Research in Pharmaceutical Sciences, February 2025; 20(1): 121-141
Received: 12-07-2023
Peer Reviewed: 01-09-2023
Revised: 03-03-2024
Accepted: 11-03-2024
Published: 14-02-2025

School of Pharmacy & Pharmaceutical Sciences
Isfahan University of Medical Sciences

Original Article

Prognostic markers and molecular pathways in primary colorectal cancer with a high potential of liver metastases: a systems biology approach

Fatemeh Bahramibanan[1], Amir Taherkhani[2,*], Rezvan Najafi[2], Neda Alizadeh[3], Hamidreza Ghadimipour[4], Nastaran Barati[5], Katayoun Derakhshandeh[6], and Meysam Soleimani[5,*]

[1]*Department of Medical Biotechnology, School of Advanced Medical Sciences and Technologies, Hamadan University of Medical Sciences, Hamadan, I.R. Iran.*
[2]*Research Center for Molecular Medicine, Institute of Cancer, Avicenna Health Research Institute, Hamadan University of Medical Sciences, Hamadan, I.R. Iran.*
[3]*Department of Anesthesiology and Critical Care, School of Medicine, Hamadan University of Medical Sciences, Hamadan, I.R. Iran.*
[4]*Department of Pathology, School of Medicine, Hamadan University of Medical Sciences, Hamadan, I.R. Iran.*
[5]*Department of Pharmaceutical Biotechnology, School of Pharmacy, Hamadan University of Medical Sciences, Hamadan, I.R. Iran.*
[6]*Department of Pharmaceutics, School of Pharmacy, Hamadan University of Medical Sciences, Hamadan, I.R. Iran.*

Abstract

Background and purpose: Colorectal cancer (CRC) holds the position of being the third most prevalent cancer and the second primary cause of cancer-related fatalities on a global scale. Approximately 65% of CRC patients survive for 5 years following diagnosis. Metastasis and recurrence frequently occur in half of CRC patients diagnosed at the late stage. This study used bioinformatics analysis to identify key signaling pathways, hub genes, transcription factors, and protein kinases involved in transforming primary CRC with liver metastasis potential. Prognostic markers in CRC were also identified.
Experimental approach: The GSE81582 dataset was re-analyzed to identify differentially expressed genes (DEGs) in early CRC compared to non-tumoral tissues. A protein interaction network (PIN) was constructed, revealing significant modules and hub genes. Prognostic markers, transcription factors, and protein kinases were determined. Boxplot and gene set enrichment analyses were performed.
Findings/Results: This study identified 1113 DEGs in primary CRC compared to healthy controls. PIN analysis revealed 75 hub genes and 8 significant clusters associated with early CRC. The down-regulation of SUCLG2 and KPNA2 correlated with poor prognosis. SIN3A and CDK6 played crucial roles in early CRC transformation, affecting rRNA processing pathways.
Conclusion and implications: This study demonstrated several pathways, biological processes, and genes mediating the malignant transformation of healthy colorectal tissues to primary CRC and may help the prognosis and treatment of patients with early CRC.

Keywords: Biomarkers; Cancer; CRC; Pathogenesis; Pathway; Prognosis.

INTRODUCTION

Colorectal cancer (CRC) ranks as the third most common form of cancer and is the second most prominent contributor to cancer-related deaths across the globe (1-4). CRC affected ~1.9 million patients, leading to 900,000 deaths around the world in 2020 (1).

*Corresponding authors:
A. Taherkhani, Tel: +98-81383815903, Fax: +98-8138276299
Email: amir.007.taherkhani@gmail.com
M. Soleimani, Tel: +98-8138381319, Fax: +98-8138276299
Email: m.soleymani@umsha.ac.ir

Access this article online	
	Website: http://rps.mui.ac.ir
	DOI: 10.4103/RPS.RPS_128_23

As the incidence and mortality of cancer increase, healthcare systems around the world are being faced with the challenge of ensuring equitable and effective care delivery for all patients, particularly in low- and middle-income countries (5,6). In late-stage CRC, metastasis and recurrence occur in approximately 50% of patients, and many of these cases develop resistance to chemotherapy (7,8,9). Although broad CRC screening approaches and new therapeutic methods have led to early detection and diminished mortality (10,11), the 5-year survival rate of CRC patients has remained at 65% (12), which is unsatisfactory. Therefore, uncovering new signaling pathways, biological processes (BPs), hub genes, prognostic markers, and associated master regulators may provide novel therapeutic targets and more appropriate strategies to combat CRC. Sayagués *et al*. conducted a study to assess the molecular alterations in primary CRC tissues compared to adjacent healthy tissues and matched liver metastases tissues, at both mRNA and miRNA levels (13). In the present study, we hypothesized that many of the differentially expressed genes (DEGs) in primary CRC tissues with a high potential of liver metastases compared with their corresponding normal specimens play a critical role in the aggressive behavior of cancer cells. Therefore it could be associated with a poor prognosis in CRC. Also, it was suggested that prognostic markers might act as hub genes in a protein interaction map (PIM) associated with the malignant transformation of normal colorectal tissues to primary CRC with a high potential of distant migration.

Metastasis of CRC to the liver confers a dismal prognosis, with a median survival of less than 2 years (14,15). Identifying patients most susceptible to liver metastases could guide personalized treatment plans and vigilant monitoring to improve outcomes (16). Previous studies have uncovered gene signatures in the primary tumor (17) and blood-based protein biomarkers (18) that correlate with the later development of liver lesions, allowing stratification of metastasis risk. Additionally, immunotherapy approaches like PD-1/PD-L1 checkpoint inhibitors have shown early promise to suppress the growth of existing liver metastases and prevent the emergence of new ones (19). Taken together, the evidence mentioned above indicates the clinical potential of prognostic biomarkers and immunotherapeutic strategies to enhance survival in CRC patients vulnerable to lethal liver spread. The subject highlighted the significance of continuous research in uncovering novel biomarkers and molecular pathways distinguishing aggressive primary colorectal tumors with high metastatic preference (20).

The present study re-analyzed the microarray GSE81582 dataset created by Sayagués *et al*., a method for exploring the molecular diversity within sporadic CRC (sCRC) tumors, who studied mRNA and miRNA levels in primary sCRC tumor samples (13). Furthermore, the previously mentioned study examined the expression of coding and non-coding RNAs in non-cancerous tissues (13). In Sayagués's study, tissue samples were collected from 23 primary sCRCs consecutive patients and 19 paired liver metastases. The median follow-up at the end of the study was 25 months. No liver metastase tissues were available for 4 patients. A total of 9 healthy tissue samples was collected at a distance of > 10 cm from the tumor region. All tissue samples were taken from patients at the Department of Surgery of the University Hospital of Salamanca (Salamanca, Spain). All tissue samples were swiftly gathered following surgical removal, frozen, and stored at -80 °C. Notably, Sayagués's study was approved by the local ethics committee at the University Hospital of Salamanca (Salamanca, Spain). Finally, the obtained dataset included invaluable information regarding the primary CRC leading to metastasis (13). Therefore, the present study harnessed the valuable dataset and pinpointed the variations in expression and hub genes, enriched pathways and Gene Ontology terms, key transcription factors, associated protein kinases, and potential prognostic biomarkers in early-stage CRC prone to liver metastasis versus healthy tissue, and validated aberrant expression of prognostic genes using bioinformatics tools. These biomarkers could be targets for therapy in early-stage CRC.

AMATERIALS AND METHODS

Data acquisition and processing

The dataset GSE81582 (13) in the Gene Expression Omnibus (21) included 23 primary colorectal tumors, 19 colorectal liver metastases, and 9 non-tumoral colorectal tissues. Transcriptome data of primary and healthy colorectal tissues based on the GPL15207 platform (Affymetrix Human Gene Expression Array) were considered for further statistical and bioinformatics analyses in this study. The online GEO2R tool identified DEGs in early CRC rather than healthy colorectal tissues. The cut-off criteria were established as a false discovery rate (FDR) of less than 0.001 and an absolute log2 of fold change (|Log2FC|) exceeding 1.

Network analysis

DEGs were imported into the STRING database (version 11.5), available at https://string-db.org (22), to identify possible interactions among proteins expressed by DEGs. After removing disconnected proteins from the graph, the connected PIM was transferred into the Cytoscape (version 3.9.1), available at https://cytoscape.org, to identify hub genes and significant modules within the network. Vertexes with the criteria of degree and betweenness above 2-fold the average of the nodes in the graph were considered hub genes (23). Furthermore, the Molecular Complex Detection (MCODE) plugin (24) demonstrated condensed regions in the PIM. Clusters with the MCODE score > 3 and the number of nodes ≥ 10 were considered significant (25).

Functional enrichment analysis

The g:Profiler web server, available at https://biit.cs.ut.ee/gprofiler/gost (26), was used to uncover signaling pathways and gene ontology (GO) annotations, including BPs, molecular functions (MFs), and cellular components (CCs) enriched in early CRC. In this regard, genes of clusters were considered for pathway and BP annotation enrichment analysis, while all DEGs were used for identifying CCs and MFs affected in primary CRC (23,25). For GO annotation analysis, the g:Profiler retrieves data from several data sources, including the Ensemble database (27), Ensemble Genomes (28), and WormBase ParaSite (29). In addition, the Kyoto Encyclopedia of Genes and Genomes (KEGG) (30), Reactome (31), and WikiPathways (32) were used by the g:Profiler for pathway enrichment analysis. This study reported significant pathways detected by the KEGG and Reactome data sources. The enriched pathways and GO annotations with the criteria of FDR less than 0.05 and the number of genes ≥ 10 (33) were assigned significantly.

Upstream regulators of the hub genes and their consensus sequences logos

The abnormal expression of transcription factors (TFs) is associated with several human disorders, including cancers. Thus, targeting critical TFs involved in tumor development may illustrate therapeutic effects in cancer patients (34). In the present study, significant TFs regulating expression of hub genes were detected implementing the iRegulon app (35) in the Cytoscape. The JASPAR database (36), available at https://jaspar.genereg.net, was investigated to achieve the consensus sequences logos of binding sites for TFs. Subsequently, the R programming (version 4.2.1) (37) was utilized to calculate the consensus sequences match scores, according to the algorithm explained by Xiong (38). Each score was defined as the probability of the consensus sequence corresponding to the binding site of TFs as $2^{(\text{match score})}$ times more than randomness.

Kinases enrichment analysis

Protein kinases mediate various biological procedures linked with cell cycle, apoptosis, and proliferation. Furthermore, over/under-expression of protein kinases could lead to cancer initiation and progression (39-41). Therefore, protein kinases are considered hot drug targets in cancer therapy due to their critical role in tumorigenesis (42). This study identified protein kinases mediating phosphorylation of TFs regulating the expression of hub genes (43); this task was accomplished using Kinase Enrichment

Analysis 3, which can be accessed at https://maayanlab.cloud/kea3 (44). Finally, a gene regulatory network (GRN) was assembled, encompassing pivotal genes, TFs, and protein kinases.

Prognostic hub genes and validation study

The Gene Expression Profiling Interactive Analysis 2 (GEPIA2) database (45), available at http://gepia2.cancer-pku.cn/#index, provides valuable Kaplan-Meier curves to evaluate the prognostic role of genes in several human cancers. This is done by integrating 2 primary RNA-Seq sources, the Genotype-Tissue Expression (GTEx) (46), and TCGA (47) data banks. Here, the prognostic role of hub genes in early CRC was studied in colon adenocarcinoma (COAD) and rectum adenocarcinoma (READ). The cut-off condition was set to log-rank test *P*-value and hazard ratio (HR) *P*-value < 0.05. Moreover, the expression patterns of prognostic markers were checked out at the transcript levels in COAD and READ using the GEPI2A database (45).

RESULTS

DEGs in primary CRC

The GEO2R determined 1113 DEGs, including 474 up- and 639 down-regulated genes, in early CRC with a high potential of liver metastases compared to the healthy colorectal tissues with the characteristics of FDR < 0.001 and |log2FC > 1| (Table S1). The volcano plot of the genes in the dataset GSE81582 (23 early CRC and 9 non-tumoral samples) was achieved by the online Shiny apps web server, available at https://huygens.science.uva.nl (48) (Fig. 1).

