Research Article GU J Sci 37(1): 150-167 (2024) DOI:10.35378/gujs.1182561

Gazi University

Journal of Science

http://dergipark.gov.tr/gujs

Investigation of Genes and Their Interactions in Liver Diseases Using Bioinformatics Algorithms

Saliha ACAR[1, *], Giyasettin OZCAN[2], Eyyup GULBANDILAR[3]

[1]*Bursa Technical University, IT Department, 16310, Bursa, Turkey*
[2]*Bursa Uludag University, Faculty of Engineering, Computer Engineering Department, 16059, Bursa, Turkey*
[3]*Eskisehir Osmangazi University, Faculty of Engineering and Architecture, Computer Engineering Department, 26040, Eskisehir, Turkey*

Highlights

- Molecular changes in the body in advanced diseases of the liver such as cirrhosis and HCC.
- A consecutive approach for analysis of microarray data.
- There is a significant change in the expression levels of 12 genes.
- Gene expression levels accelerated after the formation of cirrhosis.

Article Info

Received: 30 Sep 2022
Accepted: 20 Mar 2023

Keywords

Gene expression analysi
HCC
Cirrhosis

Abstract

In this study, we considered progression of liver diseases. Particularly we considered Hepatocellular Carcinoma Cancer, HCC, whose patients have low survival rates. For this purpose, we researched molecular structures and protein interactions involved in the initiation and progression of HCC. We exploited microarray data samples and their gene expression profiles from literature. During analysis, we implemented statistical data analysis techniques and looked for Differentially Expressed Genes during the initiation and progression of HCC. As a result of this analysis we found 12 hub genes, where 4 of them (ANLN, TOP2A, ASPM and SPINK1) were upregulated and the others (CXCL14, LINC01093, OIT3, CLEC4G, THRSP, APOF, CLTRN and FCN3) were downregulated. By performing Gene Ontology Analysis, we classified genes with increased or decreased expressions in terms of cellular component, biological process, and molecular function. Subsequently, we executed protein-protein interaction network analysis and found important interactions between the hub genes. Results of data analysis concluded that these 12 genes and their interactions play a key role in the initiation and progression of significant liver diseases and can be used as a potential biomarker for disease progression. Furthermore, gene feature analysis showed that it is becoming more difficult to compensate functional deficiencies of the proteins encoded by these genes during biological processes. In particular, Gene Ontology Analysis denoted that TOP2A gene associates with many of the biological pathways and a change in the expression of this gene can cause decent problems in many cellular functions.

1. INTRODUCTION

In the last 10 years, microarray data analysis methods have been used frequently in the investigation of the molecular mechanism of diseases such as cancer, where genetic factors are predicted to play a role in their formation and development. Microarray technology enables the detection of the expression level changes of thousands of genes on mRNA (messenger-RNA) simultaneously [1]. Many microarray studies are designed to identify genes associated with different phenotypes. In order to make microarray experiments interpretable, series of bioinformatics algorithms and statistical approaches are required [2]. As a result of these analyses; comparison of gene expression levels (usually sick and healthy) between two groups can be obtained via computational and statistical approaches. Therefore, molecular function and physiology of proteins which are products of genes can be understood in a clear form. Finally, biological and functional relationships between proteins can be obtained.

*Corresponding author, e-mail: saliha.acar@btu.edu.tr

Nowadays, there are many studies on the classification of advanced liver diseases such as cirrhosis and HCC at the molecular level and determination of their gene expression patterns. For example, Lian et al. found that the ANLN gene was expressed at a higher level in tumor tissues compared to healthy adjacent tissues and concluded that ANLN promotes tumor development by reducing the level of apoptosis (cell death) and DNA damage [3]. Satow et al., in their study for therapeutic target detection for HCC, found that the ANLN gene was expressed 3 times more in tumor tissues than in healthy adjacent tissues [4]. Wong et al. showed that high expression of TOP2A gene; and they found that it was associated with symptoms such as advanced HCC tumor histology, microvascular attack, and early-stage attack [5]. Gao et al., in their microarray analysis study, found that high expression of the TOP2A gene is a factor that reduces the survival rate for HCC patients [6]. Horvath et al. showed that by suppressing the ASPM gene, it is possible to regenerate nerve stem cells and hence prevent the development of glioblastoma cells [7]. In their study, Li et al., using tissue samples taken from individuals with HCC and HBV-infected cirrhosis, found that high expressions of FOXP1 and SPINK1 genes play an important role in canceration of cirrhosis and could be important potential biomarkers for early diagnosis of HCC [8].

Marshal et al., in their gene expression profile analysis study using tissue samples taken from different liver patients and individuals with normal liver, found that the expression level of the SPINK1 gene was quite high in the tissues of liver patients compared to that in healthy tissues, and emphasized that this gene could be a diagnostic biomarker for HCC disease [9]. Wang et al., analyzed a set of tumor and non-tumor tissue samples from 126 HCC patients. They observed that the expression level of the CXCL14 gene, which took space in 88.1% of tumor tissues, showed a significant decrease. Thus, they concluded that the CXCL14 gene has an important role for being a potential tumor suppressor for HCC [10]. Esposti et al. analyzed tissue samples of 23 individuals with normal, cirrhosis and HCC, and observed that the expression level of the LINC01093 gene in cirrhotic tissues decreased significantly compared to its expression in normal tissues [11]. Jovel et al. observed that the expression of the LINC01093 gene decreased more than 30 times in tumor tissues compared to healthy tissues in their lncRNA analysis. As a result, they stated that trying to understand the HCC mechanism through this gene would be a promising effort to manage HCC [12].

Dai et al. in their study using a total of 6 data sets; They observed that the level of the LINC01093 gene was quite low in HCC patients and concluded that LINC01093 could be a prognostic indicator for HCC, as it is one of the 3 genes that has a strong relationship with the survival rates of HCC patients [13]. Li et al. used tissue samples taken from individuals with HCC and HBV-infected cirrhosis and observed that the expression of the OIT3 gene in the tissues of individuals with HCC decreased about 2 times, compared to that of other individuals [8]. Agarwal et al., in their gene expression analysis study using tumor and neighboring tumor-free tissue samples from 423 HCC patients, observed that the expression of the OIT3 gene was one of the three most declining genes compared to tumor-free tissues [14]. Ho et al., in their gene expression analysis study using tumor and neighboring tumor-free tissue samples from 50 individuals with HCC, observed that the expression of the CLEC4G gene was the most decreasing gene compared to tumor-free tissues [15]. Similarly, Chen et al. observed in their HCC microarray gene expression analysis study observed that the expression of the CLEC4G gene was one of the most declining genes compared to tumor-free tissues, and therefore concluded that C-type lectin proteins may play a role in the pathogenesis of HCC [16]. Huang et al. stated that THRSP has an important role in the pathogenesis of NAFLD (Non-Alcoholic Fatty Liver Disease) and could be a potential therapeutic target for NAFLD treatment [17]. Cheng et al., used 75 surgical liver samples taken from 48 HCV-infected cirrhosis patients. Their pathway enrichment analysis and microarray analysis showed that the APOF gene was one of 7 genes whose expression levels were changed in healthy, cirrhotic and HCC tumor tissues. Thus, they concluded that this gene plays an important role in the progression of HCC disease [18]. Lo Re et al. observed that the expression level of the APOF gene decreased about 2 times more; compared to the expression in control group cells [19]. Chang et al. found that the expression of the FCN3 gene in tissues with tumors was significantly lower than in healthy tissues in their microarray and interactive pathway analysis study with tumor and healthy tissue samples taken from 31 HCC patients [20].

There exist several studies on miRNA, DEGs, and mRNA utilizing gene expressions. For instance, Udhaya Kumar et al. searched for the hub and core genes in ovarian cancer. Consequently, they investigated potential pathways which could take place during the molecular pathogenesis of ovarian cancer and

performed screening cancer cells. Afterwards, they did pathway enrichment and survival analysis and found 10 hub genes and four core genes [21]. On the other hand, Wan et al. aimed at molecular interactions on Noncoding RNA's and considered the roles of lncRNA-associated endogenous RNA. Their findings concluded that endogenous RNAs can help biomarker discovery and treatment strategies for prostate cancer [22]. Another study by Udhaya Kumar et al. (2020) considered Systemic lupus erythematosus and studied essential genes and molecular pathways involved in the pathogenesis. They screened differentially expressed genes between the B-cell transcriptomes patients and healthy controls. Their presented 4 upregulated and 13 downregulated genes and concluded that their results can be used as potential biomarkers or therapeutic targets [23]. Meanwhile, Fu et al. proposed a prognostic signature for the sake of lung squamous cell carcinoma prognosis. The authors focused on functional regulatory mechanisms and developing immune-related targeted therapy in the treatment. They introduced seven risk signatures for patients. In their findings, the authors presented both independent prognostic predictor, and indicator of tumor immune landscape [24].

