J Appl Biotechnol Rep. 2022 June;9(2):657-670

Journal of
Applied Biotechnology Reports

doi 10.30491/JABR.2022.316840.1473

Original Article

A Gene Expression Signature to Predict Chemotherapy Response of Colorectal Cancer Patients: Systems Biology Analysis on Transcriptomics Data and Experimental Validation

Hamed Manoochehri [1], **Hamid Tanzadehpanah** [1], **Amir Taherkhani** [1], **Akram Jalali** [1*]

[1] Research Center for Molecular Medicine, Hamadan University of Medical Sciences, Hamadan, Iran

Corresponding Author: Akram Jalali, PhD, Assistant Professor, Department of Molecular Medicine and Genetics, School of Medicine, Hamadan University of Medical Sciences, Hamadan, Iran. Tel: +98-9121324616, E-mail: aki_jalali@yahoo.com

Received November 25, 2021; **Accepted** February 1, 2022; **Online Published** June 18, 2022

Abstract

Introduction: Gene expression profiling has high potential in the identification of diagnostic, predictive, and therapeutic gene targets in human cancers such as colorectal cancer (CRC). Accordingly, in this study, an integrated systems biology analysis was done on several microarray datasets to identify key genes involved in CRC chemoresistance and also to differentiate peoples who benefit from chemotherapy. Subsequently, the findings were validated experimentally.

Materials and Methods: Datasets were retrieved from Gene Expression Omnibus (GEO). Gene expression analysis was performed using the ExAtlas software. Gene enrichment analysis was done using g: profiler. Protein-Protein Interaction Network (PPIN) was constructed in STRING and visualized/analyzed by Cytoscape 3.8.0. Significant modules were identified using the MCODE plugin in Cytoscape. The clinical importance of candidate genes was evaluated using ROC analysis and immunohistochemistry. Key candidate genes were validated using Real-Time PCR.

Results: According to findings, 26 datasets were selected. Gene expression analysis revealed 6463 Differentially Expressed Genes (DEGs), among which 4323 were unique and 2140 were related to overlapping DEGs between datasets. The overlapping DEGs with at least four shared datasets (n = 217 DEGs) were selected for further analysis. Overlapping DEGs were mainly enriched in the cellular process of response to chemicals stimulus. Most selected DEGs were enriched in KEGG pathways of cancer Benzo(a)pyrene metabolism and glucocorticoid receptor signaling. Fourteen hub genes and two significant modules were identified. Six hub genes (candidate genes) were contributed in significant modules. Among candidate genes, *LCN2*, *CXCL8*, and *EGR1* expression was significantly associated with chemotherapy response of CRC patients and chemosensitivity of CRC cell lines ($p<0.05$).

Conclusions: This study revealed three genes signature for predicting chemotherapy responsiveness and treatment decision-making in CRC patients and also for therapeutic purposes.

Keywords: Colorectal Neoplasm, Systems Biology, Protein-Protein Interaction Network, ROC Analysis, Gene Ontology, Antineoplastic Drug Resistance

Citation: Manoochehri H, Tanzadehpanah H, Taherkhani A, Jalali A. A Gene Expression Signature to Predict Chemotherapy Response of Colorectal Cancer Patients: Systems Biology Analysis on Transcriptomics Data and Experimental Validation. J Appl Biotechnol Rep. 2022;9(2):657-70. doi:10.30491/JABR.2022.316840.1473

Introduction

Colorectal Cancer (CRC), as a major global health problem, has been ranked as the third most prevalent malignancy and remains the second leading cause of cancer related deaths worldwide.[1] In 2020, about 1.9 million new cases and 935,000 deaths related to CRC was reported[2] and is expected to reach 2.2 million and 1.1 million by 2030, respectively.[3] The prevalence of CRC is 10 times more common in developed countries compared with developing countries, while mortality rates are higher in developing countries.[2]

Early stage CRC is usually asymptomatic or has nonspecific symptoms, so is typically detected at advanced stages.[3] In these stages, tumor has progressed and is surgically incurable and (neo) adjuvant chemotherapy is necessary.[4] Standard chemotherapeutic regimen for advanced CRC is based on the combination of 5-Fluorouracil (5-FU), oxaliplatin and irinotecan.[5] Also, a promising chemotherapy regimen including methotrexate, 5-FU and oxaliplatin is reported. Chemotherapy is used in order to slow tumor growth, reduce tumor size and decrease the metastasis likelihood.[5] Although most advanced CRC tumors are responsive to the combined chemotherapy regimen at first, but may be resistant later and lead to eventual tumor recurrence.[5] Generally, human cancers may be inherently resistant to therapy or later acquire drug resistance.[6] Both of chemoresistance types involve several mechanisms, such as increasing drug efflux, changing drug metabolite toxicity or potency and inhibiting cell death[5]. Understanding of basic mechanisms involving chemotherapy drug resistance is important to develop more effective therapeutic strategies[5,7] and identify a subgroup of patients who could benefit from chemotherapy.[4,8]

Gene expression profiling has high potential in the identification of diagnostic, predictive and therapeutic gene targets in human cancers, including CRC.[4,9] Microarray, as a high-throughput technology, is able to simultaneously screen the expression of thousands of genes[10] and has been widely used in biomarker discovery.[5]

However, the translation of microarrays analysis into clinical applications still remains as a challenge mainly due to variability in platforms, sampling and data analysis of different studies, that leads to lack of overlap of obtained results.[11] Also, tumor chemotherapy response is a dynamic procedure that is regulated by a complex molecular network. Therefore, interaction between genes is also important and should be considered in microarray data analysis.[12,13]

Systems biology-based network approaches evaluate the association between DEGs for bridging the gap from individual genes to disease/specific condition occurrence.[12,14] The PPIN analysis are used for identifying hub genes that are the highly connected nodes in the network and have a crucial role in essential biological processes.[15] In this study, different microarrays expression datasets were used to increase the statistical power of study. Datasets were analyzed using network-based approaches in order to identify the key candidate genes involved in chemoresistance of CRC to routine chemotherapy drugs. Also, the findings were validated experimentally. The flow chart of performed steps in the present study is shown in Figure 1.

Materials and Methods

GEO Dataset Search

A complete search was performed in GEO in order to find relevant microarray datasets using following keywords: Colorectal Cancer, Chemotherapy, Drug Resistance, Chemoresistance, Treatment Resistance, 5-FU, Tegafur, Irinotecan (Camptosar), SN-38, Methotrexate, Oxaliplatin (Eloxatin), FOLFOX and FOLFIRI. Quality of datasets was controlled using ExAtlas (https://lgsun.grc.nia.nih.gov/exatlas) online software tool. Low quality datasets were excluded from study. In addition, datasets with other treatment modalities were excluded. The information of used datasets is provided in Table 2.

Gene Expression Analysis and Identification of Overlapping DEGs

Quantile method was used in order to remove any variation from data and making them normalized and cross-comparable. ExAtlas was used to find DEGs between chemoresistant (non-responders) and chemosensitive (responders) samples or after receiving chemotherapy. Selection criteria for DEGs was fold change (FC) ≥3 and false discovery rate (FDR) ≤0.001. Datasets were excluded from study in cases with no DEGs in defined criteria. Uncharacterized, non-annotated, non-protein coding and pseudo genes were also excluded from DEGs. Overlapping DEGs between at least four individual

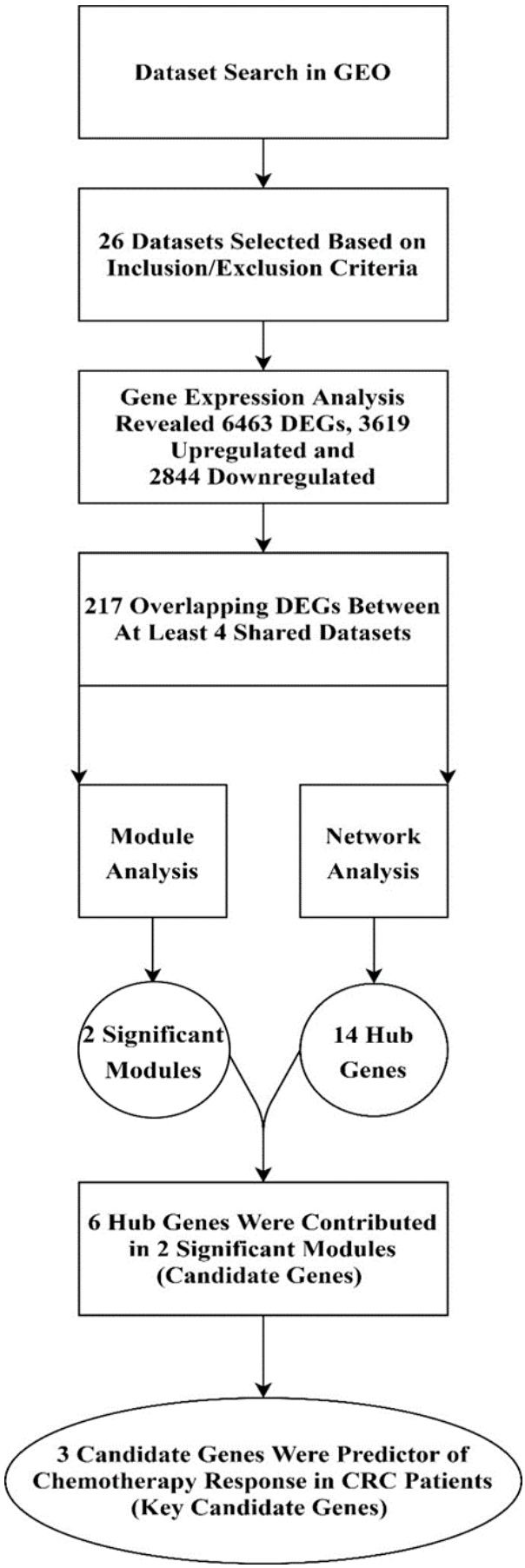


Figure 1. Flow Chart of Performed Steps.