PIM and functional analyses

After removing disconnected DEGs in the STRING database, a connected PIM (1040 genes and 7329 edges) was transferred into the Cytoscape tool for topological analyses. The MCODE plugin detected 8 considerable clusters in the PIM including modules No. 1, 2, 3, 6, 8, 13, 15, and 16 (Fig. 2).

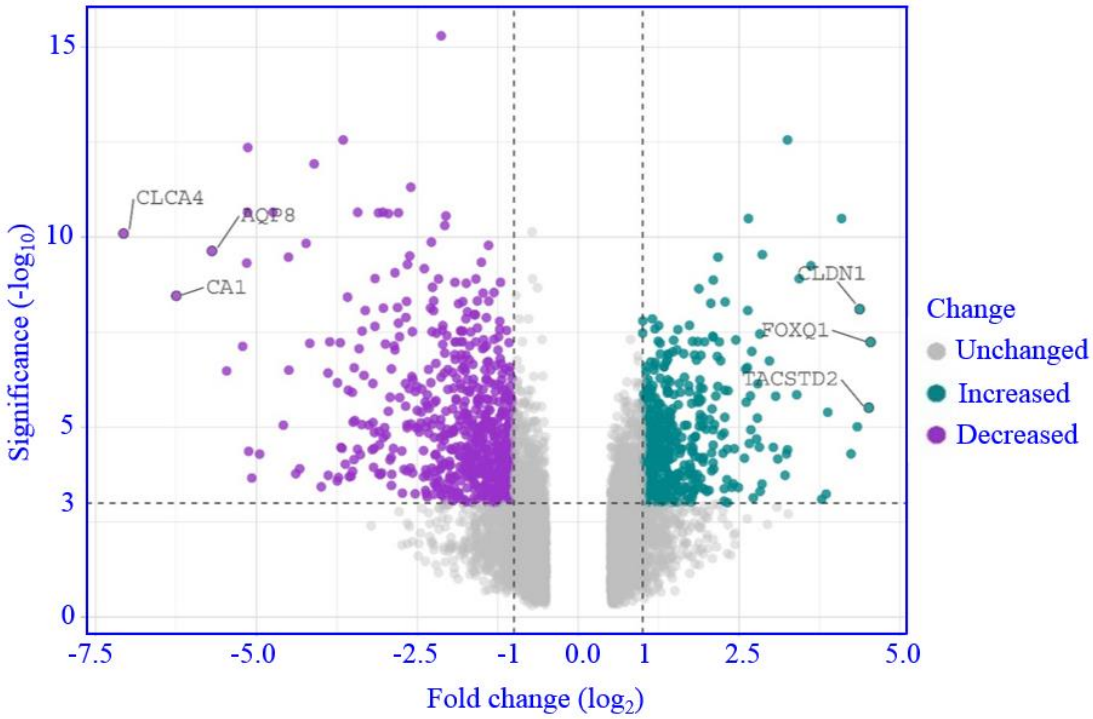


Fig. 1. Volcano plot demonstrating the log2 FC of the genes (x-axis) in primary CRC compared to normal tissue, and their corresponding significance as -log10 FDR (y-axis). CRC, colorectal cancer; FC, fold change; FDR, false discovery rate.

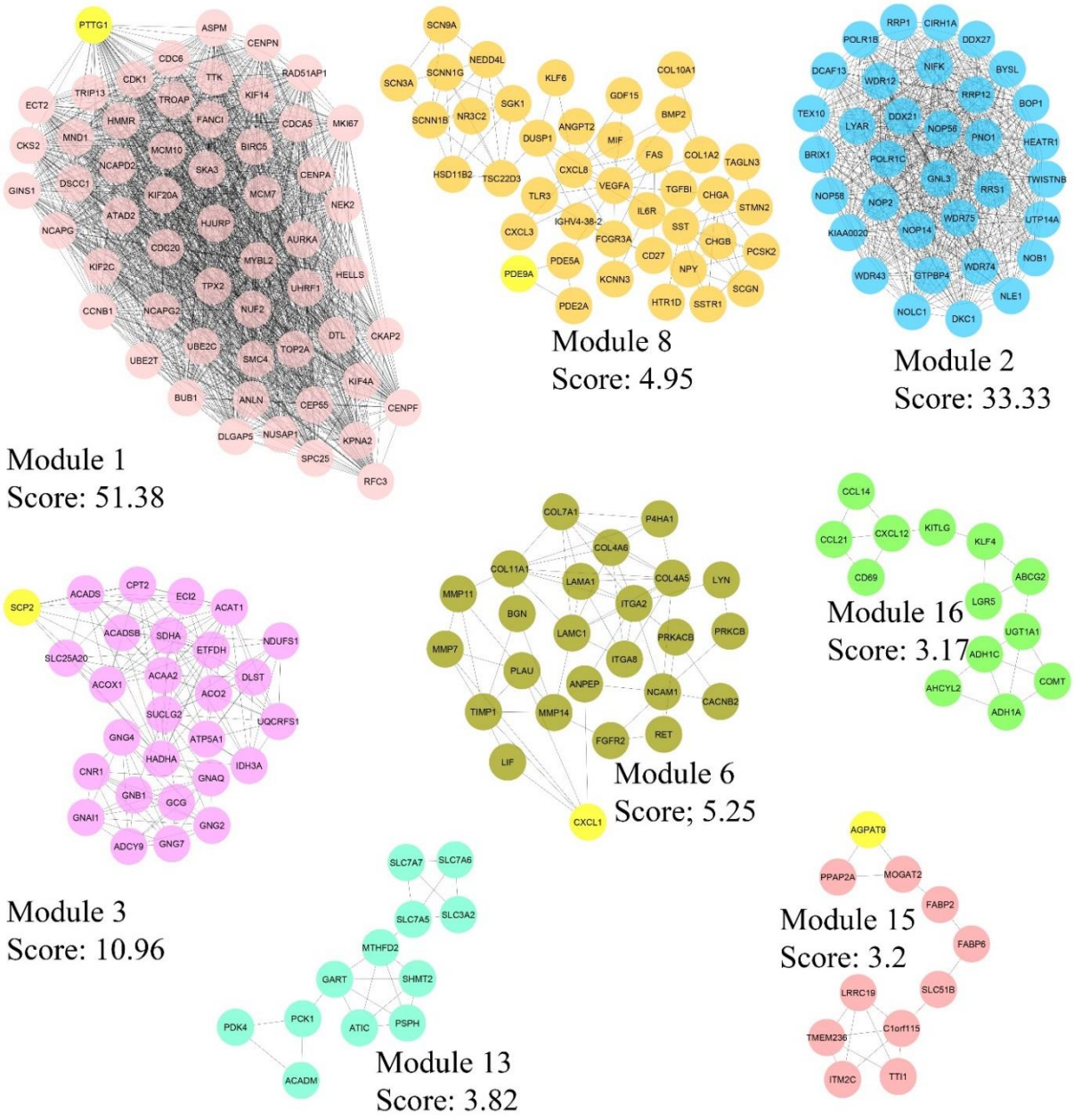


Fig. 2 . Clustering analysis. Eight substantial modules were found in the protein interaction map associated with primary colorectal cancer. MCODE plugin was used to calculate the scores of each cluster. Yellow vertexes show seed nodes.

The most notable signaling pathways impacted in primary CRC were rRNA processing in the nucleus and cytosol (REAC: R-HSA-8868773), rRNA processing (REAC: R-HSA-72312), and the cell cycle (KEGG: 04110) (Fig. 3A). Besides, cell cycle (GO: 0007049), rRNA processing (GO: 0006364), and ribosome biogenesis (GO: 0042254) demonstrated to be the most significant BPs enriched in the etiology of early CRC based on their FDR (Fig. 3B). Furthermore, protein binding (GO: 0005515), binding (GO: 0005488), and anion transmembrane transporter activity (GO: 0008509) revealed to be the most salient MFs deregulated in early-stage CRC (Fig. 3C), while intracellular organelle lumen (GO: 0070013), organelle lumen (GO: 0043233), and membrane-enclosed lumen (GO: 0031974) were the most considerable CCs enriched in the pathogenesis of primary CRC (Fig. 3D). Binding denotes the selective, non-covalent, frequently stoichiometric interaction between a molecule and one or multiple distinct sites on another molecule including polyamine binding, major histocompatibility complex binding, hormone binding, cellulosome binding, *etc* (49).

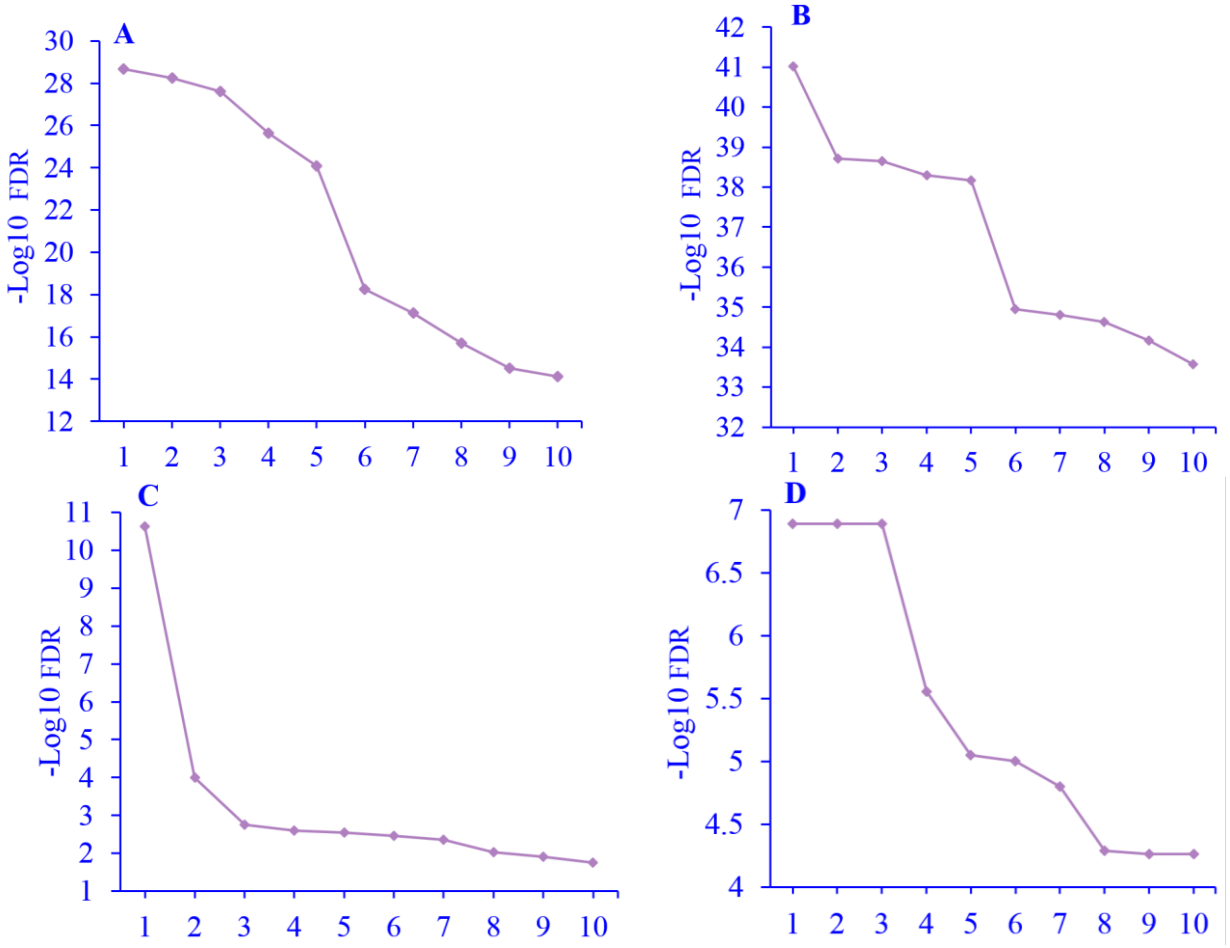


Fig. 3. Top 10 significant (A) pathways including 1, rRNA processing in the nucleus and cytosol; 2, rRNA processing; 3, cell cycle; 4, cell cycle, mitotic; 5, major pathway of rRNA processing in the nucleolus and cytosol; 6, metabolism of RNA; 7, rRNA modification in the nucleus and cytosol; 8, ribosome biogenesis in eukaryotes; 9, extracellular matrix organization; 10, M phase; (B) biological processes including 1,cell cycle; 2, rRNA processing; 3, ribosome biogenesis; 4, rRNA metabolic process; 5, cell cycle process; 6, mitotic cell cycle process; 7, chromosome segregation; 8, ribonucleoprotein complex biogenesis; 9, mitotic cell cycle; 10, nuclear division; (C) molecular functions including 1, protein binding; 2, binding; 3, anion transmembrane transporter activity; 4, Oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor ; 5, Active transmembrane transporter activity; 6, Extracellular matrix structural constituent; 7, acetyl-CoA C-acyltransferase activity; 8, oxidoreductase activity, acting on CH-OH group of donors; 9, transmembrane receptor protein serine/threonine kinase binding; 10, SnoRNA binding; (D) cellular components including 1, intracellular organelle lumen; 2, organelle lumen; 3, membrane-enclosed lumen; 4, apical part of cell; 5, apical plasma membrane; 6, cytoplasm; 7, extracellular exosome; 8, extracellular vesicle; 9, extracellular membrane-bounded organelle; 10, extracellular organelle; enriched in primary CRC tissues compared with the healthy controls. The X-axis demonstrates the name of the term. CRC, colorectal cancer; FDR, false discovery rate.

