflict-of-interest statement:** All the authors report no relevant conflicts of interest for this article.

**Data sharing statement:** No additional data are available.

**Open-Access:** This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: https://creativecommons.org/Licenses/by-nc/4.0/

**Provenance and peer review:** Unsolicited article; Externally peer reviewed.

**Peer-review model:** Single blind

**Peer-review started:** July 28, 2023

**First decision:** September 26, 2023

**Article in press:**

**Specialty type:** Oncology

**Country/Territory of origin:** China

**Peer-review report’s scientific quality classification**

Grade A (Excellent): 0

Grade B (Very good): B, B

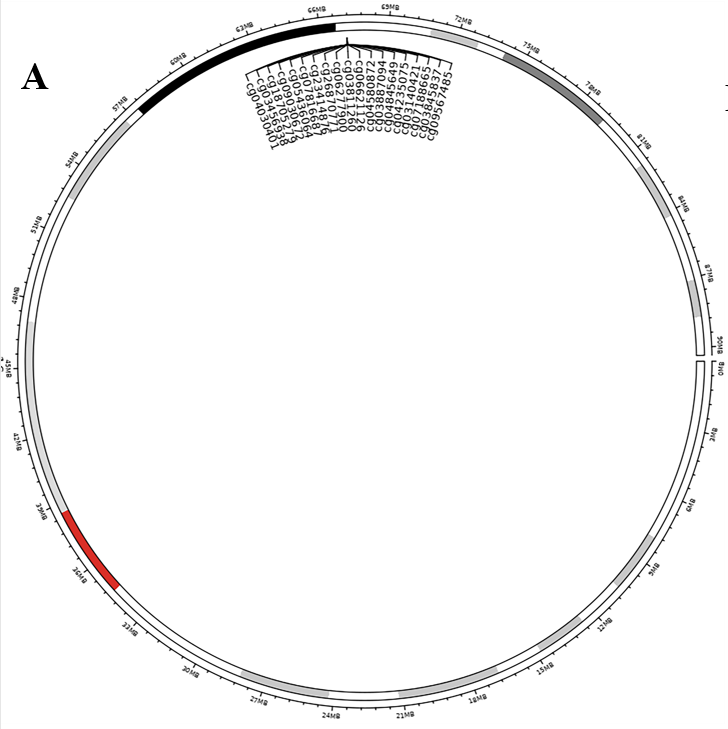
Grade C (Good): C

Grade D (Fair): 0

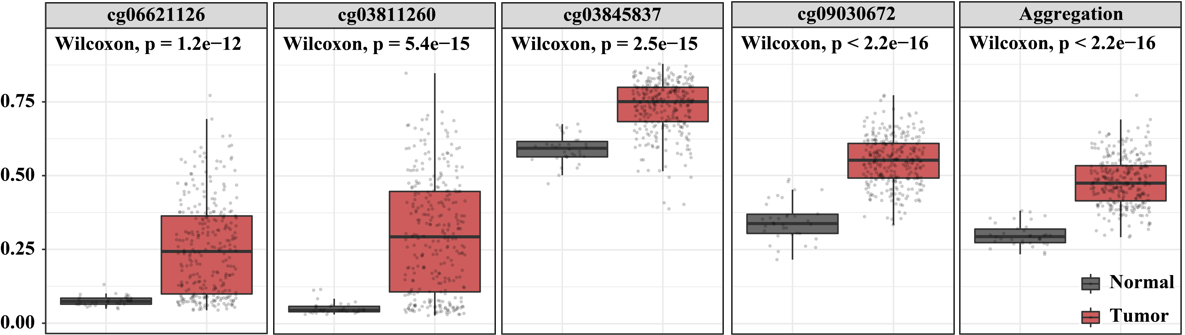
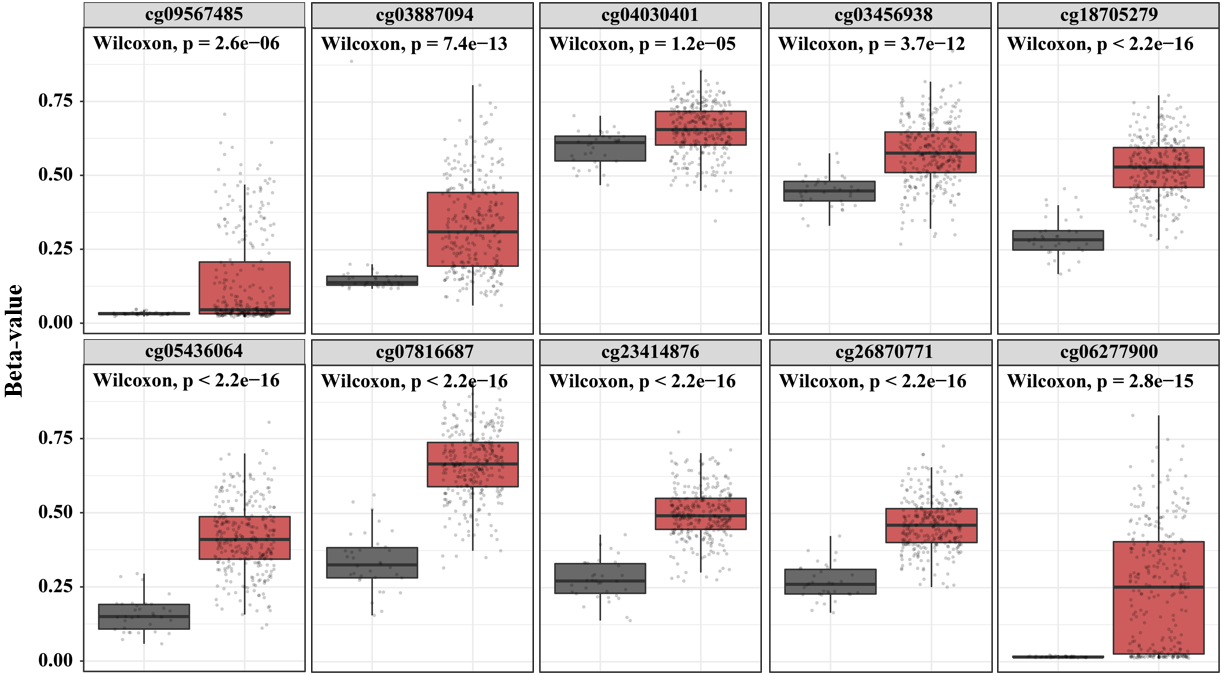
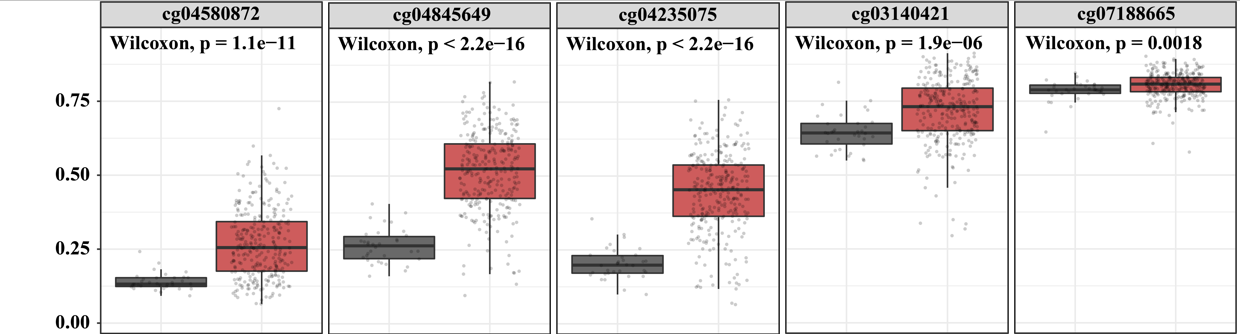
Grade E (Poor): E