datasets were selected using interactive Venn diagrams (bioinformatics.psb.ugent.be).

Gene Enrichment Analysis

Significantly enriched terms in the Gene Ontology (GO) of Biological Process (BP) and Kyoto Encyclopedia of Genes, and Genomes (KEGG) pathways were identified using g:Profiler (biit.cs.ut.ee/gprofiler/gost), a user-friendly, simple and powerful web tool for capturing GO and pathway terms. Enriched terms with p_{adj}<0.05 were considered as significant.

Protein-Protein Interaction Network (PPIN) Analysis

PPIN (physical and functional) was acquired using Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, http://www.string-db.org), a web source of known and predicted PPIs. A cutoff of 0.4 (medium confidence) was considered as minimum required interaction score and disconnected nodes were removed from the network. PPIN was visualized and analyzed via Cytoscape (version 3.8.0). Nodes with high neighborhood connectivity, closeness centrality, betweenness centrality, radiality and degree were selected using interactive Venn diagram and were considered as hub genes.

Module Selection

MCODE plugin in Cytoscape was applied to find densely interconnected regions (module screening) of the PPIN using the following criteria: degree cutoff = 2, node score cutoff = 0.2, k-core = 2, maximum depth = 100, minimum number of cluster nodes = 8.

ROC Analysis

The ROC plotter web tool (http://www.rocplot.org/crc/index) was used to link gene expression data with chemotherapy (no radiotherapy) response of 342 patients (172 responders and 170 non-responders) to commonly used drugs in CRC including 5-FU, oxaliplatin, irinotecan and capecitabine.[16] Chemotherapy response of patients was assessed by Response Evaluation Criteria in Solid Tumors (RECIST). Gene expression was measured using the Affymetrix expression microarrays method. Receiver Operating Characteristic (ROC) curve was plotted for measuring the ability of genes to discriminate between two populations of responders and non-responders. Box plot was used to compare gene expression level between responders and non-responders. Also, p<0.05 was considered significant in both of ROC curve and Box plot.

Immunohistochemistry (IHC)

Histological sections from normal and cancer tissues of colon/rectum were obtained from human protein atlas (www.proteinatlas.org). Each sample was characterized by 1 mm tissue cores. Antibody labeling was done with DAB (3,3'-diaminobenzidine) and brown color indicated positive expression of the corresponding protein (gene). The sections were counterstained with hematoxylin for observing microscopical features. Tissue microarrays were used to show antibody staining. In the present study, two stained sections of normal and cancer tissues of colon/rectum for a corresponding protein has been shown.

Cell Culture

Three colorectal cell lines of SW480, HT29 and HCT116 were cultured in 25-cm^2 T-flasks incubated at 37 ℃, 5% CO_2 with high relative humidity. SW480 and HCT116 cell lines were maintained in high glucose Dulbecco's Modified Eagle Medium (DMEM, Gipco) supplemented with 10% Fetal Bovine Serum (FBS, Gipco) and 1% penicillin-streptomycin (Gipco). The HT29 cell line was grown in Roswell Park Memorial Institute 1640 (RPMI, Gipco) supplemented with 10% FBS and 1% penicillin-streptomycin. When Cells reached 80-90% confluence, were detached using 1 ml of 0.25% trypsin/EDTA solution (Gipco) and incubation at 37 ℃ for 1 min.

MTT assay

Three colorectal cell lines (SW480, HT29, and HCT116) were tested for sensitivity to 5-FU and oxaliplatin chemotherapy drugs using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) assay. Briefly, 5×10^3 cells were seeded in a 96-well plate. After 24 h incubation in 37 °C, cells were treated with different doses of 5-FU and oxaliplatin. For period of 48 h after treating, 10 µl MTT solution (5 mg/ml) was added to each well. Plate was incubated at 37 °C for 3 h and solution of wells was replaced with 100 µl of dimethylsulfoxide (DMSO) to dissolve the formazan crystals. Finally, plate was shacked for 15 min in room temperature and its absorbance was read at 570 nm wavelength using ELISA reader. Cell viability for any dose of drugs was obtained in comparison with non-treated cells. The half maximal inhibitory concentration (IC_{50}) was calculated by plotting the cell viability against drug concentrations. IC_{50} was compared between cell lines using Mann Whitney U test. Also, p<0.05 was considered as a significant level.

RNA Extraction and cDNA Synthesis

Total RNA was isolated from cells using RNX™-PLUS solution (CinnaGen, Iran) according to manufacture instructions. The concentration and purity of extracted RNA was determined using a NanoDrop spectrophotometer by absorbance reading in 260/280 nm wavelength. Quality of RNA was validated by electrophoresis on %1 agarose gel. DNase I (Sigma, USA) was used for eliminating any potential genomic DNA contamination in RNA samples.[6] Synthesis of cDNA was performed in a final volume of 20 µl using 500 ng of total RNA by first-strand cDNA synthesis kit (Takara, Japan)

according to manufacturer's recommendations.[17]

Real-Time PCR

Real-Time PCR was done in a Roche light cycler 96 Real-Time PCR instrument using SinaSYBR Blue HS-qPCR Mix (Sinaclone, Iran) according to manufacturer instructions. Briefly, 10 µl of SinaSYBR Blue, 1 µl of each primer pair (10 µM), 2 µl of synthesized cDNA and 4 µl nuclease free water were added to a tube for each reaction. The sequence of primer pairs used in Real-Time PCR reaction is listed in Table 1. The 18s rRNA was considered as a reference gene. Real-Time PCR reaction was performed in 45 cycles after 5 min preincubation of samples at 95 °C (denaturation: 95 °C for 15 s, annealing: 50-60 °C for 30 s, extension: 72 °C for 30 s). After final cycle, melting analysis was performed in order to confirm product specificity. Product specificity was also confirmed by electrophoresis of products on 1% agarose gel. Real-Time PCR efficiency was calculated using $E = +10^{(-1/slope)}-1$ formula. Expression fold change was calculated using $F = 2^{-\Delta\Delta CT}$ formula.[17,18] Gene expression fold change was compared between cell lines using Mann Whitney U test. In addition, $p<0.05$ was considered as a significant level.

Table 1. The Sequence of Primer Pairs Used in Real-Time PCR

Gene Symbol	Primer Sequence (5´-3´)
LCN2	Forward: TGCTATGGTGTTCTTCAAG
	Reverse: GATACACTGGTCGATTGG
CXCL8	Forward: CGGAAGGAACCATCTCACTGTG
	Reverse: GGTCCACTCTCAATCACTCTCAG
EGR1	Forward: CAGCAGCAGCAGCAGCAGCACCTTC
	Reverse: GTCTCCACCAGCACCTTCTCGTTGTTCAG
18S rRNA	Forward: GTAACCCGTTGAACCCCATT
	Reverse: CCATCCAATCGGTAGTAGCG

Results

Datasets

Twenty-six microarray datasets from different experiments were selected according to study inclusion/exclusion criteria. Information of used datasets including accession, platform, drug name, sample type, number of up-regulated and downregulated DEGs is provided in Table 2.

Differentially Expressed Genes

Gene expression analysis revealed 6463 DEGs including 3619 upregulated (56.0%) and 2844 (44.0%) downregulated genes. Among the identified DEGs,4323 were unique and 2140 were related to overlapping DEGs between datasets. The frequency of overlapping DEGs is shown in Table 3. The maximum and minimum number of shared datasets for each overlapping DEG was nine (2 overlapping DEGs) and two (801 overlapping DEGs), respectively. The overlapping DEGs with at least four shared datasets (n = 217 DEGs) were selected for further analysis.