Tables S2-5 list all significant pathways, BPs, MFs, and CCs enriched in early CRC tissues with a high potential of distant metastasis. In addition to network analysis, 75 nodes revealed degree and betweenness centrality above 2-fold the average of the nodes in the PIM and, therefore, were considered hub genes. According to Jeong's study (50), hub genes within the protein-protein interaction (PPI) networks are crucial for cellular functions, and blocking them in living cells can be lethal. This study aimed to identify hub genes in the PPI network linked to the pathogenesis of primary CRC with a high potential for metastasis. Targeting these genes could hold therapeutic promise for early CRC treatment. Figure. 4 demonstrates the internal interactions among hub genes achieved by the STRING, while Table S6 presents the degree and betweenness values of the hubs. The average value of degree and betweenness were calculated as 14.09 and 0.0025, respectively.

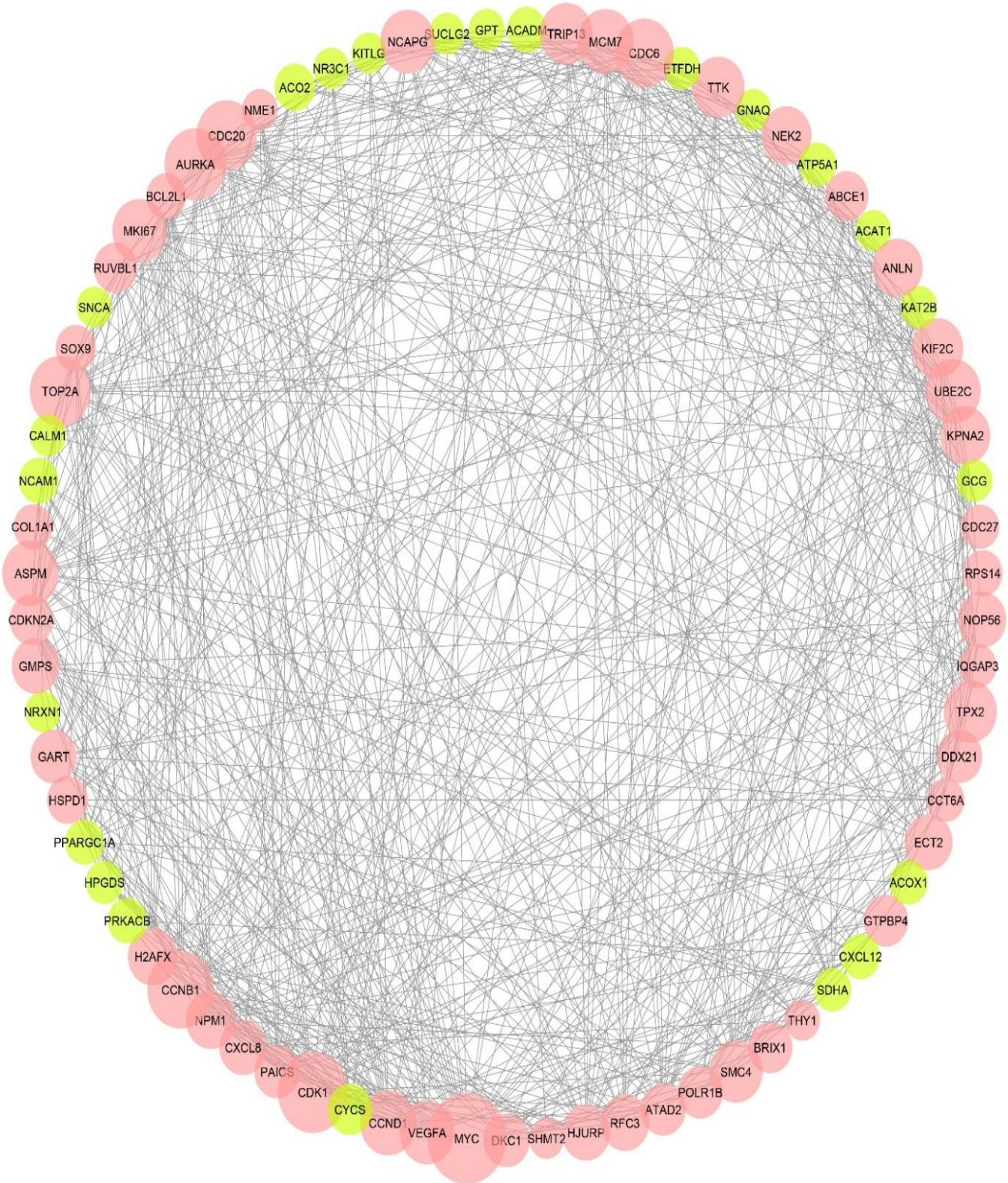


Fig. 4. Interactions between 75 hubs including 52 up-regulated (red color) and 23 down-regulated (green color) genes. The size of the nodes is positively correlated with the degree of the nodes in the main protein-protein interaction network. This network was constructed based on the attribute circle layout, in which the MYC and DKC1 demonstrated the maximum and minimum betweenness centrality values among the hubs.

Table 1. GRN study identified 18 transcription factors as upstream regulators of the hub genes.

Transcription factor	Normalized enrichment score	Targets
SIN3A	8.033	MYC, ANLN, KIF2C, NME1, MCM7, KAT2B, NCAPG, CDC20, NEK2, KPNA2, CALM1, NPM1, IQGAP3, ATAD2, TPX2, CDC6, RFC3, TTK, SMC4, PAICS, CDK1,AURKA, BRIX1, CCNB1, COL1A1, UBE2C, HJURP, HSPD1, VEGFA, DDX21, TOP2A, ASPM, NOP56, MKI67, ETFDH, H2AFX, ABCE1, SUCLG2, GTPBP4, CCND1, TRIP13, CYCS, ACOX1, ACO2, SOX9, ECT2, KITLG, BCL2L1, GART, ACAT1, NR3C1, CDC27, GPT, ATP5A1, RPS14, DKC1, RUVBL1, THY1, CCT6A, GMPS, ACADM, NCAM1, POLR1B, PRKACB
FOXM1	7.633	CCNB1, CDK1, UBE2C, BCL2L1, KPNA2, ASPM, MKI67, HJURP, TPX2, ECT2, NEK2, KIF2C, TOP2A, H2AFX, SMC4, AURKA, GTPBP4, ATAD2, ANLN, TTK, CDC20, GNAQ, MCM7, RFC3, CCND1, DKC1, COL1A1, HSPD1, NCAPG
NFYC	6.113	CDC6, GART, H2AFX, KPNA2, SOX9, SHMT2, NEK2, ECT2, UBE2C, CDK1, ASPM, NCAM1, CCND1, TPX2, COL1A1, ATAD2, MCM7, AURKA, THY1, CCT6A, TOP2A, HSPD1, KITLG, VEGFA, NR3C1, SUCLG2, NME1, SNCA, CALM1, TTK, BCL2L1, CDC20, DKC1, GNAQ, NOP56, PRKACB, HPGDS, CDKN2A, HJURP, GCG, GTPBP4, NRXN1, ETFDH
E2F4	5.956	NR3C1, RFC3, MCM7, KIF2C, NCAPG, IQGAP3, MKI67, TRIP13, CDC20, ANLN, ATAD2, ASPM, NEK2, CDK1, TOP2A, NPM1, HJURP, TTK, CDC6, UBE2C, H2AFX, TPX2, MYC, ECT2, HSPD1, CYCS, CCNB1, SMC4, KPNA2, AURKA, GMPS, CDKN2A, BRIX1
H1FX	5.501	HSPD1, MYC, MCM7, SOX9, NRXN1, CDC6, GMPS, CDK1, CDKN2A, RFC3, NCAPG, CCND1, BCL2L1, ATAD2, NR3C1, GNAQ, VEGFA, ASPM, H2AFX, KPNA2, IQGAP3, THY1, CXCL12, GART, COL1A1,NCAM1, MKI67, ABCE1, KITLG, CALM1, PAICS, ACOX1, SNCA, CDC27, SHMT2, HPGDS, KAT2B, NME1, TRIP13, SUCLG2
NFYB	5.5	COL1A1, UBE2C, GART, CDC6, H2AFX, TOP2A, SOX9, ECT2, PRKACB, NCAM1, NEK2, THY1, CDK1, MCM7, BCL2L1, ASPM, KPNA2, SHMT2, NR3C1, ATAD2, HPGDS, VEGFA, SNCA, GNAQ, KITLG, AURKA, CDC20, TTK, PPARGC1A, TPX2, SUCLG2, CALM1, HJURP, NRXN1, KAT2B, NOP56, CDC27, CCND1, ACO2, GCG, CCT6A, CYCS, CXCL12, MYC, SMC4
E2F1	5.123	GMPS, NR3C1, MCM7, MYC, CDKN2A, NRXN1, CCND1, ATAD2, NCAPG, NCAM1, PAICS, THY1, VEGFA, CDC6, SHMT2
FOXS1	4.988	CCND1, NCAM1, CXCL12, CDC6, NR3C1, SHMT2, GNAQ, MYC, SOX9, PPARGC1A, NRXN1, SUCLG2, SMC4, KITLG, GCG, KAT2B, COL1A1, THY1
MYBL2	4.664	AURKA, H2AFX, TOP2A, TPX2, HSPD1, CCNB1, HJURP, UBE2C, KPNA2, NCAPG, ANLN, ECT2, SMC4, CDC20, CDK1, ACOX1, BCL2L1, NOP56, SOX9, NEK2, ATAD2, CDC27, VEGFA, KIF2C, NME1, ACO2, PAICS, POLR1B, TTK, CYCS, ETFDH, RPS14, ASPM, GPT, CALM1, GART, MKI67
TFDP1	4.474	CDK1, NEK2, H2AFX, TOP2A, CDC6, IQGAP3, MCM7, KIF2C, ATAD2, HJURP, NCAPG, UBE2C, CCNB1, RFC3, NPM1, SMC4, TTK, ECT2, BRIX1, ASPM, CDC20, CYCS, KITLG, ETFDH, TPX2, PAICS, MYC, ANLN, CDKN2A, ATP5A1
CEBPA	3.67	SHMT2, CCND1, TPX2, KITLG, CDC27, HPGDS, NRXN1, GNAQ, MYC, GCG, PPARGC1A, SUCLG2, NCAM1
YY1	3.642	NR3C1, SMC4, MCM7, GART, PAICS, NMP1, HSPD1, ATP5A1, GNAQ, CCND1, ABCE1, NCAM1, KITLG, MYC, NOP56, CDKN2A
E2F3	3.584	SMC4, COL1A1, NPM1, SOX9, BCL2L1, MYC, KITLG, CCND1, ABCE1, GNAQ, H2AFX, ASPM, HSPD1, NRXN1, NCAM1, GCG, CXCL12, CALM1, CDC6, THY1, GART, TPX2, NR3C1, NEK2, ATAD2, NOP56, PAICS
GLI1	3.568	THY1, NR3C1, COL1A1, PRKACB, CCT6A, CALM1
MAX	3.562	NPM1, SOX9, PRKACB, HSPD1, SUCLG2, GNAQ, GCG, DDX21, BCL2L1, CYCS, GTPBP4, RFC3
MYC	3.497	PAICS, DDX21, NPM1, NME1, DKC1, ABCE1, MCM7, GART, RUVBL1, CYCS, ATP5A1, BRIX1, HSPD1, RPS14, POLR1B, CCND1, SHMT2, ACO2
MXI1	3.21	PAICS, DKC1, KPNA2, TRIP13, THY1, CCNB1, CDC20, UBE2C, NPM1, DDX21, ABCE1, CCND1, MKI67, GTPBP4, RUVBL1, BRIX1, POLR1B, NME1, NOP56, CDKN2A, HJURP, MCM7, ASPM, IQGAP3, CDK1, GART, ANLN, RFC3, HSPD1, KIF2C, CCT6A, NEK2, CYCS, ACAT1, CDC6, GMPS
ZBTB4	3.135	MYC, SOX9, NR3C1, CCND1, PPARGC1A, NCAM1