**P-Reviewer:** Duraes LC, United States; El-Arabey AA, Egypt; Rotondo JC, Italy; Shinozaki M, Japan **S-Editor:** Wang JJ **L-Editor:** A **P-Editor:**

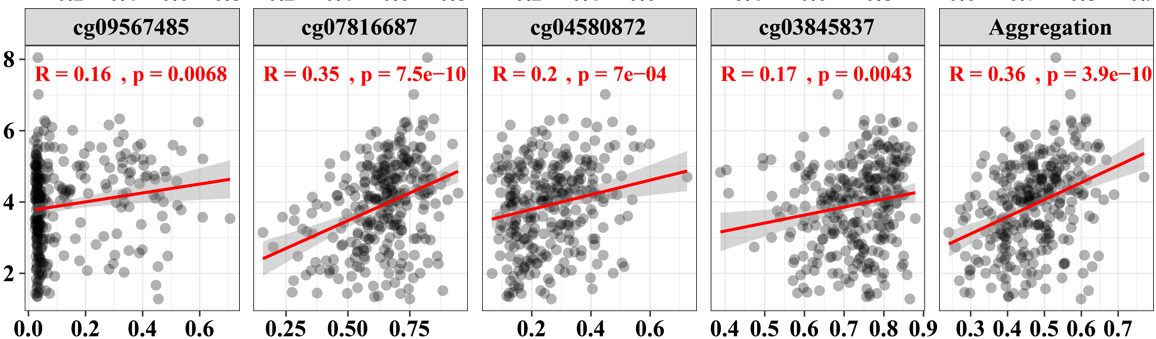
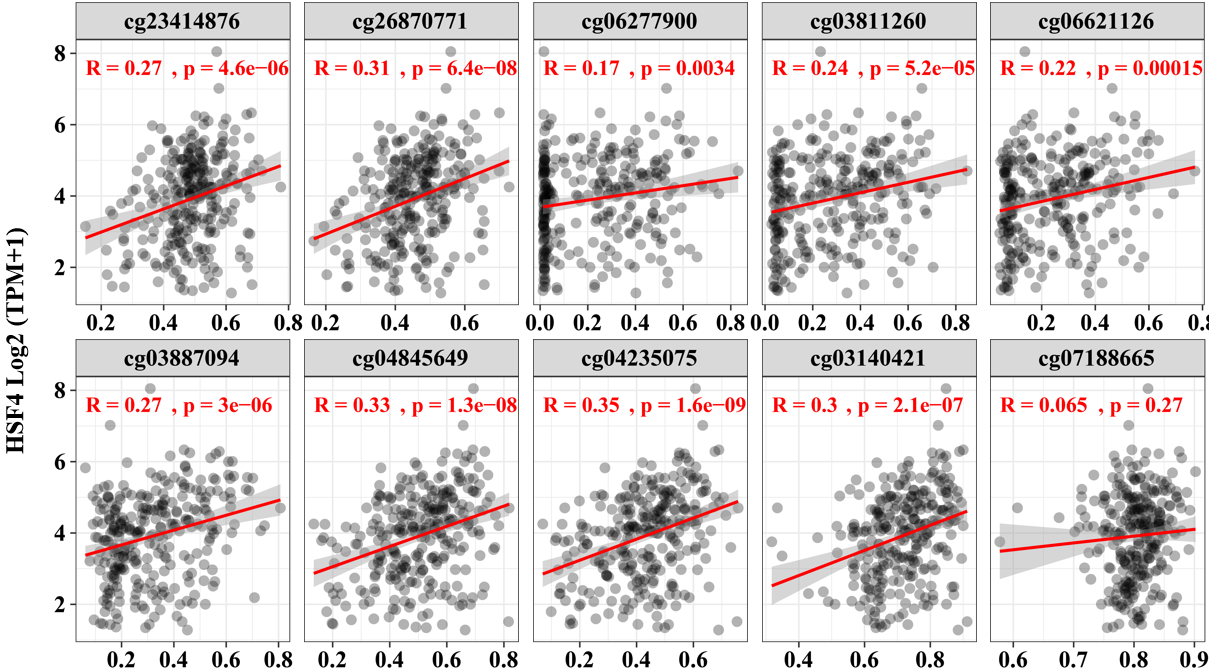
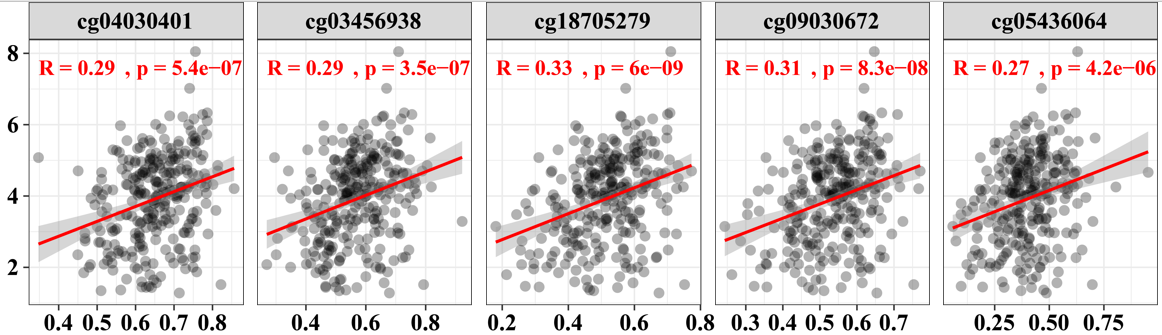
**Figure Legends**



