J. Fac. Pharm. Ankara / Ankara Ecz. Fak. Derg., 47(1): 239-249, 2023
Doi: 10.33483/jfpau.1205775

ORIGINAL ARTICLE / ÖZGÜN MAKALE

INTEGRATED BIOINFORMATIC ANALYSIS TO EVALUATE TARGET GENES AND PATHWAYS IN CHRONIC LYMPHOCYTIC LEUKEMIA

ENTEGRE BİYOİNFORMATİK ANALİZ İLE KRONİK LENFOSİTİK LÖSEMİDE HEDEF GENLERİN VE YOLAKLARIN BELİRLENMESİ

Buket ALTINOK GUNES[1]*

[1]Ankara University, Vocational School of Health Services, Department of Medical Laboratory Techniques, 06290, Ankara, Turkey

ABSTRACT

Objective: *The most common type of leukemia, chronic lymphocytic leukemia (CLL), is characterized by progressive accumulation of monoclonal B cells with a specific immunophenotype in the blood, bone marrow, and lymphoid organ. The goal of this research was to use bioinformatic analysis to comprehend the molecular mechanisms causing CLL and to investigate potential targets for the diagnosis and therapy of CLL.*

Material and Method: *Expression data from CLL patients with accession numbers GSE22529 and GSE26725 were downloaded from the GEO database for bioinformatic analysis. GSE22529 data was studied with samples from 41 CLL patients and 11 healthy groups, while GSE26725 data was studied with blood samples from 12 CLL patients and 5 healthy groups. GEO2R was used to find differentially expressed genes (DEGs) in CLL patient samples and healthy control samples. The DAVID program was used to perform GO and KEGG enrichment analyses on DEGs. Using the Cytoscape software, a protein-protein interaction (PPI) network was created, and hub genes associated with CLL were identified.*

Result and Discussion: *DEGs with p 0.05 and log2FC 0, log2FC>0 were chosen after analysis with GEO2R. In the GSE22529 dataset, 942 genes had higher expression levels in CLL patients compared with controls, while the expression of 1007 genes decreased. In the GSE26725 dataset, CLL patients had lower expression levels for 916 genes compared with controls, while 939 genes showed an increase in expression. 229 DEGs with higher expression levels and 308 DEGs with lower expression levels were found in both sets of data. It has been observed that these common genes, whose expression has changed, are enriched in protein processing in the ER, Chemokine, B-cell receptor, T-cell receptor, protein export pathways. Additionally, DDOST, RPL18, RPL18A,*

* **Corresponding Author / Sorumlu Yazar:** Buket Altinok Gunes
e-mail / e-posta: baltinok@ankara.edu.tr, **Phone / Tel.:** +903123573242/6642

Submitted / Gönderilme : 16.11.2022
Accepted / Kabul : 09.12.2022
Published / Yayınlanma : 20.01.2023

RPL19, RPL31, GNB3, GNB4, GNG11, GNGT1, NEDD8, UBE2M RBX1, FBXO21, SKP1, KLHL9 and CAND1 were identified as the most important genes. Our study's findings demonstrated that newly discovered genes and pathways may be candidates for CLL biomarkers that can be used for both the diagnosis and drug treatment of the disease.
Keywords: *Bioinformatic analysis, CLL, GEO, microarray, network module*

ÖZ

Amaç: *En yaygın lösemi türü olan kronik lenfositik lösemi (KLL), kanda, kemik iliğinde ve lenfoid organda spesifik bir immünofenotipe sahip monoklonal B hücrelerinin ilerleyici birikimi ile karakterize edilir. Bu çalışmanın amacı, biyoinformatik analiz yaparak KLL'nin altında yatan moleküler mekanizmaları araştırmak, KLL tanı ve tedavisi için potansiyel hedefleri belirlemektir.*
Gereç ve Yöntem: *Biyoinformatik analiz için GEO veri tabanından KLL hasta verilerine ait GSE22529 ve GSE26725 erişim numaralarına sahip ekspresyon dataları indirildi. GSE22529 numaralı data 41 KLL hasta ve 11 sağlıklı gruptan, GSE26725 numaralı data ise 12 KLL ve 5 sağlıklı gruptan alınan kan örnekleri ile çalışılmıştır. KLL hasta örnekleri ile sağlıklı kontrol örnekler farklı şekilde ifade edilen genleri (DEGs) bulabilmek için GEO2R ile analiz edildi. DEG'ler için DAVID programı kullanılarak GO ve KEGG zenginleştirme analizleri gerçekleştirildi. Cytoscape yazılımı kullanılarak protein-protein etkileşim (PPI) ağı oluşturuldu ve KLL ile ilişkili önemli genler tesbit edildi.*
Sonuç ve Tartışma: *GEO2R ile analiz sonrası $p<0.05$ ve $log_2FC<0$, $log2FC>0$ olan DEG'ler seçildi. GSE22529 veri setinde KLL hastalarında kontrol grubuna göre 942 genin ifadesi artmış, 1007 genin ifadesi azalmıştır. GSE26725 veri setinde ise KLL hastalarında kontrol grubuna göre 939 genin ifadesi artmış, 916 genin ifadesi azalmıştır. Her 2 veri seti için ortak olarak ifadesi artan 229, ifadesi azalan 308 DEG tanımlanmıştır. İfadesi değişen bu ortak genlerin ER'de protein işlenmesi, kemokin, B-hücre reseptör, T-hücre reseptör, protein taşıma gibi yolaklarda zenginleştiği görülmüştür. Buna ek olarak DDOST, RPL18, RPL18A, RPL19, RPL31, GNB3, GNB4, GNG11, GNGT1, NEDD8, UBE2M RBX1, FBXO21, SKP1, KLHL9 and CAND1 en önemli genler olarak belirlenmiştir. Çalışmamızın sonucu, ortaya çıkan genlerin ve yolakların KLL'nin tanısında ve ilaç tedavisinde kullanılabilecek birer biyobelirteç adayı olabileceğini göstermiştir.*
Anahtar Kelimeler: *Ağ modülü, biyoinformatik analiz, GEO, KLL, mikrodizin*