Identification of master regulators and protein kinases

The iRegulon app identified 18 TFs for hub genes (NES > 3), in which SIN3A was the most considerable upstream regulator with an NES of 8.033 and 64 downstream hub genes (Table 1). Cyclin-dependent kinase 6 (CDK6) was the most salient protein kinase involved in the

phosphorylation of TFs, according to their mean-rank value (Table 2). A GRN was built, including 71 hubs, 18 TFs, 10 protein kinases, and 636 interactions between the nodes (Fig. 5).

Binding sites logos

Among 18 TFs, the binding sites logos of NFYC, NFYB, E2F4, E2F1, FOXS1, MYBL2, CEBPA, YY1, E2F3, MAX, MYC, and MXI1 were available in the JASPAR database (Fig. 6). The minimum and maximum values for the binding sites matching scores were calculated as 14 and 22.64 for E2F1 and MYBL2, respectively.

Table 2. Top-10 ranked protein kinases involved in the phosphorylation of transcription factors regulating the hub genes.

Protein kinase	Mean rank	False discovery rate based on the STRING database	Target proteins
CDK6	13.1	8.73E-04	CEBPA, NFYB, MAX, FOXM1, GLI1, YY1, FOXS1, TFDP1, SIN3A, MYC, E2F1, E2F3, MYBL2, E2F4
CDK4	13.82	6.32E-04	CEBPA, MAX, NFYB, NFYC, FOXM1, GLI1, YY1, FOXS1, TFDP1, SIN3A, MYC, MXI1, E2F1, E2F3, MYBL2, E2F4
GSK3B	17.09	4.72E-03	CEBPA, NFYB, MAX, NFYC, FOXM1, GLI1, YY1, FOXS1, TFDP1, SIN3A, MYC, E2F1, E2F3, MYBL2, E2F4
CDK1	17.5	1.28E-03	CEBPA, MAX, NFYB, NFYC, FOXM1, GLI1, YY1, FOXS1, TFDP1, SIN3A, MYC, E2F1, E2F3, MYBL2, E2F4
CDK2	21.09	9.03E-04	CEBPA, NFYB, NFYC, MAX, FOXM1, GLI1, ZBTB4, YY1, FOXS1, TFDP1, SIN3A, MYC, E2F1, E2F3, MYBL2, E2F4
MAPK8	23.09	8.68E-03	CEBPA, MAX, FOXM1, GLI1, YY1, FOXS1, TFDP1, SIN3A, MYC, E2F1, E2F3, E2F4, MYBL2
MAPK1	25.09	0.024	CEBPA, MAX, NFYB, NFYC, FOXM1, GLI1, YY1, TFDP1, SIN3A, MYC, E2F1, E2F3, E2F4, MYBL2
PRKDC	31.09	0.013	CEBPA, MAX, NFYB, NFYC, FOXM1, GLI1, YY1, FOXS1, TFDP1, SIN3A, MYC, E2F1, E2F3, MYBL2, E2F4
MAPK3	33.55	0.043	CEBPA, MAX, FOXM1, GLI1, YY1, FOXS1, TFDP1, SIN3A, MYC, E2F1, E2F3, E2F4, MYBL2
ATM	33.91	2.61E-03	CEBPA, NFYB, MAX, NFYC, GLI1, ZBTB4, YY1, TFDP1, SIN3A, MYC, E2F1, E2F3, E2F4, MYBL2

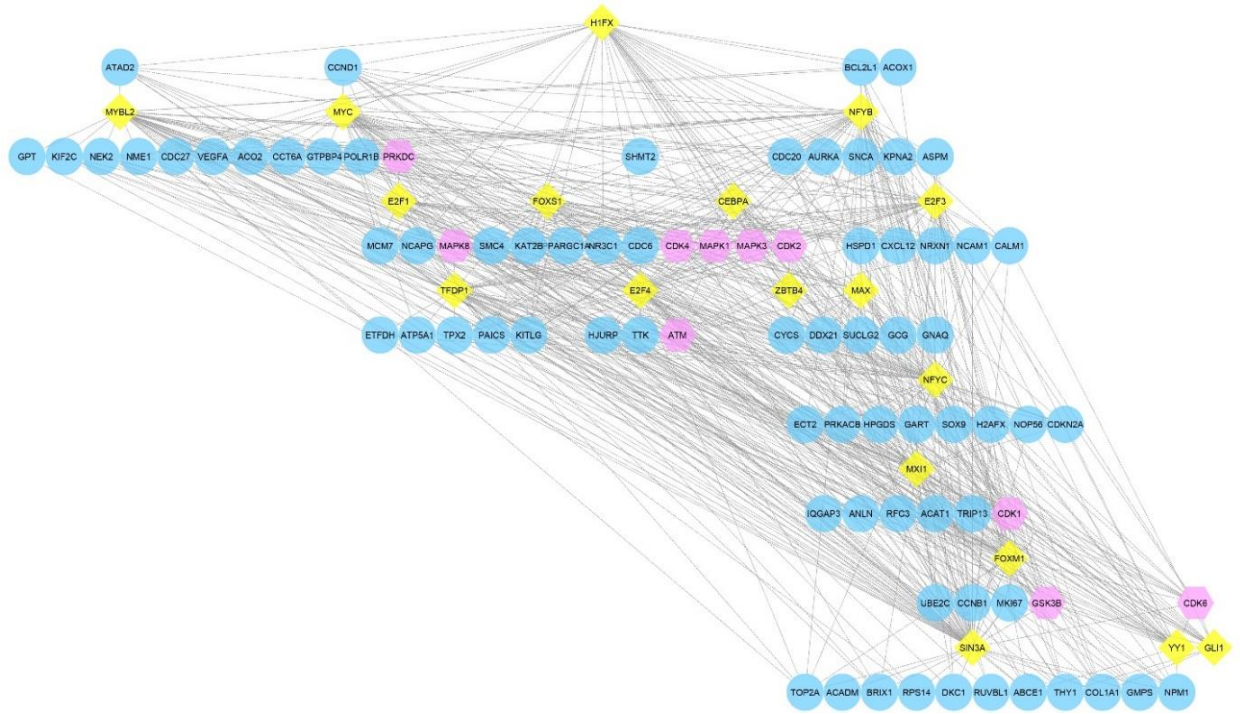


Fig. 5. A gene regulatory network for primary colorectal cancer consisted of 99 nodes including 71 hub genes (blue circles), 18 transcription factors (yellow diamonds), and 10 protein kinases (violet hexagons).

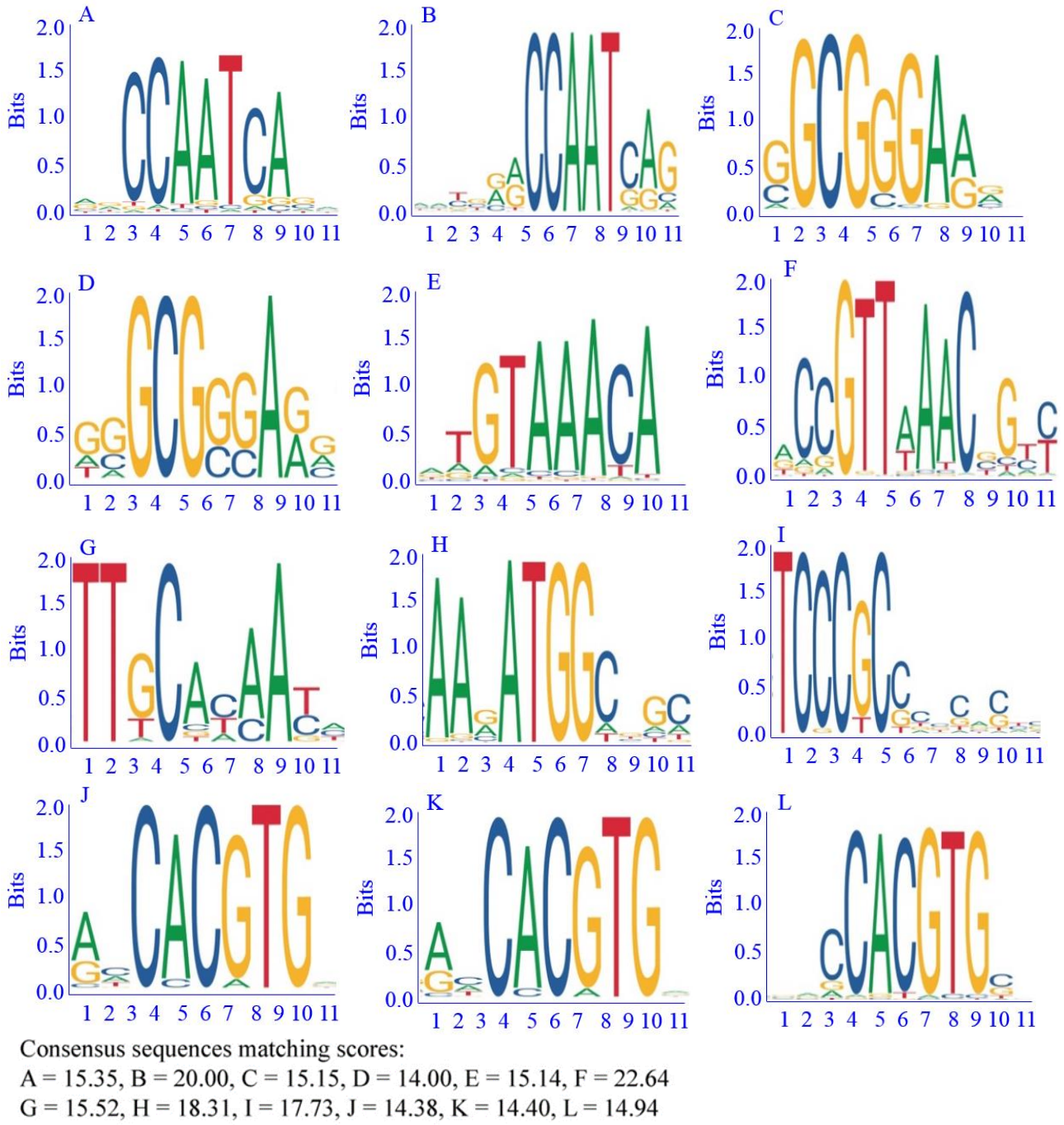


Fig. 6. Logos for the binding sites of (A) NFYC; (B) NFYB; (C) E2F4; (D) E2F1; (E) FOXS1; (F) MYBL2; (G) CEBPA; (H) YY1; (I) E2F3; (J) MAX; (K) MYC; and (L) MXI1.