In this study, we utilized microarray data and graph data structure. We implemented statistical data analysis techniques during the microarray data process. The data is obtained from liver tissue samples of 10 HCC patients and 10 healthy individuals. Microarray data obtained from a total of 37 tissue samples from 17 HCC patients and 20 individuals with cirrhotic liver. We exploited Differentially Expressed Genes (DEGs) Analysis for the detection of gene expression changes. Furthermore, we exploited Gene Ontology (GO) Analysis for functional and biological classification and Protein-Protein Interaction (PPI) Network Analysis for determining the interactions of key genes with each other.

2. MATERIAL AND METHOD

2.1. Utilized Microarray Data Sets

In this study, we utilized microarray datasets whose access numbers are GSE29721 and GSE17548. They are respectively submitted by Moshe Szyf in 2011 and by Mehmet Öztürk in 2013 to the National Center for Biotechnology Information (NCBI) of the American National Institute of Health. The datasesets are presented as free of charge to researchers in the Gene Expression Omnibus (GEO) database. Data set with access number GSE29721 is a microarray data consists of 10 healthy liver tissue samples, 10 HCC patient liver tissue [25]. We accessed to the samples on October 18, 2018. This expression data was obtained using the Affymetrix Human Genome U133 Plus 2.0 array chip and was updated in 2018. For each sample, there exists an index to handle 20 compressed. CEL files arranged consecutively. Each CEL file stores healthy and tumor tissue samples consecutively. The purpose of choosing these datasets is to determine the place of cirrhosis, which is an important turning point in the process from healthy liver to cancer.

The data set with access number GSE17548 is a microarray data, consisting of 20 liver tissue samples with cirrhosis. It also contains 17 liver tissue samples from HCC patients [25]. This expression data was obtained using the Affymetrix Human Genome U133 Plus 2.0 array chip and was updated in 2018. The data files contain cirrhosis and cancerous tissue samples and consist of 37 compressed files with. CEL extension and arranged consecutively.

2.2. Microarray Analysis

2.2.1. Implemented steps before the analysis of differentially expressed genes (DEGs)

Before moving on to the analysis to be made on the data, it is necessary to install the Bioconductor packages containing the functional instruction sets to be used in the analysis. Secondly, we read the density data files and assigned them to variables. Lastly, we did quality control and normalization of the data.

During the entire study, the codes were written in R programming language in R Studio 1.1.456 software development environment. Affy, affyio, affyQCReport, biomaRt, limma and hgu133plus2.db packages were used for DEGs analysis, and biomaRt, GO.db, Rgraphviz and topGO packages were used for GO analysis. After the relevant packages were downloaded and installed, microarray density data files were

read and transferred to a variable with the ReadAffy method and saved to the computer as a text file (txt extension). The quality assessment of the microarray chips was then performed. The affyQCReport package created for the R environment was used and the RLE and NUSE box graphics of the chips were obtained with the affyQAReport function. With the help of graphics, it is determined whether the microarray chips that may have a negative effect on the analysis to be made. For the normalization of the quality controlled raw expression data, the affy package and Robust Multi-Array Average (RMA) function was used.

2.2.2. Differentially expressed genes (DEGs) analysis

DEGs analyses were performed after quality control and normalization procedures performed separately for gene expression data in each data set (GSE29721 and GSE17548). In this study, limma analysis method has been used among the mentioned methods. For the detection of important genes, a linear model was established by loading the "limma" package into R environment. Statistical significance level (p value) was chosen as 0.05 for the varying expressions of genes in the analysis.

As a result of each DEGs analysis, a list of all genes whose expressions were changed was obtained. In terms of GSE29721, genes with an absolute value of 3 or more logarithmic fold change were selected from the list and combined in a separate table to be used as input in the next analysis (GO analysis). From the list obtained from the GSE17548 dataset, threshold value of logarithmic fold change value is determined as 2 or more. Consequently, we combined selected genes that satisfy the threshold criteria and combined in a separate table to be used as input in the next analysis (GO analysis) step.

In addition, the intersection of genes whose expressions decreased or increased by at least 2 times in the analysis of both data sets were taken and the numerical and alphanumeric data of the intersecting genes were accumulated in a separate table. Genes selected as input for GO analysis were identified by globally used gene symbols before they were tabulated.

The samples in the GSE17548 coded dataset belong to cirrhosis patients and HCC patients who have not yet turned into cancer but have no treatment options, that is, they are close to each other on the line. On the other hand, in the samples in the GSE29721 coded dataset, are more distant from each other. Data for humans and HCC patients are available and are located at the two extremes of the number line. Therefore, there is almost a 2-fold difference between the floor change rates found as a result of the studies on the two data sets. Also the highest rates of change were considered significant in order to combine (intercept) the results. Because the loss of function is higher in the genes whose expression value changes the most. Because of these observations, threshold values is decided as 2 for this study.

2.2.3. Gene ontology (GO) enrichment analysis

For each data set (GSE29721 and GSE17548), separate GO enrichment analysis was performed in order to reveal the biological meanings of genes via significantly changed expression levels and varying expression levels. With this analysis, it is aimed to find the GO terms associated with genes whose expressions changed significantly, and the corresponding classifications / processes. In the GO analysis, the weight01 algorithm [26] was used to eliminate local similarities and dependencies between GO terms, and a value of 0.05 was chosen as the level of statistical significance.

In the analysis for the GSE29721 data set, the genes that were obtained as a result of the DEGs analysis and whose expressions changed significantly (at least 3 times increased or decreased) were used as background input, and the intersection genes whose expressions significantly changed in both data sets were used as candidate list entries. In the analysis for the data set GSE17548, the genes that were obtained as a result of the DEGs analysis and whose expression changed significantly (at least 2 times increased or decreased) were used as background input, and the intersection genes whose expressions has significant changes in both data sets were used as candidate list entries.

2.2.4. Protein-protein interaction (PPI) network analysis

In order to examine the gene lists obtained by statistical analysis in the DEGs analysis, protein - protein interaction networks were established using the STRING database [27].

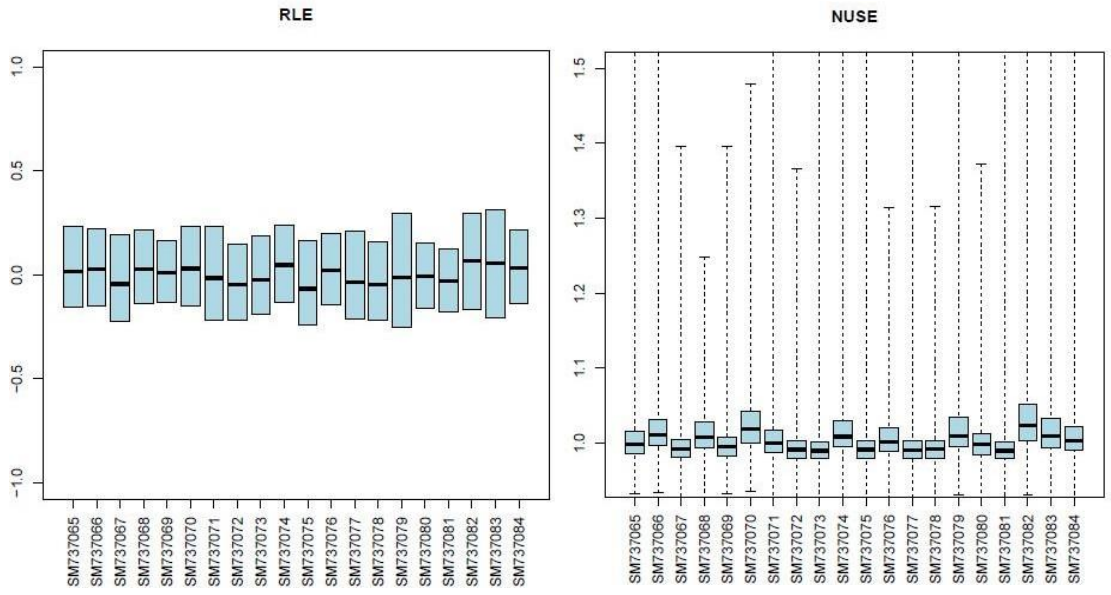
We established protein - protein interaction network after the DEGS analysis of the dataset. To be more precise, we collected 66 genes from GSE29721 whose expressions have increased or decreased at least 3 times more than the others. Furthermore, we determined 26 genes from GSE17548 whose gene expressions have increased or decreased at least 2 times more than the others. We observed that 12 genes located at the intersection of these two clusters. The protein – protein interaction network is constructed by means of these observations.