INTRODUCTION

In western countries, chronic lymphocytic leukemia is the most common type of leukemia (CLL). And more than 15,000 new cases have been reported, with approximately 4500 deaths. The average diagnosed age is 72. Men suffer more than women [1,2]. A characteristic of CLL, a clonal proliferative tumor of mature CD5-positive B lymphocytes, is an expansion of lymphocytes in the bone marrow, lymph nodes, spleen, and peripheral circulation [2]. Unfortunately, it is still unclear how CLL develops at the molecular level, which prevents early diagnosis and prompt treatment. The molecular mechanism and biomarkers of CLL development have thus been the main focus of CLL research in order to increase the probability of a disease identification and course of treatment. Many genes and cellular networks have been shown to influence the origin and progression of CLL, according to studies. Computer-based technologies, such as molecular dynamics simulation, have advanced quickly in recent years, leading to the progressive discovery of novel molecular causes of the disease. The combination of web-based systems like GEO (Gene Expression Omnibus), DAVID, data mining techniques (cancergenome.nih.gov), and computer-based analysis methods has led to the discovery of many highly sensitive and specific markers [3-6]. Gene chips are frequently used as a gene detection technology and relevant data are kept in public databases. This genomic data can be integrated and reanalyzed to potentially find specific disease-related biomarkers [7]. In clinical research applications like categorization, prognosis prediction and the identification of novel targeted therapies, the use of high-throughput platforms in gene expression has become increasingly important recently [8-10].

The original gene microarray datasets (GSE22529 and GSE26725) used in this study were obtained from the NCBI-GEO (National Center for Biotechnology Information-Gene Expression Omnibus database). Differentially expressed genes (DEGs) were identified in these samples by comparing CLL patient and control samples using R software and Bioconductor. To determine the

functional enrichment analysis of DEGs, the DAVID software program was used to analyze the Gene Ontology (GO), Kyoto gene and genome encyclopedia (KEGG) pathway and protein-protein interaction (PPI) network [11].

The novel biomarkers and pathways found in this study are predicted to provide information on possible molecular mechanisms underlying CLL. The newly discovered key targets will provide new suggestions and approaches for CLL early detection and individualized treatment.

MATERIAL AND METHOD

Affymetrix Microarray Data

We scanned the GEO database for the transcriptome datasets of patients with CLL. The National Center for Biotechnology Information's-Gene Expression Omnibus (NCBI-GEO), a freely accessible database of functional genomics, was used to download microarray datasets with accession numbers GSE22529 and GSE26725. 41 CLL patients and 11 control blood samples were included in the study with the accession number GSE22529 that was performed on the Affymetrix Human Genome U133A Array platform [HG-U133A]. Additionally, 12 CLL patients and 5 control blood samples were included in the study with accession number GSE26725 that was conducted on the Affymetrix Human Genome U133 Plus 2.0 Array platform [HG-U133 Plus 2]. The platforms used along with the patient and control numbers are listed in Table 1.

Table 1. Microarray platforms

GEO access number	CLL Patient number	Control number	Platform
GSE 22529	41	11	[HG-U133A] Affymetrix Human Genome U133A Array
GSE26725	12	5	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
Total	**53**	**16**	

Identification of DEG Analysis

To find differentially and co-expressed genes in the two datasets for the analysis of differential expression, the online analysis program GEO2R was utilized. A global public resource for high-throughput microarray and next-generation array functional genomic datasets is called GEO (http://www.ncbi.nlm.nih.gov/geo/). This data is supplied by the research community. GEO2R is an R-based web application that analyzes GEO data and allows users to compare different groups to calculate "DEG" [12]. On the preprocessed microarray data, the log2 fold change (log_2FC) was calculated. t-tests were used to compute p-values and adjusted p-values. $p< 0.05$ and $|log2FC|> 0$ criteria were considered for significant DEG screening. DEGs were screened and ranked by importance. Up-regulated genes ($log_2FC>0$) and down-regulated ($log_2FC<0$) were classified on the basis of the fold change. The intersection of two datasets was used to create a Venn diagram using the online tool Venny 2.1 and determine commonly varying DEGs. DEGs that were statistically significant in both datasets were compared, and common DEGs were used in the study's subsequent analysis.