Prognostic genes and panels in early CRC

Nineteen central genes identified in primary CRC exhibited a significant prognostic function in the context of the disease (log-rank test and HR $P < 0.05$). Over-expression of CDKN2A and down-regulation of SUCLG2, KPNA2, ABCE1, AURKA, PAICS, NPM1, GCG, DDX21, ACOX1, ACADM, GART, CYCS, NCAPG, GMPS, CXCL8, PPARGC1A, ACO2, and ETFDH were associated with a worse outcome in patients with CRC (Table 3). Besides, the under-expression of SUCLG2 and KPNA2 revealed the most considerable negative panel in CRC with the HR and log-rank test P = 2.33 and 0.00014, respectively (Table S7). The novel gene signatures promise to improve the prognosis of CRC patients in clinical settings. Physicians can assess the expression levels of SUCLG2 and KPNA2 in primary CRC patients. If the genes are downregulated, it may indicate a poor prognosis, prompting the medical team to

explore all available therapeutic options for a potentially more successful outcome. However, it's important to note that further confirmation of these findings is necessary in future research. The Kaplan-Meier curves are presented in Fig. S1.

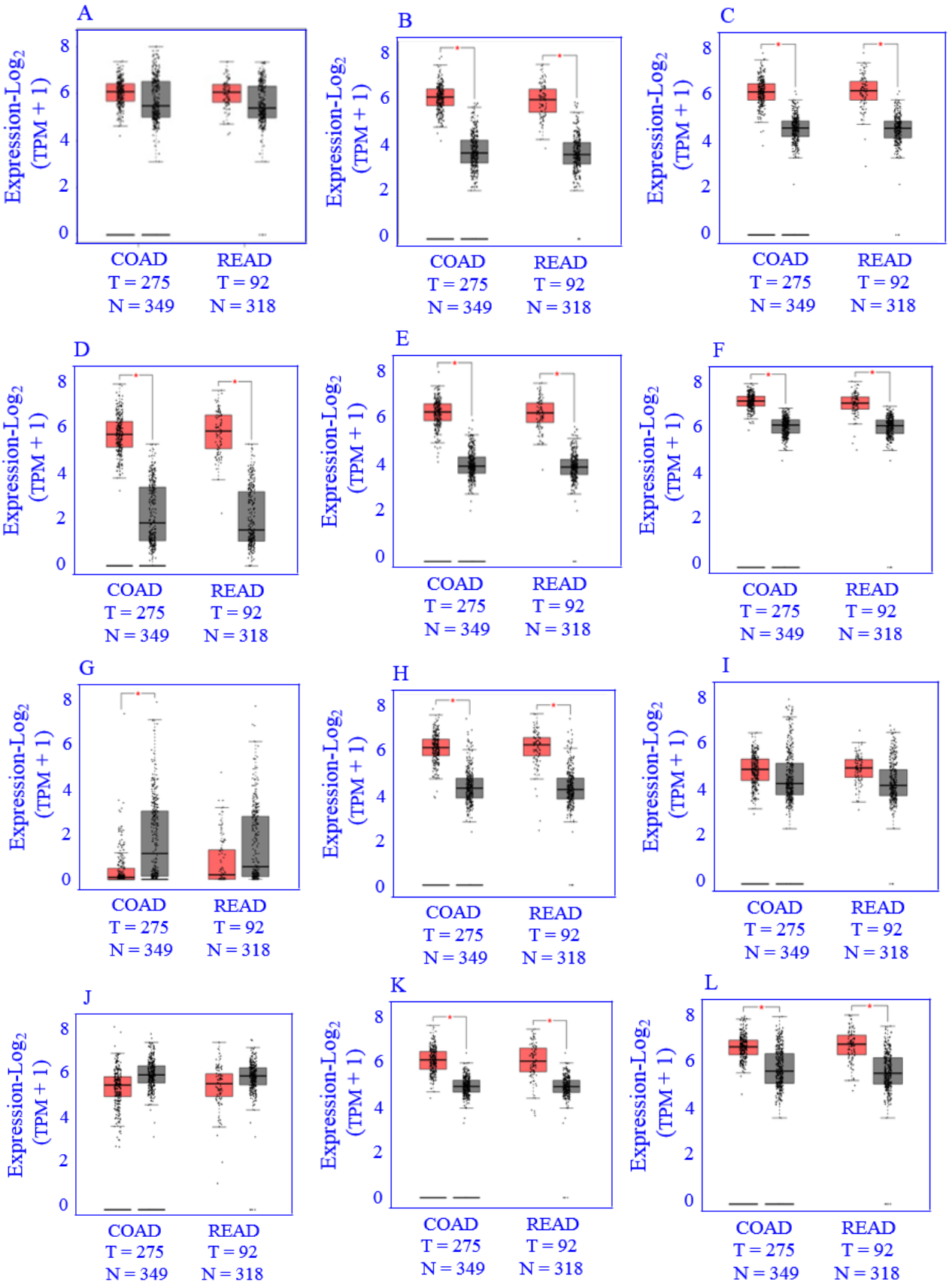
Validation study

The boxplot analysis showed that KPNA2, ABCE1, AURKA, PAICS, NPM1, DDX21, GART, NCAPG, GMPS, CXCL8, and CDKN2A were over-expressed at the mRNA levels in COAD and READ compared to the healthy colorectal tissues. GCG, ACOX1, ACADM, PPARGC1A, and ETFDH also demonstrated under-expression in COAD and READ compared to normal samples, consistent with the present findings (Fig. 7).

Table 3. A total of 19 hub genes in primary CRC demonstrated prognostic impact in colon adenocarcinoma and rectum adenocarcinoma based on the Gene Expression Profiling Interactive Analysis 2 database.

Single gene				
Gene symbol (label)	**HR (high)**	**Log-rank *P***	**P_{HR}**	**Log2 FC, primary CRC/normal**
SUCLG2 (A)	0.5	0.0016	0.002	-1.346
KPNA2 (B)	0.51	0.0024	0.0029	1.063
ABCE1 (C)	0.52	0.0043	0.005	1.009
AURKA (D)	0.53	0.0041	0.0047	1.662
PAICS (E)	0.53	0.0039	0.0045	1.447
NPM1 (F)	0.54	0.0056	0.0064	1.540
GCG (G)	0.54	0.0063	0.0073	-3.512
DDX21 (H)	0.57	0.013	0.014	1.152
ACOX1 (I)	0.57	0.011	0.012	-1.184
ACADM (J)	0.59	0.02	0.022	-1.483
GART (K)	0.6	0.023	0.025	1.140
CYCS (L)	0.6	0.022	0.024	-1.076
NCAPG (M)	0.61	0.024	0.026	1.743
GMPS (N)	0.61	0.025	0.027	1.112
CXCL8 (O)	0.62	0.032	0.034	3.252
PPARGC1A (P)	0.62	0.029	0.03	-2.067
ACO2 (Q)	0.64	0.043	0.044	-1.018
ETFDH (R)	0.65	0.047	0.049	-1.816
CDKN2A (S)	1.7	0.021	0.022	1.960
Signature				
Prognostic panel	**HR (high)**	**Log-rank *P***	**P_{HR}**	**Log2 FC, primary CRC/normal**
A+B	0.43	0.00014	0.00021	-
A to C	0.53	0.0041	0.0048	-
A to D	0.48	0.00096	0.0012	-
A to E	0.5	0.0017	0.002	-
A to F	0.49	0.0015	0.0018	-
A to G	0.48	0.00091	0.0012	-
A to H	0.5	0.002	0.0024	-
A to I	0.45	0.00047	0.00066	-
A to J	0.49	0.0016	0.002	-
A to K	0.47	0.0008	0.0011	-
A to L	0.54	0.0057	0.0066	-
A to M	0.56	0.0094	0.01	-
A to M	0.53	0.0043	0.005	-
A to O	0.44	0.00024	0.00036	-
A to P	0.44	0.00023	0.00035	-
A to Q	0.47	0.00097	0.0013	-
A to R	0.47	0.00083	0.0011	-

CRC, colorectal cancer; HR, hazard ratio; FC, fold change.


A
B
C
D
E
F
G
H
I
J
K
L
Expression-Log2 (TPM + 1)
0
2
4
6
8
COAD
T = 275
N = 349
READ
T = 92
N = 318

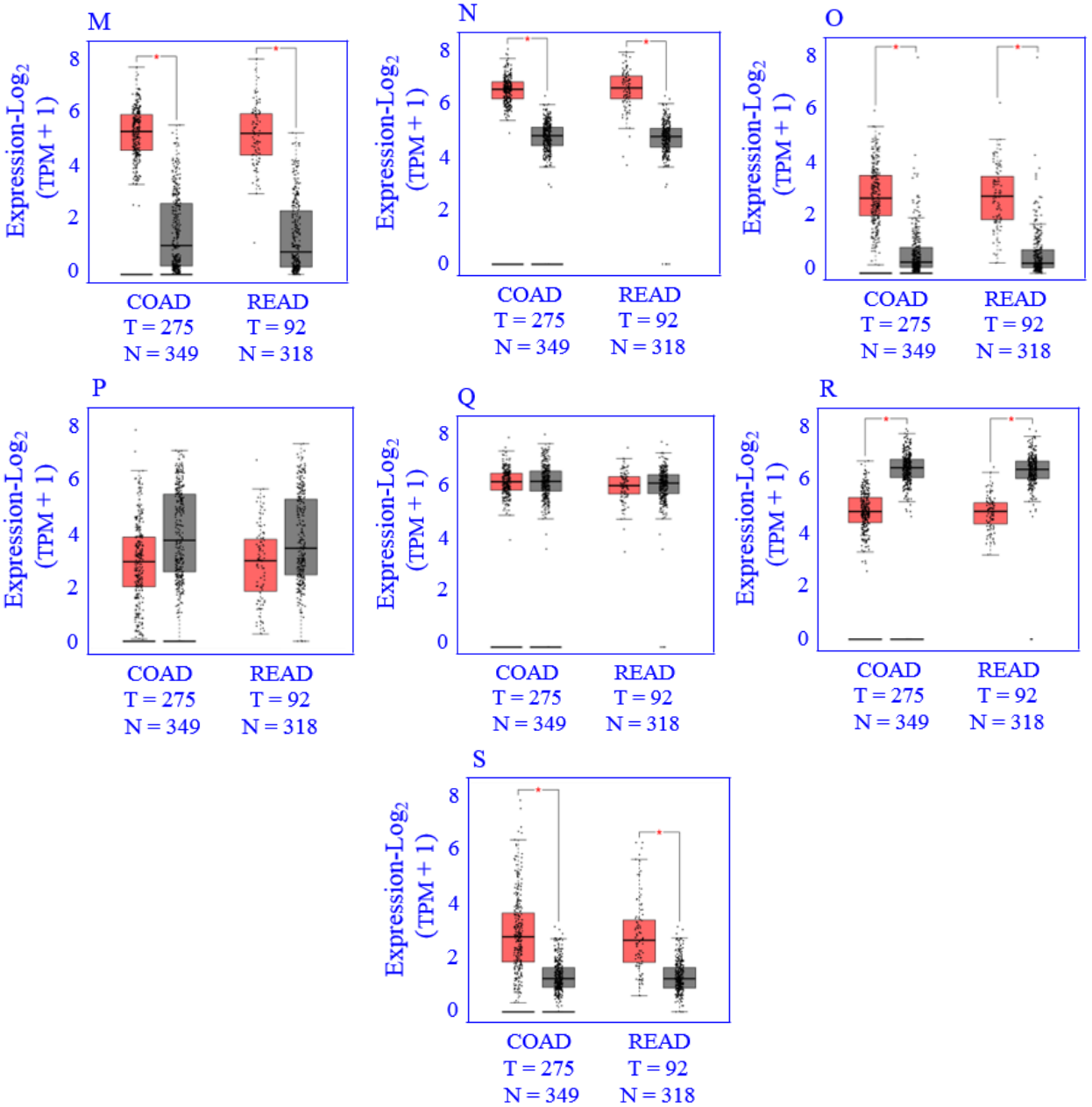


Fig. 7. Validation analysis for prognostic markers in primary CRC at the mRNA levels in COAD and READ tissues compared to the healthy samples using the GEPIA2 database. Box plots are based on 367 cancerous samples (red color) and 667 healthy controls (gray color). (A) SUCLG2; (B) KPNA2; (C) ABCE1; (D) AURKA; (E) PAICS; (F) NPM1; (G) GCG; (H) DDX21; (I) ACOX1; (J) ACADM; (K) GART; (L) CYCS; (M) NCAPG; (N) GMPS; (O) CXCL8; (P) PPARGC1A; (Q) ACO2; (R) ETFDH; and (S) CDKN2A. CRC, colorectal cancer; COAD, colon adenocarcinoma; READ, rectum adenocarcinoma.