Genes that could not be identified by the STRING database were not included in the relevant interaction network. Within the framework of known and predicted molecular interaction information, we aim to understand the relationships between genes whose expressions change significantly.

3. THE RESEARCH FINDINGS AND DISCUSSION

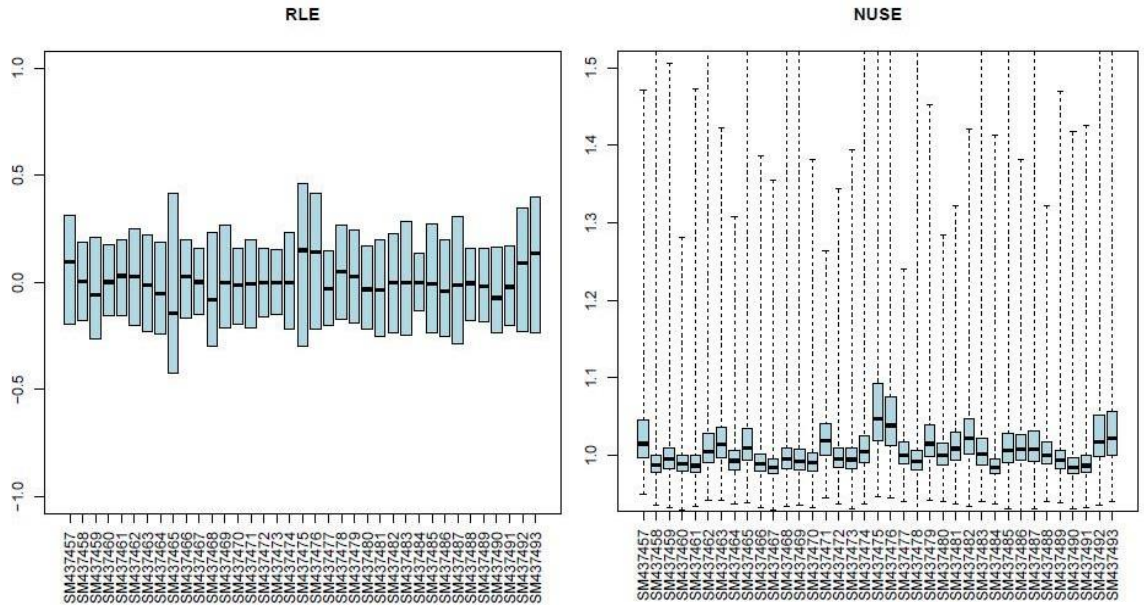
3.1. Quality Assessment Findings of Microarray Chips

RLE and NUSE plots, which are quality assessment tools for microarray chips, were obtained (Figures 1 and 2). When the RLE and NUSE graphs were examined, it was observed that the boxes belonging to the chips did not show widespread scatter and were centered around the values of 0 (zero) and 1 (one), respectively. In other words, it was determined that there was no microarray chip with a problem in terms of quality and therefore data of all chips were included in the analysis.



(a) (b)

Figure 1. *(a) RLE and (b) NUSE box graphics of the GSE29721 microarray chips*

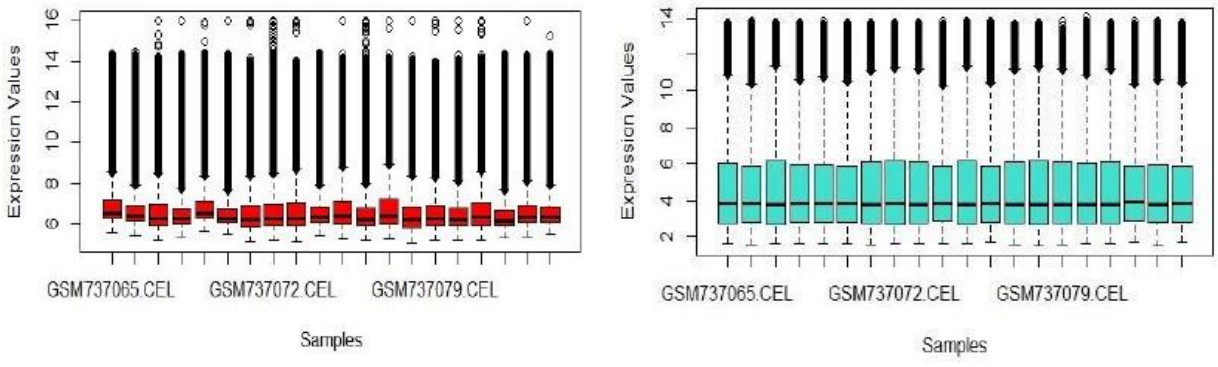


(a) (b)

Figure 2. *(a) RLE and (b) NUSE box graphics of microarray chips from GSE17548*

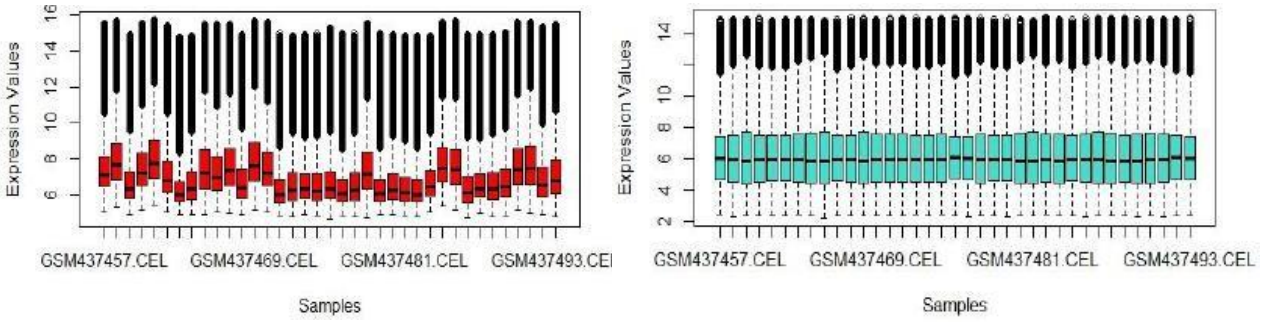
3.2. Expression Values Before and After Normalization

It was observed that, the raw expression data which were log2 transformed before the normalization process, spread at different intervals and the median values of the data were not constant. To be more precise, it was observed that the median of expression values in data set GSE29721 was fixed around 4 after the normalization process. On the other hand, median of expression values in data set GSE17548 was fixed around 6. In Figure 3 and Figure 4, box graphs denote the gene expression values of each data sets before and after normalization respectively.



(a) (b)

Figure 3. *Box plots showing the log-transformed expression values of data set GSE29721 (a) before normalization and (b) after normalization*



(a) (b)

Figure 4. *Box plots showing the log-transformed expression values of data set GSE17548 (a) before normalization and (b) after normalization*

3.3. Differentially Expressed Genes (DEGs) Analysis Findings

GSE29721 consists of microarray data taken from 10 healthy and 10 HCC tumor tissue samples. The total number of probes was 54675. After the analysis, it was found that the expression of 1368 genes increased more than one time and the expression of 1286 genes decreased more than one time. The expressions of 25 of genes increased more than 3 times, while the expressions of 60 of them decreased more than 3 times. Of these 85 genes, 19 genes were unidentified and removed. Remaining 66 genes were taken into account (Tables 1 and 2).