Gene Ontology (GO) and Pathway Enrichment Analyses

DAVID 6.8 was used to carry out the functional enrichment analysis of common DEGs. (david.abcc.ncifcrf.gov/). An online tool for high-throughput analysis of gene function derived from genomic experiments is called Annotation, Visualization and Integrated Discovery Database (DAVID) [13]. Functional categories, Gene Ontology (GO) terms, and Kyoto Encyclopedia of Genes and Genes

(KEGG) pathways are all analyzed using DAVID. To predict the functions of proteins, GO analysis was divided into three categories: biological process (BP), cellular component (CC), and molecular function (MF). To create networks of molecular interactions, reactions, and associations, A KEGG pathway analysis was utilized to locate sets of DEGs in relation to particular pathways [14]. A statistically significant difference was defined as $p<0.05$.

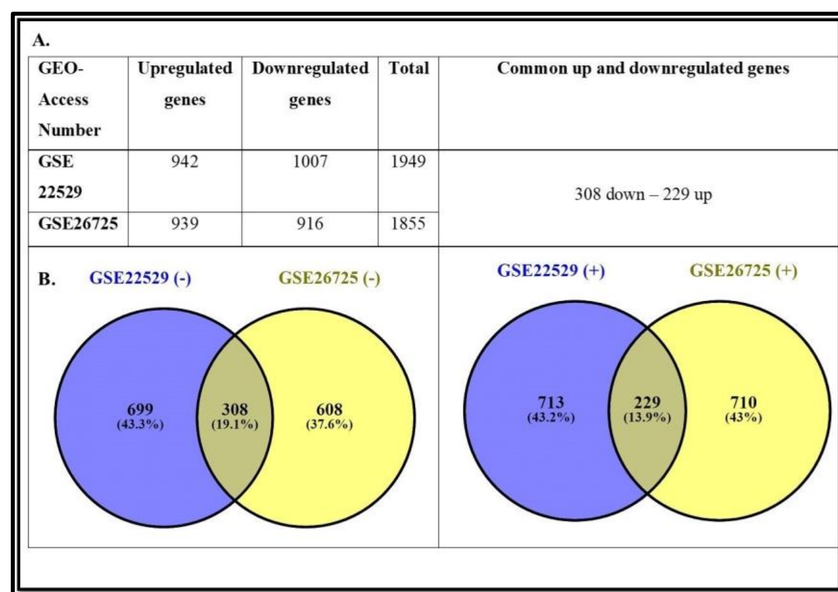
Construction of Protein-protein Interaction Network (PPI) and Hub Gene Screening

A database of confirmed and expected protein-protein connections called STRING (https://string-db.org) was used to analyze PPI networks. These connections involve functional and physical relationships, and data for these interactions are primarily derived from common expression networks, automated text mining, high-throughput experiments, and computational predictions. DEGs were mapped to the PPI network with a threshold of >0.9 interaction score. Furthermore, the PPI network was visualized and created using Cytoscape v3.6.0 software. Hub nodes were those nodes that had the most communication with their neighbors. The gene network clustering analysis was carried out using the Cytoscape software package's MCODE application to determine important PPI network modules. Cutoff degree = 2, node score cutoff = 0,2, k-core=2, and maximum depth=100 are the qualifying criteria for MCODE functional modules.

RESULT AND DISCUSSION

DEG Analysis and Identification of Common DEGs

Genes with $p< 0.05$ and $|\log2\ FC|>0$ were chosen because of transcriptome analyses in the GSE22529 and GSE26725 datasets. The expression of 1949 genes changed (942 increased-1007 decreased) in the GSE22529 dataset, while the expression of 1855 genes changed (939 increased-916 decreased) in the GSE26725 dataset. When the DEGs found in both datasets were compared, 537 genes with co-variation corresponding to protein-coding genes were discovered; 308 of them showed decreased expression and 229 showed increased expression, indicating that they were expressed in the same direction in both data sets. Figure 1A shows the most increasing and decreasing gene lists from the data of common DEGs. In both datasets, the differential expression of common DEGs, gene counts, and venn diagrams in CLL patients compared with controls are displayed in Figure 1A and 1B. ($p< 0.05$ and $|\log2\ FC| >0$). In Figure B, different color regions indicate different microarray data profiles.


A.

GEO-Access Number	Upregulated genes	Downregulated genes	Total	Common up and downregulated genes
GSE 22529	942	1007	1949	308 down – 229 up
GSE26725	939	916	1855	

Figure 1. Venn diagrams of common DEGs of GEO datasets

Tables 2 and 3 list the 20 genes with the highest increase and decrease in both the data sets.

Table 2. The top 20 upregulated and downregulated genes for GSE22529

Gen	log_2(FC)	Gen	log_2(FC)
FMOD	5.61	**IGLJ3**	-6.787
FRY	4.409	**S100A8**	-6.758
ABCA6	4.409	**FOS**	-6.241
ARHGAP44	4.321	**LYZ**	-5.874
IGFBP4	4.255	**FCN1**	-5.708
ROR1	4.092	**S100A9**	-5.589
COL9A2	3.581	**IGHV4-31**	-5.521
TTN	3.549	**HBB**	-5.358
SPATS2L	3.542	**IGHV3-23**	-5.228
SERPINF1	3.535	**IGHD**	-5.221
INPP5F	3.307	**CD14**	-5.122
PHTF1	3.222	**HBA2///HBA1**	-5.115
DMD	3.205	**MLIP**	-5.002
COL9A3	3.103	**CSGALNACT1**	-4.947
CCDC88A	3.027	**CST3**	-4.843
VIPR1	3.022	**CEBPD**	-4.76
SGPP1	2.982	**SORL1**	-4.719
NOSIP	2.94	**ITGB2**	-4.401
LEF1	2.911	**HCK**	-4.21
PDGFD	2.782	**SERPINA1**	-4.117

