Transcription Factors E2F2/3/4 As Possible Colorectal Cancer Prognostic Biomarkers

Wen-Da Wang

Department of Colorectal and Anal Surgery, Shanxi Cancer hospital

Li-Chun Wang

Department of Colorectal and Anal Surgery, Shanxi Cancer hospital

Mao-Xi Liu

Department of Colorectal and Anal Surgery, Shanxi Cancer hospital

Yan-Jun Lu

Department of Colorectal and Anal Surgery, Shanxi Cancer hospital

Bo Jiang (✉ 13834567839@163.com)

Department of Colorectal and Anal Surgery, Shanxi Cancer hospital

Research Article

Keywords: E2Fs, prognosis, colorectal cancer, bioinformatics

Posted Date: June 2nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-528292/v1>

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

[Read Full License](#)

Abstract

Background: Prognostic biomarkers remain a focus in colorectal cancer during last decades. There are some studies of E2F transcription factors in different cancer. But the utility of E2F transcription factors as cancer biomarkers in colorectal cancer has not been confirmed. In this research we analyzed multiple databases and performed experimental validation of E2Fs as biomarkers.

Methods: Bioinformatics analysis was the main part of this research. Besides, IHC was used to the human normal and tumor tissues.

Results: Our study showed that the mRNA expression levels of E2F1/3/4/5/6/7/8 were significantly upregulated in carcinoma tissues, whereas E2F2 mRNA was downregulated. High E2F3/4 expression correlated with poor prognosis, whereas high E2F2 expression was correlated with good prognosis. Pathological parameters indicated that expression levels of E2F1/3/5 varied in different tumor stages; however, expression of E2F2/4/6/7/8 did not vary with tumor stage. E2F4 expression was higher in men than women. For pathological type, E2F8 expression was higher in the mucinous adenocarcinoma group compared with the adenocarcinoma group. E2F expression levels were altered in 175 of 524 patients (33%) with colorectal cancer, and the alterations of E2Fs in mucinous adenocarcinoma were more frequent than in adenocarcinoma. Cell cycle was the core function and the most enriched pathway of E2Fs. The functions of E2F2/3/4 may be regulated by some miRNAs.

Conclusions: We proved that E2F2/3/4 can be used as prognostic biomarkers for colorectal cancer, and E2F2/3/4 mRNAs may be targeted by some miRNAs to influence “CELL CYCLE” pathway and exert their core functions in colorectal cancer.

Introduction

The E2F transcription factors comprise an eight-member family. E2Fs contain one or more conserved DNA binding domains that interact with target promoters to regulate gene expression(1, 2). E2Fs comprise two opposing functional classes: activators (E2F1–3) and repressors (E2F4-8). In addition, the activities of E2Fs are context-dependent(3, 4). Most E2Fs and their target genes coordinate the oscillatory nature of the cell cycle(5). However, E2Fs also participate in cell proliferation, tissue homeostasis, differentiation, angiogenesis, metabolism, autophagy, mitochondrial functions, DNA damage response, apoptosis, and tumorigenesis(6–11). Aberrant expression of E2Fs and alterations in E2F function coincide with poor prognosis in cancers; these findings emphasize the importance of E2Fs in the cancer phenotype(12).

Colorectal cancer (CRC) is the third most diagnosed malignant tumor and the second leading cause of mortality worldwide(13). Despite improved early diagnosis and treatment strategies (surgery, chemotherapy, radiotherapy, target therapy), the prognosis of CRC is still unsatisfactory(14). Because of CRC tumor heterogeneity, a broader and deeper understanding of the occurrence and progression mechanisms of CRC is required for individualized treatment and improvement in prognosis.

Deregulated expression of E2Fs is a common phenomenon in human cancer(15). This deregulated expression may promote or suppress tumor progression depending on cell type and tissue context. Most CRC studies have focused on only a single E2F and its important function in CRC progression(16–20). Still absent is a comprehensive analysis of the simultaneous activities of all the E2Fs in CRC progression and development. Presently, big data and bioinformatics technology have provided novel means to examine cancer mechanisms. In this study, we used public datasets and bioinformatics technologies to investigate the functions of E2Fs in CRC. Our findings provide improved understanding of the functions of E2Fs in CRC, and our report will contribute to identifying clinical implications for diagnosis, prognosis, and the treatment strategy design.

Materials And Methods

2.1 Oncomine database analysis

The Oncomine database (<http://www.oncomine.org>)(21, 22) is a web-based data mining platform that collects, standardizes, analyzes, and delivers transcriptomic cancer data for biomedical research. We used the Oncomine database to compare the mRNA expression of E2Fs in various types of cancers and their matched normal tissues. The threshold was determined as the following: $p\text{-value} \leq 1\text{E-}4$, fold change ≥ 2 .

2.2 TIMER database analysis

TIMER (<https://cistrome.shinyapps.io/timer/>) is a web-based data mining platform for systematic analysis of immune infiltration levels and gene expression(23). We used the TIMER mRNA expression data of E2Fs in CRC and compared them between cancer tissues and matched normal tissues. A $p\text{-value} < 0.01$ was considered statistically significant.