DISCUSSION

Patients with advanced-stage CRC may illustrate a dismal prognosis because of tumor recurrence. Therefore, it is necessary to identify new markers and signaling pathways involved in the initiation and development of the disease as well as biomarkers to predict the prognosis of CRC patients (51). The present study identified 75 hub genes mediating the malignant transformation of non-cancerous colorectal tissue to primary CRC based on the network biology approach. Up-regulation of 18 genes, including SUCLG2, KPNA2, ABCE1, AURKA, PAICS, NPM1, GCG, DDX21, ACOX1, ACADM, GART, CYSC, NCAPG, GMPS, CXCL8, PPARGC1A, ACO2, and ETFDH was significantly associated with a favorable prognosis in patients with CRC. In comparison, over-expression of CDKN2A was related to a poor prognosis in CRC patients (log-rank test and HR $P < 0.05$). Moreover, it

was found that combining SUCLG2 and KPNA2 led to a higher prognostic power compared to any of the genes in CRC (log-rank test $P = 0.00014$; HR = 0.4). Except for SUCLG2 and CYSC, the over/under-expression of other prognostic markers in primary CRC was validated using the GEPIA2 online tool.

The present study demonstrated significant down-regulation of Succinate-CoA ligase (GDP-forming) subunit beta, mitochondrial SUCLG2 at the early-stage CRC compared to healthy control colorectal tissues (FC = 0.39, $P = 0.000357$). Previous studies have also reported reduced activity of mitochondria in primary CRC (52,53). Literature reported the down-regulation of SUCLG2, HIG1 domain family member 1A (HIGD1A), and calcium-binding mitochondrial carrier protein SCaMC-1 (SLC25A24) at the mRNA and protein levels in patients with CRC, suggesting that 3 mitochondrial genes are involved in the initiation, progression, and prognosis of CRC. SUCLG2, HIGD1A, and SLC25A24 take part in several important BPs in the cell, including the tricarboxylic acid (TCA) cycle, apoptosis, and anaerobic environment (54-57). The SUCLG2 gene mediates succinate production in the Krebs cycle (58).

As per the findings of this current investigation, the karyopherin subunit alpha-2 (KPNA2) was significantly over-expressed in primary CRC (FC = 2.08, $P = 0.0000354$). KPNA2 is involved in transporting proteins from the cytoplasm into the nucleus (59) and performs a crucial function in the repair of DNA double-strand breaks (60). Previous studies have reported the association between KPNA2 and poor prognosis in several solid tumors, including breast cancer (61), esophageal squamous cell carcinoma (62), and gastric cancer (63). Takada *et al*. performed a study to assess the levels of protein expression in KPNA2 in CRC tissues and examined the prognostic impact of the gene on the disease using the immunohistochemistry analysis and reported that KNPA2 was over and under-expressed in approximately 75 and 25 percent of the CRC cases, respectively (64). KPNA2 up-regulation was significantly associated with lymphatic invasion ($P = 0.0245$), a dismal overall survival ($P = 0.00374$), and resistance to hyperthermochemoradiation therapy in CRC patients. Besides, up-regulation of KPNA2 was associated with a favorable prognosis in patients with CRC based on the GEPIA2 analysis (HR = 0.51, log-rank test $P = 0.0024$), suggesting that more studies are required to elucidate the exact role of KPNA2 in the tumorigenesis of CRC and its potential impact in the prognosis of the disease.

According to the present results, cyclin-dependent kinase inhibitor 2A (CDKN2A) was considerably over-expressed in early CRC (FC = 3.89, $P = 0.000429$). Drawing upon the Kaplan-Meier analysis conducted using the GEPIA2 tool, CDKN2A showed a significant association with an unfavorable prognosis in individuals diagnosed with CRC (HR = 1.7, log-rank test $P = 0.021$). CDKN2A diminishes cell growth by stopping the cell cycle at the G1 phase (65). Several previous studies have reported mutation and deletion of CDKN2A in various human carcinomas, leading to enhanced cell proliferation and cancer progression due to the down-regulation of the gene (66,67). Therefore, it may be suggested that increased CDKN2A expression in primary CRC may be a cellular defense mechanism in cancer cells to combat enhanced proliferation.

The identified hub genes such as MYC, EGFR, and MET play pivotal roles in regulating key pathways involved in colorectal carcinogenesis and progression including proliferation, apoptosis evasion, angiogenesis, metastasis, and stemness (68,69). Specifically, MYC drives uncontrolled cell cycle progression, EGFR activates MAPK/ERK and PI3K/AKT signaling to enhance survival and growth, while MET promotes EMT, migration, and metastasis. Their aberrant activation allows cells to bypass growth suppression and acquire hallmark cancer capabilities, thereby contributing to the initiation and malignant progression of CRC (70).

The present study executed more analyses to uncover upstream regulators of the hub genes and protein kinases involved in the malignant transformation of non-cancerous colorectal tissues to early-stage CRC. Accordingly, paired amphipathic helix protein Sin3a (SIN3A) and CDK6 were the most salient TFs and protein kinases, respectively.

The SIN3A/HDAC is a multi-scaffolding protein contributing to down-regulating transcription of several genes through histone deacetylation (71-73). SIN3A has shown contradictory functions in different cancers. In this regard, it has been reported that SIN3A is an inducer of proliferation and acts as an anti-apoptotic factor in lymphoma and sarcoma cell lines (74). According to another study, Ren *et al*. reported that miR-210-3p elevated cell proliferation and attenuated apoptosis in non-small cell lung cancer cells by targeting SIN3A (75). Nan *et al*. demonstrated that LINC00665 could lead to over-expression of SIN3A in CRC by sponging miR-138-5p, leading to tumor progression. Sponging refers to the ability of LINC00665, a long non-coding RNA, to bind to and soak up or "sponge" miR-138-5p microRNA molecules. This reduces the miR-138-5p available to downregulate its target gene, SIN3A. So, LINC00665 acts as a sponge or decoy for miR-138-5p, leading to increased expression of SIN3A and promoting tumor progression in CRC (76). Therefore, it may be hypothesized that SIN3A is an oncogene in CRC. Accordingly, more studies are suggested to elucidate the exact role of SIN3A in cell growth and tumor progression.

Previous studies have reported that CDK6 induces cell proliferation through the G1 phase and is involved in cancer development (77). Moreover, it regulates the catalytic activity of pyruvate kinase M2 and 6-phosphofructokinase (78). Liu *et al*. reported that miR-500a-3p acts as a tumor suppressor molecule in CRC, reducing cell proliferation and glycolysis in tumor cells. The authors demonstrated that CDK6 is a direct target of miR-500a-3p; this was done using a dual-luciferase reporter assay (79). According to the present results, as well as the effects of former studies, it may be speculated that the CDK6 functions as an oncogene in CRC.

Further analysis in this study revealed that rRNA processing was significantly involved in top-10 enriched pathways and BPs associated with early CRC. Growing evidence suggested an association between the disruption of the human gut microbiome and the development of CRC (80). In recent years, the 16S rRNA gene sequencing approach has been widely used to monitor the gut microbiome architecture in different stages of CRC (81), suggesting the impact of rRNA processing in the etiology of CRC. The analysis of human rRNA processing genes and bacterial 16S rRNA genes in CRC patients indicates ribosome function and protein synthesis dysregulation. The abnormal expression of human rRNA genes involved in ribosome production suggests disturbed ribosome biogenesis in CRC cells. Similarly, differences in the 16S rRNA genes of the CRC-associated gut microbiome imply altered microbial protein synthesis. The human gene and microbial gene analyses implicate ribosome dysfunction as a potential common mechanism linking disturbed cellular metabolism in human cells and gut bacteria to colorectal carcinogenesis (81).

Also, in another study, Salehi *et al*. proposed a framework to identify miRNA biomarkers associated with CRC metastasis to the liver (82). The authors analyzed miRNA expression profiles in primary CRC tumors with and without liver metastases. They identified a signature of 5 miRNAs (miR-203, miR-135b, miR-141, miR-125a-3p, and miR-34c-5p) that could discriminate primary tumors with liver metastasis from those without metastasis. The miRNA signature was validated in an independent cohort of CRC samples, confirming its ability to predict liver metastasis (82). Pathway analysis revealed the 5 miRNAs target genes involved in pathways deregulated in CRC, such as Wnt, TGF-β, VEGF, and MAPK signaling. Further research suggested that miRNAs promote processes involved in metastasis like angiogenesis, epithelial-mesenchymal transition, and extracellular matrix degradation. The authors proposed the 5-miRNA signature could serve as a biomarker to identify stage II/III CRC patients at high risk of developing liver metastasis, allowing personalized therapy to prevent metastasis. (82). The study shed light on the underlying biology of CRC liver metastasis and demonstrated the potential of miRNAs as promising biomarkers for improved clinical management. Only 23 cancerous and 9 healthy colorectal tissues were included in the GSE81582 dataset, which showed that the sample size was small. Additionally, the

GPL15207 platform may not present all gene symbols. Therefore, using more datasets with large sample sizes in future studies and confirming the results using experimental methods is recommended.

While this research has unearthed several promising prognostic biomarkers and pathways that were altered in primary CRC with susceptibility to liver metastasis, it is essential to acknowledge certain constraints within the study. The relatively modest sample size limited the breadth of conclusions that could be drawn. To establish the clinical utility of the proposed biomarkers definitively, a more extensive patient cohort must be incorporated for validation purposes. Furthermore, as the analysis relied on pre-existing microarray data, there was no subsequent confirmation of pivotal genes through reverse transcription-quantitative polymerase chain reaction (RT-qPCR). Therefore, follow-up investigations must employ alternative methods to substantiate the observed expression changes. Furthermore, another limitation was the absence of functional studies to elucidate the mechanistic roles of the top differentially expressed genes. Future experiments should aim to ascertain whether these genes actively contribute to liver metastasis or are merely correlated with it. Lastly, the exclusive inclusion of male patients diminished the generalizability of the findings to both genders. Furthermore, the dataset was curated using tissue samples from patients and individuals in Spain. Hence, it's essential to acknowledge that the current findings may not be applicable or generalizable to all CRC patients worldwide. Factors like mutations, epigenetics, microbiome, diet, and lifestyle may also modulate gene expression profiles in CRC patients. Large cohort studies capturing comprehensive clinical and molecular data could help assess these variables and their impacts on the CRC transcriptome and key driver genes. This could reveal additional biomarkers and refine signatures tailored to patient subpopulations. Besides, the mRNA profiles in this study were derived from the GPL15207 platform, which likely captures only a portion of the mRNA diversity. Consequently, the DEGs identified in our study may not necessarily represent all significant ones in CRC.

Notwithstanding the previously mentioned limitations, this study represented a noteworthy stride toward a more comprehensive comprehension of the molecular alterations underpinning aggressive behavior in CRC. The pathways and genes uncovered herein have the potential to form the basis for prognostic models aimed at predicting the risk of liver metastasis. Upon further validation, the proposed biomarkers may prove invaluable in guiding clinical decisions regarding chemotherapy and metastasis surveillance. Furthermore, the spotlighted genes and pathways offer fresh avenues for developing therapeutic targets, including pharmaceuticals or biologics, to impede the progression of CRC. In sum, this research laid the groundwork for an expanded exploration of novel prognostic and therapeutic strategies for CRCs, exhibiting heightened metastatic potential. Addressing the limitations above by conducting more extensive investigations involving diverse patient cohorts, validating pivotal genes, undertaking functional characterization, and incorporating both genders will optimize the clinical impact of this research trajectory.

The observed dysregulation of CDK6 and CDKN2A in CRC provided a rationale for exploring selective CDK6 inhibitors or restoring CDKN2A expression as novel therapeutic strategies, either alone or combined with chemotherapy, to suppress proliferation in CRC (83,84). Based on these findings further research into developing potent and specific CDK6 inhibitor drugs and evaluating their efficacy in CRC trials is warranted.