Gene Ontology and Pathway Analysis

The 10 top significant enriched GO terms of biological process and also significant enriched KEGG pathways terms are shown in Figure 2. Most selected DEGs were enriched in biological process of response to chemicals/organic substances/ oxygen-containing compounds, cell mobility/motility/localization, tissue development and cell differentiation. Most selected DEGs were enriched in cancer, Benzo(a)pyrene metabolism and glucocorticoid receptor KEGG pathways.

Hub Genes and Significant Modules in PPIN

A PPIN was constructed using STRING (PPI enrichment $p_{Value}<10^{-16}$). The network analyzing was performed using Cytoscape (Figure 3) and 30 top DEGs (nodes) with highest score for each of betweenness centrality, closeness centrality, degree, neighborhood connectivity and radiality parameters were selected. Four hub genes were identified with high scores in terms of betweenness centrality, closeness centrality, degree, neighborhood connectivity and radiality using interactive vendiagram (Figure 4, two asterisk in Table 4). Also, 10 hub genes were identified with high scores in terms of betweenness centrality, closeness centrality, degree and radiality using interactive vendiagram (Figure 4, one asterisk in Table 4). Two significant modules were idenified using MCODE plugin in cytoscape (Figure 5). Six hub genes contributed in significant modules were considered as candidate genes (Table 4, Figue 5).

ROC Analysis

Clinical importance of candidate genes (*PTGS2*, *CTGF*, *LCN2*, *CXCR4*, *CXCL8* and *EGR1*) were evaluated using ROC analaysis (Figures 6 and 7). Boxplotes and ROC curves revealed that only *LCN2*, *CXCL8* and *EGR1* genes are significantly capable to predict chemotherapy response of CRC patients ($p<0.05$). Non-responders had higher expression of *LCN2* and *CXCL8* genes and lower expression of *EGR1* gene. So, *LCN2*, *CXCL8* and *EGR1* genes were considered as key candidate genes.

Immunohistochemistry Findings

IHC results showed that LCN2 (Figure 8) and CXCL8 (Figure 9) proteins expression level were considerably higher in colorectal cancer tissues than colorectal normal tissues. EGR1 protein expression level was considerably lower in CRC tissues than colorectal normal tissues (Figure 10).

Chemosensitivity of Cell Lines

The IC_{50} of oxaliplatin 5-FU has been compared between HCT116, SW480, and HT29 colorectal cell lines (Figure 11). The IC_{50} value of oxaliplatin and 5-FU for HT29 cell line was significantly higher than other two cell lines ($p<0.001$). Also, IC_{50} value of oxaliplatin and 5-FU for SW480 cell line was significantly higher than HCT116 cell line

Table 2. Information of Different Microarray Datasets

No.	Accession	Platform	Drug	Sample/Cell line	DEGs
1	GSE81005	GPL15207/Affymetrix	5-FU	HCT8	↑100 ↓49
2	GSE56322	GPL14550/Agilent	5-FU	HCT116	↑199 ↓199
3	GSE97781	GPL570/Affymetrix	5-FU	Patient-derived Colonosphere	↑59 ↓14
4	GSE53617	GPL13369/Illumina	5-FU	SW620, Lovo-1	↑207 ↓237
5	GSE77180	GPL5175/Affymetrix	5-FU	HCT116	↑136 ↓14
6	GSE76489	GPL16686/Affymetrix	5-FU	LS411N, SW620	↑241 ↓199
7	GSE52735	GPL570/Affymetrix	Fluoropyrimidine	40 tissue sample of mCRC	↑9 ↓1
8	GSE59501	GPL6244/Affymetrix	Irinotecan	LoVo	↑99 ↓181
9	GSE23433	GPL4719/Spotted microarray	Irinotecan	HT-29, RIV	↑199 ↓52
10	GSE83129	GPL6244/Affymetrix	Oxaliplatin	36 CRC/ normal adjacent tissues samples	↑1 ↓3
11	GSE30011	GPL2006/Spotted oligonucleotide	Oxaliplatin	LoVo, LS513, SW1116, SW1417, SW48, LS174TT, LS411NN, HCT116, HCT15, SW480, SW948, CACO2, DLD1, HT29	↑314 ↓311
12	GSE76092	GPL21253/Agilent	Oxaliplatin	HT29	↑262 ↓219
13	GSE110425	GPL10558/Illumina	Oxaliplatin	HCT-116	↑1 ↓5
14	GSE10405	GPL2006/Spotted oligonucleotide	Oxaliplatin	HT29, LoVo, DLD1, LS513	↑69 ↓28
15	GSE19992	GPL4133/Agilent	Oxaliplatin	LoVo	↑19 ↓28
16	GSE16648	GPL570/Affymetrix	Methotrexate	HT29, CaCo2	↑167 ↓162
17	GSE11440	GPL570/Affymetrix	Methotrexate	HT29	↑215 ↓234
18	GSE9412	GPL571/Affymetrix	Methotrexate	HT29	↑211 ↓177
19	GSE16066	GPL570/Affymetrix	Methotrexate	CaCo2	↑166 ↓86
20	GSE42387	GPL16297/Agilent	Oxaliplatin/SN-38	HCT116, HT29, LoVo	↑777 ↓556
21	GSE72544	GPL10558/Illumina	Oxaliplatin/ 5-FU	A panel of 15 CRC cell lines	↑31 ↓19
22	GSE62322	GPL97/Affymetrix	FOLFIRI	57 normal colon/ primary tumor/ and liver metastasis	↑27 ↓17
23	GSE28702	GPL570/Affymetrix	FOLFOX	83 colorectal cancer tissue samples	↑14 ↓13
24	GSE3964	GPL3282/Spotted DNA/cDNA	5-FU and irinotecan	60 normal colon/ primary tumor/ and liver metastasis	↑6 ↓10
25	GSE81653	GPL16686/Affymetrix	5-FU and FOLFOX	593 colorectal cancer tissue samples	↑4 ↓0
26	GSE72969	GPL570/Affymetrix	FOLFOX, FOLFIRI	56 colorectal cancer tissue samples	↑86 ↓30

↑Upregulated, ↓Downregulated, DEGs: Differentially Expressed Genes.

Table 3. The Frequency of Overlapping DEGs between Datasets

Number of shared datasets	Frequency of overlapping DEGs	Total number of DEGs
9	2	18
8	2	16
7	9	63
6	20	120
5	53	265
4	131	524
3	275	825
2	801	1602
1	3030	3030

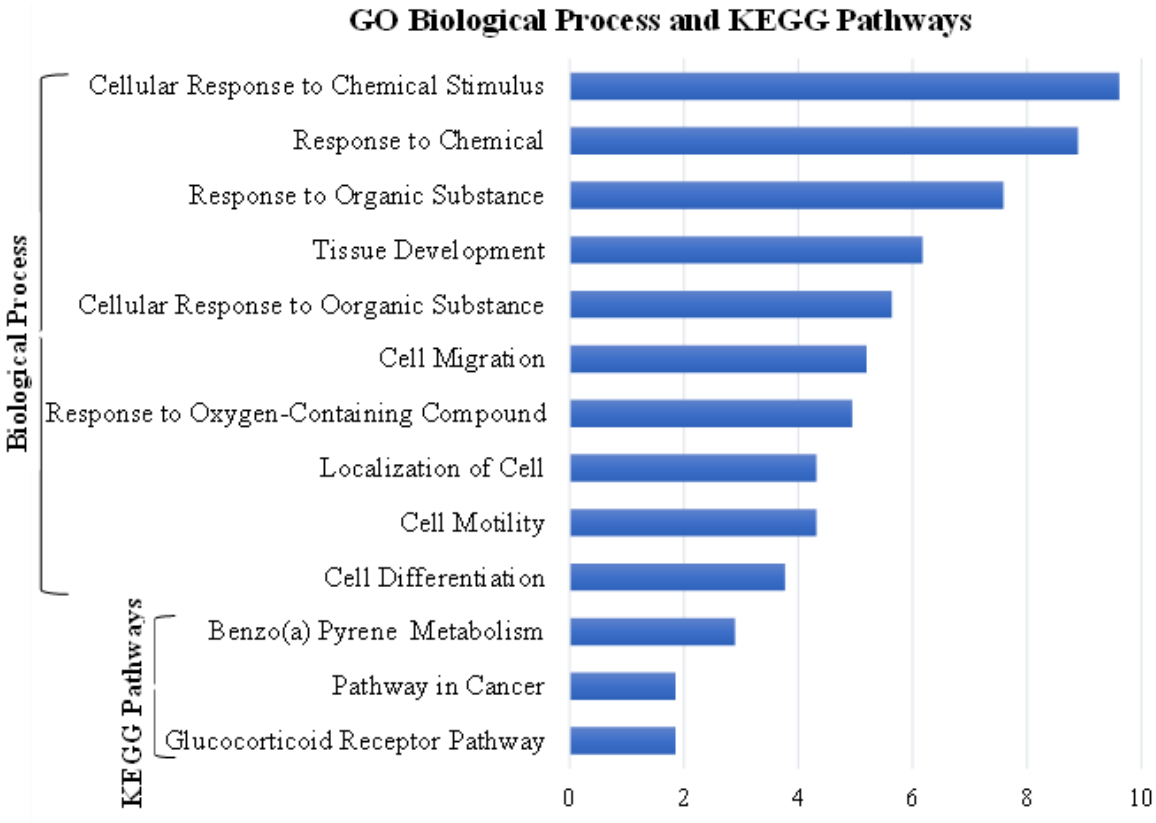


Figure 2. Significant Enriched GO Biological Process and KEGG Pathway Terms.