2.3 GEPIA database analysis

Gene Expression Profiling Interactive Analysis (GEPIA) (<http://gepia.cancer-pku.cn/index.html>) was used to generate survival curves (OS and DFS) based on RNA sequencing-determined expression from the TCGA database(24). We used GEPIA to identify differences in E2F mRNA expression in CRC and matched normal tissues. A $p\text{-value} < 0.05$ was considered statistically significant. GEPIA was also used to identify correlations between E2F mRNA expression and survival in CRC. Hazard ratio and log-rank $p\text{-value}$ were measured, and a $p\text{-value} < 0.05$ was considered statistically significant. To identify the correlation between E2Fs in CRC, we obtained the R and $p\text{-values}$; a $p\text{-value} < 0.01$ was considered statistically significant. The absolute value of R from 0-0.09 indicated no correlation, 0.1-0.3 was weak correlation, 0.3-0.5 was medium correlation, and 0.5-1.0 was strong correlation. To profile the correlation of E2F mRNA expression of E2Fs in CRC according to pathological stages, a $p\text{-value} < 0.05$ was considered statistically significant.

2.4 Immunohistochemistry staining

Two pairs of colon tissues and surrounding adjacent tissues were obtained from patients diagnosed with colon cancer in the Shanxi Cancer hospital. The Shanxi Medical University Medical Research Ethics Review Committee approved this study (2018LL288). The colon tissues were fixed by immersion in 10% neutral buffered formalin at ambient temperature for 24 h and processed for paraffin embedding. Sections of 5 µm were obtained for IHC staining. The sections were deparaffinized in xylene and rehydrated in graded alcohol baths. EDTA antigenic retrieval used 3% hydrogen peroxide in methanol to quench the endogenous peroxidase activity and incubation with 3% bovine serum albumin to block the nonspecific binding to antigen retrieval. The following antibodies were incubated with the tissues overnight at 4°C: anti-E2F1 (1:100; ab179445, Abcam), anti-E2F2 (1:100; ab235837, Abcam), anti-E2F3 (1:100; ab50917, Abcam), anti-E2F4 (1:100; ab150360, Abcam), anti-E2F5 (1:100; GTX80623, GeneTex), anti-E2F6 (1:100; ab53061, Abcam), anti-E2F7 (1:100; ab56022, Abcam), E2F8 Polyclonal Antibody (1:100; PAS-100878, Thermo). After incubation with the primary antibodies, the samples were incubated with a horseradish-peroxidase (HRP)-conjugated secondary antibody. Afterward, the expression of E2Fs in tissue slides was assessed by two pathologists. The staining outcomes were assessed as the intensity on a scale of 0 to 3, with 0 (no staining), 1 (weak staining), 2 (moderate staining), and 3 (strong staining). A semi-quantitative assessment of positive tumor cell percent was used, with a scale of 0 to 4; a score of 0 = none, 1 = 1–25%, 2 = 26–50%, 3 = 51–75%, and 4 = >75%. A histochemical score (H-score) of staining was calculated by multiplying the IHC staining intensity and tumor percent scores.

2.5 PrognoScan database analysis

PrognoScan (<http://www.abren.net/PrognoScan/>) is an online database with clinical annotation and a web-based tool for assessing the biological relationships between gene expression and prognostic information such as OS, DSS, and DFS in various types of cancers(25). In addition, this tool automatically calculates p-value, hazard ratio, and 95% confidence intervals based on a particular gene expression. We used the PrognoScan database to identify correlations between E2F mRNA expression and survival from CRC with the adjusted cox p-value < 0.05.

2.6 TCGA database analysis

The expression of E2F mRNAs in CRC was selected for further analyses with cBioPortal ([http://www.cbioportal.org/index.do?Session id=5b4c1773498eb8b3d566f7b8](http://www.cbioportal.org/index.do?Session+id=5b4c1773498eb8b3d566f7b8)). The genomic profiles included mutations, putative copy number alterations, and mRNA expression Z scores (RNA-seq v.2 RSEM). Genes co-expressed with E2Fs were calculated according to the cBioPortal's online instructions and analyses using the STRING (www.string-db.org/) database. Finally, the results of the PPI network were displayed with CentiScaPe 2.2.

2.7 GO functional annotation and KEGG pathway enrichment analyses

The Database for Annotation, Visualization, and Integrated Discovery (DAVID) v.6.8 (<https://david.ncifcrf.gov>)(26) was used to perform GO(27) functional annotation and KEGG pathway

enrichment analyses of genes co-expressed with E2Fs. The human genome was selected as the background list parameter, and a p-value < 0.05 was set as statistical significance.

2.8 StarBase analysis

StarBase v3.0 (<http://starbase.sysu.edu.cn/>) is an open-source platform for exploring mRNA target regulatory association pairs, which are verified by experiments and predicted by seven programs, PITA, RNA22, miRmap, microT, miRanda, PicTar, and TargetScan(28). In our study, the E2F2, E2F3, and E2F4 mRNA targets were examined according to the following standards: Degradome Data (low stringency), Pan-cancer (1 cancer type) and Program Number (five programs: PITA, miRmap, miRanda, PicTar and TargetScan).

2.9 GEO dataset analysis

The Gene Expression Omnibus (GEO) repository(29) distributes microarray, next-generation sequencing, and other forms of high-throughput functional genomic data, which also helps users to analyze GEO data by GEO2R, an R-based web application(30). The difference of miRNA between CRC tissues and corresponding normal tissues was analyzed by the GEO database (GSE 115513), (<http://www.ncbi.nlm.nih.gov/geo/>). A p-value < 0.05 and absolute value of fold change ≥ 0.58 ($\log_{2}FC \geq 1.5$) were set as statistical significance. The common miRNAs of the StarBase analysis and the GEO dataset analysis were recognized as validated targets.

Results

3.1 The mRNA expression levels of E2Fs in CRC

To compare E2F expression in tumor and normal tissues, we extracted from the Oncomine, TIMER, and GEPIA databases the E2F mRNA levels in multiple cancer types. In one or more of the datasets, the mRNA expression levels of E2F1, E2F3, E2F4, E2F5, E2F6, E2F7, and E2F8 were significantly upregulated in CRC patients. The mRNA expression levels of E2F2 were significantly downregulated in CRC patients based on the Oncomine and TIMER datasets. See summary in Figure 1, Figure 2A, and Figure 2B.

