Cancer Research Journal
2021; 9(4): 176-190
http://www.sciencepublishinggroup.com/j/crj
doi: 10.11648/j.crj.20210904.11
ISSN: 2330-8192 (Print); ISSN: 2330-8214 (Online)

Identification of Three Candidate Genes and Their Correlation with Drug Sensitivity in Acute Myeloid Leukemia

Fenling Zhou[1], Yuli Cao[1], Daxia Cai[1], Jiajian Liang[1], Cuilan Deng[2], Gexiu Liu[1, *], Dongmei He[1, *]

[1]Institute of Hematology, Jinan University, Guangzhou, P.R. China

[2]Department of Hematology, The First Affiliated Hospital of Jinan University, Guangzhou, China

Email address:

thedm@jnu.edu.cn (Dongmei He), tliugx@jnu.edu.cn (Gexiu Liu), zfl996225@163.com (Fenling Zhou), 13650457431@163.com (Yuli Cao), cdx6119@163.com (Daxia Cai), liangjiajianmengde@163.com (Jiajian Liang), dengcuilan1104@163.com (Cuilan Deng)

[*]Corresponding author

To cite this article:

Fenling Zhou, Yuli Cao, Daxia Cai, Jiajian Liang, Cuilan Deng, Gexiu Liu, Dongmei He. Identification of Three Candidate Genes and Their Correlation with Drug Sensitivity in Acute Myeloid Leukemia. *Cancer Research Journal*. Vol. 9, No. 4, 2021, pp. 176-190.
doi: 10.11648/j.crj.20210904.11

Received: November 19, 2021; **Accepted**: December 10, 2021; **Published**: December 24, 2021

Abstract: *Background:* Acute myeloid leukemia (AML) is a common hematopoietic tumor with extremely high morbidity and mortality. This study was designed to explore candidate genes that were related to the poor prognosis of AML patients and analyze their relationship with drug sensitivity. *Methods:* Microarray databases were performed to screen the differentially expressed genes (DEGs). DAVID 6.8 was used for further functional enrichment analysis. The protein-protein interaction (PPI) network was constructed through STRING website and Cytoscape tool. Then, we analyzed and explored the mRNA transcription level, prognosis correlation, and drug sensitivity of the candidate genes in AML via multiple acknowledged databases including the GEPIA, BloodSpot, EMBL-EBI, UALCAN, LinkedOmics, and GSCALite databases. *Results:* A total of 181 up-regulated DEGs were screened. Three candidate genes (MAP2K3, LST1, and CYTH4) related to poor outcomes of AML patients were identified. Meanwhile, the high expression levels of the three genes were verified in AML patients and AML cell lines, the expression differences of three genes at AML different subtypes were demonstrated. Drug sensitivity analysis displayed the expression levels of MAP2K3, LST1, and CYTH4 were negatively related to drug resistance, indicating that the three genes were sensitive to certain small-molecule drugs (including targeted drugs and non-targeted drugs). *Conclusion:* In summary, MAP2K3, LST1, and CYTH4 may be potential prognostic indicators for AML, and may be associated with the sensitivity of certain small molecule drugs.

Keywords: Acute Myeloid Leukemia (AML), Candidate Gene, Bioinformatics, Prognosis, Drug Sensitivity

1. Introduction

Acute myeloid leukemia (AML) is a heterogeneous hematopoietic tumor, featured by the clonal expansion of undifferentiated myeloid progenitor cells [1]. It is often accompanied by impaired hematopoietic function and bone marrow failure, leading to severe consequences for patients [2]. Meanwhile, it has been reported that various factors were involved in the molecular mechanism of AML disease progression and clinical prognosis, such as genetic, epigenetic, and proteomic alterations [3, 4]. Although considerable advances in elucidating the pathophysiology and molecular heterogeneity of AML, the development of viable therapies for patients remains a formidable challenge. Currently, the poor prognosis of AML patients remains a clinical challenge, with merely 40% of patients under 60 years of age surviving more than five years. While AML patients could be relieved with standard treatment, most patients who relapse after complete recovery survive less than five years [5, 6]. Hematopoietic stem cell transplantation (HSCT) and chemotherapy are effective treatment regimens for AML patients [7, 8]. However, drug resistance markedly decreased the efficacy of AML patients, and recurrence is still the primary cause for the failure of HSCT [7]. Therefore, it is imperative to identify reliable biomarkers that could significantly enhance the prognosis and drug treatment

sensitivity of AML patients.

The microarray database is a high-throughput tool that could generate various data, such as mRNA expression, DNA methylation, and microRNA expression [9, 10]. It also provides basic information about gene function, protein network, and regulatory pathway, which plays a key role in biomedical research [11, 12]. Currently, multiple widely recognized public databases, such as the Gene Expression Profiles Interactive Analysis (GEPIA) [13], EMBL-EBI [5], and GSCALite [14], are used to explore undiscovered tumor biomarkers and therapeutic targets via the stored microarray data.

The objective of this article was to explore the potential candidate genes that lead to the poor prognosis of AML patients through bioinformatics analysis. Except for the genes reported to have specific carcinogenic effects on AML, three up-regulated candidate genes (MAP2K3, LST1, and CYTH4) may serve as novel prognostic indicators for AML. And then, we analyzed the relation between high expression of the three genes and drug sensitivity in AML cells. This research provides initial evidence for exploring the underlying sensitive drugs for AML.

2. Materials and Methods

2.1. Processing for Microarray Data and DEGs Identification

Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) is a free public repository that collects microarray expression profile and next-generation sequencing data, from which two series (GSE65409 and GSE67936) were acquired. And GSE65409 was based on the GPL96 platform (Illumina HumanHT-12 V3.0 expression biochip, San Diego, CA, USA), which included 30 peripheral blood mononuclear cells (PBMCs) from primary AML patient samples and 8 CD34+ bone marrow (BM) cells of healthy donors. GSE67936 was based on the GPL10558 platform (Illumina HumanHT-12 V4.0 expression biochip, San Diego, CA, USA), which included 117 AML diagnosis samples and 9 CD34+ cells of normal controls. Detailed information about the above two datasets was summarized in Supplementary Figures 1 and 2. Afterward, the differentially expressed genes (DEGs) were detected through GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/) interactive web tool in NCBI [15]. The threshold was established at adjusted P-value < 0.05 and $|\log_2 FC| > 1$. Meanwhile, the overlapping DEGs were visualized by Venn diagram (http://bioinformatics.psb.ugent.be/webtools/ Venn/).