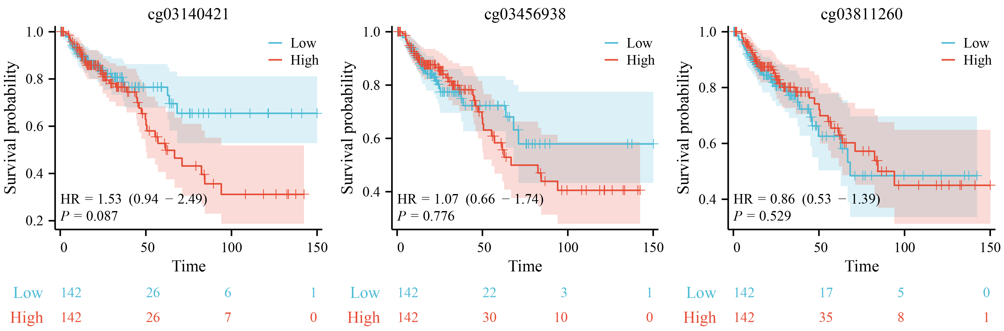
**Figure 1 Pan-cancer analysis of heat shock factor protein 4 methylation levels.** A: Schematic representation of the distribution of eat shock factor protein 4 (*HSF4*) methylated CpG loci. B: Differential analysis of the β values of 19 CpG methylation loci of *HSF4* in multiple malignancies. a*P* < 0.05, b*P* < 0.01, c*P* < 0.001, and d*P* < 0.0001; ns: No significant difference.

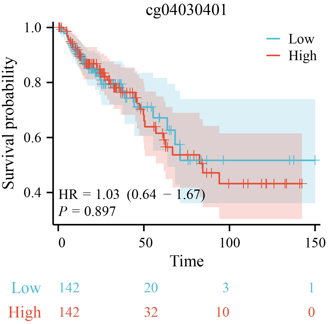
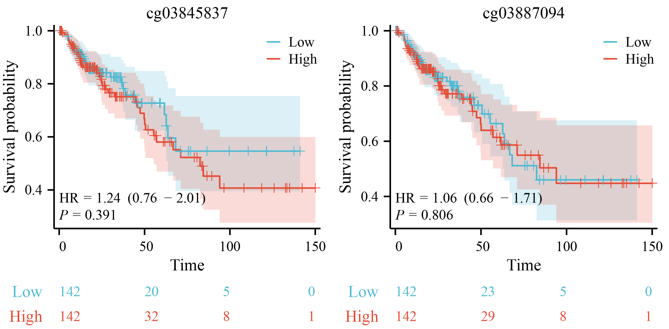


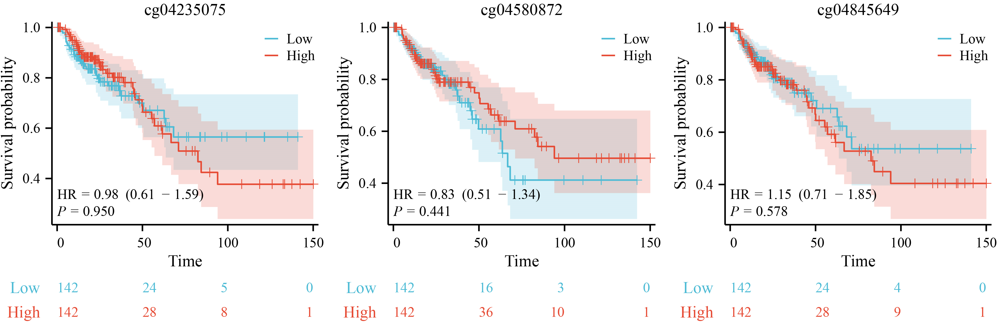
**Figure 2 Differential analysis of β values of 19 probes related to heat shock factor protein 4 methylation in** **colon adenocarcinoma and paracancerous tissues.** The β values of all 19 probes were significantly increased in the tissues of colon adenocarcinoma (COAD) patients. Black is paracancerous tissue, and red is COAD tissue.