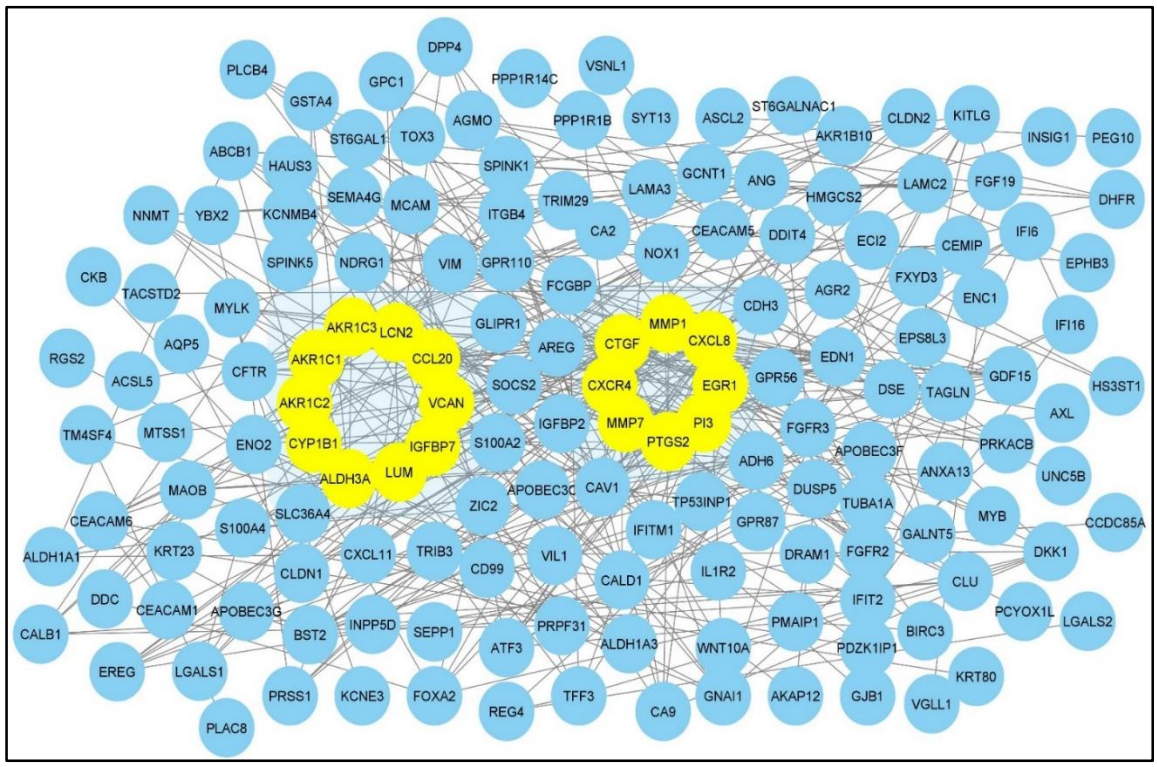


Figure 3. PPIN of Selected DEGs. Edges number: 378, average node degree: 3.58, average local clustring coefficient: 0.323, PPI enrichment $p_{Value}<10^{-16}$. Yellow nodes: Significant modules. Module 1 (Right): *MMP1, CXCL8, EGR1, PI3, PTGS2, MMP7, CXCR4, CTGF*. Module 2 (Left): *AKR1C3, LCN2, CCL20, VCAN, IGFBP7, LUM, ALDH3A, CYP1B1, AKR1C2, AKR1C1*.

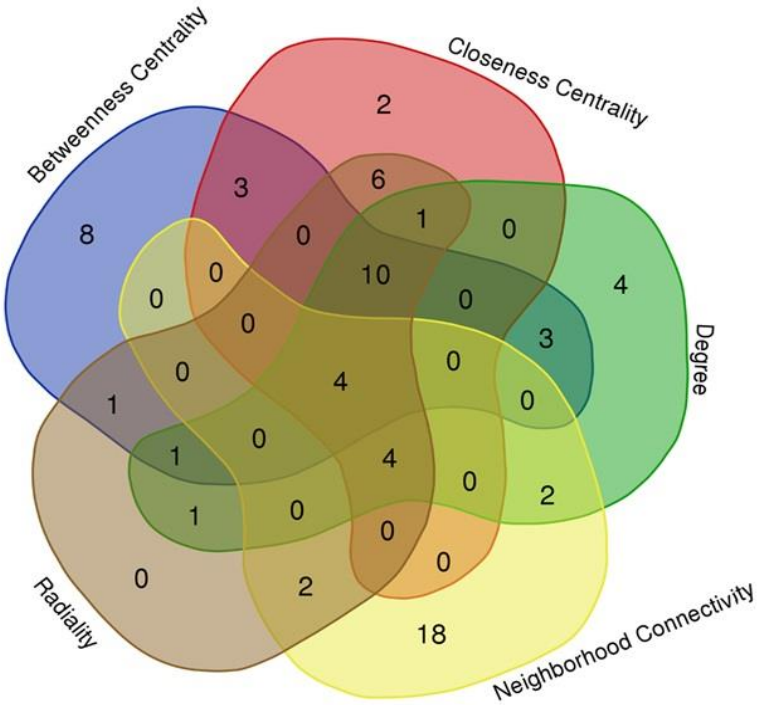


Figure 4. Venn Diagram of High Score Nodes for Each of Betweenness Centrality (n = 30), Closeness Centrality (n = 30), Degree (n = 30), Neighborhood Connectivity (n = 30) and Radiality (n = 30). Four hub genes were identified with high scores in terms of betweenness centrality, closeness centrality, degree, neighborhood connectivity and radiality. Ten hub genes were identified with high scores in terms of betweenness centrality, closeness centrality, degree and radiality.

Table 4. Identified Hub Genes in PPIN Analysis

Symbols of Hub Genes	Betweenness Centrality	Closeness Centrality	Degree	Neighborhood Connectivity	Radiality	Module Number
**EDN1	0.0718	0.3878	17	12.1760	0.9436	-
**GDF15	0.0505	0.3594	10	11.5000	0.9363	-
****PTGS2**	0.0631	0.3909	19	12.0526	0.9443	1
****CTGF**	0.0932	0.3972	20	11.7000	0.9458	1
*FOXA2	0.0716	0.3458	7	9.0000	0.9324	-
***LCN2**	0.0346	0.3542	11	9.6363	0.9348	2
***CXCR4**	0.1165	0.3983	23	9.7391	0.9460	1
*ENO2	0.0688	0.3516	9	7.2222	0.9341	-
*ATF3	0.0446	0.3434	9	10.4444	0.9317	-
*CFTR	0.1016	0.3602	13	7.2307	0.9365	-
***CXCL8**	0.2245	0.4375	28	10.1785	0.9540	1
***EGR1**	0.0920	0.3920	18	10.3888	0.9446	1
*AREG	0.0384	0.3508	12	10.5833	0.9339	-
*CAV1	0.1263	0.3994	19	10.8421	0.9463	-

*Hub genes (10 nodes) with highest score in term of betweenness centrality, closeness centrality, degree and radiality**.** **Hub genes (4 nodes) with highest score in term of betweenness centrality, closeness centrality, degree and neighborhood connectivity and radiality. Six hub genes were contributed in significant modules (candidate genes, bolded).

line (p<0.05). So, HT29 is considered as a chemoresistant cell line, HCT116 is considered as a chemosensitive cell line and SW480 is considered as a cell line with intermediate chemosensitivity.

Real-Time PCR

The expression level of key candidate genes (*LCN2*, *CXCL8*, *EGR1*) was compared between colorectal cell lines with different chemosensitivity (Figure 12). *LCN2* and *CXCL8* expression were significantly higher in HT29 (chemoresistant) and SW480 (intermediate) than HCT116 (chemosensitive). The expression level of *EGR1* in HT29 and SW480 was significantly lower than HCT116.