2.2. Integration of Protein–Protein Interaction (PPI) Network, Functional Enrichment Analysis

The STRING (Search Tool for Retrieval of Interacting Genes/Proteins, http://string-db.org/) online database [16] was used to predict interaction networks of the protein products of the up-regulated DEGs. The cut-off point was set as a confidence score of ≥0.4. Then, the Cytoscape (3.7.2) software (www.cytoscape.org/) was applied to visualize the protein interaction network relationships. In Cytoscape, CytoHubba plugin [17] provides a simple interface to explore important nodes in biological networks by numerous algorithms. To further understand the biological functions of the DEGs, the Database for Annotation, Visualization, and Integrated Discovery (DAVID; https://david.ncifcrf.gov/) was used to annotate and analyze the associated the Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of the DEGs [18]. P<0.05 was defined as the criterion for significance.

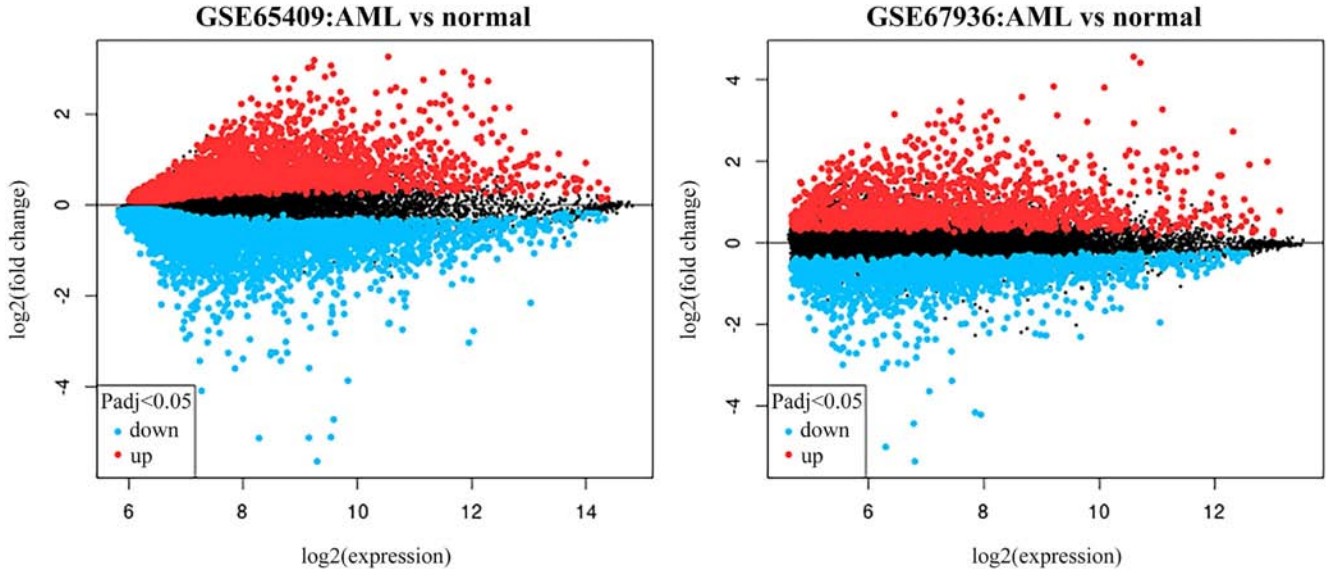


Figure 1. *Up-regulated and down-regulated genes information of the GSE65409 and GSE67936 datasets (The red dots mean all the up-regulated genes, and the blue dots mean all the down-regulated genes).*