Table 3. The top 20 upregulated and downregulated genes for GSE26725

Gen	log_2 (FC)	Gen	log_2 (FC)
ABCA6	5.113	**IGLJ3**	-6.592
PIGR	4.839	**JCHAIN**	-5.441
ARHGAP44	4.414	**SCN3A**	-5.069
PHEX	4.286	**BMS1P20**	-5.007
ENPP2	4.216	**SSPN**	-4.777
CLNK	4.034	**IGLL5**	-4.777
CCDC88A	3.879	**TSPAN13**	-4.643
IL17RB	3.863	**IGK///IGKC**	-4.486
ROR1	3.827	**IGHM**	-4.436
SPATS2L	3.478	**PPP1R14A**	-4.25
ADTRP	3.295	**MLIP**	-4.219
ABCA9	3.165	**PTPRK**	-3.934
MIR155	3.067	**CORO2B**	-3.815
FILIP1L	3.044	**NRIP1**	-3.58
AKAP12	2.748	**TSPYL5**	-3.579
LRRN1	2.736	**IGKV1OR2**	-3.554
IGSF3	2.711	**CST3**	-3.458
SGSM1	2.706	**CD1C**	-3.423
KSR2	2.658	**BHLHE41**	-3.348
RAPGEF3	2.653	**SLC38A11**	-3.336

GO and KEGG Pathway Analysis of DEGs

In our study, functional and pathway enrichment analyses of DEGs, which are common expressions in both data sets, were performed using DAVID software. Down-regulated DEGs were found in 14 clusters (enrichment Score>1.3) and were highly enriched in various aspects in the 'molecular function (MF),' 'biological process (BP),' and 'cellular component (CC)' groups, according to GO analysis. Up-regulated DEGs were found in only two clusters (enrichment Score >1.3): molecular function (MF) and biological process (BP). Table 4 shows GO terms enriched with $p<0.05$ and a list of up-regulated genes. According to KEGG pathway analyses, down-regulated genes were significantly enriched in 10 pathways, particularly "Protein processing in the endoplasmic reticulum, Chemokine signaling pathway, hematopoietic cell lineage, transcriptional misregulation in cancer," whereas upregulated genes were enriched in 3 pathways (Table 5).

Table 4. GO analysis of common DEGs in CLL

Downregulated Genes			
Category	**Pathways**	**Count**	**Pvalue**
GOTERM_BP_DIRECT	GO:0006468~protein phosphorylation	20	5.88852E-05
	GO:0045087~innate immune response	13	0.0287027
	GO:0098609~cell-cell adhesion	12	0.0026174
	GO:0002250~adaptive immune response	10	0.0003895
	GO:0006958~complement activation, classical pathway	8	0.0007098
	GO:0050776~regulation of immune response	8	0.0177295
	GO:0038095~Fc-epsilon receptor signaling pathway	8	0.0177295
	GO:0050853~B cell receptor signaling pathway	7	0.000148
	GO:0006486~protein glycosylation	6	0.0273621
	GO:0042110~T cell activation	6	0.0006589
	GO:0050871~positive regulation of B cell activation	5	0.0005624
GOTERM_CC_DIRECT	GO:0005913~cell-cell adherens junction	15	0.0003103
	GO:0005856~cytoskeleton	14	0.0033875
	GO:0034663~endoplasmic reticulum chaperone complex	4	0.0004703
	GO:0042571~immunoglobulin complex, circulating	4	0.0025327
GOTERM_MF_DIRECT	GO:0005524~ATP binding	33	0.0281698
	GO:0003779~actin binding	16	2.25E-05
	GO:0004674~protein serine/threonine kinase activity	15	0.0018041
	GO:0098641~cadherin binding involved in cell-cell adhesion	13	0.0015797
	GO:0003823~antigen binding	7	0.0047291
	GO:0016757~transferase activity, transferring glycosyl groups	7	0.0003103
	GO:0034987~immunoglobulin receptor binding	6	3.84E-05
	GO:0035091~phosphatidylinositol binding	6	0.0093082
Upregulated Genes			
Category	**Pathways**	**Count**	**Pvalue**
GOTERM_BP_DIRECT	GO:0016569~covalent chromatin modification	7	0.0025089
	GO:0035335~peptidyl-tyrosine dephosphorylation	5	0.0320941
GOTERM_MF_DIRECT	GO:0042393~histone binding	9	0.0001137
	GO:0004725~protein tyrosine phosphatase activity	5	0.0325846