Relationship between Chemosensitivity and Expression of Key Candidate Genes

The relationship between chemotherapeutic response (IC_{50}) of cell lines with the expression levels of key candidate genes was examined by Spearman's correlation test (Table 5). A significant and direct correlation was observed between IC_{50} with expression levels of *LCN2* and *CXCL8* genes (p<0.05). Also, a significant and negative correlation was observed between IC_{50} with expression levels of EGR1 gene (p<0.05). The *CXCL8* gene showed a stronger association with IC_{50}, followed by *LCN2* gene.

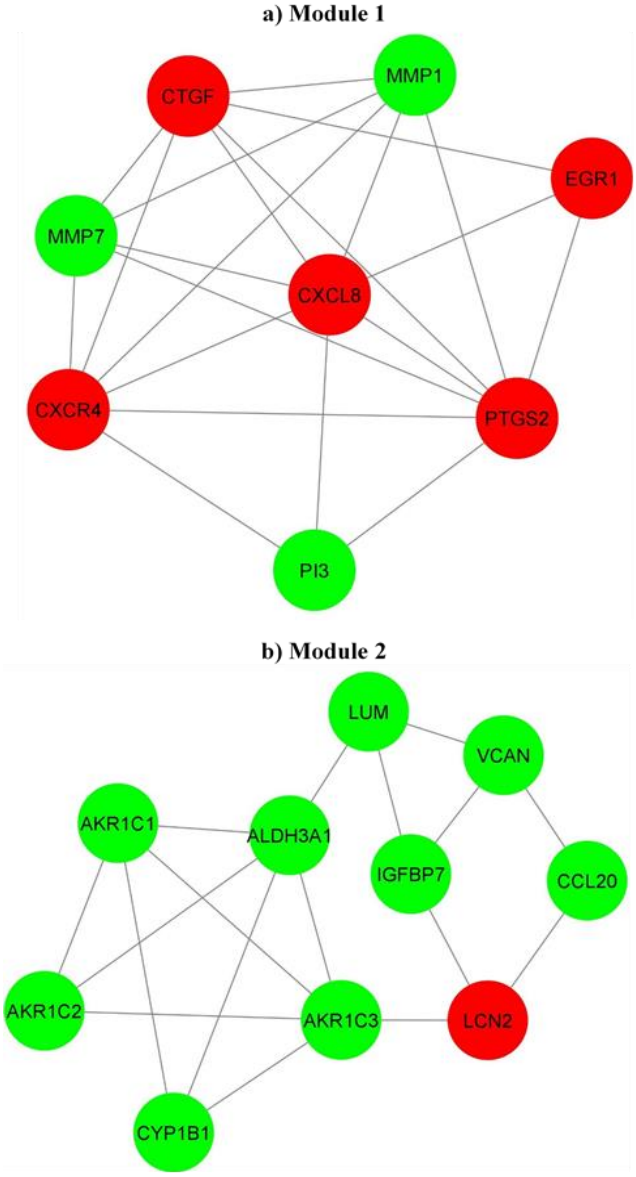


Figure 5. Significant Modules in the PPIN. Red nodes are hub genes contributed in modules (candidate genes).

Discussion

Resistance to conventional chemotherapeutic drugs is a big challenge in CRC treatment. To date, there is limited knowledge on chemoresistance mechanisms and how it can be predicted. Microarray-based gene expression analysis is a promising tool to predict responsiveness in any given therapy. An important limitation of microarray-based studies includes the low number of samples.[19] We used large number datasets in order to increase the specificity and power of study. Gene expression analysis revealed 217 overlapping DEGs shared between at least four datasets. Overlapping DEGs were mainly enriched in Benzo(a)pyrene metabolism and glucocorticoid receptor KEGG pathways. Benzo(a)pyrene metabolism pathway is involved in carcinogenesis and can also reverse the effects of chemotherapeutic drugs on cancer cells.[20] Tian et al., have shown that increased glucocorticoid receptor activity contributes in CRC development, proliferation, progression and invasion. They suggested that inhibition of glucocorticoid receptor may be helpful in combination with chemotherapy for metastatic colon cancer.[21] In this regards, GO analysis of cellular process indicated that overlapping DEGs is mainly involved in cellular process of response to chemicals. The main function of these responses is cell protection from death. There are evidences suggesting that hyperactivity of cellular response to chemicals (in cancer cells) results in acquisition of tolerance to higher doses of chemicals/drugs, which might be lethal to the cell in normal conditions.[22]

PPIN analysis revealed 14 hub genes which potentially have a key role in response to chemotherapy as they are highly connected to other genes in network. Six hub genes including *PTGS2*, *CTGF*, *LCN2*, *CXCR4*, *CXCL8* and *EGR1* were contributing in two significant modules (candidate genes). The clinical importance of these candidate genes was evaluated using ROC curve analysis. The ROC curve analysis results showed that only *CXCL8*, *LCN2* and *EGR1* genes are significantly capable to predicte the chemotherapy response of CRC patients (key candidate genes). The IHC results showed a higher expression of *LCN2* and *CXCL8* and lower expression of *EGR1* in colorectal cancer tissues compared to colorectal normal tissues. In addition, the relationship between chemosesetivity and expression level of *CXCL8*, *LCN2* and *EGR1* genes was confirmed with MTT asasy and Real-Time PCR on different colorectal cell lines.

In our study, the *CXCL8* gene showed a stronger association with IC_{50} than other two key candidate genes. Therefore, this gene can more accurately predict the chemotherapy response in patients when used as a biomarker. CXCL8, also known as IL-8 and Neutrophil-Activating Factor (NAF), is produced by epithelial cells, immune cells and endothelial cells. CXCL8 produces different biological effects by binding to three receptors of DARC, CXCR1 and CXCR2.[23] Increased CXCL8 expression and hyperactivation of its signaling induce survival, proliferation, angiogenesis, invasion, migration, Epithelial-Mesenchymal Transition (EMT) and metastasis of CRC cells.[24] Also, in Dabkeviciene et al.'s study, it has been shown that CXCL8 is upregulated in a chemoresistant subline of HCT116 CRC cells. [25] These reports are in agreement with the present study.

Vijay et al., showed that LCN2 (Lipocalin 2, also known as neutrophil gelatinase-associated lipocalin) up-regulation leads to chemo and radio resistance in colon cancer cell lines and that xenograft mouse models through T cell mediated immune pathway.[26] Also, LCN2 overexpression was associated with tumor progression, metastasis and poor disease-free survival of CRC patients in Maier et al.'s study.[27]

In contrast to these reports and our study findings, Feng et al., showed that LCN2 is an important negative regulator of

