frontiers in Oncology

ORIGINAL RESEARCH
published: 24 June 2021
doi: 10.3389/fonc.2021.597536

Systems Biomedicine of Primary and Metastatic Colorectal Cancer Reveals Potential Therapeutic Targets

Mehran Piran[1,2*], *Neda Sepahi*[3], *Afagh Moattari*[2], *Amir Rahimi*[4] *and Ali Ghanbariasad*[3]

[1] Department of Anatomy and Developmental Biology, Monash University, Melbourne, VIC, Australia, [2] Department of Bacteriology and Virology, Medical School, Shiraz University of Medical Sciences, Shiraz, Iran, [3] Noncommunicable Diseases Research Center, Fasa University of Medical Sciences, Fasa, Iran, [4] Bioinformatics and Computational Biology Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

Colorectal cancer (CRC) is one of the major causes of cancer deaths across the world. Patients' survival at time of diagnosis depends mainly on stage of the tumor. Therefore, understanding the molecular mechanisms from low-grade to high-grade stages of cancer that lead to cellular migration from one tissue/organ to another tissue/organ is essential for implementing therapeutic approaches. To this end, we performed a unique meta-analysis flowchart by identifying differentially expressed genes (DEGs) between normal, primary (primary sites), and metastatic samples (Colorectal metastatic lesions in liver and lung) in some Test datasets. DEGs were employed to construct a protein-protein interaction (PPI) network. A smaller network containing 39 DEGs was then extracted from the PPI network whose nodes expression induction or suppression alone or in combination with each other would inhibit tumor progression or metastasis. These DEGs were then verified by gene expression profiling, survival analysis, and multiple Validation datasets. We suggested for the first time that downregulation of mitochondrial genes, including ETHE1, SQOR, TST, and GPX3, would help colorectal cancer cells to produce more energy under hypoxic conditions through mechanisms that are different from "Warburg Effect". Augmentation of given antioxidants and repression of P4HA1 and COL1A2 genes could be a choice of CRC treatment. Moreover, promoting active GSK-3β together with expression control of EIF2B would prevent EMT. We also proposed that OAS1 expression enhancement can induce the anti-cancer effects of interferon-gamma, while suppression of CTSH hinders formation of focal adhesions. ATF5 expression suppression sensitizes cancer cells to anchorage-dependent death signals, while LGALS4 induction recovers cell-cell junctions. These inhibitions and inductions would be another combinatory mechanism that inhibits EMT and cell migration. Furthermore, expression inhibition of TMPO, TOP2A, RFC3, GINS1, and CKS2 genes could prevent tumor growth. Besides, TRIB3 suppression would be a promising target for anti–angiogenic therapy. SORD is a poorly studied enzyme in cancer, found to be upregulated in CRC. Finally, TMEM131 and

OPEN ACCESS

Edited by:
Yumin Li,
Lanzhou University, China

Reviewed by:
Jian Tu,
University of South China, China
Ravins Dohare,
Jamia Millia Islamia, India

***Correspondence:**
Mehran Piran
mehran.piran@monash.edu

Specialty section:
This article was submitted to Gastrointestinal Cancers, a section of the journal Frontiers in Oncology

Received: *21 October 2020*
Accepted: *31 May 2021*
Published: *24 June 2021*

Citation:
Piran M, Sepahi N, Moattari A, Rahimi A and Ghanbariasad A (2021) Systems Biomedicine of Primary and Metastatic Colorectal Cancer Reveals Potential Therapeutic Targets. Front. Oncol. 11:597536. doi: 10.3389/fonc.2021.597536

DARS genes were identified in this study whose roles have never been interrogated in any kind of cancer, neither as a biomarker nor curative target. All the mentioned mechanisms must be further validated by experimental wet-lab techniques.

Keywords: colorectal cancer, cancer progression, EMT, metastasis, therapeutic/curative targets, diagnostic biomarkers

INTRODUCTION

Colorectal cancer (CRC) is a major global medical burden worldwide (1). Approximately more than one million people are diagnosed with CRC each year, and about half of them die of CRC annually (2). Complex genetic interactions are combined with environmental factors to trigger a cell to become cancerous. Among them, aberrant growth factor signals contribute to uncontrolled cells' proliferation, which ultimately leads to metastasis. Contrary to early-stage tumor cells, malignant cells have the ability to detach from the stroma as well as acquire motility (3). This event happens during a process called Epithelial-Mesenchymal Transition (EMT), in which cells lose their epithelial characteristic, including adhesion, and subsequently dedifferentiate into mesenchymal mobile cells (4). Therefore, Investigating DEGs between primary and metastatic sites of tumors would aim to recognize key factors playing roles in cell migration. We performed the statistical analysis between primary sites and metastatic sites in one part of the analyses. While primary sites were non-malignant colon biopsies in Test datasets, CRC metastatic sites were located on the other organs.

Many molecular and pathway targets have been identified for treatment of CRC during the past decades. Besides, growing progresses have been made in development of chemotherapy and antibody drugs (5). Tyrosine kinase (TK) targeting monoclonal antibodies and small-molecule tyrosine kinase inhibitors are effective strategies (6). Targeting cancer-related inflammation biomarkers like IL-6/JAK/STAT3 pathway, which inhibits progression of solid tumors, is another beneficial therapeutic strategy (7). In addition, restraint of cytosolic β-catenin *via* disturbing hyperactive Wnt/β-catenin signaling pathway could be another treatment approach for colorectal and many other types of cancer (8–10). Inhibition of matrix metalloproteinases (MMPs) and TGFβ signaling pathways is a therapeutic approach to prevent liver metastasis (11–14). Furthermore, PI3K inhibition suppresses lung metastasis in CRC patients (15, 16). Among the known anticancer drugs, Cetuximab is one of the popular ones, which is a monoclonal antibody against epidermal growth factor receptor (EGFR) (17). Furthermore, vascular endothelial growth factor (VEGF) antibody, bevacizumab, is the standard treatment for metastatic colorectal cancer (18).

The aim of this study was to suggest multiple combinations of genetic targets that can prevent cancer progression. Therefore, We looked for the unknown key factors that partially control one or more steps of cancer progression, including cell proliferation, transformation, angiogenesis, and metastasis to the distant secondary sites. One way to identify molecular mechanism of pathogenesis in a biological context is to analyze transcriptomic data. Systematic investigation of gene expression data and cellular and molecular information in the literature for identified DEGs in normal and cancer tissues helped us to propose a number of these DEGs as therapeutic genetic targets. Once these targets were identified, we would see which ones can be targeted together in order to hinder cancer progression because each of the proposed genetic targets could control to some extent different steps of tumor formation towards malignancy. We conducted a unique meta-analysis flowchart where we separated datasets into two sets of Test and Validation datasets in order to not only recognize DE genes but also introduce them as the curative CRC targets. Furthermore, the shortest pathway scoring system for neighborhood finding around the core genes was first introduced by Seth I. Berger and his colleagues (Systems Pharmacology of Arrhythmias) on Long-QT syndrome (LQTS) PPI Network (19). They realized that the neighborhood ranking around the twelve core genes (Drivers of the LQTS Syndrome) discovered the genes that are targeted with FDA-approved drugs. Since our goal was to find therapeutic genetic targets to inhibit metastasis, we applied the algorithm on colorectal cancer PPI network for the first time. First of all, three Test datasets were constructed from three microarray studies, and DE genes were excavated for any pairwise comparison between four groups of samples. Common DEGs between similar analyses in Test datasets were regarded as final DEGs employed for PPI network assembly. Test datasets provided us with a sufficient number of common DEGs with desired log-fold change and p-value thresholds for network construction. Twelve common genes called Core genes were recognized that their expression were different between primary and metastatic sites. A smaller network called Core network was then extracted from the PPI based on a shortest-path-based scoring formula on these Core genes. To compensate for the small number of datasets in Test set, seven Validation datasets were employed from different genomic repositories to validate selected DEGs in the Core network (**Figure 1**). Besides, expression profiling and survival analysis provided more evidence about the accuracy of our results. We obtained some DEGs involved in cancer progression whose expression could be targeted (suppressed or induced) individually or in combination with one another for CRC treatment. Moreover, some of those gene expressions were proposed to be CRC biomarkers.

MATERIALS AND METHODS

Database Searching and Recognizing Pertinent Experiments

Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) and ArrayExpress (https://www.ebi.ac.uk/arrayexpress/)

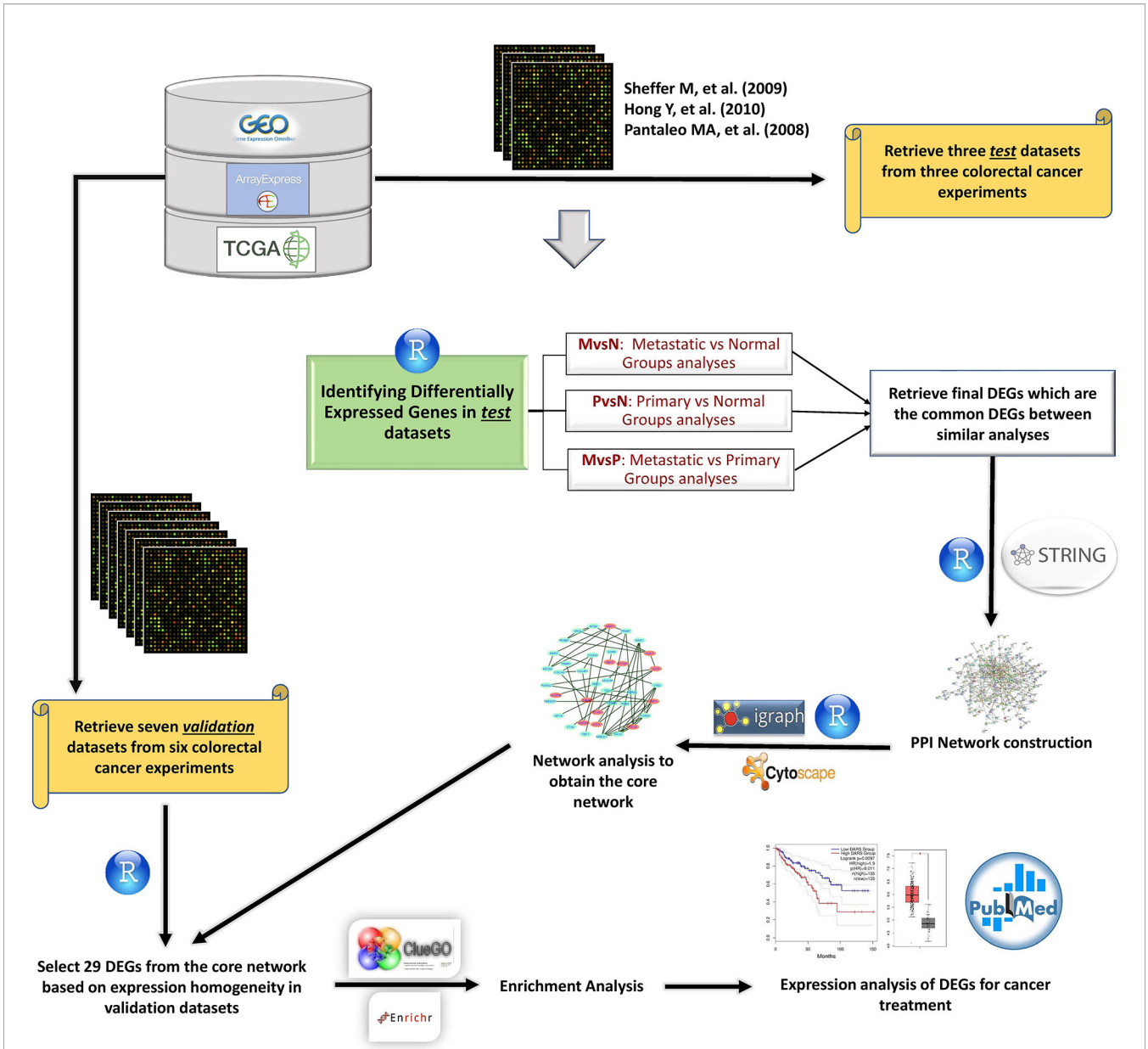

FIGURE 1 | The meta-analysis flowchart to attain therapeutic genetic targets. Gene expression datasets were extracted from different databases. Data were analyzed and visualized using R programming language. DEGs were obtained from analyzing Test datasets, then verified by Validation datasets. STRING database was utilized to construct the PPI network from DEGs. R software was used to analyze the network. Cytoscape was employed to visualize the networks, and enrichment results were obtained from ClueGO Cytoscape plugin and Enrichr online tools. Next, survival analysis and expression profiling were used for more validation of expression results. Finally, our results were compared to other studies, and molecular mechanism of validated DEGs was interrogated to propose a combination of target therapies.

repositories were searched to detect experiments containing high-quality transcriptomic samples concordance to the study design. Colorectal/colon cancer, primary, EMT, and metastasis were the keywords utilized in the search, but search was filtered for Homo sapiens. Microarray raw data (.CEL files) for GSE41258, GSE9348, and GSE10961 studies were downloaded from GEO and ArrayExpress to create Test datasets (20–22). Dataset1 and Dataset2 were constructed from samples in GSE41258 study. Dataset1 encompassed CRC liver metastasis samples, primary samples and normal samples, while Dataset2 contained CRC lung metastasis samples, primary samples and normal samples. To construct Dataset3, normal samples and primary samples were extracted from GSE9348 study, but colorectal liver metastasis samples were obtained from GSE10961 study. In all datasets, normal samples were healthy colon tissues adjacent to the primary tumors and primary samples were non-metastatic colorectal biopsies. To make Validation datasets, two new datasets were constructed from GSE41258 study whose samples

were not present in Test datasets. One dataset contained CRC liver metastasis samples, and another one contained CRC lung metastasis samples. In addition, a dataset was constructed from count RNAseq files in The Cancer Genome Atlas (TCGA) database (TCGA dataset). Three RNAseq datasets were constructed from GSE50760, GSE144259, and GSE89393 studies encompassing CRC liver metastasis samples (23–25). The last dataset was built from GSE40367 microarray study containing CRC liver metastasis samples (26). Except for TCGA dataset, all Test and Validation datasets contained three groups of metastatic, primary and normal samples.

Identifying Differential Expressed Genes in Microarray Datasets

R programming language (v3.6.2) was used to import and analyze data for each dataset separately. Preprocessing step involving background correction and probe summarization was done using RMA method in "affy" package (27). Absent probesets were identified using "mas5calls" function in this package. If a probeset contained more than two absent values in each group of samples, that probeset was regarded as absent and removed from the expression matrix. Besides, outlier samples were identified and removed using PCA and hierarchical clustering methods. Next, data were normalized using Quantile normalization approach (28). Then, standard deviation (SD) for each gene was computed, and median of all SDs was utilized as a cutoff to remove low-variant probesets. Therefore, low-variant probesets no longer influenced the significance of the high-variant genes. "Many to Many" problem (29), which is mapping multiple probesets to the same gene symbol, was solved using "nsFilter" function in "genefilter" package (30). This function selects the probeset with the highest Interquartile range (IQR) to map to the gene symbol. "limma" R package, which applies linear models on the expression matrix, was utilized to discover DE genes between all groups of samples (31). Genes with absolute log fold change larger than 0.5 and Benjamini-Hochberg adjusted p-value (32) less than 0.05 were selected as the DEGs.