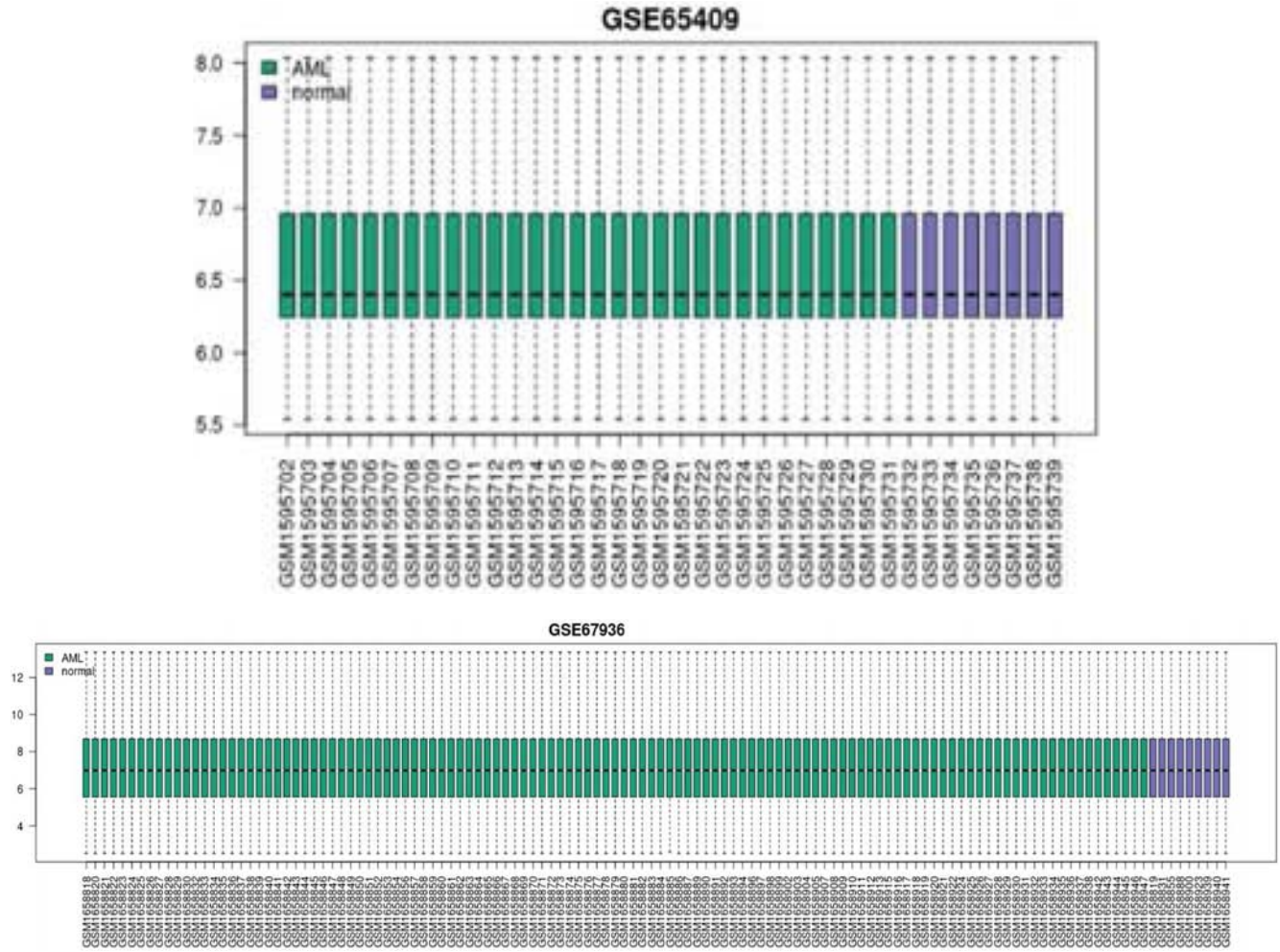


Figure 2. *Sample information of the GSE65409 and GSE67936 datasets (The green is AML samples, and the purple is normal samples).*

2.3. Screening of Survival-Related Upregulated Genes

Gene Expression Profiling Interactive Analysis (GEPIA) is an online tool based on Cancer Genome Atlas (TCGA; http://portal.gdc.cancer.gov/) and Genotype Tissue Expression (GTEx) (GTEx; https://www.gtexportal.org/) projects, which could provide key interactive functions including differential expression analysis, correlation analysis, and patient survival analysis [13]. We constructed to illuminate correlations between the DEGs expression and prognosis of AML patients by GEPIA. The cut-off value was set to 0.05. BloodSpot (http://servers.binf.ku.dk/bloodspot/) is a newly developed gene-centric mRNA expression dataset containing healthy and malignant hematopoietic cells [19]. The correlation between the candidate genes expression and genetic characteristics of AML patients were analyzed through BloodSpot.

2.4. Verification of the Expression Levels of MAP2K3, LST1, and CYTH4 in AML Patients and Cell Lines

The "differential expression analysis" module of the GEPIA2 was used to verify the expression of the candidate genes in AML patients [13]. EMBL-EBI (https://www.ebi.ac.uk) was used to further analyze the candidate genes expression in 14 common AML cell lines, A mass of cancer cell lines in this database could provide powerful clues to the gene expression in cancer subtypes of different tissue origin [20].

2.5. Expression Analysis of the Candidate Genes in Different Clinical Parameters of AML Patients

UALCAN (http://ualcan.path.uab.edu), an extensive interactive web-portal that could provide analyses based on TCGA and MET500 cohort data [21]. In our study, gene expression data were examined via the "AML" dataset and the "Expression Analysis" module of UALCAN. We use it to compare the expression differences of the 3 candidate genes in AML patients of different subtypes, ages and genders. Student's t test was used to gain a P-value. The cut-off value of P-value was 0.05.

2.6. The Correlation Analysis of MAP2K3, LST1, and CYTH4 in AML

LinkedOmics database (http://www.linkedomics.org) contains the multiple-omics data and clinical data for 11,158

patients from the TCGA project [22]. The relationships of the candidate genes were analyzed using LinkedOmics dataset and verified using GEPIA database. The P-value cutoff was 0.05.

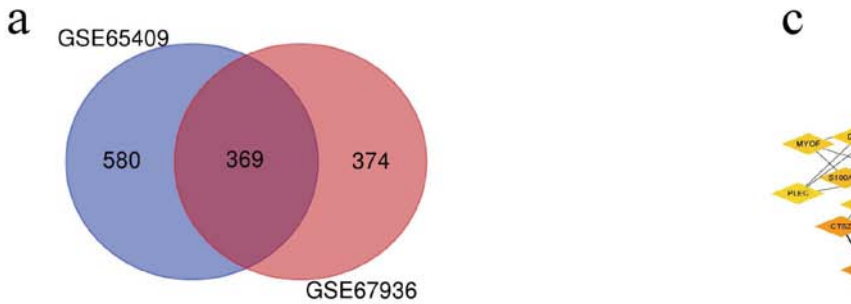
2.7. Drug Sensitivity Analysis of MAP2K3, LST1, and CYTH4 in AML Cell Lines

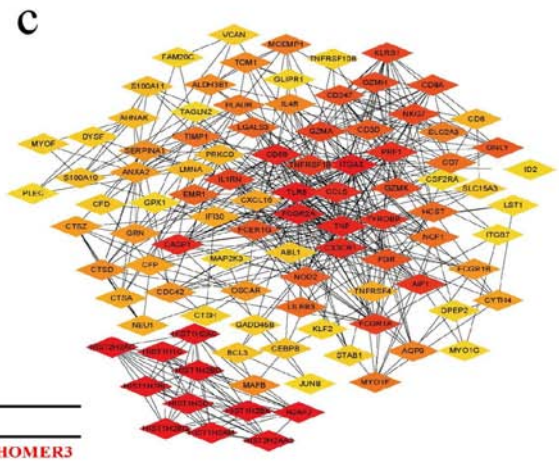
GSCALite (http://bioinfo.life.hust.edu.cn/web/GSCALite/) is a web-based platform for gene sets cancer analysis, which is the dynamic analysis and visualization of gene sets in cancer pathway activity, methylation, and drug-sensitivity analysis [14]. The Spearman correlation was performed to detect the correlation between gene expression and 265 small molecules or drugs from Cancer Drug Sensitivity Genomics (GDSC). The negative correlation means that highly expressed genes are sensitive to the drug, vise verse.