Table 5. KEGG pathway analysis of common DEGs in CLL

DEGs	Term ID	Description	Count	Pvalue
Down-regulated	hsa04141	Protein processing in endoplasmic reticulum	13	8.85E-05
	hsa04062	Chemokine signaling pathway	10	0.0092848
	hsa04640	Hematopoietic cell lineage	8	0.0013335
	hsa05202	Transcriptional misregulation in cancer	8	0.0411960
	hsa04662	B cell receptor signaling pathway	7	0.0019847
	hsa05221	Acute myeloid leukemia	6	0.0042073
	hsa04666	Fc gamma R-mediated phagocytosis	6	0.0222590
	hsa04660	T cell receptor signaling pathway	6	0.0427175
	hsa04720	Long-term potentiation	5	0.0374360
	hsa03060	Protein export	4	0.0092373
Up-regulated	hsa03013	RNA transport	7	0.0202688
	hsa04071	Sphingolipid signaling pathway	6	0.0168503
	hsa00071	Fatty acid degradation	4	0.0152482

Integrated Protein-protein Interaction (PPI) Network

In a comparison of both datasets, the PPI network structure of down-regulated and up-regulated genes in common was studied using the STRING database. The gene interaction network in down-regulated genes contains 290 nodes and 297 edges, according to the PPI analysis. DEGs are symbolized by nodes, while interactions between DEGs are symbolized by edges (Fig 2A). The STRING analysis showed that CASP10-TNFRSF10A interacted at the highest confidence interval. These genes were analyzed using Network Analyzer in Cytoscape software and the core genes were ranked based on their predicted scores. *DDOST, MAPK1, RPL18, RPL18A, GNB3, RPL19, GNB4, RPL31, GNG11, GNGT1* genes are the top 10 high order hub nodes. These genes were down-regulated. Among these genes, *DDOST* has the highest node degree (=15). PPI analysis of up-regulated genes resulted in 239 nodes and 173 edges in the gene interaction network (Fig 2B). The top 10 hub nodes with high order in the upregulated genes are the genes *RBX1, SKP1, UBE2M, FBXO21, KLHL9, JAK2, IL2, KRAS, CAND1, NEDD8.* The *RBX1* gene has the highest node degree (=19) among these genes. MCODE was then used in Cytoscape software to screen the gene interaction network modules, and Figures 3A-3B show the 4 modules with the highest MCODE score.

In this study, we compared the expression of 2 microarray datasets including CLL patients and healthy controls to identify Differentially Expressed Genes (DEG). On the basis of the gene expression profile, we discovered 537 common DEGs (229 up and 308 down). Then, we investigated these DEGs thoroughly using bioinformatics techniques, such as GO and KEGG pathway enrichment studies, PPI network analysis, and hub gene selection (Fig 2, 3).

It has been indicated that the genes whose expression changes are grouped in ER protein processing, chemokine signaling, hematopoietic cell lineage, transcriptional dysregulation in cancer, and B cell receptor signaling pathways, RNA transport, Sphingolipid signaling, and fatty acid degradation pathways.

One of the significant post-translational changes of proteins known as glycosylation is essential for many cellular biological functions and is closely linked to a variety of pathological occurrences, such as the development of tumors and inflammatory reactions. Polysaccharides or glycoproteins associated with tumors may result from changes in glycosylation, which may act as tumor markers and impact the development and prognosis of tumors. DDOST (dolichyl-diphosphooligosaccharide-protein glycosyltransferase) encodes a protein that is involved in the N-glycosylation of proteins and is a part of the oligosaccharide transferase (OST) complex [15]. Some researchers have shown that patients with skin squamous cell carcinomas that express a high percentage of DDOST have a bad prognosis [16]. Additionally, studies have revealed that DDOST is a separate prognostic factor in people with liver

cancer [17]. For the first time, DDOST, which has the highest degree of node among the hub genes with dramatically decreased expression and is enriched in the protein processing in the ER pathway, was linked to CLL in our study.