KPNA2 functions as a nuclear transport receptor that mediates the nuclear translocation of key oncoproteins. Overexpression of KPNA2 enables the nuclear import of metastasis-promoting and anti-apoptotic proteins, thereby promoting tumorigenesis, progression, and poorer prognosis in CRC. As an independent prognostic biomarker, elevated KPNA2 levels indicated worse overall and relapse-free survival in CRC patients, highlighting its potential as a prognostic predictor to guide patient management (85).

CONCLUSION

The current study determined 1113 DEGs (474 up-regulated and 639 down-regulated genes) in early CRC with a high potential of liver metastases compared to healthy colorectal specimens. Moreover, 75 genes demonstrated a salient centrality in a PIM associated with the malignant transformation of healthy colorectal tissues to primary CRC, in which 19 genes revealed a significant impact on the prognosis of the disease. The boxplot analysis confirmed the mRNA expression patterns of KPNA2, ABCE1, AURKA, PAICS, NPM1, DDX21, GART, NCAPG, GMPS, CXCL8, CDKN2A, GCG, ACOX1, ACADM, PPARGC1A, and ETFDH. Down-regulation of SUCLG2 and KPNA2 demonstrated the worst negative panel in patients with CRC. It is suggested that SIN3A and CDK6 are substantially involved in regulating the expression of hubs and phosphorylation of TFs, respectively. According to the gene set enrichment analysis, rRNA processing considerably mediated the malignant transformation of normal colorectal tissues to primary CRC with a high risk of liver metastases. These results have the potential to aid in predicting the outlook for patients with primary CRC and might contribute to the discovery of new targets for drug development in CRC therapy.

Acknowledgments

All authors would like to thank the Deputy of Research and Technology, and Research Center for Molecular Medicine, Hamadan University of Medical Sciences, Hamadan, Iran for their support. This paper was extracted from Fatemeh Bahramibanan's thesis. This work was financially supported by the Vice-Chancellor for Research, the Hamadan University of Medical Science (Grant No. 140106084713).

Conflict of interest statement

All authors declare no conflict of interest in this study.

Authors' contributions

A. Taherkhani and M. Soleimani designed the study; A. Taherkhani and F. Bahramibanan conducted and interpreted the statistical analysis, protein-protein interaction network, gene regulatory network, and enrichment analyses as well as wrote the manuscript; A. Taherkhani, M. Soleimani, F. Bahramibanan, R. Najafi, N, Alizadeh, K. Derakhshandeh, N. Barati, and H. Ghadimipour analyzed and discussed the results; M. Soleimani and R. Najafi edited the manuscript. All authors read and approved the final version of the manuscript.

Supplementary materials

The supplementary materials for this article can be found online at: https://github.com/mynewsupplementary/supplementary.

REFERENCES

1. Xi Y, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. Transl. Oncol. 2021;14(10):101174,1-7. DOI: 10.1016/j.tranon.2021.101174.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69(1):7-34. DOI: 10.3322/caac.21551.
3. 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(6):394-424. DOI: 10.3322/caac.21492.
4. Manoochehri H, Jalali A, Tanzadehpanah H, Taherkhani A, Najafi R. Aptamer-conjugated nanoliposomes containing COL1A1 siRNA sensitize CRC cells to conventional chemotherapeutic drugs. Colloids Surf B Biointerfaces. 2022;218:112714. DOI: 10.1016/j.colsurfb.2022.112714.
5. Noei A, Nili-Ahmadabadi A, Soleimani M. The enhanced cytotoxic effects of the p28-apoptin chimeric protein as a novel anti-cancer agent on breast cancer cell lines. Drug Res. 2019;69(03):144-150. DOI: 10.1055/a-0654-4952.
6. Soleimani M, Mahnam K, Mirmohammad-Sadeghi H, Sadeghi-Aliabadi H, Jahanian-Najafabadi A. Theoretical design of a new chimeric protein for the treatment of breast cancer. Res Pharm Sci. 2016;11(3):187-199. PMID: 27499788.
7. Lichtenstern CR, Ngu RK, Shalapour S, Karin M. Immunotherapy, inflammation and colorectal cancer. Cells. 2020;9(3):618,1-18. DOI: 10.3390/cells9030618.
8. Madadi S, Soleimani M. Comparison of miR-16 and cel-miR-39 as reference controls for serum miRNA normalization in colorectal cancer. J Cell Biochem. 2019;120(4):4802-4803. DOI: 10.1002/jcb.28174.

9. Bijari N, Ghobadi S, Derakhshandeh K. β-lactoglobulin-irinotecan inclusion complex as a new targeted nanocarrier for colorectal cancer cells. Res Pharm Sci. 2019;14(3):216-227. DOI: 10.4103/1735-5362.258488.
10. Wolf AMD, Fontham ETH, Church TR, Flowers CR, Guerra CE, LaMonte SJ, *et al*. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. CA Cancer J Clin. 2018;68(4):250-281. DOI: 10.3322/caac.21457.
11. Nikolaou S, Qiu S, Fiorentino F, Rasheed S, Tekkis P, Kontovounisios C. Systematic review of blood diagnostic markers in colorectal cancer. Tech Coloproctol. 2018;22(7):481-498. DOI: 10.1007/s10151-018-1820-3.
12. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RG, Barzi A, *et al*. Colorectal cancer statistics, 2017. CA Cancer J Clin. 2017;67(3):177-193. DOI: 10.3322/caac.21395.
13. Sayagués JM, Corchete LA, Gutiérrez ML, Sarasquete ME, del Mar Abad M, Bengoechea O, *et al*. Genomic characterization of liver metastases from colorectal cancer patients. Oncotarget. 2016;7(45):72908-72922. DOI: 10.18632/oncotarget.12140.
14. Van Cutsem E, Cervantes A, Nordlinger B, Arnold D. Metastatic colorectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2014;25(Suppl. 3):iii1-9. DOI: 10.1093/annonc/mdu260.
15. Kopetz S, Chang GJ, Overman MJ, Eng C, Sargent DJ, Larson DW, *et al*. Improved survival in metastatic colorectal cancer is associated with adoption of hepatic resection and improved chemotherapy. J Clin Oncol. 2009;27(22):3677-3683. DOI: 10.1200/JCO.2008.20.5278.
16. Augestad KM, Lindsetmo RO, Stulberg J, Reynolds H, Senagore A, Champagne B, *et al*. International preoperative rectal cancer management: staging, neoadjuvant treatment, and impact of multidisciplinary teams. World J Surg. 2010;34:2689-2700. DOI: 10.1007/s00268-010-0738-3.
17. Smith JJ, Deane NG, Wu F, Merchant NB, Zhang B, Jiang A, *et al*. Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. Gastroenterology. 2010;138(3):958-968. DOI: 10.1053/j.gastro.2009.11.005.
18. Weng WK, Levy R. Two immunoglobulin G fragment C receptor polymorphisms independently predict response to rituximab in patients with follicular lymphoma. J Clin Oncol. 2003;21(21):3940-3947. DOI: 10.1200/JCO.2003.05.013.
19. Overman MJ, McDermott R, Leach JL, Lonardi S, Lenz HJ, Morse MA, *et al*. Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. Lancet Oncol. 2017;18(9):1182-1191. DOI: 10.1016/S1470-2045(17)30422-9.
20. Loboda A, Nebozhyn MV, Watters JW, Buser CA, Shaw PM, Huang PS, *et al*. EMT is the dominant program in human colon cancer. BMC Med Genomics. 2011;4(1):1-10. DOI: 10.1186/1755-8794-4-9.
21. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, *et al*. NCBI GEO: archive for functional genomics data sets-update. Nucleic Acids Res. 2013;41(D1):D991-D995. DOI: 10.1093/nar/gks1193.
22. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, *et al*. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 2021;49(D1):D605-D612. DOI: 10.1093/nar/gkaa1074.
23. Bayat Z, Ahmadi-Motamayel F, Parsa MS, Taherkhani A. Potential biomarkers and signaling pathways associated with the pathogenesis of primary salivary gland carcinoma: a bioinformatics study. Genomics Inform. 2021;19(4):1-17. DOI: 10.5808/gi.21052.
24. Saito R, Smoot ME, Ono K, Ruscheinski J, Wang PL, Lotia S, *et al*. A travel guide to Cytoscape plugins. Nat Methods. 2012;9(11):1069-1076. DOI: 10.1038/nmeth.2212.
25. Bayat Z, Farhadi Z, Taherkhani A. Identification of potential biomarkers associated with poor prognosis in oral squamous cell carcinoma through integrated bioinformatics analysis: a pilot study. Gene Rep. 2021;24:101243,1-13. DOI: 10.1016/j.genrep.2021.101243.
26. Reimand J, Arak T, Adler P, Kolberg L, Reisberg S, Peterson H, *et al*. g: Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res. 2016;44(W1):W83-W89. DOI: 10.1093/nar/gkw199.
27. Hubbard T, Barker D, Birney E, Cameron G, Chen Y, Clark L, *et al*. The Ensembl genome database project. Nucleic Acids Res. 2002;30(1):38-41. DOI: 10.1093/nar/30.1.38.
28. Howe KL, Contreras-Moreira B, De Silva N, Maslen G, Akanni W, Allen J, *et al*. Ensembl Genomes 2020-enabling non-vertebrate genomic research. Nucleic Acids Res. 2020;48(D1):D689-D695. DOI: 10.1093/nar/gkz890.
29. Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. WormBase ParaSite- a comprehensive resource for helminth genomics. Mol Biochem Parasitol. 2017;215:2-10. DOI: 10.1016/j.molbiopara.2016.11.005.
30. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. Nucleic Acids Res. 2021;49(D1):D545-D551. DOI: 10.1093/nar/gkaa970.
31. Croft D, O'kelly G, Wu G, Haw R, Gillespie M, Matthews L, *et al*. Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res. 2011;39(Database issue):D691-D697. DOI: 10.1093/nar/gkq1018.