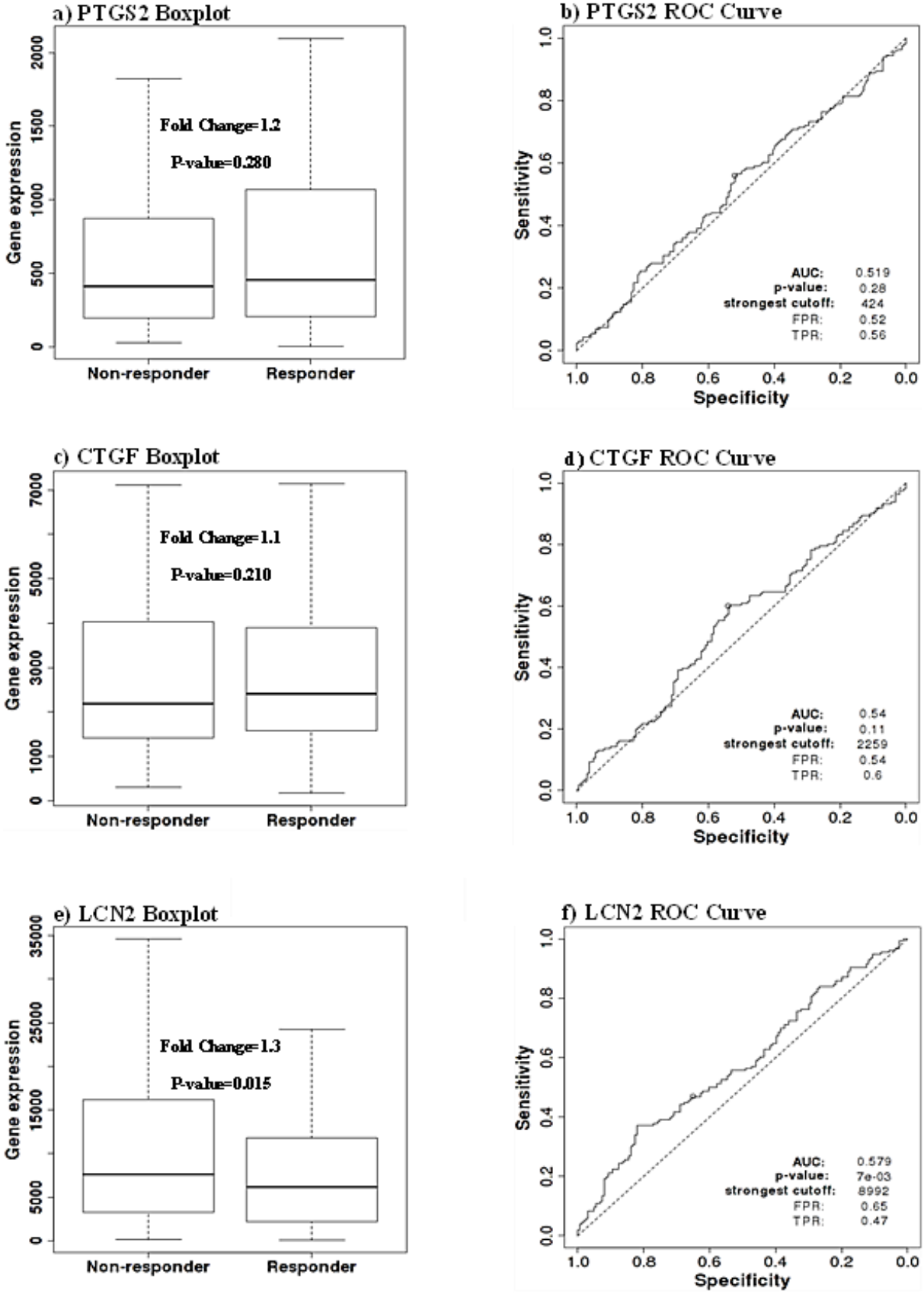


Figure 6. Boxplot and ROC Curve of PTGS2, CTGF and LCN2. a: PTGS2 Boxplot, b: PTGS2 ROC Curve, c: CTGF Boxplot, d: CTGF ROC Curve, e: LCN2 Boxplot, f: LCN2 ROC Curve.

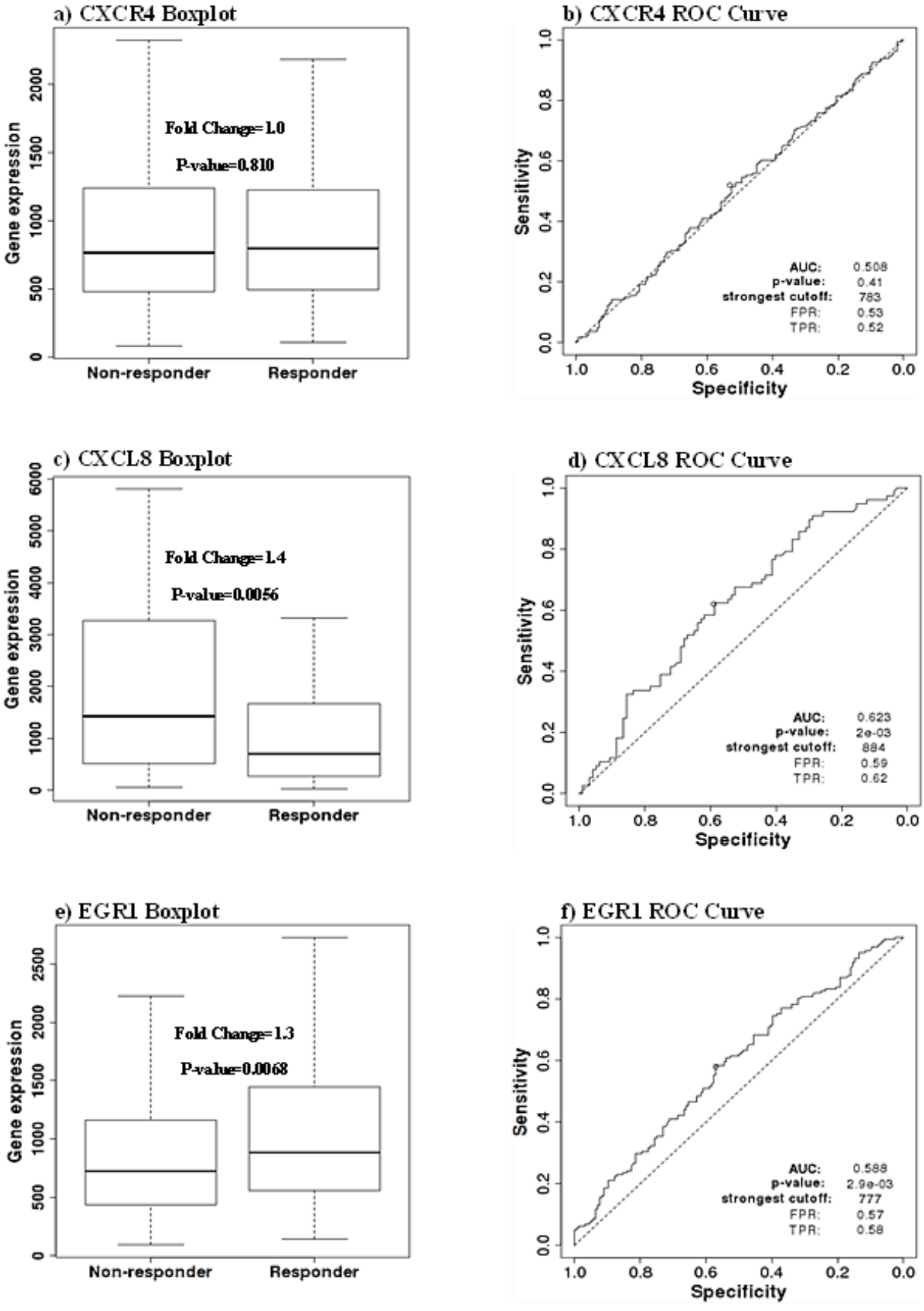
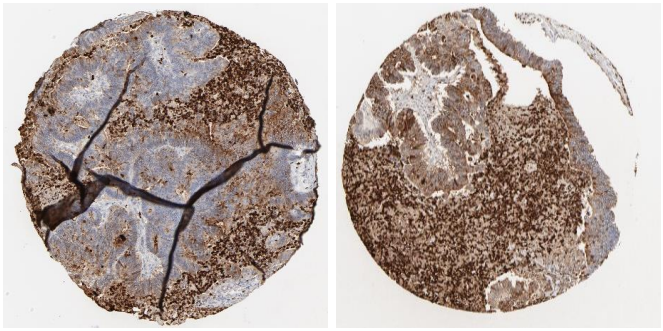


Figure 7. Boxplot and ROC Curve of CXCR4, CXCL8 and EGR1. a: CXCR4 Boxplot, b: CXCR4 ROC Curve, c: CXCL8 Boxplot, d: CXCL8 ROC Curve, e: EGR1 Boxplot, f: EGR1 ROC Curve

a) IHC of Colorectal Cancer Tissues

LCN2: High Expression



b) IHC of Colorectal Normal Tissues

LCN2: Low Expression

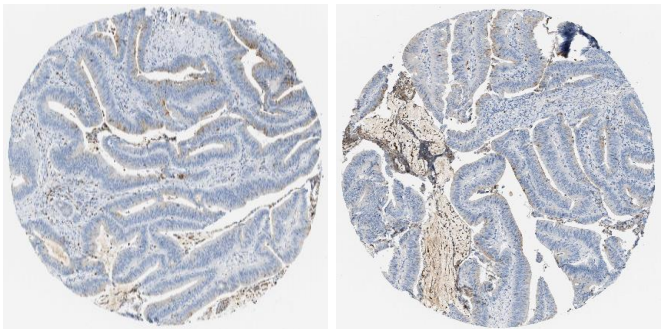
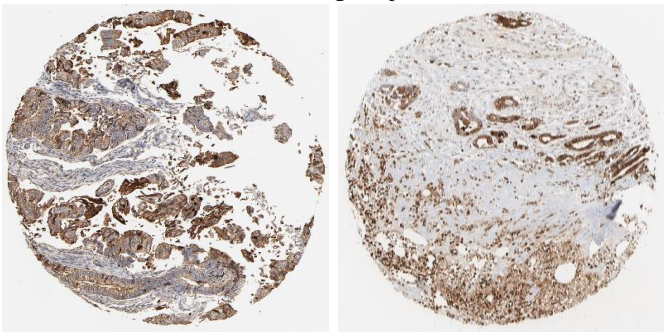


Figure 8. LCN2 Protein Expression in Normal and Cancer Tissue of Colorectal Detected by Immunohistochemistry. LCN2 significantly has higher expression in colorectal cancer tissues compared to normal colorectal tissues.