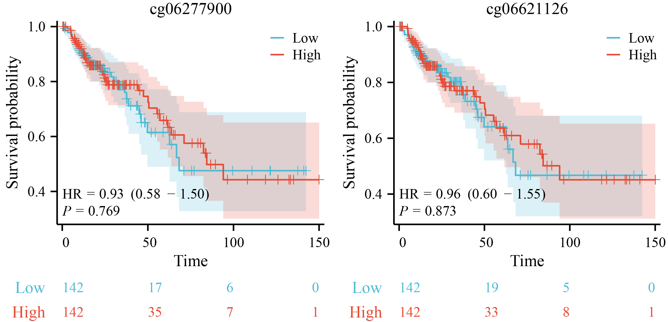
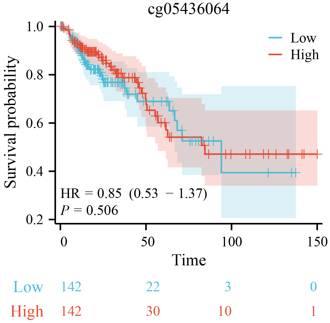


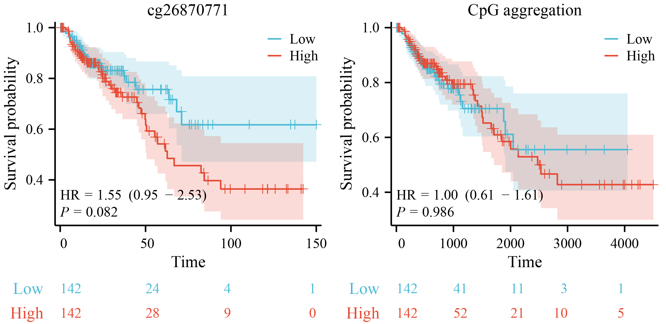
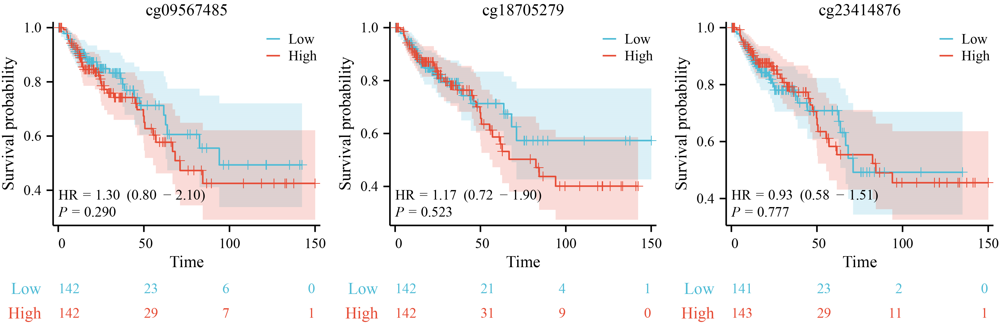
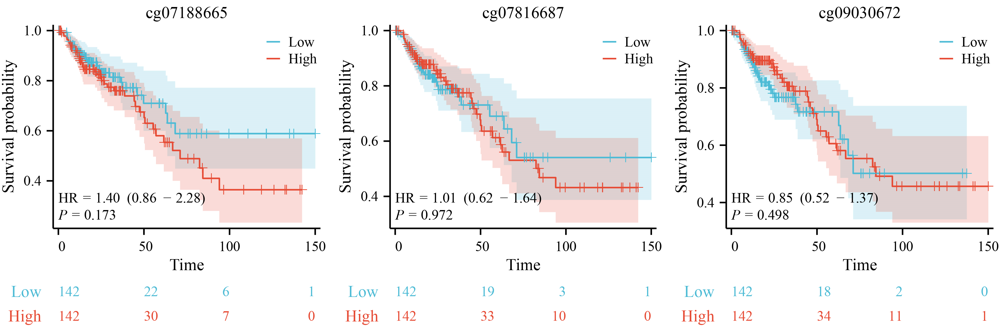
**Figure 3 Correlation analysis of heat shock factor protein 4 expression and heat shock factor protein 4 methylation levels.** The β values of all 19 probes exhibited a significant positive correlation with the expression of eat shock factor protein 4 (*HSF4*) mRNA. The x-axis is the β value of 19 probes, and y-axis is log2 (TPM + 1) of *HSF4* mRNA.



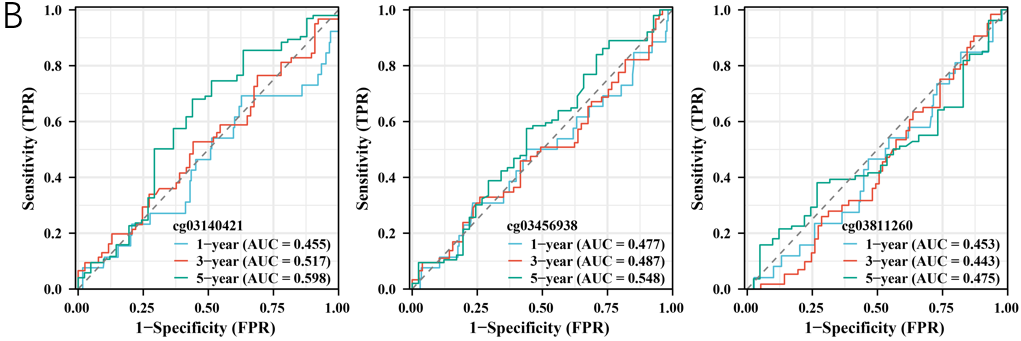
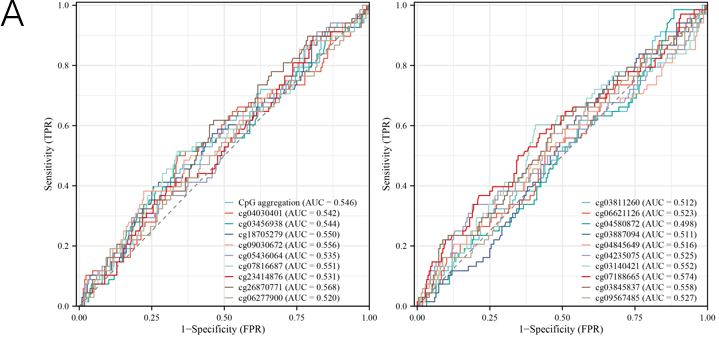