3. Results

3.1. Screening of Overlapping DEGs and Integration of PPI Network in AML

Based on the critical value of $P<0.05$ and $[\log_2 FC] > 1$, 369 overlapping DEGs were screened from GSE65409 and GSE67936 (Figure 3a). Subsequently, the 181 up-regulated DEGs were further studied (Figure 3b), and the up-regulated genes highlighted in red were reported to have specific carcinogenic effects on AML. Then, A total of 181 up-regulated DEGs were uploaded into Cytoscape software and STRING online database to further screen. Eventually, the top 100 DEGs with the high correlation were obtained by CytoHubba plugin, the PPI action network was shown in Figure 3c. The deeper the color, indicated the higher the correlation between the genes.





b

Upregulated DEGs (n=181)
ANXA2 CCNA1 MOB3A ASGR2 FCGR2A H2AFJ GZMH TYROBP OSCAR RPPH1 HOMER3
FGR ADGRE5 TMEM51 AQP9 SERTAD3 SDSL ARL4A HCST SLC22A18 LINC00936
HIST1H2AM HIST1H2BG LILRB3 GZMA TAGLN2 DYSF EVI2A GPBAR1 STAB1 LGALS3
FOXC1 SLC2A6 JUNB CYP27A1 RHOB SLC2A3 IFI30 CD6 EHD4 EMP3 HIST1H2BC
TNFRSF1B ABL1 SLC30A1 SLC16A3 CRIP2 NACC2 PLEC ZNF467 SCIMP IER5 PILRA
HIST1H2AC MAL SLC43A2 NFIL3 TNFRSF10B CD247 HIST2H2AA3 GRN PRKCD ADAP1
NEU1 SERPINA1 ITPRIPL2 ID2 FCGR1B SLC15A3 HBM MCEMP1 GZMK HIST1H2BK
LRRC25 AHNAK ITGAX IL1RN S100A10 PIM1 MAFB GRINA ITGB7 NKG7 SGSH CRIP1
VCAN VMP1 ADAP2 TNFRSF4 IL4R UBTD1 CCL5 IER3 DPEP2 RNH1 MYOF SH3BGRL3
FGFBP2 TNF GLIPR1 GAS7 BAG4 CA5B MYO1F KLF2 HIST1H3D FAM214B TLR8 CD7
HELZ2 SERTAD1 CDKN2C HSPA6 CASP1 CD8A UPP1 FCGR1A MYO1G IL13RA1 CSF2RA
PLA2G16 GLIPR2 ANKRD33 NCF1 BCL3 MIDN CDC42 CTSA HIST1H2BD FCER1G RRAS
GABARAPL1 F12 OLIG1 PRF1 LST1 TPPP3 CTSH CFP NOD2 CTSZ ALDH3B1 LMNA TOM1
PRICKLE1 HES4 FBXO6 CX3CR1 KLF10 CD3D CTSD GPX1 HIST1H1C FAM20C IFITM1
CEBPB CXCL16 TCN2 CRISPLD2 TIMP1 PLAUR MAP2K3 AIF1 LFNG LILRA5 S100A11
TYMP CFD HIST2H2AC FLVCR2 TOB1 CD68 MVP KLRB1 GADD45B HBA2 ADGRE1
PITPNM1 CYTH4 GNLY PIK3IP1

Figure 3. *Identification of consistent DEGs and PPI network analysis. (a) Venn diagram for consistent DEGs. (b) List of 181 consistent up-regulated DEGs (the up-regulated genes highlighted in red indicate genes that have been reported to have a carcinogenic effect on AML). (c) the top 100 central genes with a higher degree of connectivity. The deeper the color, indicated the higher the correlation between the genes. DEGs, differentially expressed genes.*

3.2. GO and KEGG Enrichment Analysis

To comprehend the biological value of DEGs in the interaction network, DAVID 6.8 was performed to analyze the GO terms and KEGG enrichment pathways of the top 100 upregulated DEGs. The top 15 enrichment analysis results were shown in Figure 4. GO terms describe the DEGs from three aspects, namely molecular function (MF), cellular component (CC), and biological process (BP) [23]. In the MF, the genes were primarily enriched “IgG binding”, “protein heterodimerization activity”, “receptor activity”, “protein binding”, “S100 protein binding” and so on (Figure 4a). In the CC, the genes mainly participated in “nucleosome”, “extracellular exosome”, “plasma membrane”, “extracellular space”, “nuclear nucleosome”, etc. (Figure 4b). In the BP, the

genes were principally enriched "immune response", "regulation of immune response", "Fc-gamma receptor signaling pathway involved in phagocytosis", "regulation of cell proliferation", "defense response to bacterium", etc. (Figure 4c). In KEGG enrichment analysis, the main pathways included "Systemic lupus erythematosus", "Alcoholism", "TNF signaling pathway", "Hematopoietic cell lineage", "Osteoclast differentiation", etc. (Figure 4d).

3.3. Prognostic Value Analysis of MAP2K3, LST1, and CYTH4 in AML Patients

To comprehend the possible value of the top 100 DEGs in the prognosis prediction of AML patients, survival analysis was performed using the GEPIA2 databases. At last, As shown in Figure 5, three genes (MAP2K3, LST1, and CYTH4) were related to poor prognosis of AML patients in addition to the genes reported to have specific carcinogenic effects on AML patients ($p<0.05$).