Identifying Differential Expressed Genes in RNAseq Datasets

Count files for five primary samples containing more than 90 percent tumor cells as well as five normal samples involving 100 percent normal cells were downloaded from TCGA database. Each sample was imported into R, and they were merged together to construct the TCGA expression matrix encompassing a five-sample primary group and a five-sample normal group. Genes with zero expressions in the two groups were omitted. Then, data were normalized with "DESeq2" R package (33), and DEGs were identified between the two groups. For RNAseq datasets in GEO, FPKM normalized data were downloaded and imported into R. data were log2 transformed, and using "limma" R package, DEGs were identified between the groups.

Network Construction

Final DEGs were imported into STGRING web server database and different sources of evidence were chosen to generate interactions and the Protein-Protein Interaction (PPI) network. Afterward, Interactions were downloaded and imported into R programming language in form of an annotated edgelist. Next, extra information were removed, and an undirected edgelist was obtained using "igraph" R package (34). Interaction scores were considered as weights, so a weighted PPI network was created. The giant component of the weighted PPI network was then extracted for further analyses. The weighted adjacency matrix of the giant component was created and transformed into a symmetric matrix. It was then modified into a new adjacency matrix using topological overlapping measure (TOM) function in "WGCNA" R package (35). Finally, this modified adjacency matrix was subtracted from one to create a distance adjacency matrix.

Neighborhood Ranking to the Core Genes

Using Dijkstra algorithm in R, a matrix of all shortest paths, called SP, between all pairs of nodes was constructed from the distance adjacency matrix (36). By utilizing this matrix, a distance score, DJ, for each node in the PPI network was computed. Moreover, we considered DEGs between metastatic versus primary analysis as the Core nodes in the PPI network. Dj is a scoring formula that is the average of the shortest paths from all the non-core nodes to reach the node j subtracted from the average of the shortest paths from the Core nodes to reach the node j divided by the average of the all shortest paths to reach the node j from the whole network. This scoring system implies how much close each node is to the Core nodes (19, 37).

$$D_j = \Sigma_{i \notin c} SP_{ij}/NC - \Sigma_{i \in c} SP_{ij}/C/\Sigma_i SP_{ij}/NC + C$$

C is the number of Core nodes, and NC is the number of non-core nodes. $\Sigma_i \notin c\ SP_{ij}$ is the sum of all distances in SP matrix between node j and all the non-core nodes. $\Sigma_i \in c\ SP_{ij}$ is the sum of the distances between node j and all the Core nodes. $\Sigma_i\ SP_{ij}$ is the sum of the distances between node j and all the nodes. A positive score implies that node j is closer to the Core nodes than the rest of the nodes. Nodes with positive scores were kept, but the others were removed from the network. It should be noted that D scores were calculated without imposing any threshold on edge weights. The R source codes for the network analysis are available at https://github.com/mehranpiran/Meta-Analysis.

Enrichment Analysis

Enrichment analysis was performed using ClueGO Cytoscape plugin (38). Enriched terms for biological processes were obtained from GO repository. For pathway enrichment analysis, information in KEGG (39), Reactome (40) and WikiPathways (41) databases were used. P-value were adjusted using Benjamini-Hochberg method and cut off was set on 0.05. In addition to Cytoscape, Enrichment analysis was performed using Enrichr online tool (42) as well. Enriched terms for biological processes were obtained from GO repository. For pathway enrichment analysis, WikiPathways signaling repository version 2019 for humans was used. Enriched terms with a top score and a p-value less than 0.05 were selected.

Analyzing Gene Expression Profiles

Genes were given to GEPIA2 webserver to validate identified DEGs based on datasets in TCGA genomic database (43, 44). To draw boxplots, expression profiles were compared between tumor

and normal samples in multiple colorectal adenocarcinoma (COAD) datasets. LogFC cutoff was set at 0.5 and p-value was set at 0.01. TPM normalized data were log2 transformed. To draw survival plots, "Overall Survival" option was selected and median was chosen to define the border of High and Low groups. 95% confidence interval was set for analysis. All COAD datasets with monthly expression values were selected in order to obtain survival results.

RESULTS

Data Preprocessing in Test Datasets

Each dataset was imported into R separately. Outlier sample detection was conducted using PCA and hierarchical clustering approaches. **Figure 2A** illustrates the PCA plot for samples in Dataset1. The same plot was created for the second and third datasets. Some samples in the PCA plane lay at a distance from their group, particularly along the PC1 axis, so they were considered as the outliers. To be more specific, a hierarchical clustering method introduced by Oldham MC, et al. (45) was used. To compute the distances between samples, Pearson correlation analysis was performed between them, and coefficients were subtracted from one. **Figure 2B** depicts the dendrogram for normal samples. In **Figure 2C** normal samples were plotted based on their Number-SD scores. To get this number for each sample, the average of whole distances was subtracted from the distance average of each sample. Then, results of these subtractions were normalized by standard deviation of sample distance averages (45). Samples with Number-SD less than negative two usually fall apart from their cluster set in the PCA plane. Thus, they were regarded as the outliers in our analyses. Sixteen outlier samples in GSE41258 Test dataset and three outliers in Dataset3 were recognized. **Supplementary File 1** contains information about the groups of samples.

Supplementary File 2 illustrates the average expression values for some housekeeping genes and DEGs between Primary and Metastatic samples. Common DEGs between lung-primary analysis and liver-primary analysis with absolute LogFC larger than one in GSE41258 datasets were illustrated in A. The same plot was made for the common DEGs in liver-primary analysis in the third dataset in B. Housekeeping genes

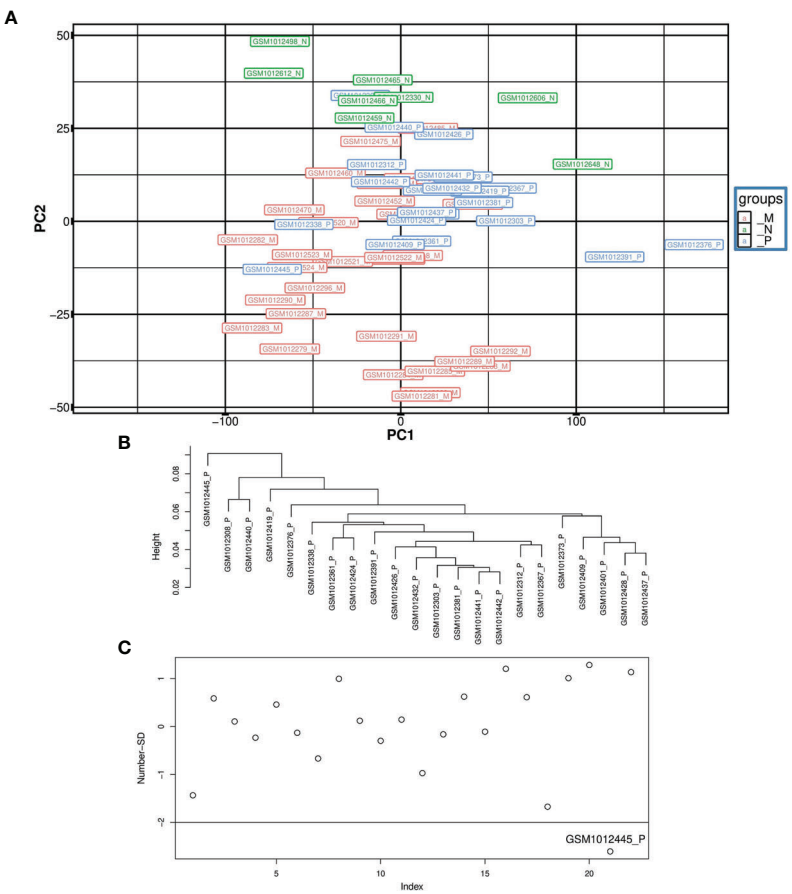

FIGURE 2 | Illustration of outlier samples in the first dataset. **(A)** is the PCA plot, **(B)** is the dendrogram for the primary samples and **(C)** is the Number-SD plot for primary samples. GSM1012445 is an outlier sample in primary group as it has located in a distance from its group in the PCA plane. In addition, it has formed a unique cluster in the dendrogram and its Number-SD score is less than negative two.

were situated on the diagonal of the plots whilst DEGs were located above or under the diagonal. Hence, the preprocessed data were of sufficient quality for downstream analyses.

Meta-Analysis and Identifying Differentially Expressed Genes

10891 unique DEGs with adjusted p-value < 0.05 and absolute log fold change > 0.5 were achieved from eight groups of DEGs yielded from eight independent analyses on three Test datasets. They included two analyses of liver metastasis versus normal, two analyses of liver metastasis versus primary, one analysis of lung metastasis versus normal, one analysis of lung metastasis versus primary, and two analyses of primary versus normal (**Table 1**). Liver metastasis contained metastatic colorectal samples taken out from liver; however, lung metastasis contained metastatic colorectal samples obtained from lung. In fact, this kind of dataset selection provided us with some DE genes that could significantly contribute to tumor progression within the primary site and towards liver and lung organs. Common DEGs between all metastasis vs normal analyses (Test datasets) were 155 genes. Common DEGs between all metastasis vs primary analyses (Test datasets) were 72 genes. Common DEGs between all primary vs normal analyses (Test datasets) were 239 genes. There were 334 unique DEGs between these three sets of analyses. Finally, from these 334 DEGs, 242 of them were identified to be in all Test and Validation analyses considered as the final DEGs. There were 12 final DEGs in primary versus normal analyses considered as the Core genes. All DE gene sets and their LogFC are presented in **Supplementary File 3**.

Undirected Protein-Protein Interaction Network

242 final DEGs were utilized to construct the Protein-Protein-Interaction (PPI) network. STRING database was employed to generate the Interactions based on seven sources of evidence, namely Neighborhood, Text mining, Experiments, Databases, Co-expression, Gene fusion, and Co-occurrence. STRING combined scores were used as the edge weights. The giant component of the weighted network with 205 nodes and 554 edges is depicted in **Figure 3**.

Determination of Core Genes Neighborhood Through Shortest Path-Based Scoring System

In this step, interactions combined score computed from all sources of evidence in STRING database were converted into weights between nodes. These weights were used as the estimation of distances in the weighted adjacency matrix. Nodes with shorter distances from the Core genes were selected, and a smaller network was extracted from the main network. Computing the shortest path score for the non-core genes led to a network of 39 nodes comprising 12 Core nodes and 27 neighbors. This multi-component graph called Core network is illustrated in **Figure 4**.

Majority of the nodes in Core network were selected for investigation based on the similarity of expression patterns in all datasets. Expression status of selected genes between any pairwise comparisons was depicted in **Table 1**. For the three Metastatic-

TABLE 1 | The practical information for the Core network DE genes in Test dataset.

	MvsN				BvsN	
DEGs	**GSE41258**		**GSE9348_GSE10961**	**DEGs**	**GSE41258**	**GSE9348_GSE10961**
	liver-normal (D1)	**lung-normal (D2)**	**liver-normal (D3)**		**primary-normal (D1/2)**	**primary-normal (D3)**
ETHE1	Down	Down	Down	SGK1	Down	Down
DARS	Up	Up	Up	EIF2S2	Up	Up
TMEM131	Down	Down	Down	TRIB3	Up	Up
TST	Down	Down	Down	DARS	Up	Up
LGALS4	Down	Down	Down	RFC3	Up	Up
TRIB3	Up	Up	Up	ETHE1	Down	Down
COL5A2	Up	Up	Up	TOP2A	Up	Up
COL4A1	Up	Up	Up	CKS2	Up	Up
PTP4A1	Down	Down	Down	SORD	Up	Up
COL1A2	Up	Up	Up	PSMA7	Up	Up
SORD	Up	Up	Up	GPX3	Down	Down
SQOR	Down	Down	Down	MAD2L1	Up	Up
OAS1	Down	Down	Down	SQOR	Down	Down
TRIM31	Down	Down	Down	TST	Down	Down
TWF1	Down	Down	Down	GINS1	Up	Up
MvsP				LGALS4	Down	Down
DEGs	GSE41258		GSE9348_GSE10961	PTP4A1	Down	Down
	liver-primary (D1)	lung-primary (D2)	liver-primary (D3)	PLAGL2	Up	Up
ETHE1	Down	Down	Down	COL1A2	Up	Up
ATF5	Up	Up	Up	CTSH	Up	Up
CDC6	Down	Down	Down	MT2A	Down	Down
TMPO	Down	Down	Down	COL5A2	Up	Up
P4HA1	Up	Up	Up			

"Up" means gene was upregulated, and "Down" means gene was downregulated. MvsN contains all the metastatic versus normal analyses, PvsN contains all the primary versus normal analyses and MvsP contains all the metastatic versus primary analyses. D stands for Dataset, so D1 means Dataset1. Some genes are present in more than one analysis.