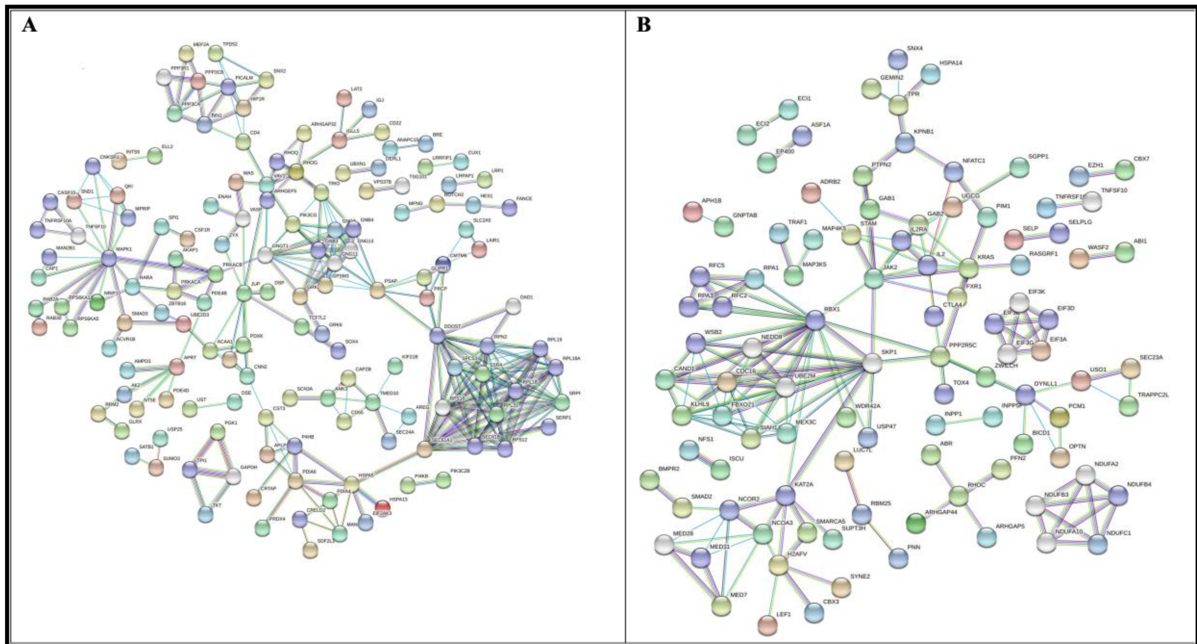


Figure 2: The protein interaction network (PPI) of identified genes.
A. Downregulated genes, Upregulated genes. Circle define DEGs, lines show protein interaction between DEGs.

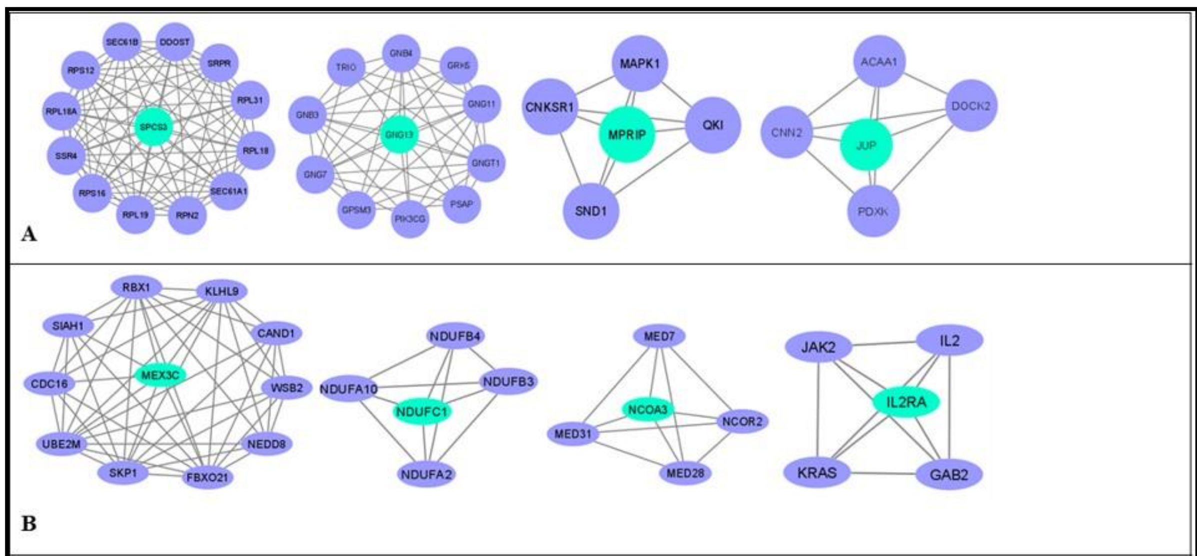


Figure 3: The top modules from the gene-gene interaction network of common DEGs.
A. Downregulated genes, B. Upregulated genes

The B cell receptor signaling pathway and related MAPK (mitogen-activated protein kinase) signaling are among the signaling pathways that are activated in the microenvironment of CLL by a

various stimulus. This illustrates the correlation between the control of the cell cycle, cell proliferation, and the instability of cancer cells, which is associated with the growth and spread of cancer cells [18]. In our research, the B cell receptor signaling pathway had enriched and differently elevated MAPK1 levels.

The ribosomal protein family (RP) is the other group in which the key genes in our study whose expression is lowered in CLL patients. Deficiencies in the ribosome biogenesis pathway are associated with an increased risk of having cancer since the ribosome is critical for the development of disorders [19]. Recent findings of somatic mutations in RP genes in hematological cancers (including T-cell acute lymphoblastic leukemia (ALL), chronic lymphocytic leukemia (CLL), and multiple myeloma) and solid tumors (such as breast cancer and melanoma) raise the potential that abnormalities in ribosome synthesis could facilitate oncogenic transformation. Additionally, RP gene deletions are extremely common in human malignancies [20]. The RPL18, RPL18A, RPL19, and RPL31 genes were discovered to be related to CLL for the first time in our study.

Another key pathway in our study was the chemokine signaling pathway. The development and longevity of CLL are largely dependent on relations with the tissue microenvironment. Nurselike cells (NLCs) produced from monocytes, mesenchymal stromal cells (MSCs), T cells, and NK cells are essential part of the microenvironment. These cells interact with CLL cells via a complex network of adhesion molecules, chemokine receptors, TNF family members, and soluble factors. In the literature, it has been established that aberrant chemokine receptor (CCR1, CCR3, CCR7, CX3CR1, CXCR3, CXCR4, and CXCR5 gene expression levels raised, SIPR1 gene expression level decreased) in leukemic cells isolated from CLL patients play a crucial part in the etiology of the disease. These findings demonstrated how the aberrant spread of leukemic cells' localized supported their survival by the microenvironment [21]. The expression of the chemokine pathway genes GNB3, GNB4, GNG11, and GNGT1 was shown to be downregulated in patients with CLL in our study.