3.2 The IHC results of E2Fs in CRC

We obtained the IHC results of E2Fs in human colon tissues. In parallel with the increased mRNA expression levels as revealed in the Oncomine, TIMER, and GEPIA datasets, the IHC expression levels of E2F1, 3, 4, 5, 6, 7 and 8 were significantly higher in carcinoma tissues (Figure 3).

Positive correlations between E2F expression levels were the following (Figure 4):

E2F1 with E2F2-E2F8

E2F2 with E2F3, E2F4, E2F7, and E2F8

E2F3 with E2F4, E2F5, E2F6, E2F7, and E2F8

E2F4 with E2F5, E2F6, E2F7, and E2F8

E2F5 with E2F6, E2F7, and E2F8

E2F6 with E2F7 and E2F8

E2F7 with E2F8.

E2F2 had a significant negative correlation with E2F5.

3.3 Prognosis values of E2Fs in CRC

The Kaplan-Meier curve and log rank test GEIPA analyses of E2Fs in CRC revealed that increased expression of E2F3 and E2F4 was associated with poor overall survival (OS) ($p < 0.05$) [Figure 5]. The prognosis values of E2Fs in CRC determined by PrognoScan revealed that increased expression of E2F1, E2F2, and E2F7 was associated with improved Disease free survival (DFS), OS, and Disease specific survival (DSS) ($p < 0.05$) [Table 1].

Table 1 The relationships between E2Fs expression and prognosis of CRC in PrognoScan

GENE	DATASET	ENDPOINT	N	P-VALUE	HR [95% CI]
E2F1	GSE17536	Disease Free Survival	145	0.012882	0.05 [0.01 - 0.49]
	GSE17536	Overall Survival	177	0.014295	0.12 [0.02 – 0.66]
	GSE17536	Disease Specific Survival	177	0.024171	0.24 [0.07 – 0.83]
E2F2	GSE17536	Disease Free Survival	145	0.030286	0.23 [0.06 – 0.87]
	GSE17537	Disease Free Survival	55	0.00569	0.00 [0.00 – 0.08]
	GSE17537	Overall Survival	55	0.011883	0.02 [0.00 – 0.43]
	GSE17537	Disease Specific Survival	49	0.039668	0.02 [0.00 – 0.83]
E2F7	GSE14333	Disease Free Survival	226	0.038041	0.67 [0.45 – 0.98]
	GSE17536	Overall Survival	177	0.044223	0.29 [0.08 – 0.97]
	GSE17537	Disease Specific Survival	49	0.000754	0.00 [0.00 – 0.03]
	GSE17537	Disease Free Survival	55	0.015361	0.01 [0.00 – 0.42]
	GSE17537	Overall Survival	55	0.021599	0.02 [0.00 – 0.55]

3.4 The correlations of the E2Fs expression and clinicopathological parameter

The correlations of E2Fs with tumor stage indicated that expression of the E2F1, E2F3, and E2F5 group varied as a function of tumor stage; however, expression of the E2F2, E2F4, E2F6, E2F7, and E2F8 groups did not vary with tumor stage [Figure 6 A]. Further, we did not find any significant differences in expression of the E2Fs between the non-metastasis and metastasis groups [Figure 6 B]. E2F4 expression was higher in the male group compared with the female group [Figure 6 C]. E2F8 expression was higher in the mucinous adenocarcinoma group compared with the adenocarcinoma group [Figure 6 D].

3.5 Alterations of E2Fs in CRC

E2Fs expression levels were varied in 175 of 524 patients with CRC (33 %), and the alterations of E2Fs in mucinous adenocarcinoma were more frequent than alterations in the adenocarcinomas. In addition, E2F1 had the highest mutation frequency with amplification [Figure 7]. The network constructed by E2Fs and the 69 most frequently altered co-expressed genes showed that cell cycle-related genes were closely associated with E2Fs alterations [Figure 8].

3.6 GO functional annotation and KEGG pathway enrichment analyses

GO analysis of the functions of E2Fs and co-expressed genes showed that the major E2F functions were regulation of cell cycle, cell division, cell proliferation, DNA replication, and DNA repair [Tables 1, 2, and 3]. KEGG pathway enrichment analysis showed Cell Cycle to be the most enriched pathway [Figure 9].

Table 2 The GO function annotation (BP) of E2Fs and genes significantly associated with E2Fs alterations of CRC analyzed by DAVID