Table 1. *The list of genes whose expressions have increased more than 3 times in the data set GSE29721*

Gene	Logarithmic Fold Change	Important Functions (Related Processes)
COCH	+3.03	Hearing and innate immune response
NUF2	+3.03	A regulatory role in chromosome segregation
CDK1	+3.08	Eukaryotic cell cycle, transferase activity and protein tyrosine kinase activity
CENPF	+3.11	Protein homodimerization activity and transcription factor binding
NEK2	+3.17	Transferase activity and protein tyrosine kinase activity
MMP12	+3.20	Embryonic development, reproduction and tissue remodeling
IGF2BP3	+3.26	Nucleic acid binding and RNA binding
CTHRC1	+3,26	Frizzled binding and Wnt-protein binding
ANLN	+3.37	Cell growth, cell migration and cytokinesis
REG3A	+3.40	Cell proliferation and/or differentiation
CCNB1	+3.40	Control of the cell cycle in mitosis
TOP2A	+3.41	RNA binding and protein heterodimerization activity
ASPM	+3.41	Mitotic spindle function in embryonic neuroblasts and regulation of neurogenesis
CDKN3	+3.41	Phosphatase activity and phosphoprotein phosphatase activity
ACSL4	+3.46	Lipid biosynthesis and fatty acid degradation
SULT1C2	+3.46	Sulfotransferase activity
CCL20	+3.72	Cytokine activity and chemokine activity
GPC3	+4.42	Heparan sulfate proteoglycan binding and peptidyl-dipeptidase inhibitor activity
SPINK1	+4.55	Serine-type endopeptidase and peptidase inhibitor activity
PEG10	+4.70	Nucleic acid binding and RNA binding

Table 2. *The list of genes whose expressions have decreased more than 3 times in the data set GSE29721*

Gene	Logarithmic Fold Change	Important Functions (Related Processes)
RDH16	-3.00	Oxidoreductase activity and NAD-retinol dehydrogenase activity
PCK1	-3.00	GTP binding and GDP binding
ETNPPL	-3.05	Pyridoxal phosphate binding and ethanolamine-phosphate phospho-lyase activity
SRD5A2	-3.07	Metabolism of steroids and Metabolism of steroid hormones
AKR1D1	-3.08	Metabolism of steroids and Synthesis of bile acids and bile salts
CLTRN	-3.10	Insulin processing and metabolism of proteins
CYP26A1	-3.12	Iron ion binding and oxidoreductase activity
MFSD2A	-3.14	Establishment of the blood-brain barrier, brain growth and function
CYP2A6	-3.15	Oxidation by cytochrome P450 and nuclear receptors
CLEC4G	-3.17	Carbohydrate binding and immune response
CLRN3	-3.19	Protein binding
LCAT	-3.21	Statin inhibition of cholesterol production and Familial hyperlipidemia type 1
BBOX1	-3.26	Carnitine synthesis and Regulation of expression of SLITs and ROBOs
GLS2	-3.30	Neurotransmitter release cycle and pyrimidine metabolism
SLC25A47	-3.31	Mitochondrial membrane
KCNN2	-3.33	Protein homodimerization activity and ion channel activity
CYP8B1	-3.35	Metabolism of steroids and arachidonic metabolism
C9	-3.40	Innate and adaptive immune response
CLEC4M	-3.41	Signaling receptor activity and calcium-dependent protein binding
AFM	-3.41	Vitamin E binding
OIT3	-3.41	Calcium ion binding
SDS	-3.42	Metabolizing serine and glycine
ADH1B	-3.44	Glucose / Energy metabolism and oxidation by cytochrome P450
GLYAT	-3.46	Amino acid conjugation and metapathway biotransformation phase I and II
TAT	-3.48	Pyridoxal phosphate binding and amino acid binding
HAO2	-3.48	Peroxisomal lipid metabolism
ALDOB	-3.48	Glycolysis and fructose metabolism
FCN2	-3.49	Calcium ion binding and calcium-dependent protein binding
CRHBP	-3.51	Signal transduction and presynaptic function of kainate receptors
GYS2	-3.53	Protein homodimerization and glycogen synthase activity, transferring glucose
CYP1A2	-3.54	Glucose / Energy metabolism and arachidonic acid metabolism
FCN3	-3.54	Carbohydrate binding and antigen binding
C3P1	-3.56	Endopeptidase inhibitor activity
NAT2	-3.59	Acetyltransferase activity and arylamine N-acetyltransferase activity
SLC22A1	-3.60	Neurotransmitter release cycle and transmembrane transporter activity
CXCL14	-3.63	Immunoregulatory and inflammatory processes
APOF	-3.68	Signaling receptor binding and lipid transporter activity
LINC01093	-3.69	Non-protein coding gene
CLEC1B	-3.71	Transmembrane signaling receptor activity and carbohydrate binding
MT1M	-3.83	Metal ion SLC transporters and cellular responses to stimuli
CNDP1	-3.93	Hydrolase and carboxypeptidase activity
TTC36	-3.94	Cell apoptosis
GBA3	-3.97	Hydrolase activity and beta-glucosidase activity
LINC00844	-4.10	Non-protein coding gene
HAMP	-4.31	Regulation of iron storage in macrophages
THRSP	-4.60	Regulation of lipogenesis

The second data set (GSE17548) consists of microarray data taken from 20 cirrhosis and 17 HCC tumor tissue samples. The total number of probes is 54675. After the analysis of the data, it was determined that the expression of 171 genes increased more than one time and the expression of 383 genes decreased more than one time. The expressions of 10 of these increased more than 2 times. The expression of 21 of them decreased more than 2 times. After the unidentified genes were removed from these 31 genes, 26 genes remained (Tables 3 and 4).

The command used for DEGs analysis was run with the condition $p<0.05$. In other words, the p value is less than 0.05 for all genes with increased/decreased expression levels obtained in the analysis results. However, we did not include this information in the table. The genes in all tables obtained as a result of the

analysis were ordered from largest to smallest according to their logarithmic fold change rates. Because we focused on the absolute magnitude of expression change rates and analyzed meaningful changes.

Table 3. *Gene list whose expressions have increased more than 2 times in the data set GSE17548*

Gene	Logarithmic Fold Change	Important Functions (Related Processes)
RRM2	+2.05	Oxidoreductase and ribonucleoside-diphosphate reductase activity
ANLN	+2.06	Cell growth, cell migration and cytokinesis
TTK	+2.10	DNA damage and cell cycle
MAGEA6	+2.13	There is no exact information
TOP2A	+2.20	RNA binding and protein heterodimerization activity
ASPM	+2.51	Mitotic spindle function in embryonic neuroblasts and regulation of neurogenesis
DUXAP10	+2.64	There is no exact information
SPINK1	+3.02	Serine-type endopeptidase and peptidase inhibitor activity

Table 4. *Gene list whose expressions have decreased more than 2 times in the data set GSE17548*

Gene	Logarithmic Fold Change	Important Functions (Related Processes)
CLTRN	-2.00	Insulin processing and metabolism of proteins
APOF	-2.01	Signaling receptor binding and lipid transporter activity
RIPOR3	-2.01	There is no exact information
SRPX	-2.03	Cell adhesion
DCN	-2.08	RNA binding and collagen binding
JCHAIN	-2.15	Peptidoglycan, phosphatidylcholine and immunoglobulin receptor binding
THRSP	-2.15	Regulation of lipogenesis
CLEC4G	-2.16	Carbohydrate binding and immune response
CYP1A2	-2.18	Glucose / Energy metabolism and arachidonic acid metabolism
OIT3	-2.25	Calcium ion binding
GREM2	-2.28	Protein homodimerization activity and heparin binding
LINC01093	-2.29	Non-protein coding gene
SLCO1B3	-2.33	Organic anion transmembrane transporter activity
CRHBP	-2.33	Signal transduction and presynaptic function of kainate receptors
FCN3	-2.35	Carbohydrate binding and antigen binding
FCN2	-2.40	Calcium ion binding and calcium-dependent protein binding
C7	-2.41	Innate and adaptive immune response
CXCL14	-2.46	Immunoregulatory and inflammatory processes

Our main purpose of the analyses is to identify target genes whose expression values change significantly in the period between the onset of physical and functional damage of the liver and the formation and spread of cancer tumors. Therefore, we selected genes that overlap in each data sets after DEGs analysis.

For this purpose, the intersection set of genes with logFC absolute value of 3 or more in the first data set (GSE29721) and genes with logFC absolute value of 2 or more in the second data set (GSE17548) were taken into account. There are a total of 12 genes in this cluster that four of them have increased expression and the remaining eight have decreased (Table 5).