To further confirm the relationship between the high expression of MAP2K3, LST1, and CYTH4 and the prognosis of AML patients, we used the BloodSpot database to explore the impact of genes epigenetic mutation on the risk stratification of AML patients. AML with t(11q23), AML with del(5q), AML with del(7q)/7q- and other chromosomal abnormalities have been widely used as clinical indicators to evaluate the poor prognosis of AML. We found that MAP2K3, LST1, and CYTH4 were almost significantly overexpressed in karyotype samples suggestive of poor prognosis, which indicated the high expression of the three genes in AML patients may suggest high-risk groups at the level of genetic stratification (Figure 6).

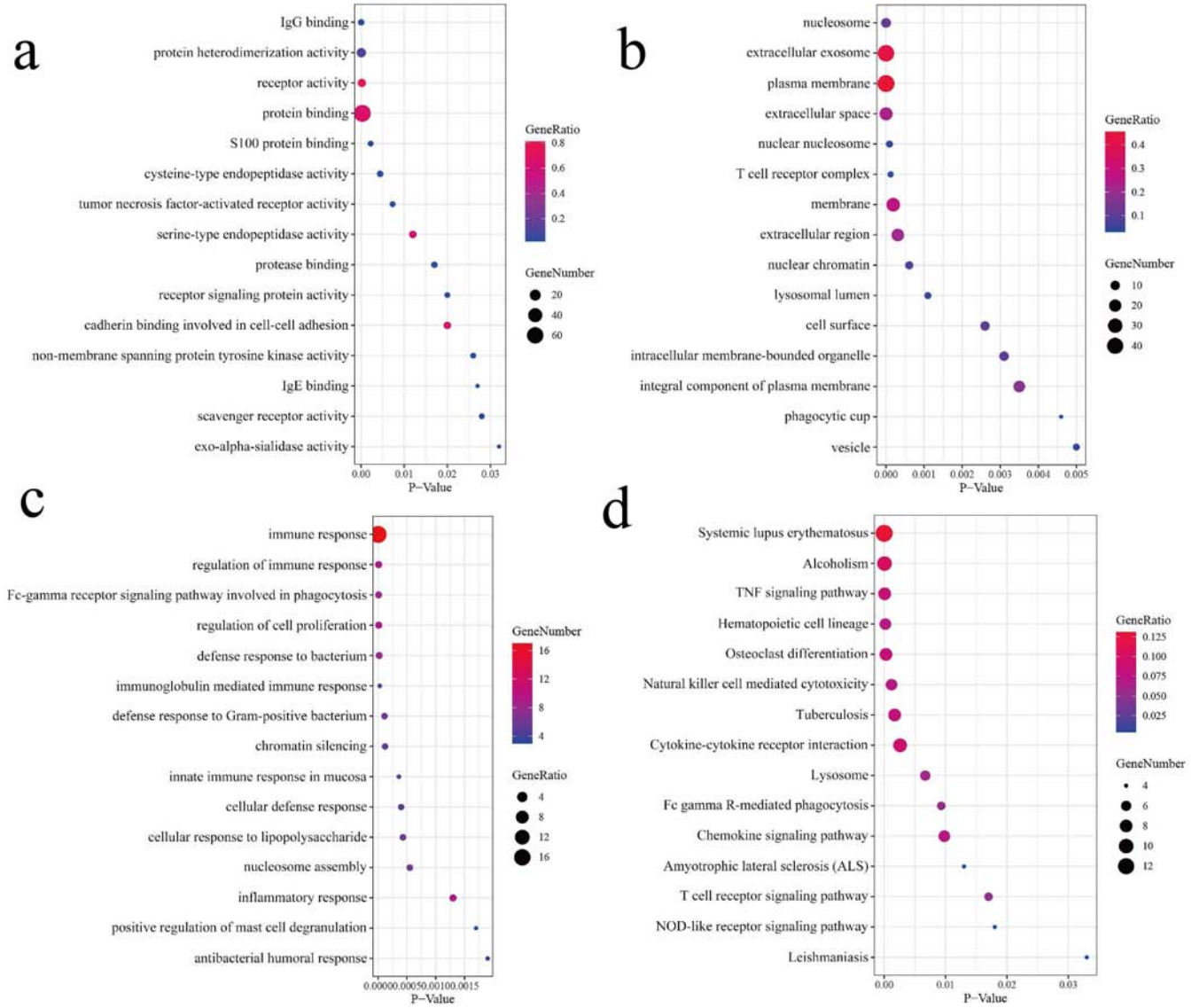


Figure 4. *The top 15 GO enrichment terms and KEGG enrichment pathways of DEGs. The Y-axis represents the enrichment term, the X-axis shows the P-value, and the color and size of the dots reflect the number and proportion of gene enrichment. (a) The top 15 Molecular functions. (b) The top 15 Cellular components. (c) The top 15 Biological processes. (d) The top 15 Functional pathways. GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.*