a) IHC of Colorectal Cancer Tissues

CXCL8: High Expression



b) IHC of Colorectal Normal Tissues

CXCL8: Low Expression

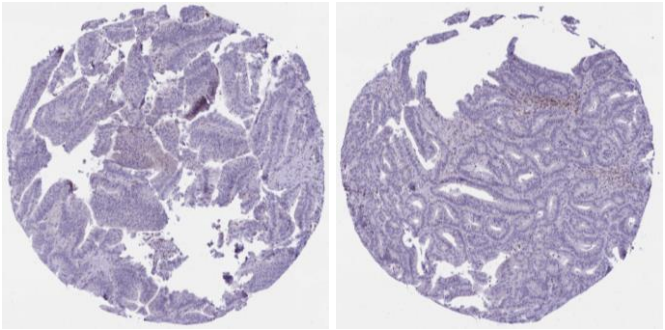
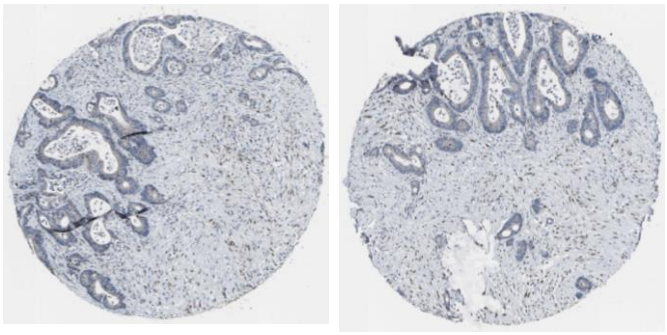


Figure 9. CXCL8 Protein Expression in Normal and Cancer Tissue of Colorectal Detected by Immunohistochemistry. CXCL8 significantly has higher expression in colorectal cancer tissues compared to normal colorectal tissues.

a) IHC of Colorectal Cancer Tissues

EGR1: Low Expression



b) IHC of Colorectal Normal Tissues

EGR1: High Expression

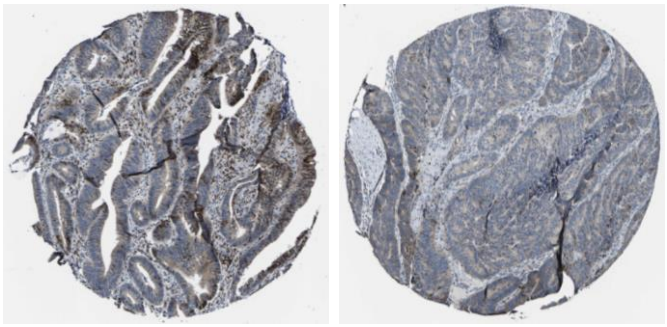


Figure 10. EGR1 Protein Expression in Normal and Cancer Tissue of Colorectal Detected by Immunohistochemistry. EGR1 significantly has lower expression in colorectal cancer tissues compared to normal colorectal tissues.

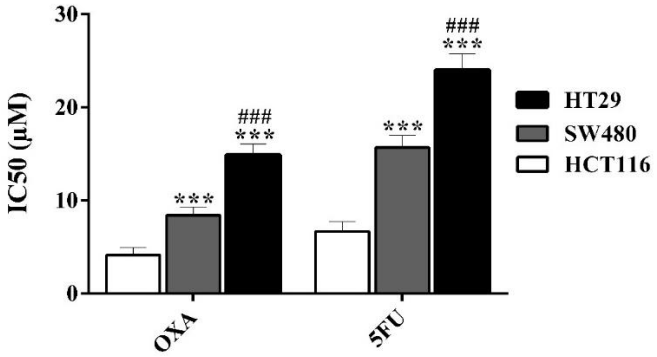


Figure 11. Comparison of IC_{50} of Oxaliplatin (OXA) and 5-FU between Colorectal Cell Lines. ***Significant difference with HCT116 at p<0.001. ###Significant difference with SW480 at p<0.001.

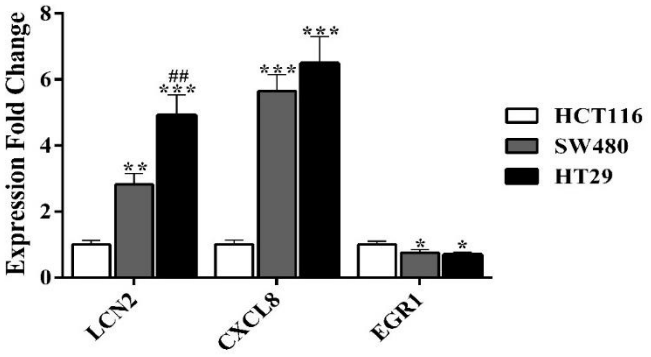


Figure 12. Comparison of expression level of LCN2, CXCL8, EGR1 Genes between Colorectal Cell Lines. *Significant difference with HCT116 at p<0.05, **Significant difference with HCT116 at p<0.01. ***Significant difference with HCT116 at p<0.001. ##Significant difference with SW480 at p<0.01.

proliferation, EMT, invasion, migration and metastasis of CRC through NF-κB pathway.[28] Reduced expression of LCN2 has been reported in drug resistance colorectal cancer cell in Xiao et al.'s study.[29]

Early Growth Response-1 (EGR-1) is a transcription factor induced by growth factors, estrogen and extracellular stimuli or stress signals, and plays an important role in controlling cell proliferation, growth, apoptosis and differentiation. EGR1 is considered to be both tumor suppressor and promoter depending on cell type and stimuli.[30-34] Moon et al., in agreement with our study showed that EGR-1 is involved

Table 5. Relationship between IC_{50} and Expression of Key Candidate Genes

Variables	IC_{50}
CXCL8 expression	Correlation coefficient=0.596 p_{Value} = 0.011
LCN2 expression	Correlation coefficient=0.574 p_{Value} = 0.014
EGR1 expression	Correlation coefficient=0.478 p_{Value} = 0.045

in growth suppression of the human colorectal cells.[35] Also, Han et al., found that knocking down EGR1 inhibited sanguinarine-induced apoptosis in HCT116 cells.[36] Contrary to our study, Myung et al., indicated that upregulation of EGR-1 is associated with CRC invasion, metastasis, tumor stage and poor clinicalc outcome.[30] In Kumar et al.'s study, EGR-1 expression in chemoresistant colorectal cancer cell lines was significantly higher than chemosensetive colorectal cancer cell lines.[37]

Conclusion

The present study revealed that *CXCL8*, *LCN2*, and *EGR1* genes are involved in biological pathways and are considered as important molecular functions related to resistance to conventional chemotraputic drugs in colorectal cancer. So, these three genes signature are potentially exceleent biomarkers for predictiong chemotherapy responsiveness and treatment decicion making of CRC patients. In addition, these genes can be inhibited (*CXCL8*, *LCN2*) or overexpressed (*EGR1*) for therapeutic purposes. However, expermimental tests showed that *CXCL8* gene potentially can more accurately predict the chemotherapy response in patients when used as a biomarker.

Authors' Contributions

All authors equally contributed to this study.

Conflict of Interest Disclosures

The authors declare that they have no conflicts interest.

Ethical Approval

The present project was approved by the Ethics Committee of Hamadan University of Medical Sciences (Ethics Code: IR.UMSHA.REC.1398.908).

Acknowledgment

This study was supported by the Vice-chancellor for Research and Technology, Hamadan University of Medical Sciences (Project Number: 9811158699).