Table 5. *Intersection set of analysis results for the two data sets (GSE 29721 and GSE 17548)*

Gene	Logarithmic Fold Change (for GSE29721)	Logarithmic Fold Change (for GSE17548)	Important Functions (Related Processes)
ANLN	+3.37	+2.06	Cell growth, cell migration and cytokinesis
TOP2A	+3.41	+2.20	RNA binding and protein heterodimerization activity
ASPM	+3.41	+2.51	Mitotic spindle function in embryonic neuroblasts and regulation of neurogenesis
SPINK1	+4.55	+3.02	Serine-type endopeptidase and peptidase inhibitor activity
CXCL14	-3.63	-2.46	Immunoregulatory and inflammatory processes
LINC01093	-3.69	-2.29	Non-protein coding gene
OIT3	-3.41	-2.25	Calcium ion binding
CLEC4G	-3.17	-2.16	Carbohydrate binding and immune response
THRSP	-4.60	-2.15	Regulation of lipogenesis
APOF	-3.68	-2.01	Signaling receptor binding and lipid transporter activity
CLTRN	-3.10	-2.00	Insulin processing and metabolism of proteins
FCN3	-3.54	-2.35	Carbohydrate binding and antigen binding

3.4. Gene Ontology (GO) Analysis Findings

For both data sets used, GO analysis was performed with genes whose expressions changed significantly according to the DEGs analysis and the genes were sorted according to logarithmic fold change rates. The results were analyzed separately and as a whole. We have found 1194 paths, during the GO analysis for the GSE29721 data set. When the results were examined, it was determined that the statistical significance of 10 GO pathways was high among them (Figure 5, Table 6).

When GO analysis considered for the GSE17548 data, we found 2198 paths. Furthermore, it was found that the statistical significance of 7 GO pathways was high among them (Figure 6, Table 7).

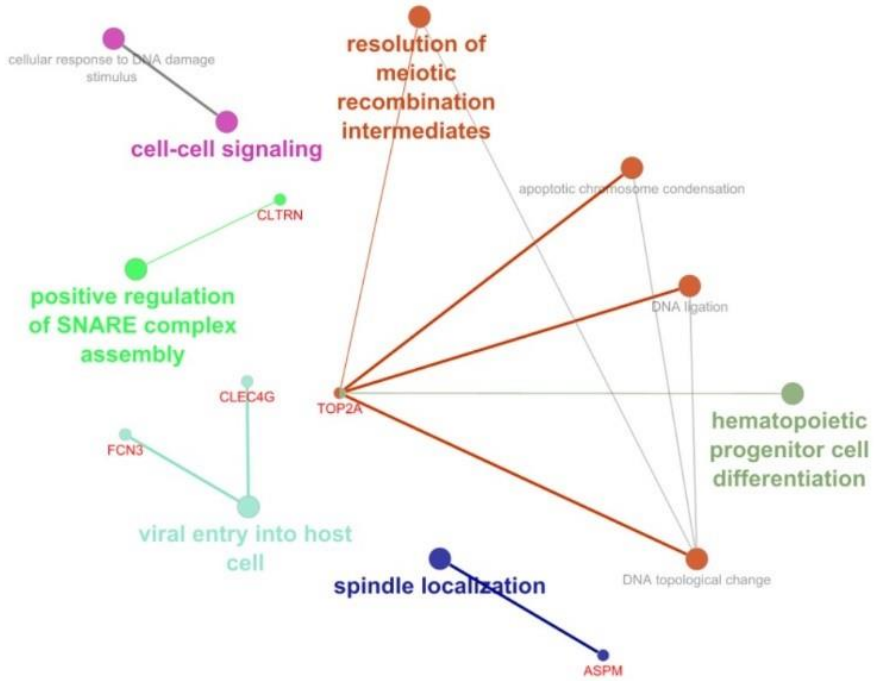


Figure 5. *The most important 10 pathways, found as a result of GO analysis for the dataset no: GSE29721*

Table 6. *The most important 10 pathways, found as a result of GO analysis for the data set GSE 29721*

Pathway ID	Pathway Annotation	Matching Proteins on the Pathway	Expression Directions of Genes
GO: 0007267	Cell-cell signaling	ASPM, CLTRN, CRHBP, CXCL14	ASPM is upregulated, the others are downregulated
GO: 0051653	Spindle localization	ASPM	Upregulated
GO: 0006974	Cellular response to DNA damage stimulus	TOP2A	Upregulated
GO: 0006266	DNA ligation	TOP2A	Upregulated
GO: 0006265	DNA topological change	TOP2A	Upregulated
GO: 0046718	Viral entry into host cell	CLEC4G, FCN3	Downregulated
GO: 0030263	Apoptotic chromosome condensation	TOP2A	Upregulated
GO: 0002244	Hematopoietic progenitor cell differentiation	ANLN, TOP2A	Upregulated
GO: 0035543	Positive regulation of SNARE complex assembly	CLTRN	Downregulated
GO: 0000712	Resolution of meiotic recombination intermediates	TOP2A	Upregulated

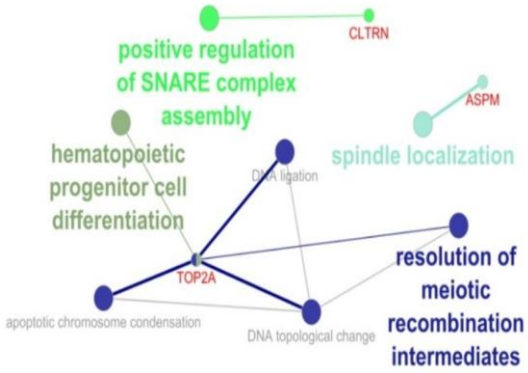


Figure 6. *The most important 7 pathways, found as a result of GO analysis for the dataset no: GSE17548*

Table 7. *The most important 7 pathways found as a result of GO analysis for the data set GSE17548*

Pathway ID	Pathway Annotation	Matching Proteins on the Pathway	Expression Directions of Genes
GO: 0051653	Spindle localization	ASPM	Upregulated
GO: 0006266	DNA ligation	TOP2A	Upregulated
GO: 0006265	DNA topological change	TOP2A	Upregulated
GO: 0030263	Apoptotic chromosome condensation	TOP2A	Upregulated
GO: 0002244	Hematopoietic progenitor cell differentiation	ANLN, TOP2A	Upregulated
GO: 0035543	Positive regulation of SNARE complex assembly	CLTRN	Downregulated
GO: 0000712	Resolution of meiotic recombination intermediates	TOP2A	Upregulated

3.5. Protein - Protein Interaction (PPI) Analysis Findings

For both data sets, protein-protein interaction analyses were carried out, with genes whose expressions changed significantly according to the DEGs analysis. By combining the analysis results of GSE29721 and GSE17548 data sets, PPI network analyses were made for the genes with the highest and lowest expressions in both data sets and interaction maps are shown in Figures 7 and 8.

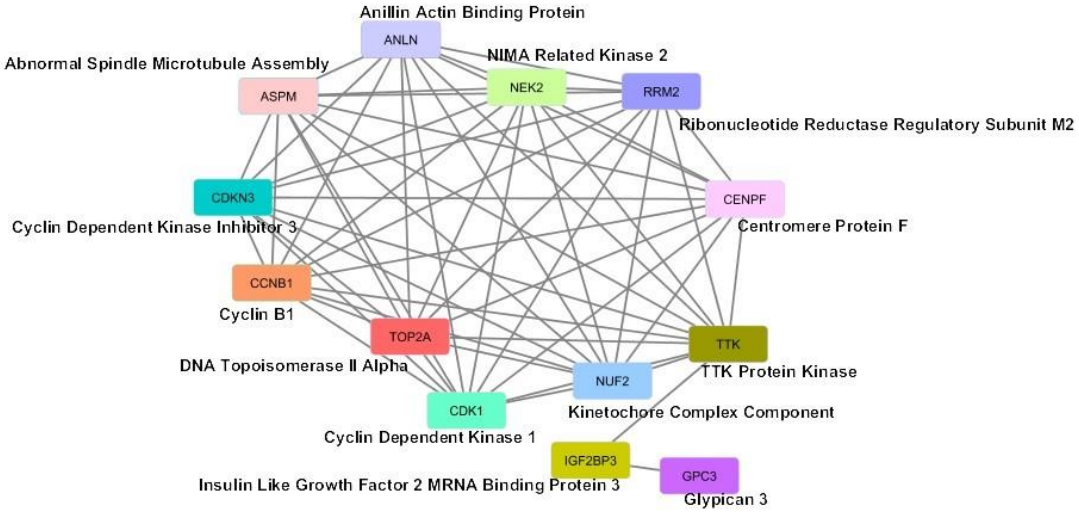


Figure 7. *Protein - protein interaction map of 23 genes whose expressions were increased the most among the two utilized data sets*