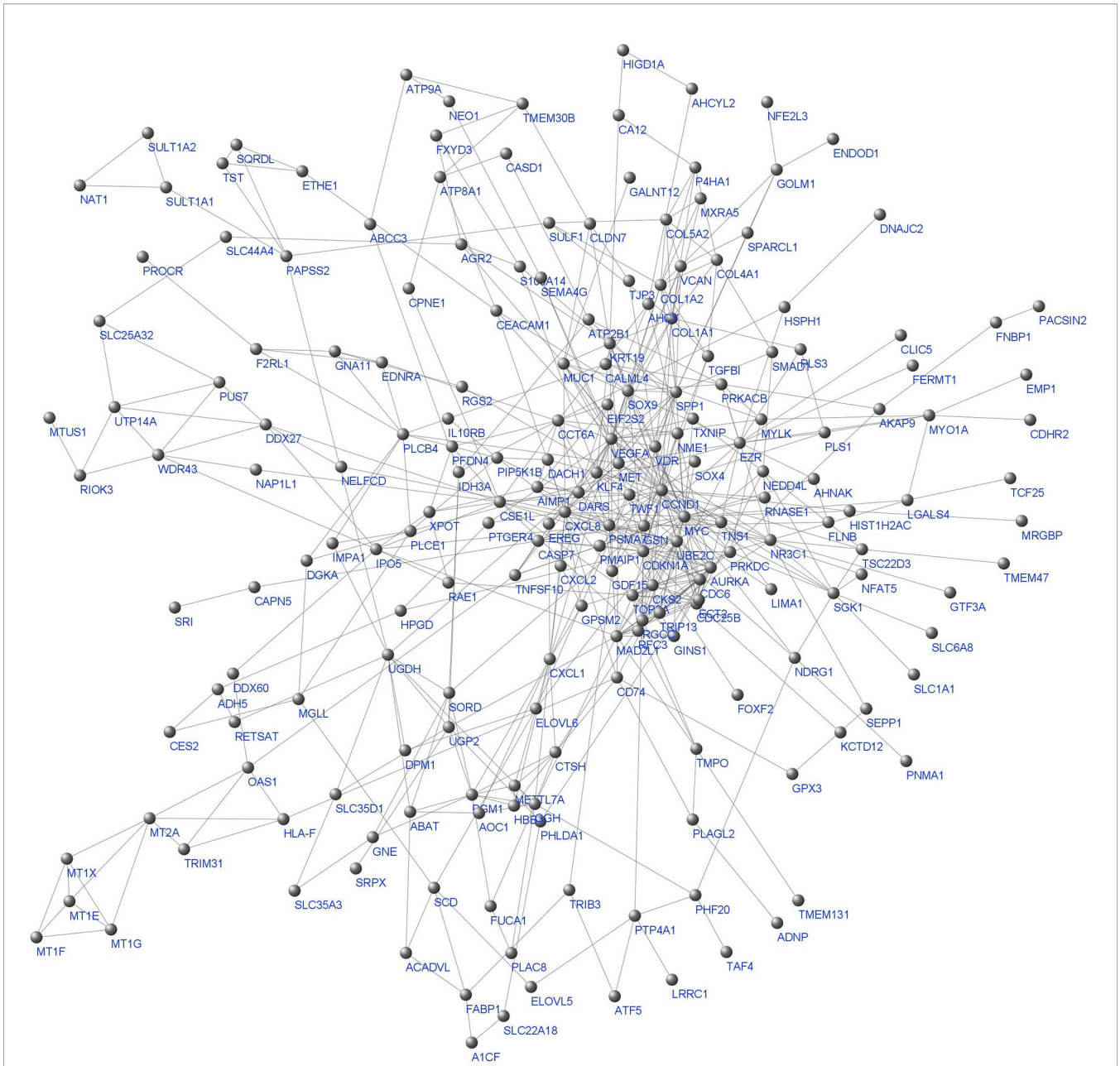

FIGURE 3 | The whole network giant component. Labels are protein/gene symbols. This is a scale free network (46) which follows a power law distribution (most of the network nodes have a low degree while there are few nodes with high degree).

Normal comparisons (MvsN), most of the nodes exhibited a similar expression pattern. The same was true for all Primary-Normal analyses (PvsN) and Metastatic-Primary comparisons (MvsP). Heatmaps were illustrated in **Figure 5** for all members of the Core network in three datasets. Clustering was performed by applying "Euclidean" distance and "Complete" method on gene expression values. Genes present in the top right corner of the three plots possessed high expression values in colon tissues. Moving from border to the center of plots, we go from Normal to primary and from primary to metastatic samples. Some genes exhibited a descending expression trend such as mitochondrial genes ETHE1, TST and SQOR. Few genes witnessed an ascending trend, such as collagen genes and SORD and P4HA1.

Network Descriptive

The giant component diameter was eight containing TRIM31, HLA-F, CD74, PLAGL2, TMPO, MAD2L1 PSMA7, AIMP1, and TWF1. Transitivity was around 60%, edge density was about 18%, and the mean distance was 3.48. Two important centralities, Degree and Betweenness, along with other centralities and the

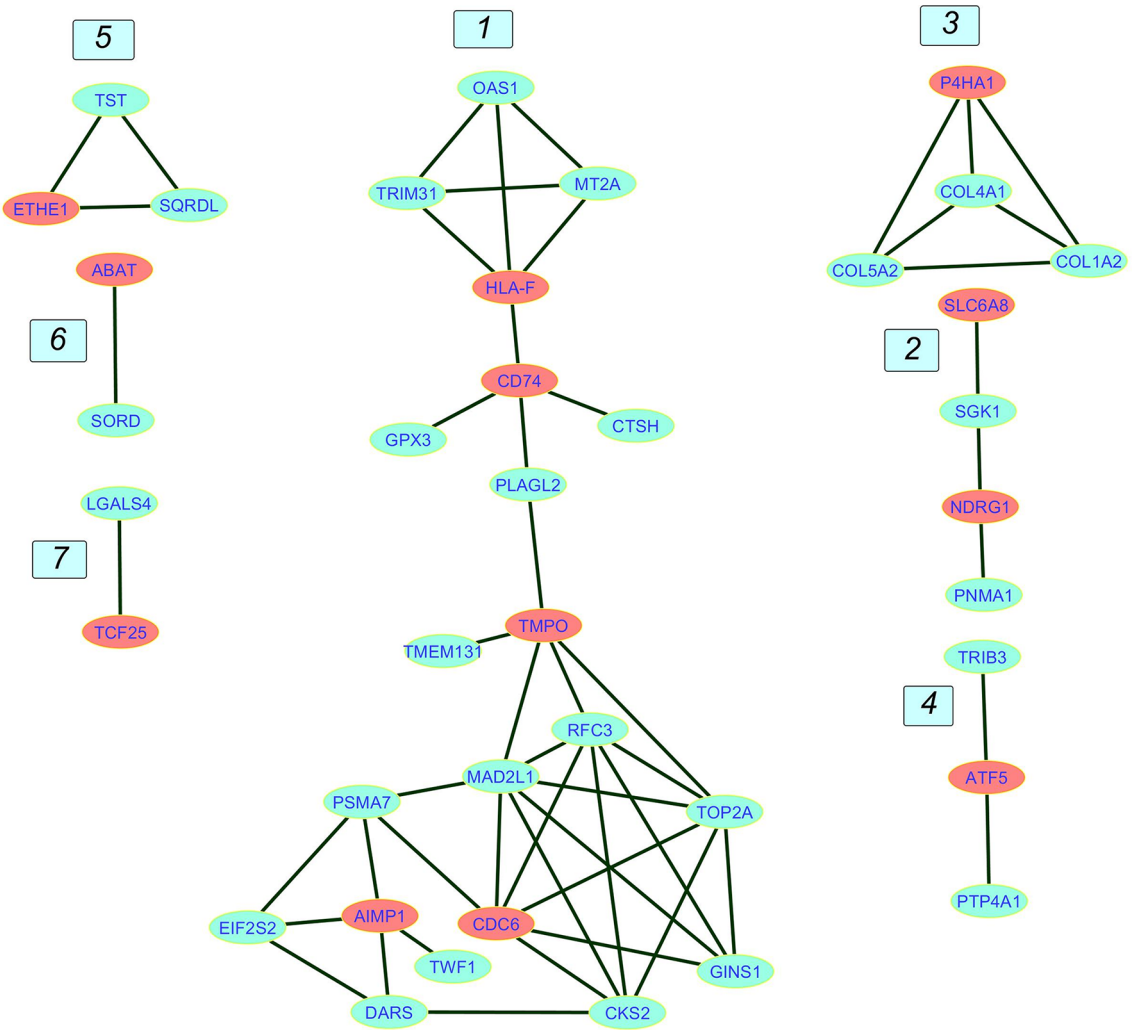

FIGURE 4 | The Core network. The network contains seven components numbered from 1 to 7. Component 1 is the giant component. Core genes are in red and non-core genes are in blue.

average distances for giant component nodes, are provided in **Supplementary File 4**. MAD2L1 had the highest Degree and a relatively high betweenness. TMPO had the highest Betweenness and a pretty high degree. Similar to TMPO, its direct neighbor, PLAGL2 had a relatively high Betweenness. This gene has linked the upper and lower parts of the PPI giant component together.

Processing Validation Datasets

Core network nodes were identified in the seven Validation datasets. They were presented in **Table 2**. Most of the DEGs illustrated similar results in both **Tables 1**, **2**, which proves the accuracy of obtained DEGs from Test datasets. Expression of genes that were totally homogeneous in each of MvsN or MvsP or PvsN analyses are presented in green, and the ones that differed only in one analysis are shown in yellow. Expression of Genes that were different in more than one dataset are in white. Absolute LogFCs less than 0.2 were not reported in **Table 2**. Expression analysis of all Validation datasets are presented in **Supplementary File 3**.

Over-Representation Analysis

Figure 6 illustrates the enrichment results for the Core network genes using ClueGO software. Three signaling databases called KEGG, Reactome, and WikiPathways were used for the pathway enrichment. Biological Function terms were enriched from GO database. Genes and terms associated with a specific cellular mechanism formed distinct components. Different pathway terms related to polymerization and degradation of collagens in extra cellular matrix (ECM) have emerged in blue, which formed a distinct component (component 3) in the Core network. In a tumor environment, different concentrations of collagen fibers are regularly secreted, degraded, and aligned together to make

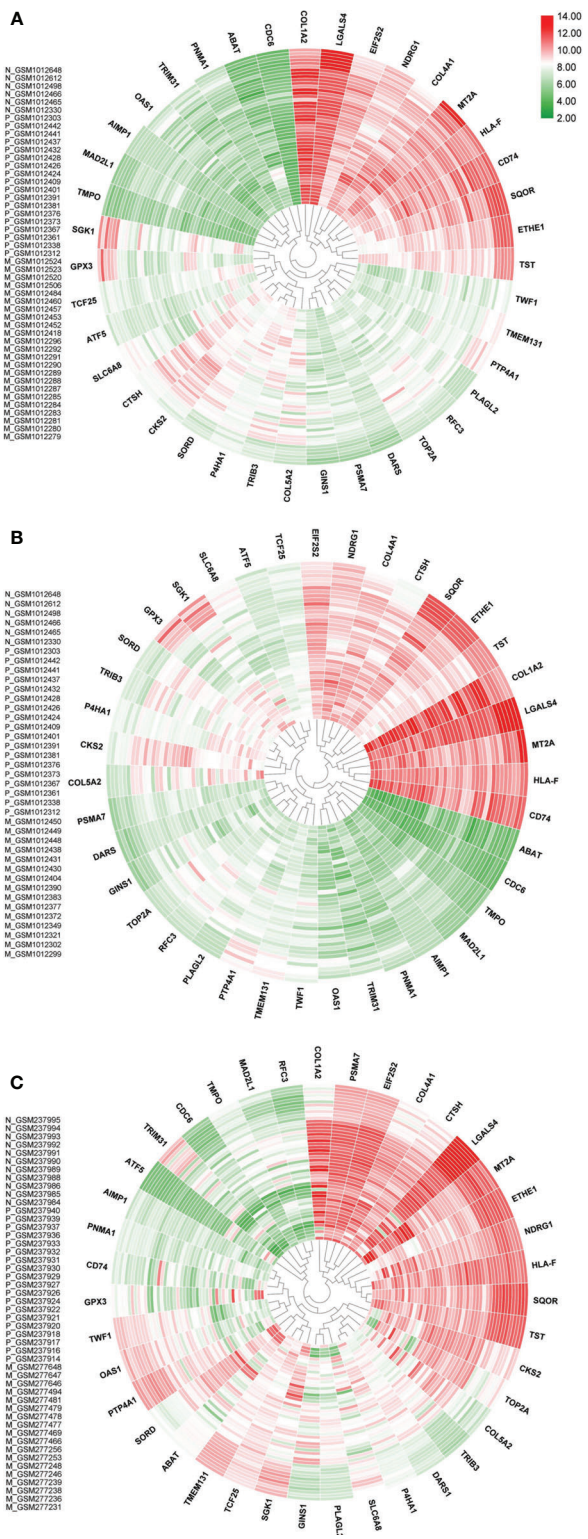

FIGURE 5 | Core network genes heatmaps. **(A)** is dataset1, **(B)** in dataset2 and **(C)** is dataset3. Values were obtained from expression matrices, Log2 transformed and normalized by quantile method. Sample IDs are placed in the left-hand side of each plot. Normal samples start with N, Primary samples start with P and liver and lung metastatic samples start with M. Outer samples are normal, middle samples are primary and inner samples are metastatic. Genes were clustered together based on hierarchical clustering.

ECM stiffness suitable for cellular migration (47, 48). Genes that were enriched for sulfide oxidation terms formed a distinct component in the Core network as well. Genes in green are engaged in interferon-gamma signaling that has dual roles in cancer. On the one hand, INF-γ has anti-proliferative functions by employing different mechanisms such as induction of p21 (49), induction of autophagy (50), regulation of EGFR/Erk1/2 and Wnt/β-catenin signalings (51), and so on. On the other hand, it enhances the outgrowth of tumor cells with invasive properties depending on cellular and microenvironmental context (52, 53).

In the enrichment analysis with "Enrichr" online tool, Gastric Cancer Network2 was of the lowest p-value containing TOP2A and RFC3 genes involved in DNA replication process. Involvement of the same genes in retinoblastoma cancer (WP2446) proposes the potential importance of these genes in different cancers. Top2A was involved in Gastric Cancer Network1 as well. All the enrichment results yielded from "Enrichr" are presented in **Supplementary File 5**.

Expression Profiling and Survival Analysis of TCGA Gene Expression Profiles

Expression of DEGs in **Tables 1**, **2** were further supported by boxplots and survival plots using GEPIA2 web server. Expression profiles were attained from 275 colorectal adenocarcinoma and 41 normal colon RNA-seq samples in TCGA database to create boxplots for each gene in **Figure 7**. Except for TWF1, all plots were in agreement with our results. In other words, if a gene was upregulated in our analysis between cancer and normal groups, the expression median for that gene in tumor samples was larger than normal samples in boxplots and vice versa. Even boxplots for expression of some genes that were later shown to be contradictory to other studies, were in favor of our findings. They were MT2A, TRIM31, CDC6, SGK1, and PTP4A1 genes.

Survival plots were also created for DEGs in **Tables 1**, **2** in different months using TCGA database. Only three genes had a significant p-value larger than 0.05 illustrated in **Figure 8**, and rest of survival plots were presented in **Supplementary File 6**. Low expression of LGALS4 is associated with poor survival rate, while high expression of COL1A2 and the new reported gene DARS is linked to poor survival rate in colorectal cancer patients. In our study, LGALS4 were downregulated in MvsN and PvsN analyses of all Test and Validation datasets while DARS and COL1A2 were upregulated in majority of MvsN and PvsN analyses. Although other survival tests were non-significant, majority of them were in agreement with our expression results.

TABLE 2 | Illustration of Core network DEGs in Validation datasets.