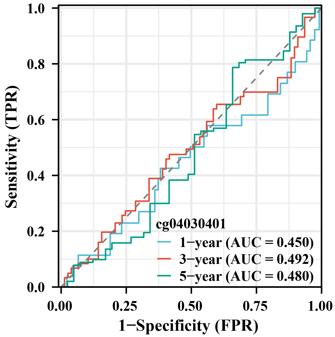
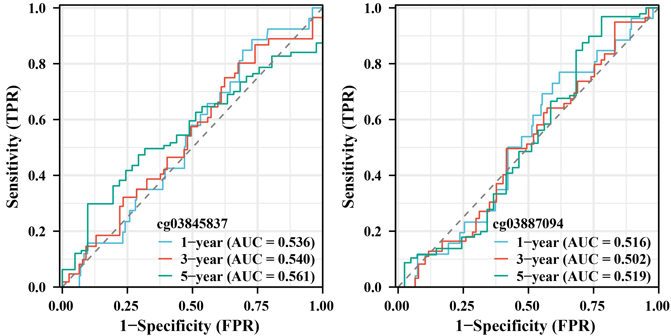


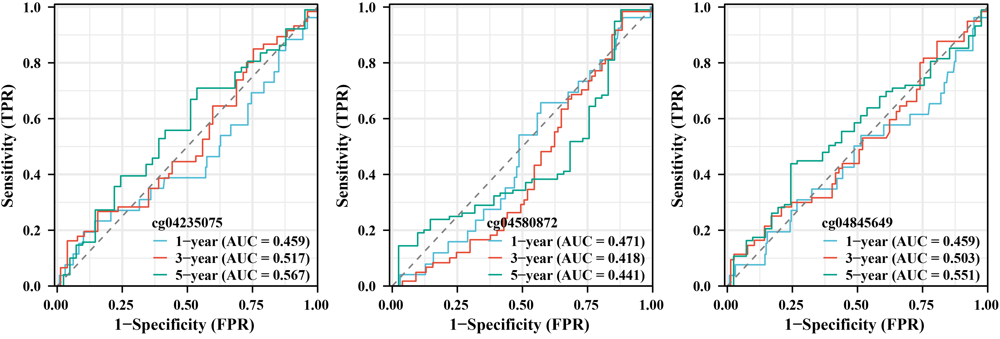


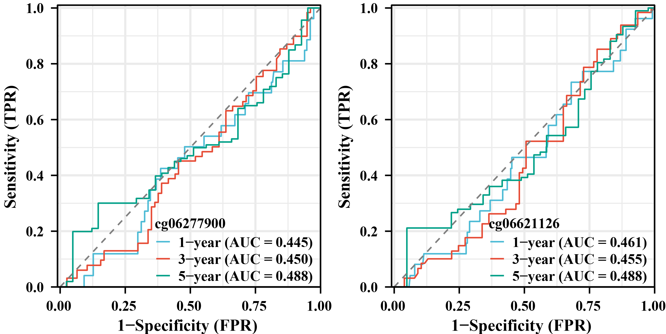
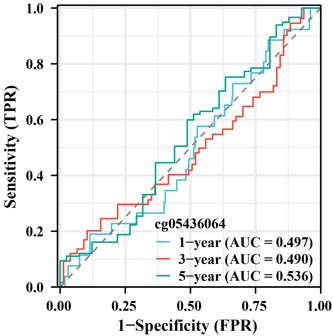


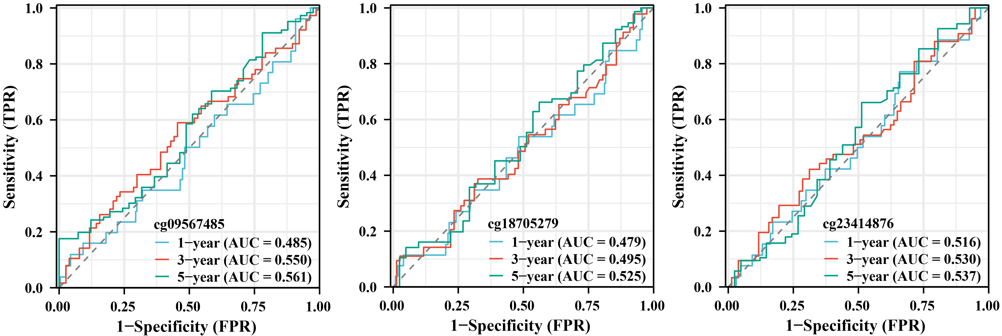
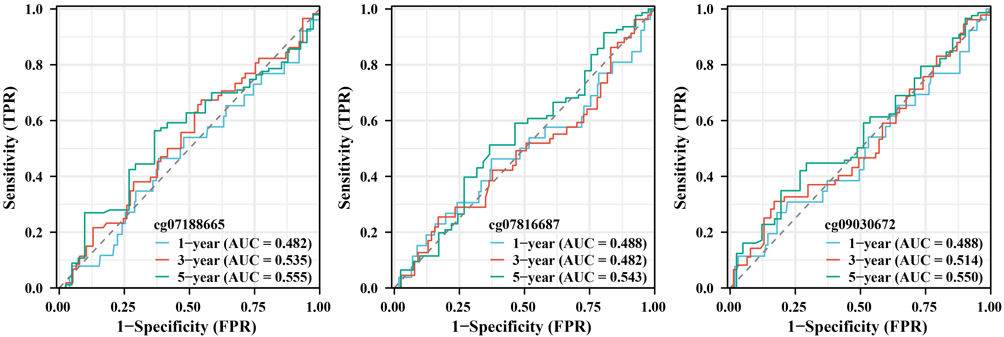
**Figure 4 Correlation analysis of heat shock factor protein 4 methylation and prognosis of patients with colon adenocarcinoma.** Kaplan Meier survival curves illustrating the survival of colon adenocarcinoma patients with high and low beta values for the 19 probes. The blue curve represents the cohort with low β values, and the red curve stands for the cohort with high β values.