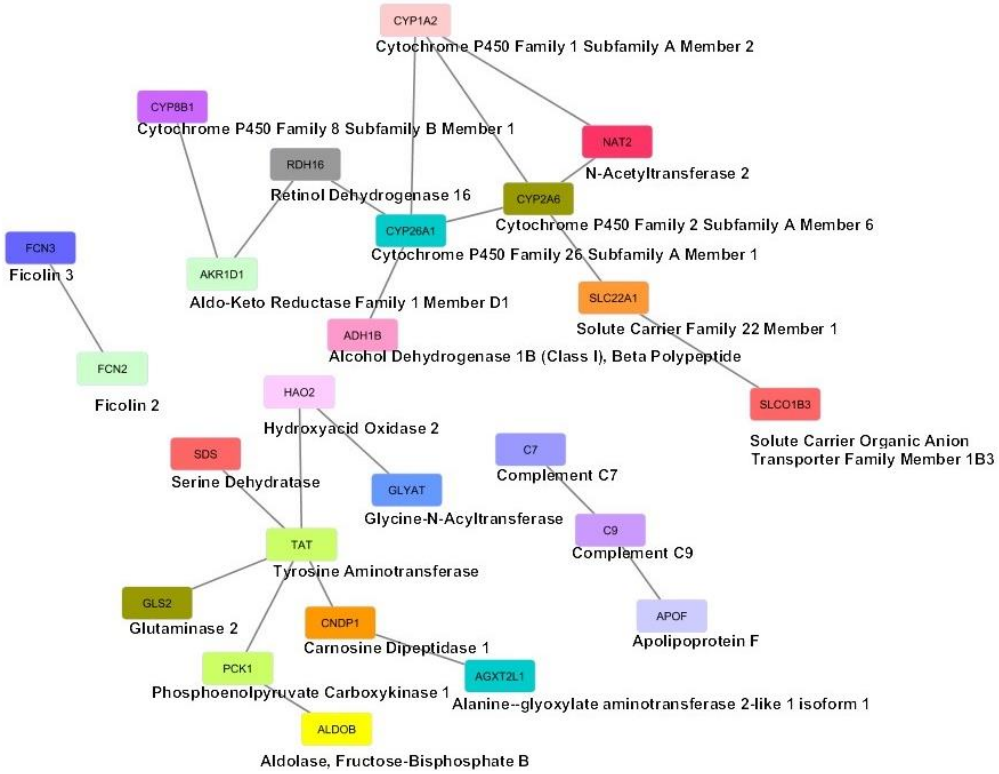


Figure 8. *Protein - protein interaction map of 47 genes with the lowest expressions among the utilized two data sets*

3.6. Intersecting Genes and Their Feature Analysis

In this study, we compared genes of healthy tissues with the tissues of cirrhotic / cancerous. We have determined a list of genes whose expressions differ significantly among the genes of individuals with cirrhosis / HCC and the genes of healthy. For this purpose, we performed series of bioinformatics analyses to identify key genes and important biological pathways that may be associated with advanced liver diseases. As a result of these analyses, it was determined that there were significant changes in the expression level of a total of 2654 genes for the first data set (GSE29721 - cancer and normal samples) and 554 genes for the second data set (GSE17548 - with cirrhosis and cancer samples). After the determination the gene list, we ranked the listed genes according to the log-fold change ratio. The genes whose expressions are changing the most positively or negatively were determined. Due to the difference in the maximum levels of log-fold change ratios, genes with an absolute value of log-fold change ratio of 3 and more in the first data set and genes with an absolute value of log-fold change ratio of 2 and more in the second data set were listed and the intersection of both lists was taken. We concluded that there are 12 genes at the intersection (Figure 9) as follows:

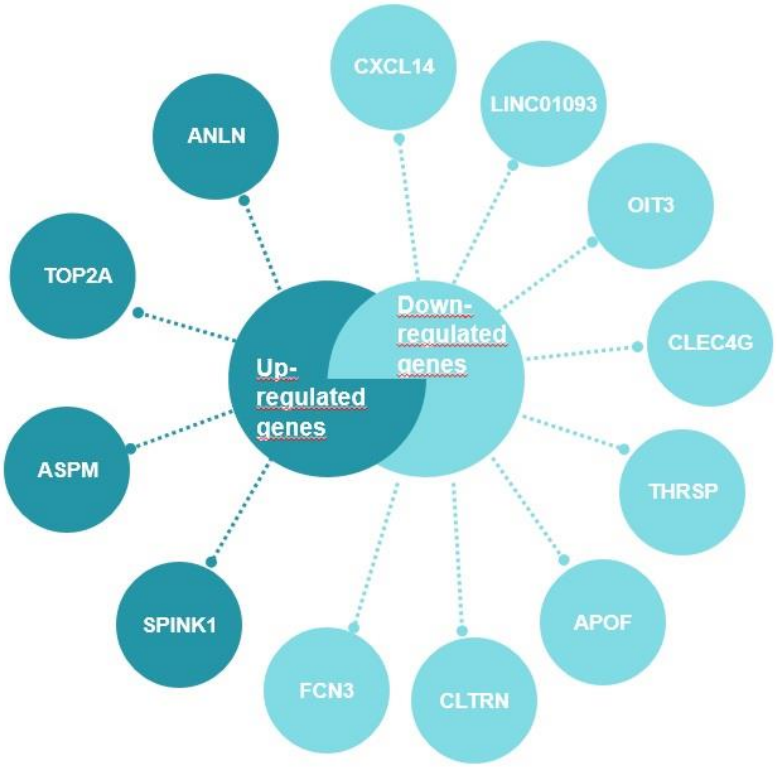


Figure 9. *The hub genes found as a result of DEGs analysis*

ANLN Gene*:* ANLN is an actin binding protein required for joining cleavage lines during cytokinesis (cell division). However, it has been shown that its expression is extremely increased in various cancer types. Expression level of the ANLN gene significantly affects the survival probability of HCC patients and it can be taken as a prognosis biomarker [3]. In this study, in parallel with the results of previous studies, it was observed that the expression of the ANLN gene in the tissues of individuals with cirrhosis and HCC increased 2 to 3 times compared to the genes in healthy tissues. Therefore, the amount of protein encoded by this gene increases in the body and it is estimated that this situation paves the way for the proliferation of unhealthy tissues in the liver.

TOP2A Gene: DNA Topoisomerase II Alpha, encodes the protein that controls and changes the topological state of DNA during the transcript. Despite its role in DNA metabolism, studies conducted in the laboratory environment have shown that the expression level of TOP2A is less dependent in case of the spread of malignant tumors [5]. As a matter of fact, according to the results obtained in this study, it was found that the expression level of the TOP2A gene significantly increased in advanced stage liver diseases. Therefore,

this and previous gene expression profile analysis studies related to liver diseases show that the high level of TOP2A expression facilitates to the formation and development of malignant cancer tumors.

***ASPM Gene**:* It is thought that ASPM gene regulates the proliferation process of neuroblasts, which are embryonic nerve cells that provide neuron formation. Recent studies denote that its expression level is highly elevated in various types of cancer, including ovarian, uterus and glioblastoma a rapidly spreading brain tumor [7]. In this study, it was determined that the ASPM gene was accelerated and expressed at a higher level in the process leading to HCC disease, compared to the ASPM genes in normal cells. This situation suggests that ASPM triggers the proliferation of diseased cells both in liver tissues with cirrhosis and in tumor tissues at a later stage.

***SPINK1 Gene**:* The protein encoded by this gene is a trypsin inhibitor secreted from pancreatic acinar cells into pancreatic juice. It is thought that it functions in the prevention of trypsin-catalyzed early zymogen activation within the pancreatic and pancreatic duct. Mutations in this gene are associated with hereditary pancreatitis and tropical calcific pancreatitis [8]. As a matter of fact, according to the results obtained in this study, it was determined that the expression level of the SPINK1 gene significantly increased in advanced stage liver diseases. Therefore, this and previous gene expression profile analysis studies on liver diseases show that the high level of SPINK1 expression facilitates the formation and development of malignant cancer tumors.

As a result of the PPI network analysis, it was determined that the **ANLN, TOP2A and ASPM genes** have expressive interaction, and it is thought that any expression level change in one gene affects the expression levels of the other two genes.