MvsN							PvsN					
DEGs	**GSE40367**	**GSE50760**	**GSE41258**		**GSE144259**	**GSE89393**	**DEGs**	**TCGA**	**GSE50760**	**GSE41258**	**GSE144259**	**GSE89393**
	liver-normal	**liver-normal**	**liver-normal**	**lung-normal**	**liver-normal**	**liver-normal**		**primary-normal**				
ETHE1	Down	Down	Down	Down	Down	Down	SGK1	Down	Down	Down	Down	Down
DARS	Up	Up	Up	Up	Up	Up	EIF2S2	Up	Up	Up	Up	Up
TMEM131	Down	Down	Down	Down	Down	Down	TRIB3	Up	Up	Up	Up	Up
TST	Down	Down	Down	Down	Down	Down	DARS	Up	Up	Up	Up	Up
LGALS4	Down	Down	Down	Down	Down	Down	RFC3	Up	Up	Up	Up	Up
TRIB3	Up	Up	Up	Up	Up	Up	ETHE1	Down	Down	Down	Down	Down
COL5A2	Down	Up	Up	Up	Up	Down	TOP2A	Up	Up	Up	Up	Up
COL4A1	Down	Up	Up	Up	Up	Down	CKS2	Up	Up	Up	Up	Up
PTP4A1	Down	Down	Down	Down	Down	Down	SORD	Up	Up	Up	Up	Up
COL1A2	Up	Up	Up	Up	Up	Up	PSMA7	Up	Up	Up	Up	Up
SORD	Up	Up	Up	Up	Up	Up	GPX3	Down	Down	Down	Down	Down
SQOR	Down	Down	Down	Down	Down	Down	MAD2L1	Up	Up	Up	Up	Up
OAS1	Down	Down	…	Down	Down	Down	SQOR	Down	Down	Down	Down	Down
TRIM31	Down	Down	Down	Down	Down	Down	TST	Down	Down	Down	Down	Down
TWF1	Down	Down	…	Down	…	Down	GINS1	Up	Up	Up	Up	Up
MvsP							LGALS4	…	Down	Down	Down	Down
DEGs	GSE40367	GSE50760	GSE41258		GSE144259	GSE89393	PTP4A1	Up	Down	Down	Down	Down
	liver-primary	liver-primary	liver-primary	lung-primary	liver-primary	liver-primary	PLAGL2	…	Up	Up	Up	Up
ETHE1	Down	Down	…	Down	Down	Down	COL1A2	Down	Up	Up	Up	Up
ATF5	…	Up	Up	…	Up	…	CTSH	…	Up	Up	Up	Up
CDC6	Down	Up	Down	Down	Down	Down	MT2A	Down	Down	Down	Down	Down
TMPO	Down	…	…	Down	Down	Down	COL5A2	Down	Up	Up	Up	Up
P4HA1	Up	Up	Up	Up	Up	Up						

Up means gene was upregulated and Down means gene was Downregulated. MvsN contains all the metastatic versus normal analyses, PvsN contains all the primary versus normal analyses and MvsP contains all the metastatic versus primary analyses. Some genes are present in more than one analysis. The expression status for the genes in green rows are similar in all datasets regardless of empty cells. In the yellow rows only one dataset is different from the others and in the white rows genes exhibited a heterogeneous expression status in different datasets. Green rows illustrate a similar expression in all datasets.

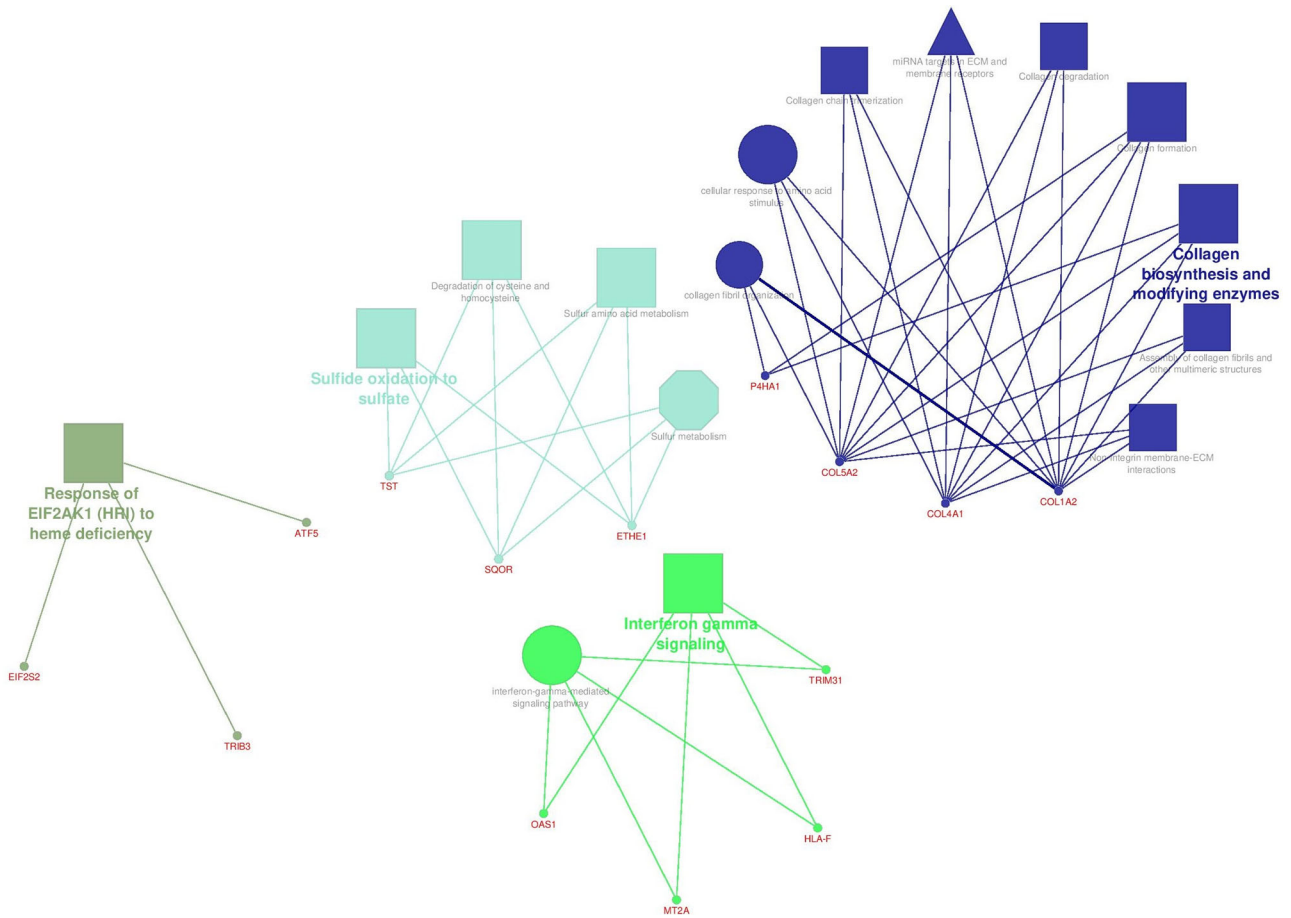

FIGURE 6 | The enrichment results for the Core network genes. Terms in the shape of octagons are from KEGG, Triangular terms are from WikiPathways, rectangular expressions are from Reactome and circular terms are from GO database. Size of the terms present their significance.

DISCUSSION

The Core network giant component is composed of an up and a down part attached *via* PLAGL2 transcription factor (TF). The lower part is engaged mainly in cell cycle and DNA replication. Components 2 and 3 contain genes involved in ECM remodeling, component 4 is composed of genes involved in transcription inhibition, and Component 5 is composed of mitochondrial genes playing essential roles in controlling cellular redox homeostasis. Here we discussed most of the genes in the Core network exhibiting more similar expression patterns which were present in **Tables 1**, **2**.

PLAGL2 is considered an oncogene in different cancers. It binds to and prevents Pirh2 proteasomal degradation, which in turn Pirh2 promotes proteasomal degradation of P53 protein (54). In glioblastoma, PLAGL2 suppresses neural stem cell differentiation by regulating Wnt/β-catenin signaling (55). Besides, PLAGL2 regulates actin cytoskeleton and cell migration through promoting actin stress fibers and focal adhesion (56). Results of PvsN analysis manifests that this gene is induced in primary tumors in colon cancer. In addition, this gene had a high betweenness centrality in the giant component (S4). Since this gene connected the two parts of the giant component, it would be a pertinent target for disturbing colon cancer network. Its induction in CRC was supported by the majority of Validation datasets in **Table 2**.

TRIM31 (a ubiquitin ligase) was downregulated in MvsN in all Test and Validation datasets. However, there are contradictory results in different studies where it was shown to be reduced in lung cancer cells (57) and stepped up in gastric (58) and colorectal cancer (59). Therefore, its downregulation in nine analyses in **Tables 1**, **2** needs to be further explored. MT2A gene is an antioxidant that protects cells against hydroxyl radicals and plays an important role in detoxifying heavy metals (60, 61). Expression inhibition of this gene results in proliferation inhibition of CRC cells (62), and its silencing promotes the effects of chemotherapy drugs on primary osteosarcoma tumors (63). However, MT2A gene expression was downregulated in PvsN analyses supported by the results in **Table 2**. Likewise, it is downregulated in pancreatic cancer as well (64). Therefore, this downregulation in primary CRC tumors has to go under more investigation. OAS1 is a protein induced by interferons that

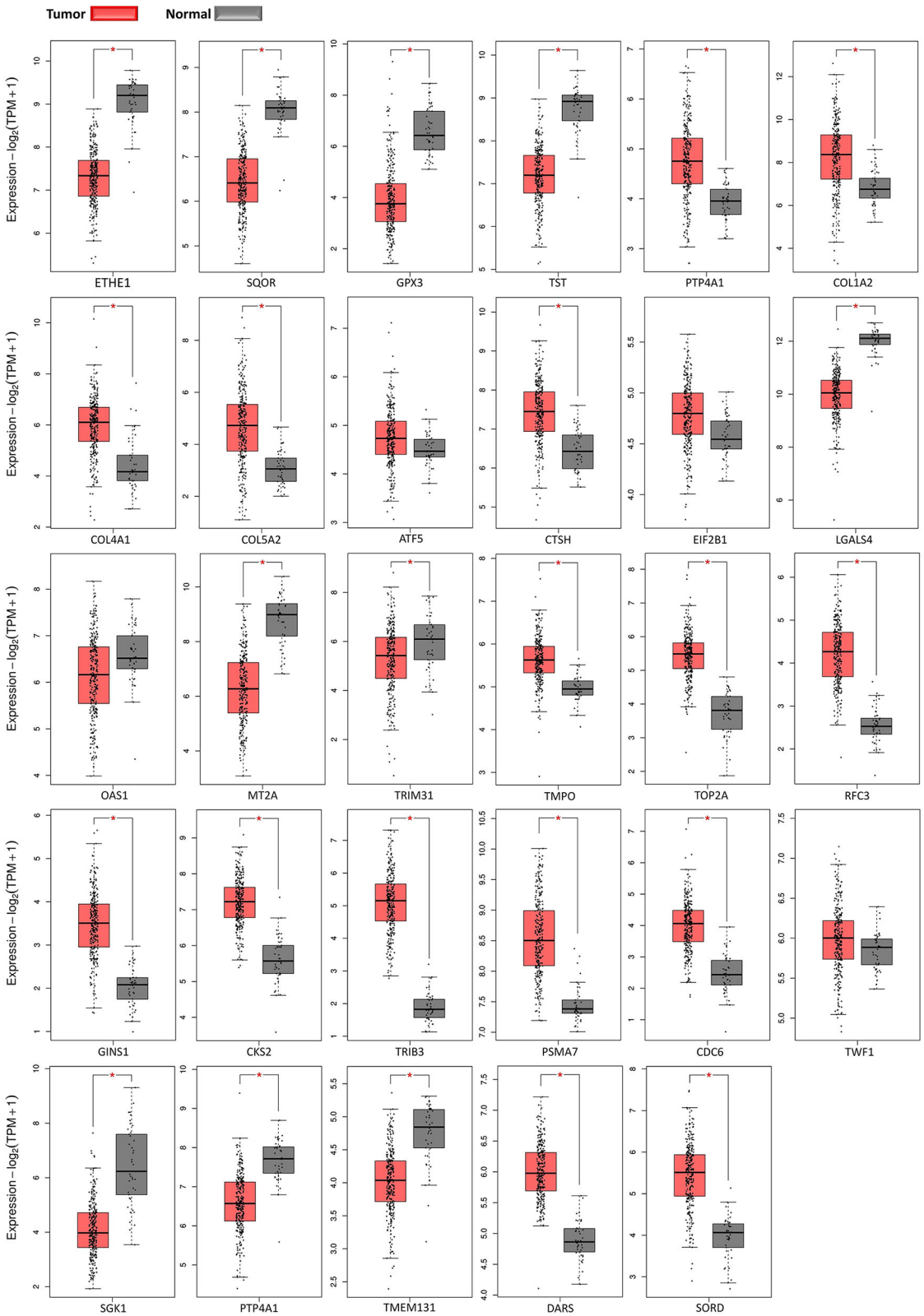

FIGURE 7 | Gene expression profiles boxplots. Red boxes present normalized expression values for tumor adenocarcinoma samples and gray boxes present values for normal colon samples.

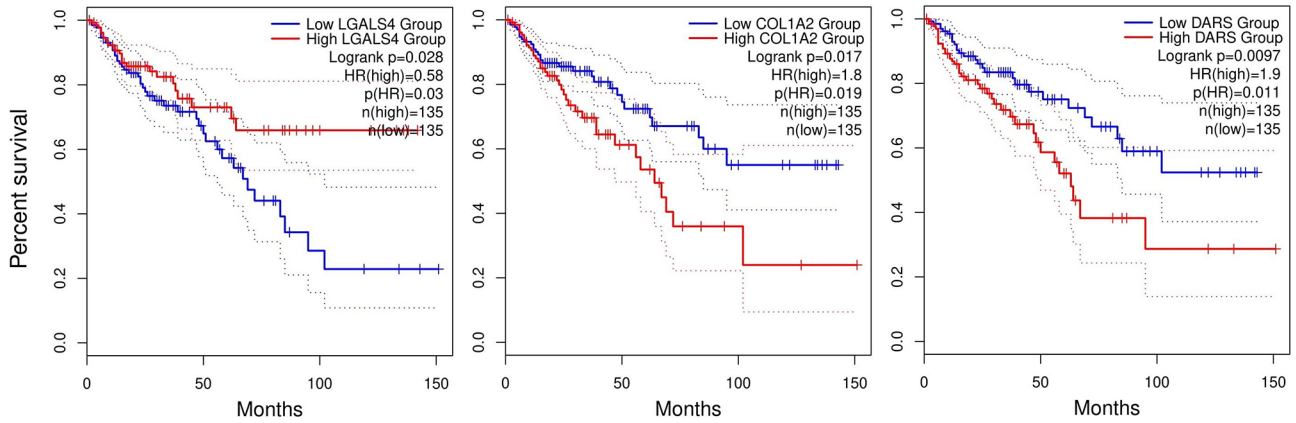

FIGURE 8 | Survival plots. Plots present the monthly survival rate of patients having high expression, red line, or low expression, blue line, of a specific gene. Patients having high expression of LGALS4 had higher survival rates compared to patients having low expression of LGALS4. Contrary, patients having low expression of COL1A2 and DARS genes had higher survival rate.

synthesizes the oligomers of adenosine from ATP. These oligomers bind to RNase L to regulate cell growth, differentiation, and apoptosis (65). Its expression is downregulated in breast ductal carcinoma and prostate cancer (PCa) at both mRNA and protein levels. In addition, OAS1 expression is negatively correlated with the progression of these cancers (65). The given information supports the downregulation of this gene in our analysis supported by Validation datasets. Consequently, expression induction of this gene might help prevent both tumor growth and cell differentiation. The mentioned three genes, TRIM31, MT2A and OAS1, were enriched for IFN-γ and all were downregulated. Although there are contradictory results in different papers, these downregulations at mRNA level would help tumor cells to defeat the anti-cancer properties of interferon gamma signaling.