32. Slenter DN, Kutmon M, Hanspers K, Riutta A, Windsor J, Nunes N, *et al*. WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. Nucleic Acids Res. 2018;46(D1):D661-D667. DOI: 10.1093/nar/gkx1064.
33. Manoochehri H, Jalali A, Tanzadehpanah H, Taherkhani A, Saidijam M. Identification of key gene targets for sensitizing colorectal cancer to chemoradiation: an integrative network analysis on multiple transcriptomics data. J Gastrointest Cancer. 2022;53(3):649-668. DOI: 10.1007/s12029-021-00690-2.
34. Tao W, Radstake TR, Pandit A. RegEnrich gene regulator enrichment analysis reveals a key role of the ETS transcription factor family in interferon signaling. Commun Biol. 2022;5(1):1-12. DOI: 10.1038/s42003-021-02991-5.
35. Janky Rs, Verfaillie A, Imrichová H, Van de Sande B, Standaert L, Christiaens V, *et al*. iRegulon: from a gene list to a gene regulatory network using large motif and track collections. PLoS Comput Biol. 2014;10(7):e1003731,1-19. DOI: 10.1371/journal.pcbi.1003731.
36. Bryne JC, Valen E, Tang MHE, Marstrand T, Winther O, da Piedade I, *et al*. JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. Nucleic Acids Res. 2008;36(Database issue): D102-D106. DOI: 10.1093/nar/gkm955.
37. Gentleman R. R programming for bioinformatics (Chapman & Hall/CRC Computer Science & Data Analysis). 1st Edition. New York:Chapman and Hall/CRC;2009.
38. Xiong J. Essential bioinformatics. 1st edition. New York:Cambridge University Press; 2006.
39. Köstler WJ and Zielinski CC. Targeting receptor tyrosine kinases in cancer. In: Receptor tyrosine kinases: structure, functions and role in human disease. New York: Springer; 2015. pp. 225-278. DOI: 10.1007/978-1-4939-2053-2_10.
40. Kittler H, Tschandl P. Driver mutations in the mitogen-activated protein kinase pathway: the seeds of good and evil. Br J Dermatol. 2018;178:26-27. DOI: 10.1111/bjd.16119.
41. Maurer G, Tarkowski B, Baccarini M. Raf kinases in cancer-roles and therapeutic opportunities. Oncogene. 2011;30(32):3477-3488. DOI: 10.1038/onc.2011.160.
42. Bhullar KS, Lagarón NO, McGowan EM, Parmar I, Jha A, Hubbard BP, *et al*. Kinase-targeted cancer therapies: progress, challenges and future directions. Mol Cancer. 2018;17(1):48,1-20. DOI: 10.1186/s12943-018-0804-2.
43. Mahfuz A, Zubair-Bin-Mahfuj A, Podder DJ. A network-biology approach for identification of key genes and pathways involved in malignant peritoneal mesothelioma. Genomics Inform. 2021;19(2),1-14. DOI: 10.5808/gi.21019.
44. Lachmann A, Ma'ayan A. KEA: kinase enrichment analysis. Bioinformatics. 2009;25(5):684-686. DOI: 10.1093/bioinformatics/btp026.
45. 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(W1):W98-W102. DOI: 10.1093/nar/gkx247.
46. Consortium G. The Genotype-Tissue Expression (GTEx) pilot analysis: multi tissue gene regulation in humans. Science. 2015;348(6235):648-660. DOI: 10.1126/science.1262110.
47. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol. 2015;19(1A):A68-A77.
48. Goedhart J, Luijsterburg MS. VolcaNoseR is a web app for creating, exploring, labeling and sharing volcano plots. Sci Rep. 2020;10(1):1-5. DOI: 10.1038/s41598-020-76603-3.
49. Binns D, Dimmer E, Huntley R, Barrell D, O'donovan C, Apweiler R. QuickGO: a web-based tool for Gene Ontology searching. Bioinformatics. 2009;25(22):3045-3406. DOI: 10.1093/bioinformatics/btp536.
50. Jeong H, Mason SP, Barabási AL, Oltvai ZN. Lethality and centrality in protein networks. Nature. 2001;411(6833):41-42. DOI: 10.1038/35075138.
51. Gao QZ, Qin Y, Wang WJ, Fei BJ, Han WF, Jin JQ, *et al*. Overexpression of AMPD2 indicates poor prognosis in colorectal cancer patients *via* the Notch3 signaling pathway. World J Clin Cases. 2020;8(15):3197-3208. DOI: 10.12998/wjcc.v8.i15.3197.
52. Haberman Y, Karns R, Dexheimer PJ, Schirmer M, Somekh J, Jurickova I, *et al*. Ulcerative colitis mucosal transcriptomes reveal mitochondriopathy and personalized mechanisms underlying disease severity and treatment response. Nat Commun. 2019;10(1):1-13. DOI: 10.1038/s41467-018-07841-3.
53. Jackson DN, Theiss AL. Gut bacteria signaling to mitochondria in intestinal inflammation and cancer. Gut microbes. 2020;11(3):285-304. DOI: 10.1080/19490976.2019.1592421.
54. Zhang W, Lin L, Xia L, Cai W, Dai W, Zou C, *et al*. Multi-omics analyses of human colorectal cancer revealed three mitochondrial genes potentially associated with poor outcomes of patients. J Transl Med. 2021;19(1):1-16. DOI: 10.1186/s12967-021-02939-7.
55. Nagao T, Shintani Y, Hayashi T, Kioka H, Kato H, Nishida Y, *et al*. Higd1a improves respiratory function in the models of mitochondrial disorder. FASEB J. 2020;34(1):1859-1571. DOI: 10.1096/fj.201800389R.
56. Chinopoulos C, Batzios S, van den Heuvel LP, Rodenburg R, Smeets R, Waterham HR, *et al*. Mutated SUCLG1 causes mislocalization of SUCLG2 protein, morphological alterations of mitochondria and an early-onset severe neurometabolic disorder. Mol Genet Metab. 2019;126(1):43-52. DOI: 10.1016/j.ymgme.2018.11.009.

57. Writzl K, Maver A, Kovačič L, Martinez-Valero P, Contreras L, Satrustegui J, *et al*. *De novo* mutations in SLC25A24 cause a disorder characterized by early aging, bone dysplasia, characteristic face, and early demise. Am J Hum Genet. 2017;101(5):844-855. DOI: 10.1016/j.ajhg.2017.09.017.
58. Dobolyi A, Bagó AG, Gál A, Molnár MJ, Palkovits M, Adam-Vizi V, *et al*. Localization of SUCLA2 and SUCLG2 subunits of succinyl CoA ligase within the cerebral cortex suggests the absence of matrix substrate-level phosphorylation in glial cells of the human brain. J Bioenerg Biomembr. 2015;47(1-2):33-41. DOI: 10.1007/s10863-014-9586-4.
59. Lange A, Mills RE, Lange CJ, Stewart M, Devine SE, Corbett AH. Classical nuclear localization signals: definition, function, and interaction with importin α. J Biol Chem. 2007;282(8):5101-5105. DOI: 10.1074/jbc.R600026200.
60. Tseng SF, Chang CY, Wu KJ, Teng SC. Importin KPNA2 is required for proper nuclear localization and multiple functions of NBS1. J Biol Chem. 2005;280(47):39594-39600. DOI: 10.1074/jbc.M508425200.
61. Gluz O, Wild P, Meiler R, Diallo-Danebrock R, Ting E, Mohrmann S, *et al*. Nuclear karyopherin α2 expression predicts poor survival in patients with advanced breast cancer irrespective of treatment intensity. Int J Cancer Res. 2008;123(6):1433-1438. DOI: 10.1002/ijc.23628.
62. Sakai M, Sohda M, Miyazaki T, Suzuki S, Sano A, Tanaka N, *et al*. Significance of karyopherin-α 2 (KPNA2) expression in esophageal squamous cell carcinoma. Anticancer Res. 2010;30(3):851-856. PMID: 20393006.
63. Altan B, Yokobori T, Mochiki E, Ohno T, Ogata K, Ogawa A, *et al*. Nuclear karyopherin-α2 expression in primary lesions and metastatic lymph nodes was associated with poor prognosis and progression in gastric cancer. Carcinogenesis. 2013;34(10):2314-2321. DOI: 10.1093/carcin/bgt214.
64. Takada T, Tsutsumi S, Takahashi R, Ohsone K, Tatsuki H, Suto T, *et al*. KPNA2 over-expression is a potential marker of prognosis and therapeutic sensitivity in colorectal cancer patients. J Surg Oncol. 2016;113(2):213-217. DOI: 10.1002/jso.24114.
65. Hannou SA, Wouters K, Paumelle R, Staels B. Functional genomics of the CDKN2A/B locus in cardiovascular and metabolic disease: what have we learned from GWASs? Trends Endocrinol Metab. 2015;26(4):176-184. DOI: 10.1016/j.tem.2015.01.008,
66. Chen Z, Guo Y, Zhao D, Zou Q, Yu F, Zhang L, Xu L. Comprehensive analysis revealed that CDKN2A is a biomarker for immune infiltrates in multiple cancers. Front Cell Dev Biol. 2021; 9:808208,1-13. DOI: 10.3389/fcell.2021.808208.
67. Royds JA, Pilbrow AP, Ahn A, Morrin HR, Frampton C, Russell IA, *et al*. The rs11515 polymorphism is more frequent and associated with aggressive breast tumors with increased ANRIL and decreased p16 INK4a expression. Front Oncol. 2016;5:306,1-9. DOI: 10.3389/fonc.2015.00306.
68. He TC, Sparks AB, Rago C, Hermeking H, Zawel L, Da Costa LT, *et al*. Identification of c-MYC as a target of the APC pathway. Science. 1998;281(5382):1509-1512. DOI: 10.1126/science.281.5382.1509.
69. Lievre A, Bachet JB, Boige V, Cayre A, Le Corre D, Buc E, *et al*. KRAS mutations as an independent prognostic factor in patients with advanced colorectal cancer treated with cetuximab. J Clin Oncol. 2008;26(3):374-379. DOI: 10.1200/JCO.2007.12.5906.
70. Takeuchi H, Bilchik A, Saha S, Turner R, Wiese D, Tanaka M, *et al*. c-MET expression level in primary colon cancer: a predictor of tumor invasion and lymph node metastases. Clin Cancer Res. 2003;9(4):1480-1488. PMID: 12684423.
71. Grzenda A, Lomberk G, Zhang JS, Urrutia R. Sin3: master scaffold and transcriptional corepressor. Biochim Biophys Acta. 2009;1789(6-8):443-450. DOI: 10.1016/j.bbagrm.2009.05.007.
72. Hassig CA, Fleischer TC, Billin AN, Schreiber SL, Ayer DE. Histone deacetylase activity is required for full transcriptional repression by mSin3A. Cell. 1997;89(3):341-347. DOI: 10.1016/S0092-8674(00)80214-7.
73. Jafari H, Hussain S, Campbell MJ. Nuclear receptor coregulators in hormone-dependent cancers. Cancers. 2022;14(10):2402,1-30. DOI: 10.3390/cancers14102402.
74. Dannenberg J-H, David G, Zhong S, Van Der Torre J, Wong WH, DePinho RA. mSin3A corepressor regulates diverse transcriptional networks governing normal and neoplastic growth and survival. Genes Dev. 2005;19(13):1581-1595. DOI: 10.1101/gad.1286905.
75. Ren J, Li X, Dong H, Suo L, Zhang J, Zhang L, *et al*. miR-210-3p regulates the proliferation and apoptosis of non-small cell lung cancer cells by targeting SIN3A. Exp Ther Med. 2019;18(4):2565-2573. DOI: 10.3892/etm.2019.7867.
76. Nan S, Zhang S, Jin R, Wang J. LINC00665 up-regulates SIN3A expression to modulate the progression of colorectal cancer via sponging miR-138-5p. Cancer Cell Int. 2022;22(1):1-12. DOI: 10.1186/s12935-021-02176-4.
77. Nebenfuehr S, Kollmann K, Sexl V. The role of CDK6 in cancer. Int J Cancer Res. 2020;147(11):2988-2995. DOI: 10.1002/ijc.33054.
78. Wang H, Nicolay BN, Chick JM, Gao X, Geng Y, Ren H, *et al*. The metabolic function of cyclin D3-CDK6 kinase in cancer cell survival. Nature. 2017;546(7658):426-430. DOI: 10.1038/nature22797.
79. Liu Y, Tang W, Ren L, Liu T, Yang M, Wei Y, *et al*. Activation of miR-500a-3p/CDK6 axis suppresses aerobic glycolysis and colorectal cancer progression. J Transl Med. 2022;20(1):106,1-12. DOI: 10.1186/s12967-022-03308-8.

80. Drewes JL, White JR, Dejea CM, Fathi P, Iyadorai T, Vadivelu J, *et al*. High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. NPJ Biofilms Microbioms. 2017;3:34,1-12.
DOI: 10.1038/s41522-017-0040-3.
81. Osman MA, Neoh HM, Ab Mutalib NS, Chin SF, Jamal R. 16S rRNA gene sequencing for deciphering the colorectal cancer gut microbiome: current protocols and workflows. Front Microbiol. 2018;9:767,1-10.
DOI: 10.3389/fmicb.2018.00767.
82. Salehi Z, Haddad P, Tavallaei O. Prediction of biomarker miRNAs signature in colorectal cancer metastasis to liver cancer.Electron J Gen Med. 2018;16(1):1-11.
DOI: 10.29333/ejgm/93467.
83. Malumbres M, Barbacid M. Cell cycle, CDKs and cancer: a changing paradigm. Nat Rev Cancer. 2009;9(3):153-166.
DOI: 10.1038/nrc2602.
84. Bai J, Li Y, Zhang G. Cell cycle regulation and anticancer drug discovery. Cancer Biol Med. 2017;14(4):348-362.
DOI: 10.20892/j.issn.2095-3941.2017.0033.
85. Wang CI, Wang CL, Wang CW, Chen CD, Wu CC, Liang Y, *et al.* Importin subunit alpha-2 is identified as a potential biomarker for non-small cell lung cancer by integration of the cancer cell secretome and tissue transcriptome. Int J Cancer. 2011;128(10):2364-2372.
DOI: 10.1002/ijc.25568.