CXCL14 Gene: For inflammation, infection, tissue damage, allergy, cardiovascular diseases, as well as cell migration and activation mediated by chemokines involved in malignant tumor pathophysiology, these molecules must bind to their specific receptors on the cell. CXCL14 gene is also one of these receptors [28]. The CXCL14 protein encoded by this gene is a member of the CXC chemokine family that regulates the development of new blood vessels [29]. CXCL14 gene expressed in many normal cells; It is either not expressed at all or expressed at a very low level in the tissues of individuals with diseases such as head and neck cancer (HNSCC), prostate and pancreatic cancer [10]. Similarly, in this study, it was determined that the expression level of the CXCL14 gene is significantly reduced in liver diseases such as cirrhosis and cancer. Thus, it can be concluded that the protein encoded by this gene has lost its function and failed to prevent the spread of the tumor.

LINC01093 Gene: This gene is located in a long, non-protein coding RNA region, and the expression level changes in this gene play an important role in the development phase of HCC disease [11]. In this study, it was determined that the expression level of the LINC01093 gene decreased significantly in liver diseases such as cirrhosis and cancer.

***OIT3 gene**:* This gene performs calcium ion binding and protein binding functions in molecular terms and encodes a protein that may be a potential negative biomarker specific to the liver [30]. In this study, it was determined that the expression level of the OIT3 gene decreased significantly in liver diseases such as cirrhosis and cancer. Therefore, the protein encoded by this gene is insufficient to perform its function, and this leads to the proliferation of unhealthy tissues in the liver.

***CLEC4G gene**:* CLEC4G is one of the 3 important C-type lectin protein types that have the ability to assist in recognizing viral and bacterial pathogens [31]. C-type lectins are calcium-dependent glycan-binding proteins and perform various immune functions, including inflammation and immunity of tumor / virally infected cells [15]. However, in this study, it was determined that the expression level of the CLEC4G gene in the tissues of individuals with cirrhosis and HCC was quite low, compared to healthy individuals. Therefore, the protein encoded by this gene is insufficient to perform its function, and this leads to the proliferation of unhealthy tissues in the liver.

THRSP gene: THRSP is a nuclear protein that is primarily important for the regulation of lipid (fat) metabolism. It is primarily expressed in tissues synthesizing fatty acids such as breast and liver [32]. In this study, it was determined that the expression level of the THRSP gene is significantly decreased in liver diseases such as cirrhosis and cancer. This observation indicates that the liver deviates from the healthy functioning of lipid (fat) metabolism in advanced

diseases.

***APOF gene**:* APOF, which is a lipid transfer inhibitory protein, is an important cholesteryl ester (esterified cholesterol) transfer protein activity regulator [33]. In this study, it was observed that the expression level of the APOF gene in cirrhotic and tumor cells decreased about 2 to 3 times compared to healthy cells. This is a sign that lipid metabolism and accordingly cholesterol regulation may be impaired in advanced stage diseases of the liver.

CLTRN gene: This gene encodes the protein that performs amino acid transfer regulation function [32]. In this study, it was determined that the CLTRN gene in the cells of individuals with cirrhosis and HCC is expressed at a very low level compared to the CLTRN genes in normal cells. Therefore, significantly low expression of this gene is thought to contribute to the progression of the disease in advanced stage liver diseases.

***FCN3 gene**:* FCN3, which is accepted as a serum lectin protein, specifically mediates the execution of apoptotic (programmed) cell death process [34]. In this study, it was determined that FCN3 gene, located in the cells of individuals with cirrhosis and HCC, is expressed at a very low level compared to FCN3 genes in normal cells. This suggests that FCN3 protein may be inadequate to fulfill its function, leading to an increase in tumor cells in the advanced stage, thus causing the disease to progress.

In the first data set (GSE29721 - cancerous and normal samples), it was observed that the log-fold change ratios increased to the range of 4.5 - 5, but in the second data set (GSE17548 - cirrhosis and cancer samples) these rates were mostly in the range of 2.5 - 3. This result shows that the expression levels of the genes of people with liver disease are changed more after they reach the cirrhosis stage.

4. RESULTS AND DISCUSSION

In this study, we intended to evaluate diseases that occur in the human liver and definitive treatment is not possible. We also tried to understand the phases of liver failure via gene expressions. We wanted to discover the gene expressions alterations during the phases of healthy tissue, HBV, HCV, cirrhosis and HCC patients. The results can be utilized as a biomarker for predicting the phase of liver disease. Also presenting overexpressed/underexpressed genes can pave path for targeted therapy future studies.

In this study, we analyzed molecular changes occurring in the body in advanced stage diseases such as cirrhosis and HCC. We aim at cellular, biological and physiological problems caused by these changes and target proteins in order to prevent them. For this purpose, the microarray method, which is frequently used in genomic data analysis, was preferred. We research significant gene expression differences between the patient and control groups and aim to determine the problems caused by these differences in the related pathways.

We need to mention that there exist other researches that utilize GSE29721 or GSE17548 datasets [35-38]. However, our study differs from these studies primarily in terms of purpose and perspective. Our goal is not only focus on a specific disease. But also, we search genes that contributes to various liver diseases. Furthermore, we consider gene expression and gene interactions during liver disease phases of healthy tissue, HBV, HCV, cirrhosis and HCC.

As a result of the findings, it was determined that gene expression level alterations occur after the onset of cirrhosis. In order to monitor the treatment and prognosis of the patient, it is very important to intervene in the course of the disease before HCC tumor formation in the liver tissues.

Studies on therapeutic targets that cause impairment of liver functions can be used to improve treatment success, and clinical studies should be conducted to verify the findings obtained from microarray data analysis. So that we can ensure their applicability in diagnosis and treatment. Since the analyses of this study are carried out in a computer environment using microarray data, the results of the analysis should also be tested and verified biologically.

CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

[1] Hayran, A., "Content-based search in gene expression databases", Master's Thesis, Baskent University Institute of Science and Engineering, Ankara, 1-10, (2014).

[2] Atis, S., "Microarray analysis of gene expression data in breast cancer", Master's Thesis, Uludağ University Institute of Health Sciences, Bursa, 1-11, (2014).

[3] Lian, Y.F., Huang, Y.L., Wang, J.L., Deng, M.H., Xia, T.L., Zeng, M.S., Chen, M.S., Wang, H.B., and Huang, Y.H., "Anillin is required for tumor growth and regulated by miR-15a/miR-16-1 in HBV-related hepatocellular carcinoma", Aging, 10(8): 1884-1901, (2018).

[4] Satow, R., Shitashige, M., Kanai, Y., Takeshita, F., Ojima, H., Jigami, T., Honda, K., Kosuge, T., Ochiya, T., Hirohashi, S., and Yamada, T., "Combined functional genome survey of therapeutic targets for hepatocellular carcinoma", Clinical Cancer Research, 16(9): 2518-2528, (2010).

[5] Wong, N., Yeo, W., Wong, W.L., Wong, N.L.Y., Chan, K.Y.Y., Mo, F.K.F., Koh, J., Chan, S.L., Chan, A.T.C., Lai, P.B.S., Ching, A.K.K., Tong, J.H.M., Ng, H.K., Johnson, P.J., and To, K.F., "TOP2A overexpression in hepatocellular carcinoma correlates with early age onset shorter patients survival and chemoresistance", International Journal of Cancer, 124: 644–652, (2009).

[6] Gao, X., Wang, X., and Zhang, S., "Bioinformatics identification of crucial genes and pathways associated with hepatocellular carcinoma", Bioscience Reports, 38(6): 1-8, (2018).

[7] Horvath, S., Zhang, B., Carlson, M., Lu, K.V., Zhu, S., Felciano, R.M., Laurance, M.F., Zhao, W., Qi, S., Chen, Z., Lee, Y., Scheck, A.C., Liau, L.M., Wu, H., Geschwind, D.H., Febbo, P.G., Kornblum, H.I., Cloughesy, T.F., Nelson, S.F., and Mischel, P.S., "Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target", Pnas, 103(46): 17402-17407, (2006).

[8] Li, F., Liu, T., Xiao, C.Y., Yu, J.X., Lu, L.G., and Xu, M.Y., "FOXP1 and SPINK1 reflect the risk of cirrhosis progression to HCC with HBV infection", Biomedicine & Pharmacotherapy, 72: 103-108, (2015).

[9] Marshall, A., Lukk, M., Kutter, C., Davies, S., Alexander, G., and Odom, D.T, "Global gene expression profiling reveals SPINK1 as a potential hepatocellular carcinoma marker", Plos One, 8(3): e59459, (2013).