CTSH gene is a lysosomal cysteine protease upregulated in PvsN. This protease plays an important role in cytoskeletal protein Talin maturation. Talin promotes integrin activation and focal adhesion formation leading to cell migration (66). Validation datasets more verified upregulation of this gene in CRC. As a result, suppression of CTSH expression could be a choice of metastasis inhibition. Glutathione peroxidase 3 (GPX3) is an important antioxidant enzyme that protects cells against Reactive Oxygen Species (ROS), downregulated in many cancers. For instance, its expression is suppressed in human colon carcinoma Caco2 cell lines, resulting in augmented ROS production (67). It reduces H_2O_2 and lipid peroxides to water and lipid alcohols, respectively, and in turn, oxidizes glutathione to glutathione disulfide (68). Downregulation of GPX3 happened in PvsN analyses, leading to ascending of H_2O_2 level, which is positively correlated with tumor progression (69). Its downregulation was further supported by all datasets in **Table 2**. As a result, induction of GPX gene families would be a therapeutic approach.

TMPO gene had the greatest Betweenness centrality illustrating a reduced expression trend in MvsP analyses supported by Validation datasets. This gene produces different protein isoforms *via* alternative splicing (70, 71). The proteins are located in the nucleus of the cells, which help form nuclear lamina and maintain nucleus membrane structure (72). TMPO prevents the depolymerization of nuclear laminas and excessive activation of the mitosis pathway. Therefore, its downregulation would prevent an excessive mitotic cycle.

TMEM131 is a transmembrane protein that was downregulated in MvsN analyses in all datasets in **Tables 1**, **2**. No documentation was found to connect this gene to a specific cancer. Therefore, this gene might be biomarker of CRC. Furthermore, Enrichment analysis using "Enrichr" online tool showed that this gene was also involved in interferon-gamma signaling (S5). A recent study has discovered that amino termini of human TMEM131 recruit monomers of collagens for assembly. Carboxy termini of this gene guide collagen cargo machinery from Endoplasmic Reticulum towards Golgi apparatus, contributing to collagen maturation and secretion. Moreover, TMEM131 deficiency diminishes collagen production, maturation, and secretion in Caenorhabditis Elegans (73). This gene is also important for ER processing of cuticle collagen cargos and apical ECM (aECM) formation in Drosophila melanogaster. These findings highlight the conserved role of this gene in collogen biosynthesis (74). The methylation rate of this gene is reduced in T-Cells and peripheral blood cells in Down syndrome patients (Trisomy 21). This gene also marks lymphocyte precursor cells for lineage specification (75).

TOP2A gene was upregulated in PvsN analyses entirely endorsed by the validation results. In breast cancer (BC) HER-2 and TOP2A are the molecular targets for several anticancer medicines that are bolstered together (76). Moreover, Copy Number Variations (CNVs) in TOP2A gene have been identified as biomarkers of colorectal cancer (77). This enzyme controls DNA topological structure, and its upregulation is a hallmark of aberrant cell growth in tumors (78). TOP2A mRNA expression is an independent prognostic factor in patients with (Estrogen Receptor) ER-positive breast cancer and could be useful in the assessment of breast cancer risk (79).

Therefore, in addition to being a possible target for CRC therapy, this gene could be either a possible prognostic or diagnostic marker of CRC.

Replication Factor C subunit 3 (RFC3) plays a role in DNA replication, DNA damage repair, and cell cycle checkpoint control. Hepatocellular carcinoma (HCC) and cell proliferation of ovarian tumors are suppressed by shRNA-mediated silencing of RFC3 gene (80, 81). This gene was upregulated in PvsN analyses and is upregulated in Triple-negative breast cancer (TNBC) as well (82). Validation datasets more supported its upregulation. Since expression inhibition of this gene at both mRNA and protein levels suppresses the migratory and invasive ability of MCF-7 cell lines (83), this gene would be a therapeutic target for colorectal cancer treatment. Moreover, TOP2A and RFC3 were shown to be engaged in the Gastric Cancer Network2 pathway in the enrichment analysis by "Enrichr" (S5), indicating the importance of these two genes in cancer progression.

Mitotic Arrest Deficient 2 Like1 (MAD2L1) is a mitotic spindle assembly checkpoint molecule upregulated in PvsN in both Test and Validation analyses. It is responsible for preventing anaphase initiation until precise and complete metaphase alignment of all chromosomes takes place. An increase in the level of MAD2L1 transcripts is detected in a large number of samples with ductal breast carcinoma (84). Its upregulation in our analysis would provide evidence that cancerous cells were dealing with mitotic deficiencies. The GINS complex is a DNA replication machinery component in the eukaryotes and is an essential tool for initiating and progressing DNA replication forks (85). GINS1 (PSF1) mRNA level is positively correlated with tumor size in CRC patients and is a prognostic marker of CRC (86). This gene has been recently introduced as a targeted oncogenic agent for inhibition of synovial sarcoma (87). It was totally upregulated in PvsN analyses in **Tables 1**, **2**. Therefore, its expression inhibition would be a potential target for inhibition of tumor growth by disturbing DNA replication machinery.

CDC6, one of the Core genes, plays a critical role in regulation of the eukaryotic DNA replication onset, and its downregulation has been demonstrated in prostate cancer (88). It is a regulator of cell cycle in S phase, and its expression is regulated by E2F Transcription factor and androgen receptors (AR) in PCa cells (89). Transfection of CDC6 siRNA leads to not only decreased level of ovarian cancer cell proliferation but also increased apoptosis rates (90). Cdc6 and Cdt1 are highly expressed in aggressive BC and therefore is considered a potent therapeutic target in BC patients (91). Results for this gene in MvsP analyses were contradictory to the BC results, but it is similar to prostate cancer. The majority of Validation datasets depicted downregulation of this gene in CRC. No study directly measured the expression level of this gene in CRC samples; therefore, it is worth investigating to see whether it could be a CRC biomarker or a curative target.

CKS2 protein interacts with the catalytic subunit of the cyclin-dependent kinases, and its downregulation contributes to suppression of p-Akt and p-mTOR. Therefore, one of CSK2 oncogenic roles is played by Akt/mTOR oncogenic pathway (92). CKS2 is expressed at a high level in CRC tissues, and it has revealed that increased CKS2 expression is highly correlated with enhanced metastatic stage (93). Importantly, CKS2 is considered a potential biomarker and therapeutic target for the BC treatment due to the fact that its inhibition suppresses cell proliferation and invasion *in vitro* and *in vivo* (94). In the PvN analyses, this gene was upregulated, which would be a therapeutic target for CRC treatment because validation results completely supported this upregulation.

PSMA7 gene encodes a protein that is one of the essential subunits of 20S proteasome complex (95). Overexpression of PSMA7 both at the mRNA and protein levels has been reported in gastric cancer (96). Depletion of PSMA7 by shRNA-transfected RKO CRC cell lines mediates inhibition of cell growth and migration. Consequently, inhibition of PSMA7 could be a beneficial therapeutic strategy for colorectal cancer patients (97). This gene was upregulated in PvsN analyses in test and Validation datasets.

DARS encodes the cytosolic aspartyl-tRNA synthetase found to be upregulated in MvsN and PvsN analyses in all Test and Validation datasets (a total of 16 analyses). This gene encodes a member of a multi-enzyme complex that its role has been proved in mediating attachment of amino acids to their cognate tRNAs. Some studies have reported that DARS-AS1 gene (encoding a long noncoding RNA) act as an oncogene (98) and is positively associated with the pathological stages in thyroid and ovarian cancer by targeting mir-129 and mir-532-3p, respectively (99, 100). Moreover, this gene is directly upregulated by HIF1 gene, which stabilizes RBM39 protein in Myeloma (101). Mutations in this gene have been previously reported in neuroinflammatory diseases and Leukodystrophies (102, 103). However, there is not enough evidence in the literature that associates DARS1 (DARS) gene to different cancers. Moreover, patients having a lower expression of this gene have a higher survival rate in **Figure 8**.

EIF-2 consists of alpha, beta, and gamma subunits. EIF2B or EIF2S2 acts in the early steps of protein synthesis. GTP-bound EIF-2 transfers Met-tRNAi to the 40S ribosomal subunit to start protein synthesis. The hydrolysis of GTP to GDP takes place at the end of the initiation process that leads to release of the inactive eIF2·GDP from ribosome. Exchange of GDP for GTP is performed by beta subunit so that active EIF-2 is ready for another round of initiation (104). In one study, EIF2B was proposed as a potential therapeutic target in lung cancer (76). Moreover, elimination of natural killer cell cytotoxicity *via* promoted expression of natural killer (NK) cell ligands is done by pSer535-eIF2B following the expression of pSer9-GSK-3β (inactive GSK3β) and generation of ROS, which promotes breast cancer growth and metastasis (105). Since Tyr216-GSK-3β (Active GSK3β) has inhibitory effects on the EMT process by interfering with TNF-alpha signaling (106), induction of active GSK-3β together with suppression of EIF2B would be a therapeutic approach to prevent EMT (107). EIF2B stepped up in PvsN analyses which was supported by validation results.

TWF1 gene encodes Twinfilin, an actin monomer-binding protein that promotes EMT in pancreatic cancer tissues (108). TWF1 siRNA dramatically inhibits F-actin organization and focal adhesions formation, promoting the mesenchymal-to-epithelial

transition (MET) in MDA-MB-231 cell lines. Besides, The responsiveness of these cell lines to anti-cancer drugs such as doxorubicin and paclitaxel is augmented by siRNA inhibition of TWF1 expression (109). Furthermore, expression levels of EMT markers, VIM and SNAI2, are reduced due to miR-30c action on TWF1 mRNA (109). However, in MvsN analyses, this gene witnessed a decreased expression in both Test and Validation datasets. As a result, Its upregulation in CRC has to be further explored.

SGK1, a member of component 2, and AKT are two families of AGC protein superfamily. SGK1 is a serine/threonine kinase that activates particular potassium, sodium, and chloride channels (110). SGK1 is a downstream effector of PI3K, which runs pathways independent of pathways shared with AKT. The two kinases are phosphorylated and activated by PDK1 and mTORC2 complex (111, 112). In general, PI3K-dependent survival signals can be mediated by either Akt or SGK1 that inactivates the pro-apoptotic proteins Bad and FKHRL1 (113). A study on A498 kidney cancer cells found that survival signals promoted by IL-2 are mediated by SGK1 activation (114). Moreover, the promoter of SGK1 is under tight control of the p53 protein (115). SGK1 has been shown to mediate cell survival and drug resistance to platinoid and taxane compounds in breast cancer patients (116). On the contrary, this gene was totally downregulated in PvsN analyses in all Validation and Test datasets. These overall downregulations might be specific to CRC, so it could be a diagnostic hallmark of CRC and should go under more interrogation.

Component 3 contains collagen (COL1A2, COL5A2, and COL4A1) and P4HA1 (a collagen hydroxylase) genes interconnected in the process of ECM remodeling based on the enrichment results. All members witnessed an ascending trend in expression from normal samples to metastatic samples in **Figure 5** panels. In Test datasets, collagen genes presented an upregulation trend in MvsN and PvsN analyses, while their expression followed a mixed trend in Validation datasets. P4HA1 one of the Core genes upregulated in MvsP in all Test and Validation datasets. Expression of COL1A2 followed a homogeneous upregulating trend in both Test and Validation datasets which is a marker of EMT (117). P4HA1 is engaged in breast and pancreatic metastasis (118, 119). Under hypoxic tumor conditions, HIF-1 induces expression of genes that encodes collagen prolyl (P4HA1 and P4HA2) and lysyl (PLOD2) hydroxylases. P4HA1 and P4HA2 are required for collagen deposition, whereas PLOD2 is required for ECM stiffening and collagen fiber alignment (120). These changes in ECM triggered by HIF-1 are necessary for motility and invasion because, in focal adhesion junctions, actin cytoskeleton is connected to ECM through attachment of integrins to collagens (121). Besides, there is a positive feedback between P4HA1 and HIF-1 in modulation of ECM. As a result, targeting P4HA1 and P4HA2 expressions would inhibit the progression of cell migration *via* HIF1-Collagen pathway.

PTP4A1 a member of component 4, is a protein phosphatase engaged in p21-activated kinase (PAK) signaling pathway. Inhibition of PTP4A1 gene in MDA-MB-231 breast cancer cell lines by an increase in miR-944 expression impairs cell invasion (122). However, this gene was downregulated in MvsN and PvsN in all Test datasets and most Validation datasets. This downregulation would be a biomarker for CRC, and its molecular role in CRC needs to be interrogated. BCL-2 is a target of ATF5, one of the Core genes (123). ATF5 was upregulated in MvsP analyses in Test and Validation datasets. There are pieces of evidence that link the role of ATF5 in mitochondrial dysfunction in cancer progression (124). In malignant glioma, metastatic cells take advantage of survival signals triggered by ATF5 gene, which is essential to ignore anchorage-dependent and niche-dependent cell death signals (125). Thus, expression inhibition of ATF5 would hinder the survival signals in CRC cells. TRIB3 is a prognosis hallmark of colorectal cancer, activated under hypoxic conditions (126). TRIB3 silencing suppresses VEGF−A expression in gastric cancer cells inhibiting endothelial cell migration and vessel formation. This gene was upregulated in MvsN analyses in all Test and Validation datasets. Therefore, it would be a promising target for anti−angiogenic therapy (127).