References

1. Wong MC, Huang J, Lok V, Wang J, Fung F, Ding H, et al. Differences in incidence and mortality trends of colorectal cancer worldwide based on sex, age, and anatomic location. Clin Gastroenterol Hepatol. 2021; 19(5):955-66. doi:10.1016/j.cgh.2020.02.026
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3): 209-49. doi:10.3322/caac.21660
3. Xu H, Ma Y, Zhang J, Gu J, Jing X, Lu S, Fu S, Huo J. Identification and verification of core genes in colorectal cancer. Biomed Res Int. 2020;2020:8082697. doi:10.1155/2020/8082697
4. Estevez-Garcia P, Rivera F, Molina-Pinelo S, Benavent M, Gymez J, Limyn ML, et al. Gene expression profile predictive of response to chemotherapy in metastatic colorectal cancer. Oncotarget. 2015;6(8):6151-9. doi:10.18632/oncotarget.3152
5. Zheng Y, Zhou J, Tong Y. Gene signatures of drug resistance predict patient survival in colorectal cancer. Pharmacogenomics J. 2015;15(2):135-43. doi:10.1038/tpj.2014.45
6. Khoshinani HM, Afshar S, Pashaki AS, Mahdavinezhad A, Nikzad S, Najafi R, et al. Involvement of miR-155/FOXO3a and miR-222/PTEN in acquired radioresistance of colorectal cancer cell line. Jpn J Radiol. 2017;35(11):664-72. doi:10.1007/s11604-017-0679-y
7. Manoochehri H, Sheykhhasan M, Samadi P, Pourjafar M, Saidijam M. System biological and experimental validation of miRNAs target genes involved in colorectal cancer radiation response. Gene Rep. 2019;17:100540. doi:10.1016/j.genrep.2019.100540
8. Khoei SG, Manoochehri H, Saidijam M. Systemic biological study for identification of miR-299-5p target genes in cancer. Meta Gene. 2020;24:100655. doi:10.1016/j.mgene.2020.100655
9. Brettingham-Moore KH, Duong CP, Heriot AG, Thomas RJ, Phillips WA. Using gene expression profiling to predict response and prognosis in gastrointestinal cancers—the promise and the perils. Ann Surg Oncol. 2011;18(5):1484-91. doi:10.1245/s10434-010-1433-1
10. Sonachalam M, Shen J, Huang H, Wu X. Systems biology approach to identify gene network signatures for colorectal cancer. Front Genet. 2012;3:80. doi:10.3389/fgene.2012.00080
11. Chu CM, Yao CT, Chang YT, Chou HL, Chou YC, Chen KH, et al. Gene expression profiling of colorectal tumors and normal mucosa by microarrays meta-analysis using prediction analysis of microarray, artificial neural network, classification, and regression trees. Dis Markers. 2014;2014:634123. doi:10.1155/2014/634123
12. Yuan J, Tan L, Yin Z, Tao K, Wang G, Shi W, et al. Bioinformatics analysis identifies potential chemoresistance-associated genes across multiple types of cancer. Oncol Lett. 2019;18(3):2576-83. doi:10.3892/ol.2019.10533
13. Manoochehri H, Asadi S, Tanzadehpanah H, Sheykhhasan M, Ghorbani M. CDC25A is strongly associated with colorectal cancer stem cells and poor clinical outcome of patients. Gene Rep. 2021;25:101415. doi:10.1016/j.genrep.2021.101415
14. 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. 2021:1-20. doi:10.1007/s12029-021-00690-2
15. Lv J, Li L. Hub genes and key pathway identification in colorectal cancer based on bioinformatic analysis. Biomed Res Int. 2019;2019:1545680. doi:10.1155/2019/1545680
16. Fekete JT, Győrffy B. ROCplot. org: Validating predictive biomarkers of chemotherapy/hormonal therapy/anti-HER2 therapy using transcriptomic data of 3,104 breast cancer patients. Int J Cancer. 2019;145(11):3140-51. doi:10.1002/ijc.32369
17. Manoochehri H, Gheitasi R, Pourjafar M, Yazdi A, Sheykhhasan M, Amini R. Investigating the relationship between the severity of coronary artery disease and expression level of TRAF3IP2. Gene Rep. 2021;23: 101041. doi:10.1016/j.genrep.2021.101041
18. Gharib A, Komaki A, Khoshinani HM, Saidijam M, Barkley V, Sarihi A, et al. Intrahippocampal 5-HT1A

receptor antagonist inhibits the improving effect of low-frequency stimulation on memory impairment in kindled rats. Brain Res Bull. 2019;148:109-17. doi:10.1016/j.brainresbull.2019.03.007
19. Abu N, Othman N, W Hon K, Nazarie WF, Jamal R. Integrative meta-analysis for the identification of hub genes in chemoresistant colorectal cancer. Biomarkers Med. 2020;14(7):525-37. doi:10.2217/bmm-2019-0241
20. Dzobo K, Hassen N, Senthebane DA, Thomford NE, Rowe A, Shipanga H, et al. Chemoresistance to Cancer Treatment: Benzo-α-Pyrene as Friend or Foe?. Molecules. 2018;23(4):930. doi:10.3390/molecules23040930
21. Tian D, Tian M, Han G, Li JL. Increased glucocorticoid receptor activity and proliferation in metastatic colon cancer. Sci Rep. 2019;9(1):11257. doi:10.1038/s41598-019-47696-2
22. Orrenius S, Nicotera P, Zhivotovsky B. Cell death mechanisms and their implications in toxicology. Toxicol Sci. 2011;119(1):3-19. doi:10.1093/toxsci/kfq268
23. Bie Y, Ge W, Yang Z, Cheng X, Zhao Z, Li S, et al. The crucial role of CXCL8 and its receptors in colorectal liver metastasis. Dis Markers. 2019;2019:8023460. doi:10.1155/2019/8023460
24. Lee YS, Choi I, Ning Y, Kim NY, Khatchadourian V, Yang D, et al. Interleukin-8 and its receptor CXCR2 in the tumour microenvironment promote colon cancer growth, progression and metastasis. Br J Cancer. 2012;106(11):1833-41. doi:10.1038/bjc.2012.177
25. Dabkeviciene D, Jonusiene V, Zitkute V, Zalyte E, Grigaitis P, Kirveliene V, et al. The role of interleukin-8 (CXCL8) and CXCR2 in acquired chemoresistance of human colorectal carcinoma cells HCT116. Med Oncol. 2015;32(12):258. doi:10.1007/s12032-015-0703-y
26. Srinivasalu VK, Chaudhary N, Bhagyashree., Dwivedi N, PK S, Thorat R, Batulla C, Dhar SK, Dalal SN, Das M. LCN2 and colon cancer—Have we hit the jackpot. J Clin Oncol. 2020;38(15 suppl):e15608. doi:10.1200/jco.2020.38.15_suppl.e15608
27. Maier HT, Aigner F, Trenkwalder B, Zitt M, Vallant N, Perathoner A, et al. Up-regulation of neutrophil gelatinase-associated lipocalin in colorectal cancer predicts poor patient survival. World J. Surg. 2014;38(8):2160-7. doi:10.1007/s00268-014-2499-x
28. Feng M, Feng J, Chen W, Wang W, Wu X, Zhang J, et al. Lipocalin2 suppresses metastasis of colorectal cancer by attenuating NF-κB-dependent activation of snail and epithelial mesenchymal transition. Mol Cancer. 2016;15(1):77. doi:10.1186/s12943-016-0564-9
29. Guo X, Li Q, Wang YF, Wang TY, Chen SJ, Tian ZW. Reduced lipocalin 2 expression contributes to vincristine resistance in human colon cancer cells. Recent Pat Anti-Cancer Drug Discov. 2018;13(2):248-54. doi:10.2174/1574892813666171221120504
30. Myung DS, Park YL, Kim N, Chung CY, Park HC, Kim JS, et al. Expression of early growth response-1 in colorectal cancer and its relation to tumor cell proliferation and apoptosis. Oncol Rep. 2014;31(2):788-94. doi:10.3892/or.2013.2884
31. Shajahan-Haq AN, Boca SM, Jin L, Bhuvaneshwar K, Gusev Y, Cheema AK, et al. EGR1 regulates cellular metabolism and survival in endocrine resistant breast cancer. Oncotarget. 2017;8(57):96865-84. doi:10.18632/oncotarget.18292
32. Wei F, Jing H, Wei M, Liu L, Wu J, Wang M, et al. Ring finger protein 2 promotes colorectal cancer progression by suppressing early growth response 1. Aging (Albany NY). 2020;12(24):26199-220. doi:10.18632/aging.202396
33. Kim SH, Park YY, Cho SN, Margalit O, Wang D, DuBois RN. Kruppel-like factor 12 promotes colorectal cancer growth through early growth response protein 1. PLoS One. 2016;11(7):e0159899. doi:10.1371/journal.pone.0159899
34. Lee SH, Bahn JH, Choi CK, Whitlock NC, English AE, Safe S, et al. ESE-1/EGR-1 pathway plays a role in tolfenamic acid-induced apoptosis in colorectal cancer cells. Mol Cancer Ther. 2008;7(12):3739-50. doi:10.1158/1535-7163.MCT-08-0548
35. Moon YS, Cui LG, Yang H. Involvement of early growth response gene 1 (EGR-1) in growth suppression of the human colonic tumor cells by apigenin and its derivative isovitexin. J Life Sci. 2007;17(1):110-5. doi:10.5352/JLS.2007.17.1.110
36. Han MH, Kim GY, Yoo YH, Choi YH. Sanguinarine induces apoptosis in human colorectal cancer HCT-116 cells through ROS-mediated Egr-1 activation and mitochondrial dysfunction. Toxicol lett. 2013;220(2):157-66. doi:10.1016/j.toxlet.2013.04.020
37. Kumar SS, Tomita Y, Wrin J, Bruhn M, Swalling A, Mohammed M, et al. High early growth response 1 (EGR1) expression correlates with resistance to anti-EGFR treatment in vitro and with poorer outcome in metastatic colorectal cancer patients treated with cetuximab. Clin Transl Oncol. 2017;19(6):718-26. doi:10.1007/s12094-016-1596-8