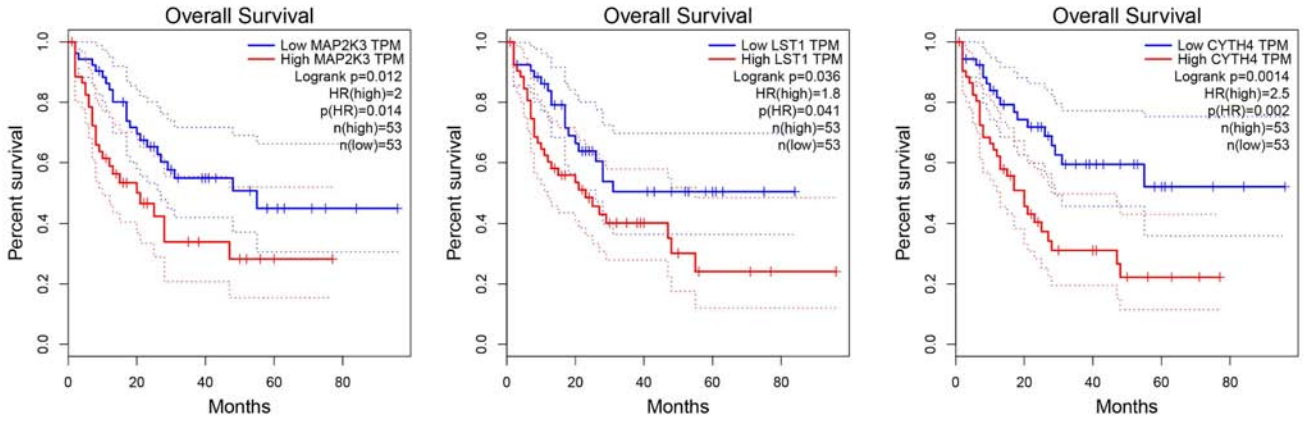


Figure 5. *Prognostic value of MAP2K3, LST1, and CYTH4 in AML patients, analyzed by GEPIA.*

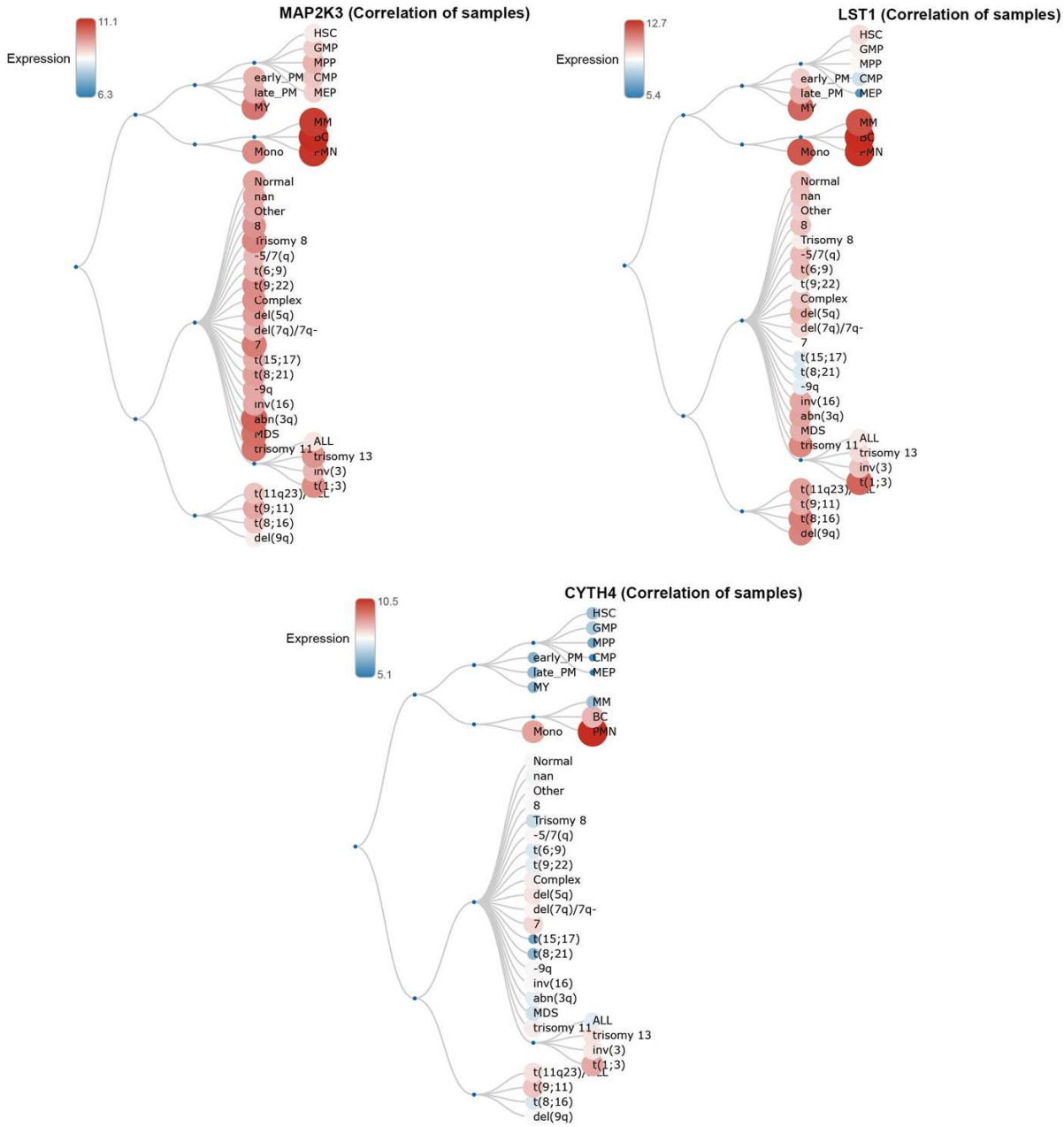


Figure 6. *The relationship between the 3 candidate genes expression and genetic characteristics of AML patients, analyzed by BloodSpot. Detailed descriptions of the data were presented in Table 1.*

Table 1. *Detailed descriptions of the data.*

Short	Abbreviation
HSC	Hematopoietic stem cell
MPP	Multipotential progenitors
CMP	Common myeloid progenitor cell
GMP	Granulocyte monocyte progenitors
MEP	Megakaryocyte-erythroid progenitor cell
early_PM	Early Promyelocyte
Late_PM	Late Promyelocyte
BC	Band cell
MM	Metamyelocytes
MY	Myelocyte
Mono	Monocytes
PMN	Polymorphonuclear cells
Normal	AML with Normal karyotype
Complex	AML with Complex karyotype
inv (16)	AML with inv (16)
t (15;17)	AML with t (15;17)
t (8;21)	AML with t (8;21)
t(11q23)/MLL	AML with t (11q23)/MLL
MDS	MDS
nan	AML with no karyotype information
Trisomy 8	AML with Trisomy 8
del (5q)	AML with del (5q)
del(7q)/7q-	AML with del (7q)/7q-
t (9;11)	AML with t (9;11)
Other	AML with Other abnormalities
7	AML with +7
Complex_ del(5q)	AML with Complex del(5q)
Complex_ untypical	AML with Complex untypical karyotype
ALL	ALL
inv (3)	AML with inv (3)
trisomy 11	AML with trisomy 11
trisomy 13	AML with trisomy 13
t (6;9)	AML with t (6;9)
t (8;16)	AML with t (8;16)
del (9q)	AML with del (9q)
t (1;3)	AML with t (1;3)
-5/7(q)	AML with -5/7(q)
-9q	AML with -9q
8	AML with +8
t (9;22)	AML with t (9;22)
abn (3q)	AML with abn (3q)