Genes in component 5 are mitochondrial which their role in cancer progression has not been sufficiently investigated so far. All three genes were downregulated in our analysis in both Validation and Test datasets. They also exhibited a reducing trend from normal to primary and from primary to metastatic in **Figure 5** panels. These genes are highly expressed in normal colon tissue compared to other tissues due to the presence of anaerobic bacteria in the digestive tract (128). These findings are supported by the RNA-seq expression information in the Gene database of NCBI (129). ETHE1 (persulfide dioxygenase) and SQOR are antioxidants that convert hydrogen sulfide (H2S) to persulfide, then to sulfite. Hence, they protect cells against toxic concentrations of sulfide. ETHE1 gene was downregulated in the three analyses while SQOR was downregulated in MvsN and PvsN analyses. All these expressions were totally verified by the Validation datasets. Their downregulation is essential for cancer cells proliferation and survival. Under the hypoxic environment of CRC tumor, sulfide is a supplementary tool that provides electron for mitochondrial electron transport chain (ETC) to generate ATP (130). These mechanisms, along with Warburg effect help tumor cells to survive from the hypoxic environment. As a result, helping expression induction or activation of ETHE1 and SQOR proteins will increase sulfide scavenging and this would hinder CRC tumor growth. TST thiosulfate sulfurtransferase encodes a protein that is localized to the mitochondria and catalyzes the conversion of thiosulfate and cyanide to thiocyanate and sulfite respectively. Therefore, like the previous two mitochondrial enzymes, it acts in Hydrogen sulfide metabolism (131).

SORD (Sorbitol dehydrogenase) is another element of component 6 upregulated in MvsN, and PvsN analyses. Little is known about the connection between SORD and cancer. This enzyme hydrogenates Fructose to Sorbitol in Fructose catabolism pathway. Subsequently, sorbitol is dehydrogenated to glucose *via* AKR1B1 enzyme providing fuel for cells (132). In addition, excess glucose promotes EMT through autocrine TGFβ stimulation (133). Expression suppression of either enzyme reduces EMT in human lung cancer cells and EMT-driven colon cancer mouse model (133). Two studies demonstrated that SORD is an androgen-regulated gene in prostate cancer (134, 135).

siRNA inhibition of this gene leads to proliferation and migration inhibition of A549 lung cancer cells (136). Since SORD exhibited an ascending trend in all Validation and Test datasets in **Figure 5**, it might be a potential target and biomarker to prevent EMT and cell growth in CRC. LGALS4 is implicated in regulating cell-cell and cell-matrix interactions, so its induction might have positive curative impacts on CRC cells. This gene is primarily expressed in small intestine, colon, and rectum, which is suppressed in CRC (137). It was downregulated in MvsN and PvsN analyses in Validation and Test datasets. It is also a blood marker of CRC (138).

In summary, we illustrated some therapeutic targets and biomarkers for CRC. A combination of these targets would beneficially disturb progression of colorectal cancer. Generally, the discovered antioxidants were downregulated in different stages of CRC, namely ETHE1, SQOR, TST, and GPX3. We proposed that these downregulations under hypoxic conditions would help cancer cells to produce more energy for cell proliferation. In addition, the hypoxic environment alters ECM suitable for cell migration by induction of P4HA1 gene through HIF-1 signaling pathway and induction of COL1A2. Boxplots (expression profiling) in **Figure 7** supported our results for all these genes. In addition, survival plot in **Figure 8** demonstrated that there is a higher death probability for CRC patients expressing a high level of COL1A2 than patients having a low level of this gene. Consequently, colorectal cancer cells would take advantages of explained mechanisms along with Warburg effect to not only survive from the hypoxic environment of tumors but also proliferate faster and migrate better. Therefore, induction of mentioned antioxidants and suppression of P4HA1 and COL1A2 genes would be a choice of CRC treatment.

Induction of active GSK-3β together with suppression of EIF2B would prevent EMT in CRC. Induction of OAS1 to increase the anti-cancer effects of interferon gamma, suppression of CTSH to hinder formation of focal adhesions, expression inhibition of ATF5 gene to make cancer cells sensitive to anchorage-dependent death signals, and induction of LGALS4 gene (supported by survival analysis) to recover cell-cell junctions would be the combination of genetic targets that prevent EMT and cell migration. In addition, expression inhibition of TMPO, TOP2A, RFC3, GINS1, and CKS2 genes could prevent tumor growth and TRIB3 expression suppression would be a favorable target for anti–angiogenic therapy. PSMA7 gene was a previously reported target for CRC treatment that was also found in our study. Results for expression of all these genes were supported by expression profiling.

MT2A and TRIM31 which were engaged in IFN-γ signaling, CDC6, SGK1 and PTP4A1 genes, presented a homogeneous expression pattern in both test and Validation datasets, although our results were contradictory to other studies in different cancers. Nevertheless, we used 10 different datasets from different technologies to ensure the accuracy of the results. Besides, expression profiling supported expression of these genes. However, they have to be further interrogated in colorectal cancer progression.

TMEM131 and DARS genes had specific uniform expression trends as analyses went from normal to metastatic. DARS expression inhibition would increase the survival rate in CRC patients based on **Figure 8**. Therefore, this gene might be a CRC prognostic marker or a curative target. Downregulation of TMEM131 might be associated with the amount of collagen secretion in ECM to make the environment suitable for migration (Stiffness of ECM). SORD is a poorly studied gene in cancer that its expression reduction might prevent cell proliferation and EMT in CRC. The relation of these three genes to colorectal cancer progression has been reported for the first time in this study. More investigation is required to find their molecular mechanism causing colorectal cancer promotion.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

MP: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, writing – Original Draft Preparation, Writing – Review and Editing. NS: Formal Analysis, Review and Editing. AM: Review and Editing. AR: Review and Editing. AG: Review and Editing. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

This manuscript has been released as a pre-print at bioRxiv, Piran, M,. Piran, M,. Sepahi, N,. Ghanbariasad, A,. Rahimi, A,. Systems biomedicine of primary and metastatic colorectal cancer reveals potential therapeutic targets [Preprint] (2020). Available at: https://www.biorxiv.org/content/10.1101/2020.01.25.919415v9 (Accessed August 24, 2020).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.597536/full#supplementary-material

Supplementary File 1 | Contains the information about all samples in all datasets.

Supplementary File 2 | Contains the quality plots.

Supplementary File 3 | Contains all DEGs between all pairwise groups of samples.

Supplementary File 4 | Contains network centralities.

Supplementary File 5 | Contains enrichment information.

Supplementary File 6 | Contains survival plots.

REFERENCES

1. Favoriti P, Carbone G, Greco M, Pirozzi F, Pirozzi REM, Corcione F. Worldwide Burden of Colorectal Cancer: A Review. *Updates Surg* (2016) 68 (1):7–11. doi: 10.1007/s13304-016-0359-y
2. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Patterns and Trends in Colorectal Cancer Incidence and Mortality. *Gut* (2017) 66(4):683–91. doi: 10.1136/gutjnl-2015-310912
3. Valkenburg KC, de Groot AE, Pienta KJ. Targeting the Tumour Stroma to Improve Cancer Therapy. *Nat Rev Clin Oncol* (2018) 15(6):366–81. doi: 10.1038/s41571-018-0007-1
4. Lamouille S, Xu J, Derynck R. Molecular Mechanisms of Epithelial–Mesenchymal Transition. *Nat Rev Mol Cell Biol* (2014) 15(3):178–96. doi: 10.1038/nrm3758
5. Al-Sohaily S, Biankin A, Leong R, Kohonen-Corish M, Warusavitarne J. Molecular Pathways in Colorectal Cancer. *J Gastroenterol Hepatol* (2012) 27 (9):1423–31. doi: 10.1111/j.1440-1746.2012.07200.x
6. Sierra JR, Cepero V, Giordano S. Molecular Mechanisms of Acquired Resistance to Tyrosine Kinase Targeted Therapy. *Mol Cancer* (2010) 9 (1):75. doi: 10.1186/1476-4598-9-75
7. Wang S-W, Sun Y-M. The IL-6/JAK/STAT3 Pathway: Potential Therapeutic Strategies in Treating Colorectal Cancer. *Int J Oncol* (2014) 44(4):1032–40. doi: 10.3892/ijo.2014.2259
8. Dihlmann S, von Knebel Doeberitz M. Wnt/β-Catenin-Pathway as a Molecular Target for Future Anti-Cancer Therapeutics. *Int J Cancer* (2005) 113(4):515–24. doi: 10.1002/ijc.20609
9. Vilchez V, Turcios L, Marti F, Gedaly R. Targeting Wnt/β-Catenin Pathway in Hepatocellular Carcinoma Treatment. *World J Gastroenterol* (2016) 22 (2):823. doi: 10.3748/wjg.v22.i2.823
10. Liu D, Chen L, Zhao H, Vaziri ND, Ma S-C, Zhao Y-Y. Small Molecules From Natural Products Targeting the Wnt/β-Catenin Pathway as a Therapeutic Strategy. *Biomed Pharmacother* (2019) 117:108990. doi: 10.1016/j.biopha.2019.108990
11. Katz LH, Li Y, Chen J-S, Muñoz NM, Majumdar A, Chen J, et al. Targeting TGF-β Signaling in Cancer. *Expert Opin Ther Targets* (2013) 17(7):743–60. doi: 10.1517/14728222.2013.782287
12. Pickup M, Novitskiy S, Moses HL. The Roles of Tgfβ in the Tumour Microenvironment. *Nat Rev Cancer* (2013) 13(11):788–99. doi: 10.1038/nrc3603
13. Cathcart J, Pulkoski-Gross A, Cao J. Targeting Matrix Metalloproteinases in Cancer: Bringing New Life to Old Ideas. *Genes Dis* (2015) 2(1):26–34. doi: 10.1016/j.gendis.2014.12.002
14. Jabłońska-Trypuć A, Matejczyk M, Rosochacki S. Matrix Metalloproteinases (Mmps), The Main Extracellular Matrix (ECM) Enzymes in Collagen Degradation, as a Target for Anticancer Drugs. *J Enzyme Inhibit Med Chem* (2016) 31(sup1):177–83. doi: 10.3109/14756366.2016.1161620
15. Rychahou P, Bae Y, Reichel D, Zaytseva YY, Lee EY, Napier D, et al. Colorectal Cancer Lung Metastasis Treatment With Polymer–Drug Nanoparticles. *J Controlled Release* (2018) 275:85–91. doi: 10.1016/j.jconrel.2018.02.008
16. Rychahou P, Bae Y, Zaytseva Y, Lee EY, Weiss HL, Evers BM. Colorectal Cancer Pulmonary Metastasis Treatment With Lung-Selective Delivery of PAN-Class I PI3K Inhibitors. *AACR* (2016)76(14):16–20. doi: 10.1158/1538-7445.AM2016-390
17. Jonker DJ, O'Callaghan CJ, Karapetis CS, Zalcberg JR, Tu D, Au H-J, et al. Cetuximab for the Treatment of Colorectal Cancer. *New Engl J Med* (2007) 357(20):2040–8. doi: 10.1056/NEJMoa071834
18. Tol J, Koopman M, Cats A, Rodenburg CJ, Creemers GJ, Schrama JG, et al. Chemotherapy, Bevacizumab, and Cetuximab in Metastatic Colorectal Cancer. *New Engl J Med* (2009) 360(6):563–72. doi: 10.1056/NEJMoa0808268
19. Berger SI, Ma'ayan A, Iyengar R. Systems Pharmacology of Arrhythmias. *Sci Signaling* (2010) 3(118):ra30–ra. doi: 10.1126/scisignal.2000723
20. Sheffer M, Bacolod MD, Zuk O, Giardina SF, Pincas H, Barany F, et al. Association of Survival and Disease Progression With Chromosomal Instability: A Genomic Exploration of Colorectal Cancer. *Proc Natl Acad Sci* (2009) 106(17):7131–6. doi: 10.1073/pnas.0902232106
21. Hong Y, Downey T, Eu KW, Koh PK, Cheah PY. A 'Metastasis-Prone'Signature for Early-Stage Mismatch-Repair Proficient Sporadic Colorectal Cancer Patients and its Implications for Possible Therapeutics. *Clin Exp metastasis* (2010) 27(2):83–90. doi: 10.1007/s10585-010-9305-4
22. Pantaleo M, Astolfi A, Nannini M, Paterini P, Piazzi G, Ercolani G, et al. Gene Expression Profiling of Liver Metastases From Colorectal Cancer as Potential Basis for Treatment Choice. *Br J Cancer* (2008) 99(10):1729–34. doi: 10.1038/sj.bjc.6604681
23. Kim S-K, Kim S-Y, Kim J-H, Roh SA, Cho D-H, Kim YS, et al. A Nineteen Gene-Based Risk Score Classifier Predicts Prognosis of Colorectal Cancer Patients. *Mol Oncol* (2014) 8(8):1653–66. doi: 10.1016/j.molonc.2014.06.016
24. Ji Q, Zhou L, Sui H, Yang L, Wu X, Song Q, et al. Primary Tumors Release ITGBL1-Rich Extracellular Vesicles to Promote Distal Metastatic Tumor Growth Through Fibroblast-Niche Formation. *Nat Commun* (2020) 11 (1):1–18. doi: 10.1038/s41467-020-14869-x
25. Goryca K, Kulecka M, Paziewska A, Dabrowska M, Grzelak M, Skrzypczak M, et al. Exome Scale Map of Genetic Alterations Promoting Metastasis in Colorectal Cancer. *BMC Genet* (2018) 19(1):1–9. doi: 10.1186/s12863-018-0673-0
26. Roessler S, Lin G, Forgues M, Budhu A, Hoover S, Simpson RM, et al. Integrative Genomic and Transcriptomic Characterization of Matched Primary and Metastatic Liver and Colorectal Carcinoma. *Int J Biol Sci* (2015) 11(1):88. doi: 10.7150/ijbs.10583
27. Gautier L, Cope L, Bolstad BM, Irizarry RA. Affy—Analysis of Affymetrix Genechip Data at the Probe Level. *Bioinformatics* (2004) 20(3):307–15. doi: 10.1093/bioinformatics/btg405
28. Do JH, Choi D-K. Normalization of Microarray Data: Single-Labeled and Dual-Labeled Arrays. *Molecules Cells (Springer Sci Business Media BV).* (2006) 22(3):254–261
29. Ramasamy A, Mondry A, Holmes CC, Altman DG. Key Issues in Conducting a Meta-Analysis of Gene Expression Microarray Datasets. *PloS Med* (2008) 5(9):e184. doi: 10.1371/journal.pmed.0050184
30. Gentleman R, Carey V, Huber W, Hahne F. Genefilter: Methods for Filtering Genes From Microarray Experiments. *R Package version.* (2011) 1(0).
31. Smyth GK. Limma: Linear Models for Microarray Data. *Bioinf Comput Biol Solutions Ysing R Bioconductor: Springer* (2005) p. 397–420. doi: 10.1007/0-387-29362-0_23
32. Ferreira J, Zwinderman A. On the Benjamini–Hochberg Method. *Ann Stat* (2006) 34(4):1827–49. doi: 10.1214/009053606000000425
33. Love MI, Huber W, Anders S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data With Deseq2. *Genome Biol* (2014) 15(12):550. doi: 10.1186/s13059-014-0550-8
34. Luke DA. A User's Guide to Network Analysis in R. *Springer* (2015) 72. doi: 10.1007/978-3-319-23883-8
35. Langfelder P, Horvath S. WGCNA: An R Package for Weighted Correlation Network Analysis. *BMC Bioinf* (2008) 9(1):559. doi: 10.1186/1471-2105-9-559
36. N Jasika, N Alispahic, A Elma, K Ilvana, L Elma and N Nosovic eds. *Dijkstra's Shortest Path Algorithm Serial and Parallel Execution Performance Analysis. 2012 Proceedings of the 35th International Convention MIPRO*. United States IEEE (2012).
37. Azimzadeh Jamalkandi S, Mozhgani S-H, Gholami Pourbadie H, Mirzaie M, Noorbakhsh F, Vaziri B, et al. Systems Biomedicine of Rabies Delineates the Affected Signaling Pathways. *Front Microbiol* (2016) 7:1688. doi: 10.3389/fmicb.2016.01688
38. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. Cluego: A Cytoscape Plug-in to Decipher Functionally Grouped Gene Ontology and Pathway Annotation Networks. *Bioinformatics* (2009) 25 (8):1091–3. doi: 10.1093/bioinformatics/btp101
39. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* (2000) 28(1):27–30. doi: 10.1093/nar/28.1.27
40. 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* (2010) 39(suppl_1):D691–D7. doi: 10.1093/nar/gkq1018
41. Pico AR, Kelder T, Van Iersel MP, Hanspers K, Conklin BR, Evelo C. Wikipathways: Pathway Editing for the People. *PloS Biol* (2008) 6(7):e184. doi: 10.1371/journal.pbio.0060184
42. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: A Comprehensive Gene Set Enrichment Analysis Web Server 2016 Update. *Nucleic Acids Res* (2016) 44(W1):W90–W7. doi: 10.1093/nar/gkw377

43. Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: An Enhanced Web Server for Large-Scale Expression Profiling and Interactive Analysis. *Nucleic Acids Res* (2019) 47(W1):W556–W60. doi: 10.1093/nar/gkz430
44. Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, et al. The Cancer Genome Atlas Pan-Cancer Analysis Project. *Nat Genet* (2013) 45(10):1113. doi: 10.1038/ng.2764
45. Oldham MC, Horvath S, Konopka G, Iwamoto K, Langfelder P, Kato T, et al. Identification and Removal of Outlier Samples Supplement for": Functional Organization of the Transcriptome in Human Brain. *Nat Neurosci* (2013) 11 (11):1271-82. doi: 10.1038/ng.2764
46. Khanin R, Wit E. How Scale-Free are Biological Networks. *J Comput Biol* (2006) 13(3):810–8. doi: 10.1089/cmb.2006.13.810
47. Frantz C, Stewart KM, Weaver VM. The Extracellular Matrix at a Glance. *J Cell Sci* (2010) 123(24):4195–200. doi: 10.1242/jcs.023820
48. Huang S, Ingber DE. Cell Tension, Matrix Mechanics, and Cancer Development. *Cancer Cell* (2005) 8(3):175–6. doi: 10.1016/j.ccr.2005.08.009
49. Gooch JL, Herrera RE, Yee D. The Role of P21 in Interferon Gamma-Mediated Growth Inhibition of Human Breast Cancer Cells. *Cell Growth Diff* (2000) 11(6):335.
50. Li P, Du Q, Cao Z, Guo Z, Evankovich J, Yan W, et al. Interferon-Gamma Induces Autophagy With Growth Inhibition and Cell Death in Human Hepatocellular Carcinoma (HCC) Cells Through Interferon-Regulatory Factor-1 (IRF-1). *Cancer Lett* (2012) 314(2):213–22. doi: 10.1016/j.canlet.2011.09.031
51. Wang L, Wang Y, Song Z, Chu J, Qu X. Deficiency of Interferon-Gamma or Its Receptor Promotes Colorectal Cancer Development. *J Interferon Cytokine Res* (2015) 35(4):273–80. doi: 10.1089/jir.2014.0132
52. Zaidi MR, Merlino G. The Two Faces of Interferon-Γ in Cancer. *Clin Cancer Res* (2011) 17(19):6118–24. doi: 10.1158/1078-0432.CCR-11-0482
53. Mojic M, Takeda K, Hayakawa Y. The Dark Side of IFN-Γ: Its Role in Promoting Cancer Immunoevasion. *Int J Mol Sci* (2018) 19(1):89. doi: 10.3390/ijms19010089
54. Zheng G, Ning J, Yang Y-C. PLAGL2 Controls the Stability of Pirh2, an E3 Ubiquitin Ligase for P53. *Biochem Biophys Res Commun* (2007) 364(2):344–50. doi: 10.1016/j.bbrc.2007.10.003
55. Zheng H, Ying H, Wiedemeyer R, Yan H, Quayle SN, Ivanova EV, et al. PLAGL2 Regulates Wnt Signaling to Impede Differentiation in Neural Stem Cells and Gliomas. *Cancer Cell* (2010) 17(5):497–509. doi: 10.1016/j.ccr.2010.03.020
56. Sekiya R, Maeda M, Yuan H, Asano E, Hyodo T, Hasegawa H, et al. PLAGL2 Regulates Actin Cytoskeletal Architecture and Cell Migration. *Carcinogenesis* (2014) 35(9):1993–2001. doi: 10.1093/carcin/bgu081
57. Li H, Zhang Y, Zhang Y, Bai X, Peng Y, He P. TRIM31 Is Downregulated in non-Small Cell Lung Cancer and Serves as a Potential Tumor Suppressor. *Tumor Biol* (2014) 35(6):5747–52. doi: 10.1007/s13277-014-1763-x
58. Sugiura T. The Cellular Level of TRIM31, An RBCC Protein Overexpressed in Gastric Cancer, Is Regulated by Multiple Mechanisms Including the Ubiquitin—Proteasome System. *Cell Biol Int* (2011) 35(7):657–61. doi: 10.1042/CBI20100772
59. Wang H, Yao L, Gong Y, Zhang B. TRIM31 Regulates Chronic Inflammation *Via* NF-κb Signal Pathway to Promote Invasion and Metastasis in Colorectal Cancer. *Am J Trans Res* (2018) 10(4):1247.
60. Ma H, Su L, Yue H, Yin X, Zhao J, Zhang S, et al. HMBOX1 Interacts With MT2A to Regulate Autophagy and Apoptosis in Vascular Endothelial Cells. *Sci Rep* (2015) 5:15121. doi: 10.1038/srep15121
61. Liu Y, Liu H, Chen W, Yang T, Zhang W. EOLA1 Protects Lipopolysaccharide Induced IL-6 Production and Apoptosis by Regulation of MT2A in Human Umbilical Vein Endothelial Cells Molecular and Cellular Biochemistry. *Mol Cell Biochem* (2014) 395: (1-2):45–51. doi: 10.1007/s11010-014-2110-7
62. Xu Y, Jin H, Tan X, Liu X, Ding Y. Tea Polyphenol Inhibits Colorectal Cancer With Microsatellite Instability by Regulating the Expressions of HES1, JAG1, MT2A and MAFA. *J Chin Integr Med* (2010) 8(9):870–6. doi: 10.3736/jcim20100911
63. Mangelinck A, da Costa MEM, Stefanovska B, Bawa O, Polrot M, Gaspar N, et al. MT2A Is an Early Predictive Biomarker of Response to Chemotherapy and a Potential Therapeutic Target in Osteosarcoma. *Sci Rep* (2019) 9(1):1–12. doi: 10.1038/s41598-019-48846-2
64. Hui B, Xu Y, Zhao B, Ji H, Ma Z, Xu S, et al. Overexpressed Long Noncoding RNA TUG1 Affects the Cell Cycle, Proliferation, and Apoptosis of Pancreatic Cancer Partly Through Suppressing RND3 and MT2A. *OncoTargets Ther* (2019) 12:1043. doi: 10.2147/OTT.S188396
65. Maia CJ, Rocha SM, Socorro S, Schmitt F, Santos CR. Oligoadenylate Synthetase 1 (OAS1) Expression in Human Breast and Prostate Cancer Cases, and its Regulation by Sex Steroid Hormones. *Adv Modern Oncol Res* (2016) 2(2):97–110. doi: 10.18282/amor.v2.i1.70
66. Jevnikar Z, Rojnik M, Jamnik P, Doljak B, Fonović UP, Kos J. Cathepsin H Mediates the Processing of Talin and Regulates Migration of Prostate Cancer Cells. *J Biol Chem* (2013) 288(4):2201–9. doi: 10.1074/jbc.M112.436394
67. Barrett CW, Ning W, Chen X, Smith JJ, Washington MK, Hill KE, et al. Tumor Suppressor Function of the Plasma Glutathione Peroxidase Gpx3 in Colitis-Associated Carcinoma. *Cancer Res* (2013) 73(3):1245–55. doi: 10.1158/0008-5472.CAN-12-3150
68. Sarsour EH, Kumar MG, Chaudhuri L, Kalen AL, Goswami PC. Redox Control of the Cell Cycle in Health and Disease. *Antioxidants Redox Signaling* (2009) 11(12):2985–3011. doi: 10.1089/ars.2009.2513
69. JN Moloney and TG Cotter. eds. *ROS Signalling in the Biology of Cancer. Seminars in Cell & Developmental Biology*. United Kingdom: Elsevier (2018).
70. Holmer L, Worman H. Inner Nuclear Membrane Proteins: Functions and Targeting. *Cell Mol Life Sci CMLS* (2001) 58(12-13):1741–7. doi: 10.1007/PL00000813
71. Harris CA, Andryuk PJ, Cline S, Chan HK, Natarajan A, Siekierka JJ, et al. Three Distinct Human Thymopoietins are Derived From Alternatively Spliced Mrnas. *Proc Natl Acad Sci* (1994) 91(14):6283–7. doi: 10.1073/pnas.91.14.6283
72. Huang W, Su X, Yan W, Kong Z, Wang D, Huang Y, et al. Overexpression of AR-Regulated Lncrna TMPO-AS1 Correlates With Tumor Progression and Poor Prognosis in Prostate Cancer. *Prostate* (2018) 78(16):1248–61. doi: 10.1002/pros.23700
73. Zhang Z, Bai M, Barbosa GO, Chen A, Wei Y, Luo S, et al. Broadly Conserved Roles of TMEM131 Family Proteins in Intracellular Collagen Assembly and Secretory Cargo Trafficking. *Sci Adv* (2020) 6(7):eaay7667. doi: 10.1126/sciadv.aay7667
74. Zheng SL, Adams JG, Chisholm AD. Form and Function of the Apical Extracellular Matrix: New Insights From Caenorhabditis Elegans, Drosophila Melanogaster, and the Vertebrate Inner Ear. *Faculty Rev* (2020) 9. doi: 10.12703/r/9-27
75. Do C, Xing Z, Yu YE, Tycko B. Trans-Acting Epigenetic Effects of Chromosomal Aneuploidies: Lessons From Down Syndrome and Mouse Models. *Epigenomics* (2017) 9(2):189–207. doi: 10.2217/epi-2016-0138
76. Simon R, Atefy R, Wagner U, Forster T, Fijan A, Bruderer J, et al. HER-2 and TOP2A Coamplification in Urinary Bladder Cancer. *Int J Cancer* (2003) 107 (5):764–72. doi: 10.1002/ijc.11477
77. Vang Nielsen K, Ejlertsen B, Møller S, Trøst Jørgensen J, Knoop A, Knudsen H, et al. The Value of TOP2A Gene Copy Number Variation as a Biomarker in Breast Cancer: Update of DBCG Trial 89D. *Acta Oncol* (2008) 47(4):725–34. doi: 10.1080/02841860801995396
78. De Resende MF, Vieira S, Chinen LTD, Chiappelli F, da Fonseca FP, Guimarães GC, et al. Prognostication of Prostate Cancer Based on TOP2A Protein and Gene Assessment: TOP2A in Prostate Cancer. *J Trans Med* (2013) 11(1):36. doi: 10.1186/1479-5876-11-36
79. Rody A, Karn T, Ruckhäberle E, Müller V, Gehrmann M, Solbach C, et al. Gene Expression of Topoisomerase II Alpha (TOP2A) by Microarray Analysis is Highly Prognostic in Estrogen Receptor (ER) Positive Breast Cancer. *Breast Cancer Res Treat* (2009) 113(3):457–66. doi: 10.1007/s10549-008-9964-x
80. Shen H, Xu J, Zhao S, Shi H, Yao S, Jiang N. ShRNA-Mediated Silencing of the RFC3 Gene Suppress Ovarian Tumor Cells Proliferation. *Int J Clin Exp Pathol* (2015) 8(8):8968.
81. Yao Z, Hu K, Huang H, Xu S, Wang Q, Zhang P, et al. ShRNA-Mediated Silencing of the RFC3 Gene Suppresses Hepatocellular Carcinoma Cell Proliferation. *Int J Mol Med* (2015) 36(5):1393–9. doi: 10.3892/ijmm.2015.2350
82. He Z-Y, Wu S-G, Peng F, Zhang Q, Luo Y, Chen M, et al. Up-Regulation of RFC3 Promotes Triple Negative Breast Cancer Metastasis and Is Associated With Poor Prognosis *Via* EMT. *Trans Oncol* (2017) 10(1):1–9. doi: 10.1016/j.tranon.2016.10.004
83. Zhou J, Zhang W-W, Peng F, Sun J-Y, He Z-Y, Wu S-G. Downregulation of Hsa_Circ_0011946 Suppresses the Migration and Invasion of the Breast