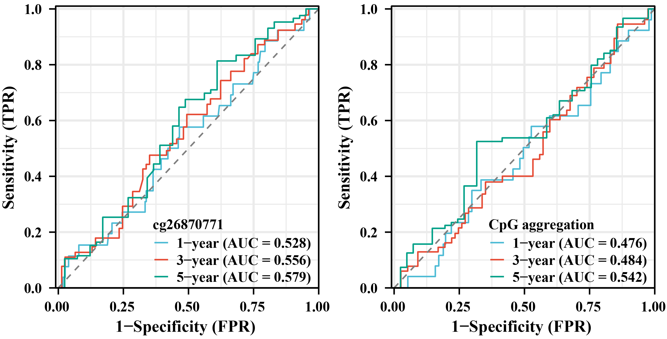




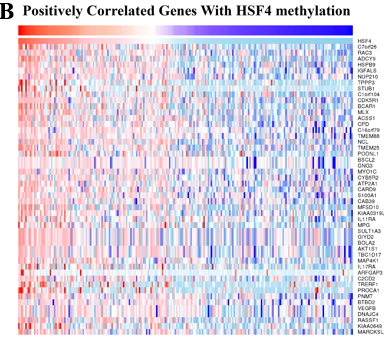
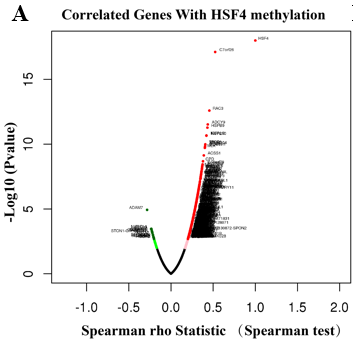


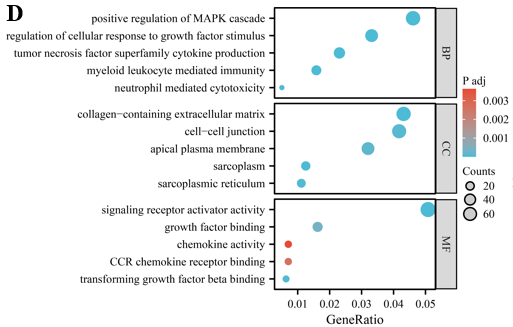
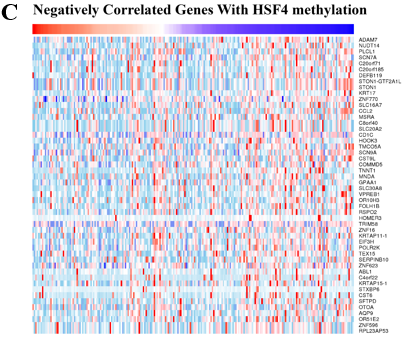


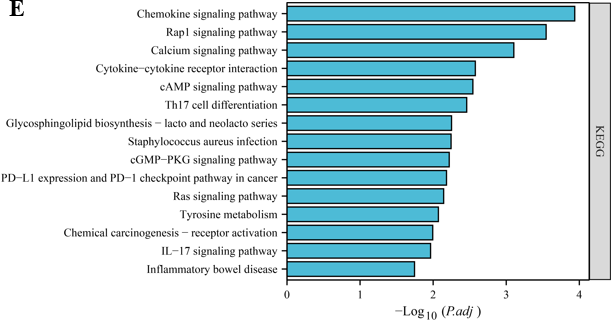




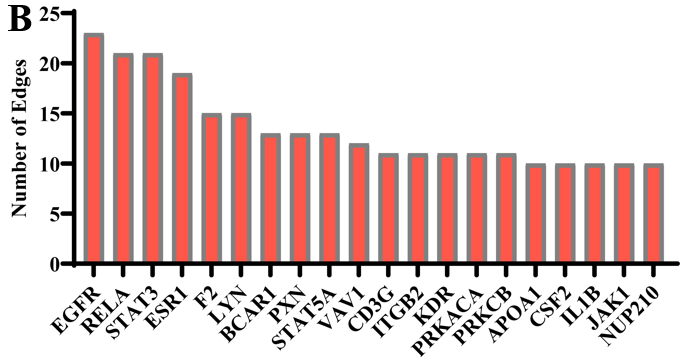
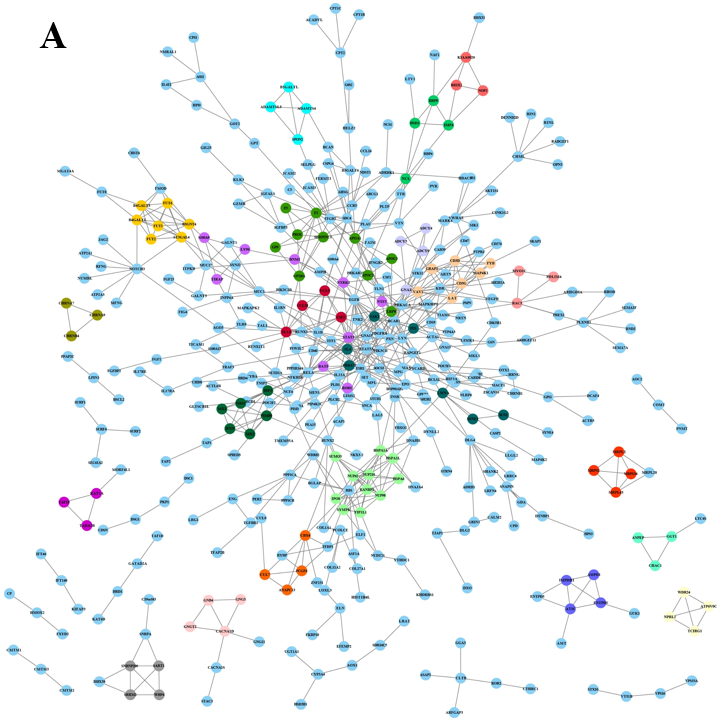
**Figure 5 Correlation analysis of heat shock factor protein 4 methylation and diagnosis in patients with colonic adenocarcinoma.** A: Receiver operating characteristic (ROC) curves exhibiting the diagnostic value of 19 heat shock factor protein 4 (*HSF4*) methylation-associated probes in colon adenocarcinoma patients. B: Time-dependent ROC curves displaying the area under the curve of *HSF4* methylation at 1, 3 and 5 years. AUC: Area under the curve; TPR: True positive rate; FPR: False positive rate.