3.4. Verification of High Expression Levels of the 3 Candidate Genes in AML Patients and Cell Lines

Using the GEPIA dataset, this study reanalyzed the expression levels of the above three genes in AML patients (Figure 7). MAP2K3, LST1, and CYTH4 were significantly overexpressed in AML patients compared with normal controls ($P<0.05$). Meanwhile, we further analyzed the gene expression in 14 common AML cell lines using the EMBL-EBI bioinformatics website, and the results showed that MAP2K3, LST1, and CYTH4 were increased in most common AML cell lines (Figure 8).

3.5. Expression of the 3 Candidate Genes in AML Different Subtypes and Their Correlation with Patients' Age

Then, UALCAN database was performed to analyze the expression differences of MAP2K3, LST1, and CYTH4 in AML patients of different subtypes, ages and genders (Figure 9). As shown in Figure 9a, MAP2K3, LST1, and CYTH4 showed significant expression differences in AML of different subtypes.

Compared with other subtypes, the three genes showed lower expression in M3 type. Besides, the expression of three candidate genes was highest in elderly patients (especially patients over 60 years old), but the difference between genders was not statistically significant (Figures 9b, c).

3.6. The Correlation Analysis of MAP2K3, LST1, and CYTH4 in AML

To further explore the potential role of the 3 candidate genes, we analyzed the association between MAP2K3, LST1, and CYTH4 via the LinkedOmics database. And it was found that MAP2K3 was both positively connected with LST1 ($R = 0.4198$, $p < 0.05$), and CYTH4 ($R = 0.2721$, $p < 0.05$). LST1 was positively connected with CYTH4 ($R = 0.6343$, $p < 0.05$) (Figure 10a). Next, we verified their relationship using the GEPIA database. Expectedly, the results also suggested that there was a positive correlation between MAP2K3, LST1, and CYTH4 in AML (Figure 10b).

Table 2. *Detailed information of common clinical small-molecule drugs.*

Drug_name	Synonyms	Drug targets pathway	Drug targets
Phenformin	DBI	Other	Biguanide agent
Daporinad	APO866, FK866, FK866	Metabolism	NAMPT
AICA Ribonucleotide	AICAR, N1-(b-D-Ribofuranosyl)-5-aminoimidazole-4-carboxamide	Metabolism	AMPK agonist
Vorinostat	Zolinza, SAHA, suberanilohydroxamic acid, suberoylanilide hydroxamic acid, MK-0683	Chromatin histone acetylation	HDAC inhibitor Class I, IIa, IIb, IV
VNLG/124	HDAC inhibitor XV	Chromatin histone acetylation	HDAC, RAR
Tubastatin A	-	Chromatin histone acetylation	HDAC1, HDAC6, HDAC8
CUDC-101	CUDC 101	Other	HDAC1-10, EGFR, ERBB2
CAY10603	-	Chromatin histone acetylation	HDAC1, HDAC6
Belinostat	PXD101, PXD-101	Chromatin histone acetylation	HDAC1
AR-42	HDAC-42, AR 42, AR42	Chromatin histone acetylation	HDAC1
Dacinostat	NVP-LAQ824, LAQ824	Chromatin histone acetylation	HDAC1
NPK76-II-72-1	-	Cell cycle	PLK3
MPS-1-IN-1	-	Mitosis	MPS1
Ispinesib Mesylate	SB-715992	Mitosis	KSP
GSK1070916	GSK-1070916	Mitosis	AURKA, AURKC
Genentech Cpd 10	-	Mitosis	AURKA, AURKB
ZM447439	ZM-447439, ZM 447439	Mitosis	AURKA, AURKB
Tozasertib	MK 0457, MK-0457, MK-045, VX-680 VX 680 VX-68	Mitosis	AURKA, AURKB, AURKC, others
I-BET-762	GSK525762A	Chromatin other	BRD2, BRD3, BRD4
JQ1	JQ-1, (+)-JQ-1	Chromatin other	BRD2, BRD3, BRD4, BRDT
ZSTK474	KIN001-167, ZSTK-474, ZSTK 474	PI3K/MTOR signaling	PI3K (class 1)
PIK-93	PIK 93, PIK93	PI3K/MTOR signaling	PI3Kgamma
PI-103	PI-103, PI103, PI 103	Other, kinases	PI3Kalpha, DAPK3, CLK4, PIM3, HIPK2
KIN001-244	PDK1 inhibitor 7	Metabolism	PDK1 (PDPK1)
AKT inhibitor VIII	Akti-1/2, KIN001-102	PI3K/MTOR signaling	AKT1, AKT2, AKT3
Omipalisib	GSK2126458, GSK-2126458, EX-8678, GSK458	PI3K/MTOR signaling	PI3K (class 1), MTORC1, MTORC2
Idelalisib	CAL-101, Zydelig	PI3K/MTOR signaling	PI3Kdelta
BX-912	-	Metabolism	PDK1 (PDPK1)
GSK690693	GSK 690693, GSK-690693	PI3K/MTOR signaling	AKT1, AKT2, AKT3
AS605240	KIN001-173, AS-605240	PI3K/MTOR signaling	PI3Kgamma
XMD14-99	-	Other, kinases	ALK, CDK7, LTK, others
Cabozantinib	BMS-907351, XL-184, Cometriq	RTK signaling	VEGFR, MET, RET, KIT, FLT1, FLT3, FLT4, TIE2, AXL
WZ3105	-	Other	SRC, ROCK2, NTRK2, FLT3, IRAK1, others
Fedratinib	TG101348, TG-101348, SAR302503, SAR-302503	Other, kinases	JAK2
Sunitinib	Sutent, Sunitinib Malate, SU-11248	RTK signaling	PDGFR, KIT, VEGFR, FLT3, RET, CSF1R
Sorafenib	Nexavar, 284461-73-0, BAY 43-9006	RTK signaling	PDGFR, KIT, VEGFR, RAF
QL-XI-92	-	Cytoskeleton	DDR1
OSI-930	OSI 930 OSI930	RTK signaling	KIT
Masitinib	AB1010, Masivet	RTK signaling	KIT, PDGFRA, PDGFRB