[10] Wang, W., Huang, P., Zhang, L., Wei, J., Xie, Q., Sun, Q., Zhou, X., Xie, H., Zhou, L., and Zheng, S., "Antitumor efficacy of C-X-C motif chemokine ligand 14 in hepatocellular carcinoma in vitro and in vivo", Cancer Science, 104(11): 1523-1531, (2013).

[11] Esposti, D.D., Vargas, H.H., Voegele, C., Jimenez, N.F., Forey, N., Bancel, B., Calvez-Kelm, F.L., McKay, J., Merle, P., and Herceg, Z., "Identification of novel long non-coding RNAs deregulated in hepatocellular carcinoma using RNA-sequencing", Oncotarget, 7(22): 31862-31877, (2016).

[12] Jovel, J., Lin, Z., O'keefe, S., Willows, S., Wang, W., Zhang, G., Patterson, J., Moctezuma-Velazquez, C., Kelvin, D.J., Wong, G.K.S., and Mason, A.L., "A survey of molecular heterogeneity in hepatocellular carcinoma", Hepatology Communications, 2(8): 945-959, (2018).

[13] Dai, M., Chen, S., Wei, X., Zhu, X., Lan, F., Dai, S., and Qin, X., "Diagnosis prognosis and bioinformat ics analysis of lncRNAs in hepatocellular carcinoma", Oncotarget, 8(56): 95799-95809, (2017).

[14] Agarwal, R., Narayan, J., Bhattacharyya, A., Saraswat, M., and Tomar, A.K., "Gene expression profiling, pathway analysis and subtype classification reveal molecular heterogeneity in hepatocellular carcinoma and suggest subtype specific therapeutic targets", Cancer Genetic, 216-217: 37-51, (2017).

[15] Ho, D.W.H., Kai, A.K.L., and Ng, I.O.L., "TCGA whole-transcriptome sequencing data reveals significantly dysregulated genes and signaling pathways in hepatocellular carcinoma", Frontiers of Medicine, 9(3): 322-330, (2015).

[16] Chen, J., Qian, Z., Li, F., Li, J., and Lu., Y., "Integrative analysis of microarray data to reveal regulation patterns in the pathogenesis of hepatocellular carcinoma", Gut and Liver, 11(1): 112-120, (2017).

[17] Huang, Y.Y., Gusdon, A.M., and Qu, S., "Cross-talk between the thyroid and liver: A new target for nonalcoholic fatty liver disease treatment", World Journal of Gastroenterology, 19(45): 8238- 8246, (2013).

[18] Cheng, P., Cheng, Y., Su, M.X., Li, D., Zhao, G.Z., Gao, H., Li, Y., Zhu, J.Y., Li, H., and Zhang, T., "Bicluster and pathway enrichment analysis of HCV-induced cirrhosis and hepatocellular carcinoma", Asian Pacific Journal of Cancer Prevention, 13: 3741-3745, (2012).

[19] Re, L.O., Douet, J., Buschbeck, M., Fusilli, C., Pazienza, V., Panebianco, C., Castracani, C.C., Mazza, T., Volti, G.L., and Vinciguerra, M., "Histone variant macroH2A1 rewires carbohydrate and lipid metabolism of hepatocellular carcinoma cells towards cancer stem cells", Epigenetics, 13(8): 829-845, (2018).

[20] Chang, Q., Chen, J., Beezhold, K.J., Castranova, V., Shi, X., and Chen, F., "JNK1 activation predicts the prognostic outcome of the human hepatocellular carcinoma", Molecular Cancer, 8(64): 1-14, (2009).

[21] Udhaya Kumar, S., Thirumal Kumar, D., Siva, R., George Priya Doss, C., and Zayed, H., "Integrative bioinformatics approaches to map potential novel genes and pathways involved in ovarian cancer", Frontiers in Bioengineering and Biotechnology, 7(391): (2019).

[22] Wan, J., Jiang S., Jiang, Y., Ma, W., Wang, X., and Cui, R., "Data mining and expression analysis of differential lncRNA ADAMTS9-AS1 in prostate cancer", Frontiers in Genetics, 10(1377): (2020).

[23] Udhaya Kumar, S., Thirumal Kumar, D., Siva, R., George Priya Doss, C., Younes, S., Younes, N., Sidenna, M., and Zayed, H., "Dysregulation of signaling pathways due to differentially expressed genes from the b-cell transcriptomes of systemic lupus erythematosus patients - A bioinformatics approach", Frontiers in Bioengineering and Biotechnology, 8(276): (2020).

[24] Fu, D., Zhang, B., Yang, L., Huang, S., and Xin, W., "Development of an immune-related risk signature for predicting prognosis in lung squamous cell carcinoma", Frontiers in Genetics, 11(978): (2020).

[25] https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi. Access date: 19.11.2018

[26] Alexa, A., Rahnenführer, J. and Lengauer, T., "Improved scoring of functional groups from gene expression data by decorrelating GO graph structure", Bioinformatics, 22(13): 1600-1607, (2006).

[27] https://string-db.org Access date: 27.11.2019

[28] Citak, F.E., Citak, E.C., and Karadeniz, C., "Chemokines and diseases", Turkiye Klinikleri Journal of Medical Sciences, 22(2): 210-216, (2002).

[29] Strieter, R.M., Burdick, M.D., Mestas, J., Gomperts, B., Keane, M.P., and Belperio, J.A., "Cancer CXC chemokine networks and tumour angiogenesis", European Journal of Cancer, 42(6): 768-778, (2006).

[30] Xu, Z.G., Du, J.J., Zhang, X., Cheng, Z.H., Ma, Z.Z., Xiao, H.S., Yu, L., Wang, Z.Q., Li, Y.Y., Huo, K.K., and Han, Z.G., "A novel liver-specific zona pelluc ida domain containing protein that is expressed rarely in hepatocellular carcinoma", Hepatology, 38(3): 735-744, (2003).

[31] Huang, Y.W., and Meng, X.J., "Identification of a porcine DC-SIGN-related C-type lectin, porcine CLEC4G (LSECtin), and its order of intron removal during splicing: Comparative genomic analyses of the cluster of genes CD23/CLEC4G/DC-SIGN among mammalian species", Developmental and Comparative Immunology, 33(6): 747-760, (2009).

[32] Kuemmerle, N.B., and Kinlaw, W.B., "THRSP (thyroid hormone responsive)", Atlas of Genetics and Cytogenetics in Oncology and Haematology, 15(6): 480-482, (2011).

[33] Morton, R.E., Gnizak, H.M., Greene, D.J., Cho, K.H., and Paromov, V.M., "Lipid transfer inhibitor protein (apolipoprotein F) concentration in normolipidemic and hyperlipidemic subjects", Journal of Lipid Research, 49: 127-135, (2008).

[34] Yu, S.J., Kim, H., Min, H., Sohn, A., Cho, Y.Y., Yoo, J.J., Lee, D.H., Cho, E.J., Lee, J.H., Gim, J., Park, T., Kim, Y.J., Kim, C.Y., Yoon, J.H., and Kim, Y., "Targeted proteomics predicts a sustained complete-response after transarterial chemoembolization and clinical outcomes in patients with hepatocellular carcinoma: a prospective cohort study", Journal of Proteome Research, 16(3): 1239-1248, (2017).

[35] Yue, C., Ren, Y., Ge, H., Liang, C., Xu, Y., Li, G., and Wu, J., "Comprehensive analysis of potential prognostic genes for the construction of a competing endogenous RNA regulatory network in hepatocellular carcinoma", OncoTargets and Therapy, 12: 561-576, (2019).

[36] Jiang, C.H., Yuan, X., Li, J.F., Xie, Y.F., Zhang, A.Z., Wang, X.L., Yang, L., Liu, C.X., Liang, W.H., Pang, L.J., Zou, H., Cui, X.B., Shen, X.H., Qi, Y., Jiang, J.F., Gu, W.Y., Li, F., and Hu, J.M., "Bioinformatics-based screening of key genes for transformation of liver cirrhosis to hepatocellular carcinoma", Journal of Translational Medicine, 18(40): (2020).

[37] Chen, Q.F., Xia, J.G., Li, W., Shen, L.J., Huang, T., and Wu, P., "Examining the key genes and pathways in hepatocellular carcinoma development from hepatitis B virus-positive cirrhosis", Molecular Medicine Reports, 18(6): 4940-4950, (2018).

[38] He, B., Yin, J., Gong, S., Gu, J., Xiao, J., Shi, W., Ding, W., and He, Y., "Bioinformatics analysis of key genes and pathways for hepatocellular carcinoma transformed from cirrhosis", Medicine, 96(25): (2017).