ID	DESCRIPTION	GENES	P-VALUE
GO:0051726	regulation of cell cycle	CCL7, CCL8, CXCL11	1.18E-05
GO:0006260	DNA replication	E2F2, E2F5, DTL, FOXM1, CCNF, PKMYT1, MYBL2	4.17E-05
GO:0051301	cell division	RECQL4, DTL, RRM1, BRIP1, MCM10, ORC1, CDK2	9.08E-05
GO:0006977	DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	KIF2C, CDCA8, NCAPH, SGO1, CCNF, KATNB1, BUB1B, UBE2C, CDK2	1.22E-04
GO:0007067	mitotic nuclear division	KIF2C, SGO1, CCNF, KIF15, PKMYT1, BUB1B, CDK2	5.38E-04
GO:0007062	sister chromatid cohesion	KIF2C, CDCA8, SGO1, KIF18A, BUB1B	8.50E-04
GO:0060707	trophoblast giant cell differentiation	PLK4, E2F7, E2F8	0.001069
GO:0000086	G2/M transition of mitotic cell cycle	PLK4, FOXM1, PKMYT1, MELK, CDK2	0.00243234563684612
GO:0008283	cell proliferation	XRCC5, KIF2C, E2F8, KIF15, BUB1B, MCM10, MELK	0.00388806554484318
GO:0000083	regulation of transcription involved in G1/S transition of mitotic cell cycle	E2F4, E2F6, ORC1	0.0039817619547411
GO:0007018	microtubule-based movement	KIF23, KIF2C, KIF15, KIF18A	0.004506
GO:0006890	retrograde vesicle-mediated transport, Golgi to ER	KIF23, KIF2C, KIF15, KIF18A	0.004664
GO:0019886	antigen processing and presentation of exogenous peptide antigen via MHC class II	KIF23, KIF2C, KIF15, KIF18A	0.006424
GO:0032877	positive regulation of DNA endoreduplication	E2F7, E2F8	0.008202
GO:0000082	G1/S transition of mitotic cell cycle	PKMYT1, MCM10, ORC1, CDK2	0.00853075482116703
GO:0001890	placenta development	E2F7, E2F8, CCNF	0.009072
GO:0007080	mitotic metaphase plate congression	KIF2C, CDCA8, KIF18A	0.010101

GO:0007049	cell cycle	E2F2, E2F3, FOXM1, SUV39H1, CCAR1	0.0121896039954393
GO:0071930	negative regulation of transcription involved in G1/S transition of mitotic cell cycle	E2F1, E2F7	0.012277
GO:0000398	mRNA splicing, via spliceosome	RALY, PRPF4B, RNPS1, CPSF3, CCAR1	0.013159
GO:0032508	DNA duplex unwinding	XRCC5, RECQL4, BRIP1	0.014085
GO:0006281	DNA repair	RECQL4, FANCI, FOXM1, RAD54L, CDK2	0.015907
GO:1990086	lens fiber cell apoptotic process	E2F1, E2F2	0.016337
GO:0070345	negative regulation of fat cell proliferation	E2F1, E2F3	0.02038
GO:0060718	chorionic trophoblast cell differentiation	E2F7, E2F8	0.02038
GO:0006351	transcription, DNA-templated	E2F1, XRCC5, RALY, UTP4, E2F3, E2F4, E2F5, E2F6, FOXM1, E2F7, E2F8, SUV39H1, RNPS1, ZBTB24, CCAR1	0.026128
GO:0032466	negative regulation of cytokinesis	E2F7, E2F8	0.028417
GO:0006302	double-strand break repair	XRCC5, RECQL4, BRIP1	0.030147
GO:0051436	negative regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle	BUB1B, UBE2C, CDK2	0.034473
GO:0000733	DNA strand renaturation	RECQL4, RAD54L	0.036388
GO:0007019	microtubule depolymerization	KIF2C, KIF18A	0.04035
GO:0070365	hepatocyte differentiation	E2F7, E2F8	0.04035
GO:0006310	DNA recombination	XRCC5, RECQL4, RAD54L	0.045755
GO:0006396	RNA processing	PNPT1, U2SURP, SSB	0.06037
GO:0006406	mRNA export from nucleus	NUP153, RNPS1, CPSF3	0.063686
GO:0007099	centriole replication	PLK4, CDK2	0.063785
GO:0000122	negative regulation of transcription from RNA polymerase II promoter	E2F1, E2F6, E2F7, FOXM1, E2F8, SUV39H1, DNAJA3	0.074986

GO:0051276	chromosome organization	CDCA8, RAD54L	0.075291
GO:0016925	protein sumoylation	NUP153, CDCA8, SENP6	0.083562
GO:0051439	regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle	UBE2C, CDK2	0.090414
GO:0071398	cellular response to fatty acid	XRCC5, E2F1	0.090414
GO:0007126	meiotic nuclear division	RAD54L, CDK2	0.094157
GO:0002040	sprouting angiogenesis	E2F7, E2F8	0.097885

Table 3 The GO function annotation (CC) of E2Fs and genes significantly associated with E2Fs alterations of CRC analyzed by DAVID