Cancer Cell Line MCF-7 by Targeting RFC3. *Cancer Manage Res* (2018) 10:535. doi: 10.2147/CMAR.S155923
84. Scintu M, Vitale R, Prencipe M, Gallo AP, Bonghi L, Valori VM, et al. Genomic Instability and Increased Expression of BUB1B and MAD2L1 Genes in Ductal Breast Carcinoma. *Cancer Lett* (2007) 254(2):298–307. doi: 10.1016/j.canlet.2007.03.021
85. Labib K, Gambus A. A Key Role for the GINS Complex at DNA Replication Forks. *Trends Cell Biol* (2007) 17(6):271–8. doi: 10.1016/j.tcb.2007.04.002
86. Wei H-B, Wen J-Z, Wei B, Han X-Y, Zhang S. Expression and Clinical Significance of GINS Complex in Colorectal Cancer. *Chin J Gastrointestinal Surg* (2011) 14(6):443–7.
87. Tang L, Yu W, Wang Y, Li H, Shen Z. Anlotinib Inhibits Synovial Sarcoma by Targeting GINS1: A Novel Downstream Target Oncogene in Progression of Synovial Sarcoma. *Clin Trans Oncol* (2019) 1–1021(12):1624–33. doi: 10.1007/s12094-019-02090-2
88. Robles LD, Frost AR, Davila M, Hutson AD, Grizzle WE, Chakrabarti R. Down-Regulation of Cdc6, a Cell Cycle Regulatory Gene, in Prostate Cancer. *J Biol Chem* (2002) 277(28):25431–8. doi: 10.1074/jbc.M201199200
89. Liu Y, Gong Z, Sun L, Li X. FOXM1 and Androgen Receptor Co-Regulate CDC6 Gene Transcription and DNA Replication in Prostate Cancer Cells. *Biochim Biophys Acta (BBA)-Gene Regul Mech* (2014) 1839(4):297–305. doi: 10.1016/j.bbagrm.2014.02.016
90. Sun T-Y, Xie H-J, Li Z, He H, Kong L-F. Expression of CDC6 in Ovarian Cancer and Its Effect on Proliferation of Ovarian Cancer Cells. *Int J Clin Exp Med* (2016) 9(6):10544–50.
91. Mahadevappa R, Neves H, Yuen SM, Bai Y, McCrudden CM, Yuen HF, et al. The Prognostic Significance of CDc6 and CDt1 in Breast Cancer. *Sci Rep* (2017) 7(1):985. doi: 10.1038/s41598-017-00998-9
92. Xu J, Wang Y, Xu D. CKS2 Promotes Tumor Progression and Metastasis and is an Independent Predictor of Poor Prognosis in Epithelial Ovarian Cancer. *Eur Rev Med Pharmacol Sci* (2019) 23(8):3225–34. doi: 10.26355/eurrev_201904_17681
93. Yu M-H, Luo Y, Qin S-L, Wang Z-S, Mu Y-F, Zhong M. Up-Regulated CKS2 Promotes Tumor Progression and Predicts a Poor Prognosis in Human Colorectal Cancer. *Am J Cancer Res* (2015) 5(9):2708.
94. Huang N, Wu Z, Hong H, Wang X, Yang F, Li H. Overexpression of CKS2 Is Associated With a Poor Prognosis and Promotes Cell Proliferation and Invasion in Breast Cancer. *Mol Med Rep* (2019) 19(6):4761–9. doi: 10.3892/mmr.2019.10134
95. Coux O, Tanaka K. Goldberg Aljarob. *Structure Functions 20S 26S Proteasomes* (1996) 65(1):801–47. doi: 10.1146/annurev.bi.65.070196.004101
96. Xia S, Tang Q, Wang X, Zhang L, Jia L, Wu D, et al. Overexpression of PSMA7 Predicts Poor Prognosis in Patients With Gastric Cancer. *Oncol Lett* (2019) 18(5):5341–9. doi: 10.3892/ol.2019.10879
97. Hu X-T, Chen W, Zhang F-B, Shi Q-L, Hu J-B, Geng S-M, et al. Depletion of the Proteasome Subunit PSMA7 Inhibits Colorectal Cancer Cell Tumorigenicity and Migration. *Oncol Rep* (2009) 22(5):1247–52. doi: 10.3892/or_00000561
98. Zhu J, Han S. Downregulation of Lncrna DARS-AS1 Inhibits the Tumorigenesis of Cervical Cancer *Via* Inhibition of IGF2BP3. *OncoTargets Ther* (2021) 14:1331. doi: 10.2147/OTT.S274623
99. Zheng W, Tian X, Cai L, Shen Y, Cao Q, Yang J, et al. Lncrna DARS-AS1 Regulates MicroRNA-129 to Promote Malignant Progression of Thyroid Cancer. *Eur Rev Med Pharmacol Sci* (2019) 23:10443–52. doi: 10.26355/eurrev_201912_19683
100. Huang K, Fan W, Fu X, Li Y, Meng Y. Long Noncoding RNA DARS-AS1 Acts as an Oncogene by Targeting Mir-532-3p in Ovarian Cancer. *Eur Rev Med Pharmacol Sci* (2019) 23(6):2353–9.
101. Tong J, Xu X, Zhang Z, Ma C, Xiang R, Liu J, et al. Hypoxia-Induced Long non-Coding RNA DARS-AS1 Regulates RBM39 Stability to Promote Myeloma Malignancy. *Haematologica* (2020) 105(6):1630. doi: 10.3324/haematol.2019.218289
102. Taft RJ, Vanderver A, Leventer RJ, Damiani SA, Simons C, Grimmond SM, et al. Mutations in DARS Cause Hypomyelination With Brain Stem and Spinal Cord Involvement and Leg Spasticity. *Am J Hum Genet* (2013) 92 (5):774–80. doi: 10.1016/j.ajhg.2013.04.006
103. Muthiah A, Housley GD, Klugmann M, Fröhlich D. The Leukodystrophies HBSL and LBSL—Correlates and Distinctions. *Front Cell Neurosci* (2021) 14:478. doi: 10.3389/fncel.2020.626610
104. Kimball SR. Eukaryotic Initiation Factor Eif2. *Int J Biochem Cell Biol* (1999) 31(1):25–9. doi: 10.1016/S1357-2725(98)00128-9
105. Jin F, Wu Z, Hu X, Zhang J, Gao Z, Han X, et al. The PI3K/Akt/GSK-3β/ROS/Eif2b Pathway Promotes Breast Cancer Growth and Metastasis *Via* Suppression of NK Cell Cytotoxicity and Tumor Cell Susceptibility. *Cancer Biol Med* (2019) 16(1):38. doi: 10.20892/j.issn.2095-3941.2018.0253
106. Wang H, Wang H-S, Zhou B-H, Li C-L, Zhang F, Wang X-F, et al. Epithelial–Mesenchymal Transition (EMT) Induced by TNF-α Requires AKT/GSK-3β-Mediated Stabilization of Snail in Colorectal Cancer. *PloS One* (2013) 8(2):e56664. doi: 10.1371/journal.pone.0056664
107. Goto D, Tanaka I, Sato M, Kato T, Miyazawa A, Hase T, et al. P1. 03-13 Eif2β, A Subunit of Translation-Initiation Factor Eif2, as a Potential Therapeutic Target for Non-Small Cell Lung Cancer. *J Thoracic Oncol* (2018) 13(10):S516–S7. doi: 10.1016/j.jtho.2018.08.694
108. Jin X, Pang W, Zhang Q, Huang H. Microrna-486-5p Improves Nonsmall-Cell Lung Cancer Chemotherapy Sensitivity and Inhibits Epithelial–Mesenchymal Transition by Targeting Twinfilin Actin Binding Protein 1. *J Int Med Res* (2019) 47(8):3745–3756. doi: 10.1177/0300060519850739
109. Bockhorn J, Dalton R, Nwachukwu C, Huang S, Prat A, Yee K, et al. Microrna-30c Inhibits Human Breast Tumour Chemotherapy Resistance by Regulating TWF1 and IL-11. *Nat Commun* (2013) 4:1393. doi: 10.1038/ncomms2393
110. Arencibia JM, Pastor-Flores D, Bauer AF, Schulze JO, Biondi RM. AGC Protein Kinases: From Structural Mechanism of Regulation to Allosteric Drug Development for the Treatment of Human Diseases. *Biochim Et Biophys Acta (BBA)-Proteins Proteomics* (2013) 1834(7):1302–21. doi: 10.1016/j.bbapap.2013.03.010
111. Di Cristofano A. *SGK1: The Dark Side of PI3K Signaling. Current Topics in Developmental Biology* Vol. 123. United States: Elsevier (2017). p. 49–71.
112. Park J, Leong ML, Buse P, Maiyar AC, Firestone GL, Hemmings BA. Serum and Glucocorticoid-Inducible Kinase (SGK) Is a Target of the PI 3-Kinase-Stimulated Signaling Pathway. *EMBO J* (1999) 18(11):3024–33. doi: 10.1093/emboj/18.11.3024
113. Brunet A, Park J, Tran H, Hu LS, Hemmings BA, Greenberg ME. Protein Kinase SGK Mediates Survival Signals by Phosphorylating the Forkhead Transcription Factor FKHRL1 (FOXO3a). *Mol Cell Biol* (2001) 21(3):952–65. doi: 10.1128/MCB.21.3.952-965.2001
114. Amato R, Menniti M, Agosti V, Boito R, Costa N, Bond HM, et al. IL-2 Signals Through Sgk1 and Inhibits Proliferation and Apoptosis in Kidney Cancer Cells. *J Mol Med* (2007) 85(7):707–21. doi: 10.1007/s00109-007-0205-2
115. Maiyar AC, Huang AJ, Phu PT, Cha HH, Firestone GL. P53 Stimulates Promoter Activity of the Sgk Serum/Glucocorticoid-Inducible Serine/Threonine Protein Kinase Gene in Rodent Mammary Epithelial Cells. *J Biol Chem* (1996) 271(21):12414–22. doi: 10.1074/jbc.271.21.12414
116. Wu W, Chaudhuri S, Brickley DR, Pang D, Karrison T, Conzen SD. Microarray Analysis Reveals Glucocorticoid-Regulated Survival Genes That Are Associated With Inhibition of Apoptosis in Breast Epithelial Cells. *Cancer Res* (2004) 64 (5):1757–64. doi: 10.1158/0008-5472.CAN-03-2546
117. Zeisberg M, Neilson EG. Biomarkers for Epithelial-Mesenchymal Transitions. *J Clin Invest* (2009) 119(6):1429–37. doi: 10.1172/JCI36183
118. Cao X, Cao Y, Li W, Zhang H, Zhu Z. P4HA1/HIF1α Feedback Loop Drives the Glycolytic and Malignant Phenotypes of Pancreatic Cancer. *Biochem Biophys Res Commun* (2019)516(3):606–12. doi: 10.1016/j.bbrc.2019.06.096
119. Xu R. P4HA1 is a New Regulator of the HIF-1 Pathway in Breast Cancer. *Cell Stress* (2019) 3(1):27. doi: 10.15698/cst2019.01.173
120. Gilkes DM, Bajpai S, Chaturvedi P, Wirtz D, Semenza GL. Hypoxia-Inducible Factor 1 (HIF-1) Promotes Extracellular Matrix Remodeling Under Hypoxic Conditions by Inducing P4HA1, P4HA2, and PLOD2 Expression in Fibroblasts. *J Biol Chem* (2013) 288(15):10819–29. doi: 10.1074/jbc.M112.442939
121. Burridge K, Fath K, Kelly T, Nuckolls G, Turner C. Focal Adhesions: Transmembrane Junctions Between the Extracellular Matrix and the Cytoskeleton. *Annu Rev Cell Biol* (1988) 4(1):487–525. doi: 10.1146/annurev.cb.04.110188.002415
122. Flores-Pérez A, Marchat LA, Rodríguez-Cuevas S, Bautista VP, Fuentes-Mera L, Romero-Zamora D, et al. Suppression of Cell Migration Is Promoted by Mir-944 Through Targeting of SIAH1 and PTP4A1 in Breast Cancer Cells. *BMC Cancer* (2016) 16(1):379. doi: 10.1186/s12885-016-2470-3

123. Dluzen D, Li G, Tacelosky D, Moreau M, Liu DX. BCL-2 Is a Downstream Target of ATF5 That Mediates the Prosurvival Function of ATF5 in a Cell Type-Dependent Manner. *J Biol Chem* (2011) 286(9):7705–13. doi: 10.1074/jbc.M110.207639
124. Greene LA, Lee HY, Angelastro JM. The Transcription Factor ATF5: Role in Neurodevelopment and Neural Tumors. *J Neurochem* (2009) 108(1):11–22. doi: 10.1111/j.1471-4159.2008.05749.x
125. Sheng Z, Evans SK, Green MR. An Activating Transcription Factor 5-Mediated Survival Pathway as a Target for Cancer Therapy. *Oncotarget* (2010) 1(6):457. doi: 10.18632/oncotarget.180
126. Miyoshi N, Ishii H, Mimori K, Takatsuno Y, Kim H, Hirose H, et al. Abnormal Expression of TRIB3 in Colorectal Cancer: A Novel Marker for Prognosis. *Br J Cancer* (2009) 101(10):1664. doi: 10.1038/sj.bjc.6605361
127. Dong S, Xia J, Wang H, Sun L, Wu Z, Bin J, et al. Overexpression of TRIB3 Promotes Angiogenesis in Human Gastric Cancer. *Oncol Rep* (2016) 36 (4):2339–48. doi: 10.3892/or.2016.5017
128. Helmy N, Prip-Buus C, Vons C, Lenoir V, Abou-Hamdan A, Guedouari-Bounihi H, et al. Oxidation of Hydrogen Sulfide by Human Liver Mitochondria. *Nitric Oxide* (2014) 41:105–12. doi: 10.1016/j.niox.2014.05.011
129. Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, et al. Analysis of the Human Tissue-Specific Expression by Genome-Wide Integration of Transcriptomics and Antibody-Based Proteomics. *Mol Cell Proteomics* (2014) 13(2):397–406. doi: 10.1074/mcp.M113.035600
130. Szabo C, Ransy C, Módis K, Andriamihaja M, Murghes B, Coletta C, et al. Regulation of Mitochondrial Bioenergetic Function by Hydrogen Sulfide. Part I. Biochemical and Physiological Mechanisms. *Br J Pharmacol* (2014) 171(8):2099–122. doi: 10.1111/bph.12369
131. Gibbins MTG. *Metabolic and Vascular Effects of Thiosulfate Sulfurtransferase Deletion* Edinburgh Medical School Thesis and Dissertation Collection. (2018).
132. Krause N, Wegner A. Fructose Metabolism in Cancer. *Cells* (2020) 9 (12):2635. doi: 10.3390/cells9122635
133. Schwab A, Siddiqui A, Vazakidou ME, Napoli F, Böttcher M, Menchicchi B, et al. Polyol Pathway Links Glucose Metabolism to the Aggressiveness of Cancer Cells. *Cancer Res* (2018) 78(7):1604–18. doi: 10.1158/0008-5472.CAN-17-2834
134. Szabó Z, Hämäläinen J, Loikkanen I, Moilanen A-M, Hirvikoski P, Väisänen T, et al. Sorbitol Dehydrogenase Expression Is Regulated by Androgens in the Human Prostate. *Oncol Rep* (2010) 23(5):1233–9. doi: 10.3892/or_00000755
135. Capaia M, Granata I, Guarracino M, Petretto A, Inglese E, Cattrini C, et al. A Hnrnp K–AR-Related Signature Reflects Progression Toward Castration-Resistant Prostate Cancer. *Int J Mol Sci* (2018) 19(7):1920. doi: 10.3390/ijms19071920
136. Lu H, Zhu H. Effect of SiRNA–Mediated Gene Silencing of Transketolase on A549 Lung Cancer Cells. *Oncol Lett* (2017) 14(5):5906–12. doi: 10.3892/ol.2017.6916
137. Liao A, Xiao Y. Clinic Significance of Downexpression of LGALS4 Gene in Colorectal Cancer. *Cancer Res Clinic* (2011) 23(10):661–3.
138. Rodia MT, Solmi R, Mattia L. TSPAN8 and LGALS4 Combination as Blood Biomarkers for Colorectal Cancer Detection. *Cancer Cell Microenvironment.* (2016) 3(3).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Piran, Sepahi, Moattari, Rahimi and Ghanbariasad. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.