One of the molecular mechanisms that strictly regulates hematopoiesis is the ubiquitin proteasome system. It is a type of post-translational modification, which refers to a change in a protein after a cell makes it. Malignancies in leukemia types can result from this mechanism dysregulation. A recent study has revealed that CLL has an increased in ubiquitin-like alterations, with numerous proteins in pertinent pathways being implicated [22]. Studies have shown that the prognostic biomarkers for AML are NEDD8, UBE2M and RBX1 [23]. In our study, we discovered that CLL patients had significantly higher expression levels of the genes for the ubiquitin-like proteins NEDD8, UBE2M, RBX1, FBXO21, SKP1, KLHL9, and CAND1.

One of the most studied and significant pathways in carcinogenesis and cancer maintenance is JAK/Signal Transducer and Activator of Transcription (STAT) signaling. This pathway participates in the control of differentiation, proliferation, migration, survival, or cytotoxicity in response to more than 50 cytokines, growth factors, and hormones, depending on cell type, developmental stage, and microenvironment. Recent research has demonstrated that CLL cells had overexpressed JAK2 compared with healthy B lymphocytes [24]. The results of our investigation were consistent with the literature in that we discovered that JAK2 gene expression levels considerably increased in patients with CLL. The main cytokine IL-2, which has immunomodulatory effects on the immune system and induces STAT4 and JAK2 activation in NK cells, was also found to be overexpressed in the CLL patients in our research.

The KRAS gene, a member of the RAS family, is another key gene in our study. The Ras genes are among the oncogenes that are most frequently activated in cancer. On the inner side of the plasma membrane, RAS proto-oncogenes like KRAS, NRAS, and HRAS serve as molecular switches that transmit extracellular signals to the nucleus. They regulate the growth and differentiation of numerous different cell types [25]. RAS mutations have been extensively investigated in many leukemia types, but the study with patients with AML indicated that KRAS expression increased in AML independently of the mutation. In our study, it was observed that KRAS expression levels were elevated in CLL patients.

In conclusion, this study's findings suggest that DDOST, RPL18, RPL18A, RPL19, RPL31, GNB3, GNB4, GNG11, GNGT1 genes might be linked to CLL. The current research may be useful in

understanding the molecular mechanisms underlying CLL and identifying new biomarkers that may be drug targets. Although drug response rates are high in patients treated with immune therapy and targeted agents in CLL, the mechanisms leading to the relapse of the disease in a significant proportion of patients are still not fully elucidated. Therefore, more studies are needed on the key genes and pathways associated with CLL.

AUTHOR CONTRIBUTIONS

Concept: B.A.G.; Design; - ; Control: - ; Sources: B.A.G.; Materials: B.A.G.; Data Collection and/or Processing: B.A.G.; Analysis and/or Interpretation: B.A.G.; Literature Review: B.A.G.; Manuscript Writing: B.A.G.; Critical Review: B.A.G.; Other: B.A.G.

CONFLICTS OF INTEREST

The authors declare that there is no real, potential, or perceived conflict of interest for this article.

ETHICS COMMITTEE APPROVAL

The authors declare that the ethics committee approval is not required for this study

REFERENCES

1. Yan, H., Tian, S., Kleinstern, G., Wang, Z., Lee, J.H., Boddicker, N.J., Cerhan, J.R., Kay, N.E., Braggio, E., Slager, S.L. (2020). Chronic lymphocytic leukemia (CLL) risk is mediated by multiple enhancer variants within CLL risk loci. Human Molecular Genetics, 29(16), 2761-2774. [CrossRef]
2. Hallek, M. (2019). Chronic lymphocytic leukemia: 2020 update on diagnosis, risk stratification and treatment. Am J Hematology, 94(11), 1266-1287. [CrossRef]
3. Gao, C., Zhou, C., Zhuang, J., Liu, L., Wei, J., Liu, C., Huayao, Li, H., Sun, C. (2019). Identification of key candidate genes and miRNA-mRNA target pairs in chronic lymphocytic leukemia by integrated bioinformatics analysis. Molecular Medicine Reports, 19, 362-374. [CrossRef]
4. Luo, P., Yang, Q., Cong, L.L., Wang, X.F., Li, Y.S., Zhong, X.M., Xie, R.T., Jia, C.Y., Yang, H.Q., Li, W.P., Cong, X.L., Xia, Q., Fu, D., Zeng, Q.H., Ma, Y.S. (2017). Identification of miR124a as a novel diagnostic and prognostic biomarker in nonsmall cell lung cancer for chemotherapy. Molecular Medicine Reports, 16, 238-246. [CrossRef]
5. Kamaraj, B., Gopalakrishnan, C., Purohit, R. (2014). In silico analysis of miRNA-mediated gene regulation in OCA and OA genes. Cell Biochemistry and Biophysics, 70,1923-1932. [CrossRef]
6. Zhang, Y., Han, X., Wu, H., Zhou, Y. (2017). Bioinformatics analysis of transcription profiling of solid pseudopapillary neoplasm of the pancreas. Molecular Medicine Reports, 16, 1635-1642. [CrossRef]
7. Yan, H., Zheng, G., Qu, J., Liu, Y., Huang, X., Zhang, E., Cai, Z. (2019). Identification of key candidate genes and pathways in multiple myeloma by integrated bioinformatics analysis. Journal of Cellular Physiology, 234(12), 23785-23797. [CrossRef]
8. Bustin, S.A., Dorudi, S. (2004). Gene expression profiling for molecular staging and prognosis prediction in colorectal cancer. Expert Review of Molecular Diagnostics, 4, 599-607. [CrossRef]
9. Nannini, M., Pantaleo, M.A., Maleddu, A., Astolfi, A., Formica, S., Biasco, G. (2009). Gene expression profiling in colorectal cancer using microarray technologies: results and perspectives. Cancer Treatment Reviews, 35, 201-209. [CrossRef]
10. Kulasingam, V., Diamandis, E.P. (2008). Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. Nature Clinical Practice Oncology, 5, 588-599. [CrossRef]
11. Gao, C., Zhou, C., Zhuang, J., Liu, L., Wei, J., Liu, C., Li, H., Sun, C. (2019). Identification of key candidate genes and miRNA-mRNA target pairs in chronic lymphocytic leukemia by integrated bioinformatics analysis. Molecular Medicine Reports, 19(1), 362-374. [CrossRef]
12. Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C.L., Serova, N., Davis, S., Soboleva, A. (2013). NCBI GEO: archive for functional genomics data sets-update. Nucleic Acids Research, 39(1), D1005-D1010. [CrossRef]
13. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C.,

Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature Genetics. 25, 25-29. [CrossRef]

14. Kanehisa M., Goto S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Research, 28, 27-30. [CrossRef]
15. Chang, X., Pan, J., Zhao, R., Yan, T., Wang, X., Guo, C., Yang, Y., Wang, G. (2022). DDOST Correlated with Malignancies and Immune Microenvironment in Gliomas. Frontiers in Immunology, 13, 1-13. [CrossRef]
16. Shapanis, A., Lai, C., Smith, S., Coltart, G., Sommerlad, M., Schofield, J., Parkinson, E., Skipp, P.,Healy, E. (2021). Identification of Proteins Associated With Development of Metastasis From Cutaneous Squamous Cell Carcinomas (cSCCs) via Proteomic Analysis of Primary cSCCs. British Journal of Dermatology, 184(4), 709-721. [CrossRef]
17. Zhu, C., Xiao, H., Jiang, X., Tong R., Guan, J. (2021). Prognostic Biomarker DDOST and Its Correlation With Immune Infiltrates in Hepatocellular Carcinoma. Frontiers in Genetics, 12, 1-13. [CrossRef]
18. Gao, C., Zhou, C., Zhuang, J., Liu, L., Wei, J., Liu, C., Li, H., Sun, C. (2019). Identification of key candidate genes and miRNA-mRNA target pairs in chronic lymphocytic leukemia by integrated bioinformatics analysis. Molecular Medicine Reports, 19, 362-374. [CrossRef]
19. Sbarrato, T., Horvilleur, E., Pöyry, T., Hill, K., Chaplin, L.C., Spriggs, R.V., Stoneley, M., Wilson, L., Jayne, S., Vulliamy, T., Beck, D., Dokal, I., Dyer, M.J.S., Yeomans, A.M., Packham, G., Bushell, M., Wagner, S.D., Willis, A.E. (2016). A ribosome-related signature in peripheral blood CLL B cells is linked to reduced survival following treatment. Cell Death and Disease, 7(6), e2249. [CrossRef]
20. Goudarzi, K.M., Lindström, M.S. (2016) Role of ribosomal protein mutations in tumor development (Review). International Journal of Oncology, 48(4), 1313-1324. [CrossRef]
21. Burger, J.A. (2010). Chemokines and chemokine receptors in chronic lymphocytic leukemia (CLL): From understanding the basics towards therapeutic targeting. Seminars in Cancer Biology 20, 424-430. [CrossRef]
22. Arenas, V., Castaño, J.L., Dominguez-Garcia, J.J., Yañez, L., Pipao´n, C. (2021). A Different View for an Old Disease: NEDDylation and Other Ubiquitin Like Post-Translational Modifications in Chronic Lymphocytic Leukemia. Frontiers in Oncology, 11, 729550. [CrossRef]
23. Chen, Y., Sun., L. (2022). Inhibition of NEDD8 NEDDylation induced apoptosis in acute myeloid leukemia cells via p53 signaling pathway. Bioscience Reports 42(8), 1-16. [CrossRef]
24. Severin, F., Frezzato, F., Visentin, A., Martini, V., Trimarco, V., Carraro, S., Tibaldi, E., Brunati,A.M., Piazza, F., Semenzato, G., Facco, M., Trentin, L. (2019). In chronic lymphocytic leukemia the JAK2/STAT3 pathway is constitutively activated and its inhibition leads to CLL cell death unaffected by the protective bone marrow microenvironment. Cancers (Basel). 11(12), 1939. [CrossRef]
25. Zhou, J.D., Yao, D.M., Li, X.X., Zhang, T.J., Zhang, W., Ma, J.C., Guo, H., Deng, Z.Q., Lin, J., Qian, J. (2017). KRAS overexpression independent of RAS mutations confers an adverse prognosis in cytogenetically normal acute myeloid leukemia. Oncotarget, 8(39), 66087-66097. [CrossRef]