ID	DESCRIPTION	GENES	P-VALUE
GO:0005654	nucleoplasm	E2F1, XRCC5, KIF23, GPN1, E2F2, UTP4, E2F3, E2F4, PRPF4B, E2F5, E2F7, FOXM1, U2SURP, PKMYT1, COPS8, MYBL2, MCM10, GTSE1, SENP6, CDCA8, FANCI, EIF3E, ORC1, CCAR1, NUP153, DTL, SUV39H1, BRIP1, RNPS1, UBE2C, RAD54L, CDK2, ADRM1, SGO1, RRM1, NUTF2, CPSF3, WDR43	4.96E-13
GO:0005634	nucleus	E2F1, RALY, XRCC5, KIF23, E2F3, E2F4, E2F5, E2F6, E2F7, FOXM1, E2F8, U2SURP, KATNB1, COPS8, MCM10, KIF2C, CDCA8, NCAPH, NUDCD1, EIF3E, ORC1, DNAJA3, RECQL4, SPATA33, DTL, CCNF, SUV39H1, KIF18A, BRIP1, SSB, RNPS1, RAD54L, CDK2, ZBTB24, ARMC1, EMC8, ADRM1, SGO1, CPNE3, MELK, ATAD2B	2.61E-06
GO:0005667	transcription factor complex	E2F2, E2F3, E2F4, E2F5, E2F6, E2F7, E2F8, CDK2	1.01E-05
GO:0000776	kinetochore	KIF2C, SGO1, KIF18A, BUB1B, WDR43	2.81E-04
GO:0005813	centrosome	KIF23, PLK4, DTL, SGO1, KIF15, KATNB1, WDR43, CDK2	0.001332
GO:0000775	chromosome, centromeric region	KIF2C, CDCA8, SGO1, SUV39H1	0.001426
GO:0005730	nucleolus	XRCC5, UTP23, UTP4, NUP153, PLK4, CDCA8, E2F5, DTL, MCM10, WDR43, ORC1	0.001716
GO:0005829	cytosol	SPATA33, KIF23, XRCC5, KIF15, KIF18A, PKMYT1, RNPS1, UBE2C, GTSE1, CDK2, KIF2C, NCAPH, PLK4, CDCA8, SGO1, EIF3E, RRM1, CTU2, BUB1B, CPNE3, NUTF2, ARHGAP11A, ORC1, DNAJA3	0.002805
GO:0005737	cytoplasm	GPN1, E2F3, E2F5, FOXM1, PNPT1, KATNB1, COPS8, MCM10, ACP1, SENP6, NUDCD1, FANCI, EIF3E, CTU2, ORC1, DNAJA3, RECQL4, NUP153, DTL, KIF18A, BRIP1, RNPS1, UBE2C, CDK2, ARMC1, EMC8, ADRM1, SGO1, RRM1, BUB1B, CPNE3, SHCBP1	0.004871
GO:0005874	microtubule	KIF23, KIF2C, KIF15, KIF18A, KATNB1, WDR43	0.007286
GO:0005694	chromosome	RECQL4, UTP4, PRPF4B, DTL	0.007999
GO:0030496	midbody	KIF23, CDCA8, KATNB1, SHCBP1	0.013959
GO:0005871	kinesin complex	KIF23, KIF2C, KIF18A	0.018145
GO:0001650	fibrillar center	UTP4, WDR43	0.023152
GO:0016020	membrane	XRCC5, RECQL4, PNPT1, KIF15, KATNB1, PKMYT1, GTSE1, ADRM1, KIF2C, EMC8, NCAPH, FANCI, EIF3E,	0.04265

		MELK, TMEM30A	
GO:0000777	condensed chromosome kinetochore	KIF2C, SGO1, BUB1B	0.045277
GO:0044613	nuclear pore central transport channel	NUP153, NUTF2	0.049496
GO:0031965	nuclear membrane	NUP153, DTL, BRIP1, NUTF2	0.060074
GO:0051233	spindle midzone	CDCA8, BUB1B	0.071518
GO:0005739	mitochondrion	E2F1, GPN1, EMC8, SLC25A32, BRI3BP, PNPT1, RPUSD1, CTU2, DNAJA3, ARMC1	0.072749
GO:0005819	spindle	KIF23, KIF15, SHCBP1	0.080833
GO:0005680	anaphase-promoting complex	BUB1B, UBE2C	0.085919
GO:0000784	nuclear chromosome, telomeric region	XRCC5, SSB, ORC1	0.091333
GO:0015630	microtubule cytoskeleton	KIF2C, KIF18A, KATNB1	0.099761

Table 4 The GO function annotation (MF) of E2Fs and genes significantly associated with E2Fs alterations of CRC analyzed by DAVID

ID	DESCRIPTION	GENES	P-VALUE
GO:0005515	protein binding	RALY, KIF23, XRCC5, E2F1, E2F2, E2F3, E2F4, PRPF4B, E2F5, E2F6, E2F7, E2F8, U2SURP, PKMYT1, COPS8, MCM10, UHRF1BP1, GTSE1, SENP6, KIF2C, CDCA8, NUDCD1, FANCI, CTU2, ORC1, CCAR1, DTL, KIF15, CCNF, UBE2C, CDK2, ZBTB24, EMC8, ADRM1, SGO1, RRM1, BUB1B, CPSF3, WDR43, MELK, SHCBP1, GPN1, UTP4, FOXM1, PNPT1, MYBL2, ACP1, NCAPH, EIF3E, DNAJA3, TMEM30A, RECQL4, UTP23, NUP153, VAC14, SUV39H1, KIF18A, BRIP1, SSB, RNPS1, RAD54L, ARMC1, PLK4, NUTF2, CPNE3	2.49E-11
GO:0001047	core promoter binding	E2F1, E2F2, E2F3, E2F7, E2F8, CCAR1	7.69E-06
GO:0005524	ATP binding	RECQL4, KIF23, XRCC5, PRPF4B, KIF15, KIF18A, PKMYT1, BRIP1, UBE2C, RAD54L, CDK2, KIF2C, PLK4, RRM1, BUB1B, ORC1, MELK, DNAJA3, ATAD2B	4.06E-05
GO:0046983	protein dimerization activity	E2F1, E2F2, E2F3, E2F4, E2F5, E2F6	4.49E-04
GO:0044822	poly(A) RNA binding	XRCC5, RALY, UTP23, UTP4, PRPF4B, PNPT1, U2SURP, SSB, RNPS1, EIF3E, CPNE3, WDR43, CCAR1	0.002773
GO:0003677	DNA binding	XRCC5, E2F1, E2F2, NUP153, E2F3, E2F4, E2F5, E2F6, FOXM1, E2F7, KIF15, BRIP1, RAD54L, ZBTB24, FANCI, ORC1	0.004163
GO:0003777	microtubule motor activity	KIF23, KIF2C, KIF15, KIF18A	0.004833
GO:0004674	protein serine/threonine kinase activity	PLK4, PRPF4B, PKMYT1, BUB1B, CPNE3, MELK, CDK2	0.005327
GO:0003723	RNA binding	RALY, PNPT1, RPUSD1, U2SURP, SSB, RNPS1, THUMPD2, CPSF3	0.008637
GO:0008017	microtubule binding	KIF23, KIF15, KIF18A, KATNB1, WDR43	0.012002
GO:0000049	tRNA binding	CTU2, SSB, THUMPD2	0.019992
GO:0003700	transcription factor activity, sequence-specific DNA binding	E2F1, E2F2, E2F3, E2F4, E2F5, E2F6, E2F7, FOXM1, E2F8, MYBL2	0.020575
GO:0008134	transcription factor binding	E2F1, E2F2, E2F4, E2F5, DNAJA3	0.033089
GO:0036310	annealing helicase activity	RECQL4, RAD54L	0.033623
GO:0016887	ATPase activity	KIF23, KIF2C, KIF15, ATAD2B	0.043371