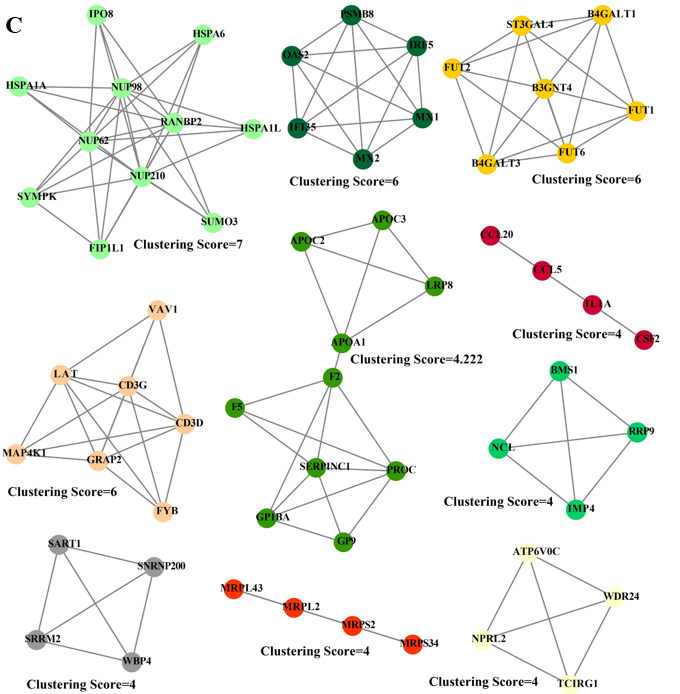




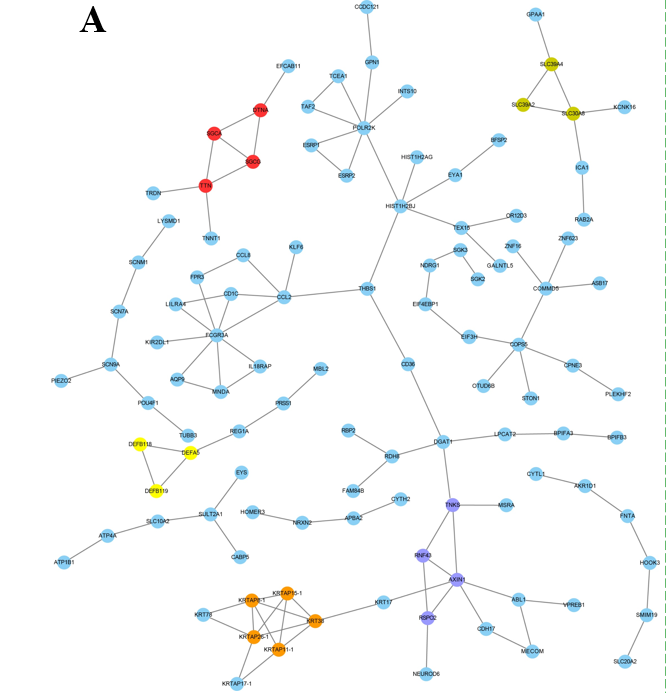


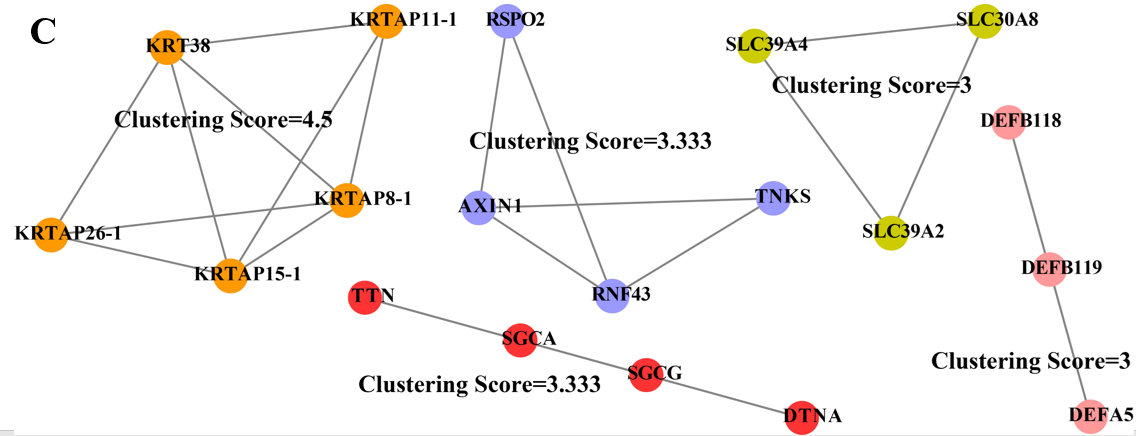
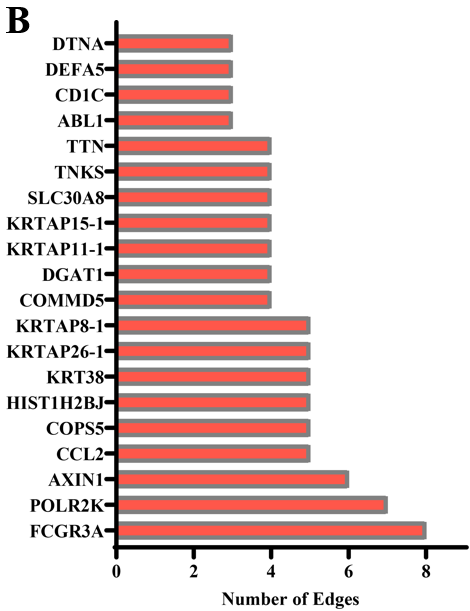
**Figure 6 Identification of heat shock factor protein 4 methylation-related genes and their enrichment analysis in colorectal cancer.** A: Volcano plot showing genes positively and negatively associated with heat shock factor protein 4 (*HSF4*) methylation in colorectal cancer. B, C: Expression profiles of the top 50 genes ranked by absolute correlation coefficient of *HSF4* methylation-related genes. B is the expression profile of genes positively associated with *HSF4* methylation; C is the expression profile of genes negatively associated with *HSF4* methylation. D: Bubble plots exhibiting the GO enrichment results of all *HSF4* methylation-related genes. E: Possible pathways involved in *HSF4* methylation-related genes obtained by Kyoto Encyclopedia of Genes and Genomes enrichment analysis. HSF4: Heat shock factor protein 4; BP: Biological process; CC: Cell component; MF: Molecular function; PD-L1: Programmed cell death-Ligand 1; PD-1: Programmed death 1; IL-17: Interleukin-17.