GO:0051010	microtubule plus-end binding	KIF2C, KIF18A	0.045942
GO:0003714	transcription corepressor activity	E2F6, E2F7, E2F8, CCAR1	0.05588
GO:0004672	protein kinase activity	PRPF4B, PKMYT1, BUB1B, MELK, CDK2	0.06716
GO:0045502	dynein binding	KATNB1, WDR43	0.070118

3.7 The mRNA-target of E2F2, E2F3, and E2F4

The E2F2, E2F3, and E2F4 mRNAs were identified as-targets of miRNAs (Figure 10) according to the standards described in Methods (Figure 9 and 10).

3.8 Differential miRNA expression between CRC tissues and corresponding normal tissues

To assess the potential targeting of miRNAs to E2F2, E2F3, and E2F4, we compared the miRNAs between CRC tissues and corresponding normal tissues in the GEO database (GSE 115513). We obtained 148 miRNAs, 90 upregulated and 58 downregulated, between CRC tissues and corresponding normal tissues. The miRNAs to potentially target E2F2 included hsa-miR-17-5p, hsa-miR-93-5p, hsa-miR-20b-5p, hsa-miR-106b-5p, and hsa-miR-20a-5p. The results also showed that E2F3 was a target of hsa-miR-10b-5p and hsa-miR-497-5p. There were no miRNAs to target E2F4 [Table 5].

Table 5 The miRNA-target expression of E2F2/E2F3 between CRC tissues and corresponding normal tissues in GSE115513

GENE	miRNA	adj. P-VALUE	logFC	CANCER	NORMAL
E2F2	has-miR-17-5p	6.92E-148	1.547963	750	654
	hsa-miR-93-5p	2.02E-101	1.169552	750	654
	hsa-miR-20b-5p	1.32E-93	1.465394	750	654
	hsa-miR-106b-5p	8.20E-58	1.024036	750	654
	hsa-miR-20a-5p	4.93E-131	1.605937	750	654
E2F3	hsa-miR-10b-5p	2.19E-20	-0.59542	750	654
	hsa-miR-497-5p	1.30E-26	-0.84106	750	654

Discussion

The E2Fs are encoded by eight genes whose protein products form a core transcriptional axis crucial for coordinating the oscillatory nature of the cell cycle, proliferation, tissue homeostasis, differentiation, angiogenesis, metabolism, autophagy, mitochondrial functions, DNA damage response, apoptosis, and tumorigenesis(6–11). Deregulated expression of E2Fs has been observed in many types of cancers(31–33). In addition, restoring the balance between E2F1 and E2F7 is a therapeutic strategy in head and neck squamous cell carcinomas(34). Although the functions of single E2F genes in tumorigenesis and progression of CRC have been partially confirmed(16–20), a bioinformatics analysis of these transcription factors in CRC had not been performed. Thus, we investigated the mRNA and protein expression, prognostic values, and potential biological functions of E2Fs in CRC that will contribute to identifying clinical implications for diagnosis, prognosis, and treatment strategy.

We found that the mRNA and protein levels of E2F1, E2F3, E2F4, E2F5, E2F6, E2F7, and E2F8 were upregulated in CRC tissues, whereas the expression level of E2F2 was downregulated. In addition, the expression of E2Fs showed complex, intertwined positive correlations with each other; however, E2F2 showed a negative correlation with E2F5. We also found the E2Fs expression level were altered in CRC tissues. These results may indicate that synergy or mutual antagonism by E2Fs promote or suppress CRC tumor progression. The expression levels of different E2Fs were related to the OS and clinical pathological parameters, and E2F expression varied with tumor stage. These results suggest that E2Fs could serve as markers for CRC progression.

In colon cancer, E2F1 is the most highly examined gene of the eight E2F genes. CDCA3 mediates p21-dependent proliferation by regulating E2F1 expression in colorectal cancer(35). For E2F2, miR-155 regulates the proliferation and cell cycle of colorectal carcinoma cells by targeting E2F2 mRNA(36). For E2F3, miR-503 inhibits cell proliferation and induces apoptosis in colorectal cancer cells by targeting E2F3 mRNA(37). Micro RNA-449b inhibits proliferation of SW1116 colon cancer stem cells by downregulating CCND1 and E2F3 expression(38). CircPRMT5 circular RNA promotes proliferation of colorectal cancer through sponging miR-377 to induce E2F3 expression(39). For E2F5, miRNA-34a targets FMNL2 and E2F5 and suppresses the progression of colorectal cancer(40). MicroRNA-32 inhibits the proliferation, migration, and invasion of human colon cancer cell lines by targeting E2F5 mRNA(41). In view of the activities of miRNAs to target E2Fs in CRC, we found that several miRNAs may target E2F2, E2F3 and E2F4. Therefore, these miRNAs may exert their functions on E2F2, E2F3 and E2F4 mRNAs to regulate CRC tumors, but the mechanism needs further study.

Some studies have shown that E2Fs may exert function by regulating cell signaling pathways. Knockdown of E2F8 suppressed CRC cell proliferation through the NF- κ B pathway(42). Upregulated miR-1258 regulates cell cycle and inhibits cell proliferation by directly targeting E2F8 in CRC(43). In our research, the network constructed by E2Fs and the 69 most frequently altered co-expressed genes showed that the cell cycle-related genes were associated with E2Fs mutations.

The GO functional annotation and KEGG pathway enrichment analyses showed that the major functions of E2Fs were regulation of cell cycle, cell division, and proliferation, DNA replication and repair. The cell

cycle was the most enriched pathway. We also found new functions and pathways for E2Fs in CRC, such as angiogenesis, dynein binding, annealing helicase activity, and anaphase-promoting complex; these new associations are worthy of further study.

We have systematically analyzed the relationship between E2F transcription factors and CRC. We suggest that these transcription factors could be markers of CRC. Although functions of some E2Fs in CRC have been reported, the mutual regulation of these factors is rarely reported for CRC, and the regulatory issue should be further investigated. Our bioinformatics analysis showed potential miRNA targets for E2F2, E2F3, and E2F4; these particular factors are related mainly to cell cycle signaling pathways.

Abbreviations

CNAs: copy number alterations; CRC: colorectal cancer; DBDs: DNA binding domains; DFS: Disease free survival; DSS: Disease specific survival; GEO: Gene expression omnibus; HR: Hazard ratio; HRP: Horseradish-peroxidase; H-score: Histochemical score; NGS: Next-generation sequencing; OS: Overall survival

Declarations

Acknowledgments

The authors thank AiMi Academic Services (www.aimieditor.com) for English language editing and review services.

Author Contributions

Wen-Da Wang, Li-Chun Wang carried out data analysis. Wen-Da Wang drafted the manuscript; Li-Chun Wang completed the IHC experiments. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of Shanxi (grant numbers 201901D111398).

Availability of data and materials

All of the data and material details are showed in the Methods part. And the related data was download from those websites.

Ethics approval and consent to participate

This article related human tumor tissues were obtained from Shanxi Cancer Hospital Sample library. Shanxi Medical Ethics Committee was approved and consented to participate in this research. Study was

approved by can carried out in accordance with institutional review board of Shan Xi Medical University. Informed consent was obtained from participants/legal guardians of participants involved in the study.

Consent for publication

Not applicable.

Conflict of Interest Statement

All authors declare no conflict of interest in this study.

Author details

¹Department of Colorectal and Anal Surgery, Shanxi Cancer hospital, Shanxi 030013 China

References

1. Trimarchi JM, Lees JA: Sibling rivalry in the E2F family. *Nat Rev Mol Cell Biol* 2002;3:11-20
2. Attwooll C, Lazzerini Denchi E, Helin K: The E2F family: specific functions and overlapping interests. *EMBO J* 2004;23:4709-4716
3. Chong JL, Tsai SY, Sharma N, et al.: E2f3a and E2f3b contribute to the control of cell proliferation and mouse development. *Mol Cell Biol* 2009;29:414-424
4. Chong JL, Wenzel PL, Sáenz-Robles MT, et al.: E2f1-3 switch from activators in progenitor cells to repressors in differentiating cells. *Nature* 2009;462:930-934
5. Chen HZ, Tsai SY, Leone G: Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nat Rev Cancer* 2009;9:785-797
6. Thurlings I, de Bruin A: E2F Transcription Factors Control the Roller Coaster Ride of Cell Cycle Gene Expression. *Methods Mol Biol* 2016;1342:71-88
7. Emmrich S, Pützer BM: Checks and balances: E2F-microRNA crosstalk in cancer control. *Cell Cycle* 2010;9:2555-2567
8. Pei X, Du E, Sheng Z, Du W: Rb family-independent activating E2F increases genome stability, promotes homologous recombination, and decreases non-homologous end joining. *Mech Dev* 2020;162:103607
9. Jiang H, Martin V, Gomez-Manzano C, et al.: The RB-E2F1 pathway regulates autophagy. *Cancer Res* 2010;70:7882-7893
10. Benevolenskaya EV, Frolov MV: Emerging links between E2F control and mitochondrial function. *Cancer Res* 2015;75:619-623
11. Bertoli C, Herlihy AE, Pennycook BR, Kriston-Vizi J, de Bruin R: Sustained E2F-Dependent Transcription Is a Key Mechanism to Prevent Replication-Stress-Induced DNA Damage. *Cell Rep* 2016;15:1412-1422