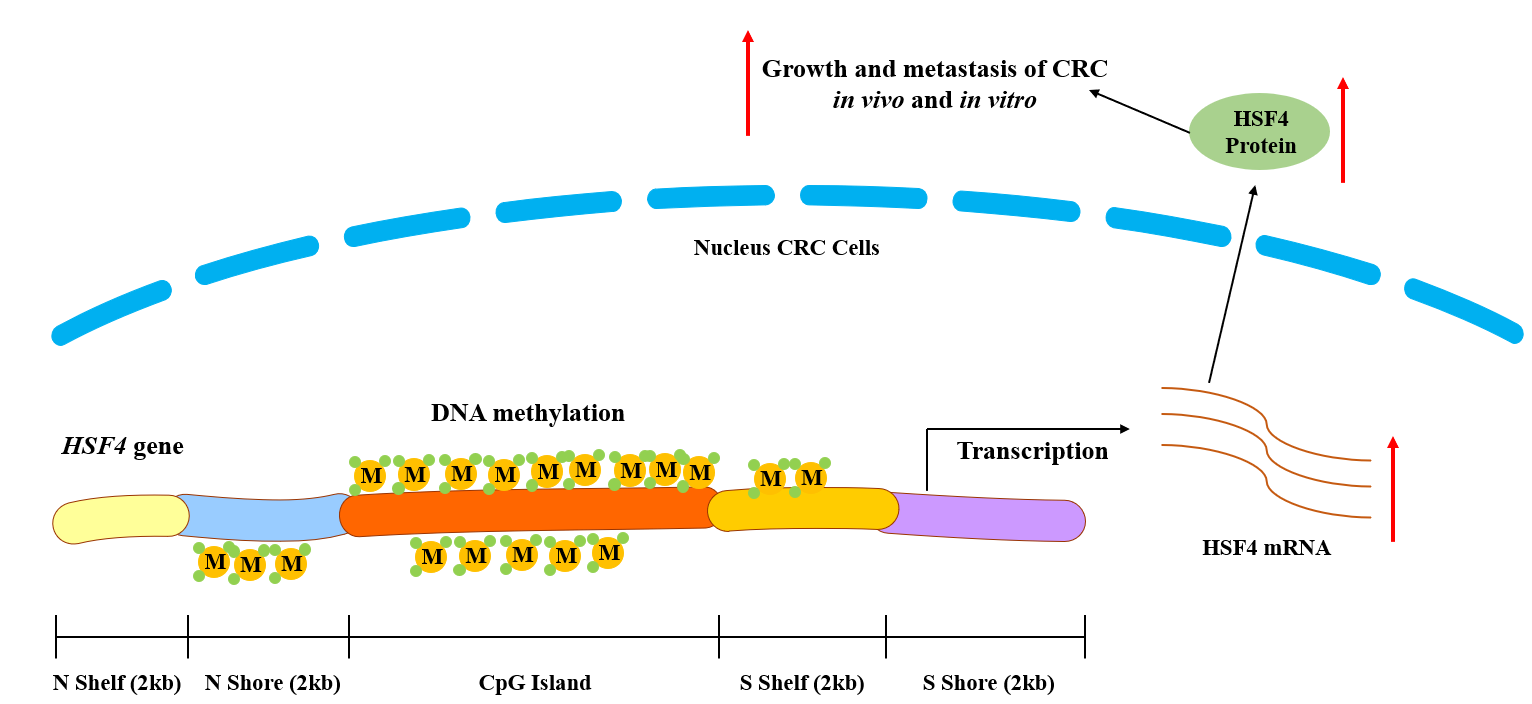


**Figure 7 Protein-protein interaction network construction of h****eat shock factor protein 4 methylation positively associated genes.** A: Protein-protein interaction (PPI) network of heat shock factor protein 4 methylation positively related genes constructed based on String database and MCODE algorithm. B: Top 20 genes in PPI network in terms of edge number. C: The top 10 clustering networks in terms of clustering scores obtained by the MCODE algorithm.





**Figure 8 Protein-protein interaction network construction of heat shock factor protein 4 methylation negatively associated genes.** A: Representative images of the protein-protein interaction (PPI) network of heat shock factor protein 4 methylation negatively associated genes. B: The bar chart displaying edge number of each gene in the PPI network. C: Gene composition and interactions of clustering networks obtained by the MCODE algorithm.



**Figure 9 Diagrammatic representation of this study.** CRC: Colorectal cancer; HSF4: Heat shock factor protein 4.

**Table 1 Basic information of the dataset in this study**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Web** | **Sample source** | **Sample type** | **Platform** | **Samples number** |
| SMART | TCGA\_COAD | Tissue | Methylation 450K | Normal = 34, tumor = 288 |
| SMART | TCGA\_32 cancer types | Tissue | Methylation 450K | Normal = 676, tumor = 8604 |
| LinkedOmics | TCGA\_COADREAD | Tissue | Methylation 27K | Tumor = 233 |