12. Wong JV, Dong P, Nevins JR, Mathey-Prevot B, You L: Network calisthenics: control of E2F dynamics in cell cycle entry. *Cell Cycle* 2011;10:3086-3094
13. 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
14. Bagaria SP, Heckman MG, Diehl NN, Parker A, Wasif N: Delay to Colectomy and Survival for Patients Diagnosed with Colon Cancer. *J Invest Surg* 2019;32:350-357
15. Evangelou K, Havaki S, Kotsinas A: E2F transcription factors and digestive system malignancies: how much do we know. *World J Gastroenterol* 2014;20:10212-10216
16. Shamir ER, Devine WP, Pekmezci M, et al.: Identification of high-risk human papillomavirus and Rb/E2F pathway genomic alterations in mutually exclusive subsets of colorectal neuroendocrine carcinoma. *Mod Pathol* 2019;32:290-305
17. Yamaguchi K, Yamaguchi R, Takahashi N, et al.: Overexpression of cohesion establishment factor DSCC1 through E2F in colorectal cancer. *PLoS One* 2014;9:e85750
18. Pitroda SP, Khodarev NN, Huang L, et al.: Integrated molecular subtyping defines a curable oligometastatic state in colorectal liver metastasis. *Nat Commun* 2018;9:1793
19. Bielski CM, Zehir A, Penson AV, et al.: Genome doubling shapes the evolution and prognosis of advanced cancers. *Nat Genet* 2018;50:1189-1195
20. Dosil MA, Navaridas R, Mirantes C, et al.: Tumor suppressive function of E2F-1 on PTEN-induced serrated colorectal carcinogenesis. *J Pathol* 2019;247:72-85
21. Rhodes DR, Kalyana-Sundaram S, Mahavisno V, et al.: Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia* 2007;9:166-180
22. Rhodes DR, Yu J, Shanker K, et al.: ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia* 2004;6:1-6
23. Li T, Fu J, Zeng Z, et al.: TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res* 2020;48:W509-509W514
24. 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:W98-98W102
25. Mizuno H, Kitada K, Nakai K, Sarai A: PrognosScan: a new database for meta-analysis of the prognostic value of genes. *BMC Med Genomics* 2009;2:18
26. Dennis G Jr, Sherman BT, Hosack DA, et al.: DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 2003;4:P3
27. Ashburner M, Ball CA, Blake JA, et al.: Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000;25:25-29
28. Li JH, Liu S, Zhou H, Qu LH, Yang JH: starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res* 2014;42:D92-97

29. Barrett T, Edgar R: Gene expression omnibus: microarray data storage, submission, retrieval, and analysis. *Methods Enzymol* 2006;411:352-369
30. Barrett T, Wilhite SE, Ledoux P, et al.: NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res* 2013;41:D991-995
31. Yao H, Lu F, Shao Y: The E2F family as potential biomarkers and therapeutic targets in colon cancer. *PeerJ* 2020;8:e8562
32. Lan W, Bian B, Xia Y, et al.: E2F signature is predictive for the pancreatic adenocarcinoma clinical outcome and sensitivity to E2F inhibitors, but not for the response to cytotoxic-based treatments. *Sci Rep* 2018;8:8330
33. Manicum T, Ni F, Ye Y, Fan X, Chen BC: Prognostic values of E2F mRNA expression in human gastric cancer. *Biosci Rep* 2018;38
34. Henry D, Brumaire S, Hu X: Involvement of pRb-E2F pathway in green tea extract-induced growth inhibition of human myeloid leukemia cells. *Leuk Res* 2019;77:34-41
35. Qian W, Zhang Z, Peng W, et al.: CDCA3 mediates p21-dependent proliferation by regulating E2F1 expression in colorectal cancer. *Int J Oncol* 2018;53:2021-2033
36. Li T, Yang J, Lv X, et al.: miR-155 regulates the proliferation and cell cycle of colorectal carcinoma cells by targeting E2F2. *Biotechnol Lett* 2014;36:1743-1752
37. Chang SW, Yue J, Wang BC, Zhang XL: miR-503 inhibits cell proliferation and induces apoptosis in colorectal cancer cells by targeting E2F3. *Int J Clin Exp Pathol* 2015;8:12853-12860
38. Fang Y, Gu X, Li Z, Xiang J, Chen Z: miR-449b inhibits the proliferation of SW1116 colon cancer stem cells through downregulation of CCND1 and E2F3 expression. *Oncol Rep* 2013;30:399-406
39. Yang B, Du K, Yang C, et al.: CircPRMT5 circular RNA promotes proliferation of colorectal cancer through sponging miR-377 to induce E2F3 expression. *J Cell Mol Med* 2020;24:3431-3437
40. Lu G, Sun Y, An S, et al.: MicroRNA-34a targets FMNL2 and E2F5 and suppresses the progression of colorectal cancer. *Exp Mol Pathol* 2015;99:173-179
41. Yang F, Chen L, Wang ZJ: MicroRNA-32 inhibits the proliferation, migration and invasion of human colon cancer cell lines by targeting E2F transcription factor 5. *Eur Rev Med Pharmacol Sci* 2019;23:4156-4163
42. Yan PY, Zhang XA: Knockdown of E2F8 Suppresses Cell Proliferation in Colon Cancer Cells by Modulating the NF- κ B Pathway. *Ann Clin Lab Sci* 2019;49:474-480
43. Zhang Z, Li J, Huang Y, et al.: Upregulated miR-1258 regulates cell cycle and inhibits cell proliferation by directly targeting E2F8 in CRC. *Cell Prolif* 2018;51:e12505

Figures

Figure 1

The mRNA expression levels of E2Fs in different types of cancers (Oncomine).

Figure 2

Expression of E2F mRNAs in CRC tumor and normal tissue. (A) Expression assessed in TIMER (see Methods). (B) Expression assessed in GEPIA).

Figure 3

Immunohistochemistry of paired colon tissues and surrounding adjacent tissues.

Figure 4

The correlations of the expression of E2Fs with each other in CRC and assessed by GEIPA.

Figure 5

The CRC prognosis values of E2Fs measured by GEIPA.

Figure 6

Correlations of E2F expression with CRC clinicopathological parameters from the TCGA dataset. (A) Correlations with tumor stage. (B) Correlations with metastasis. (C) Correlations with sex. (D) Correlations with pathological types.

Figure 7

The CRC mutation analysis of E2Fs by cBioPortal.

Figure 8

The PPI network for E2Fs and the 69 most frequently altered co-expressed genes.

Figure 9

The cell cycle signaling pathway regulated by the E2F and alteration in CRC.

Figure 10

Venn diagram of overlapping mRNA target of E2F2, E2F3, and E2F4.