Drug_name	Synonyms	Drug targets pathway	Drug targets
Linifanib	ABT-869, ABT 869	RTK signaling	VEGFR1, VEGFR2, VEGFR3, CSF1R, FLT3, KIT
Foretinib	GSK1363089, XL-880, EXEL-2880, GSK089	RTK signaling	MET, KDR, TIE2, VEGFR3/FLT4, RON, PDGFR, FGFR1, EGFR
Lestaurtinib	CEP-701, SP-924, SPM-924, A-154475, KT-555	Other, kinases	FLT3, JAK2, NTRK1, NTRK2, NTRK3
NVP-BHG712	BHG712	RTK signaling	EPHB4
Quizartinib	AC220, AC 220, AC-220, Asp-2689	RTK signaling	FLT3
Alectinib	CH5424802, CH 542802, Alecensa	RTK signaling	ALK
UNC0638	UNC-0638, UNC 0683	Chromatin histone methylation	G9a and GLP methyltransferases
THZ-2-49	-	Cell cycle	CDK9
THZ-2-102-1	-	Cell cycle	CDK7
PHA-793887	PHA793887, PHA 793887	Cell cycle	CDK2, CDK7, CDK5
AT-7519	AT7519	Cell cycle	CDK1, CDK2, CDK4, CDK6, CDK9
XMD13-2	-	Apoptosis regulation	RIPK1
TPCA-1	-	Other, kinases	IKK2
TL-1-85	-	Other, kinases	TAK
STF-62247	STF62247	Other	Autophagy inducer
SNX-2112	SNX 2112	Protein stability and degradation	HSP90
QL-XII-61	-	Other, kinases	BMX, BTK
QL-XII-47	-	Other, kinases	BTK, BMX
QL-X-138	-	Other, kinases	BTK
NG-25	NG25	Other, kinases	TAK1, MAP4K2
KIN001-260	Bayer IKKb inhibitor, ACHP	Other, kinases	IKKB
KIN001-236	-	RTK signaling	Angiopoietin-1 receptor
JW-7-24-1	-	Other, kinases	LCK
CX-5461	CX5461, CX 5461	Other	RNA Polymerase 1
BX795	BX-795	Other, kinases	TBK1, PDK1 (PDPK1), IKK, AURKB, AURKC
BMS-345541	BMS345541, IKK Inhibitor 3	Other, kinases	IKK1, IKK2
BIX02189	BIX 02189	ERK MAPK signaling	MEK5, ERK5
BAY-61-3606	Syk Inhibitor, BAY-613606	Other, kinases	SYK
Tretinoin	ATRA, Vesanoid, Renova, Atralin, Tretin-X, Avita	Other	Retinoic acid
ZG-10	-	JNK and p38 signaling	JNK1
YM201636	YM-201636, YM 201636	PI3K/MTOR signaling	PIKFYVE
XMD8-92	XMD 8-92	ERK MAPK signaling	MAPK7
Ruxolitinib	INCB-18424, Ruxolitinib Phosphate, Jakafi	Other, kinases	JAK1, JAK2
Enzastaurin	LY317615	Other, kinases	PKCB
DMOG	Dimethyloxalylglcine	Metabolism	HIF-PH
XMD15-27	-	Other, kinases	CAMK2
Navitoclax	ABT-263, ABT263, ABT 263	Apoptosis regulation	BCL2, BCL-XL, BCL-W
PAC-1	GTPL5238	Apoptosis regulation	Procaspase-3, Procaspase-7
OSI-027	A-1065-5	PI3K/MTOR signaling	MTORC1, MTORC2
AZD8055	AZD-8055	PI3K/MTOR signaling	MTORC1, MTORC2
CP466722	CP-466722, CP 466722, 1080622-86-1	Genome integrity	ATM
AZD7762	SN1031853762	Cell cycle	CHEK1, CHEK2
TL-2-105	-	Other	not defined
FR-180204	FR 180204, FR180204, ERK Inhibitor II	ERK MAPK signaling	ERK1, ERK2
FMK	KIN001-242	Other, kinases	RSK
Trametinib	GSK1120212, Mekinist	ERK MAPK signaling	MEK1, MEK2
Ponatinib	AP24534, AP-24534, KIN001-192, Iclusig	Other, kinases	ABL, PDGFRA, VEGFR2, FGFR1, SRC, TIE2, FLT3
Nilotinib	Tasigna, AMN 107	ABL signaling	ABL
Y-39983	-	Cytoskeleton	ROCK

Drug_name	Synonyms	Drug targets pathway	Drug targets
GSK429286A	-	Cytoskeleton	ROCK1, ROCK2
TAK-715	KIN001-201, TAK 715	JNK and p38 signaling	p38alpha, p38beta
Methotrexate	Abitrexate, Amethopterin, Rheumatrex, Trexall, Folex	DNA replication	Antimetabolite
5-Fluorouracil	5-FU	Other	Antimetabolite (DNA & RNA)
Pelitinib	EKB-569	EGFR signaling	EGFR

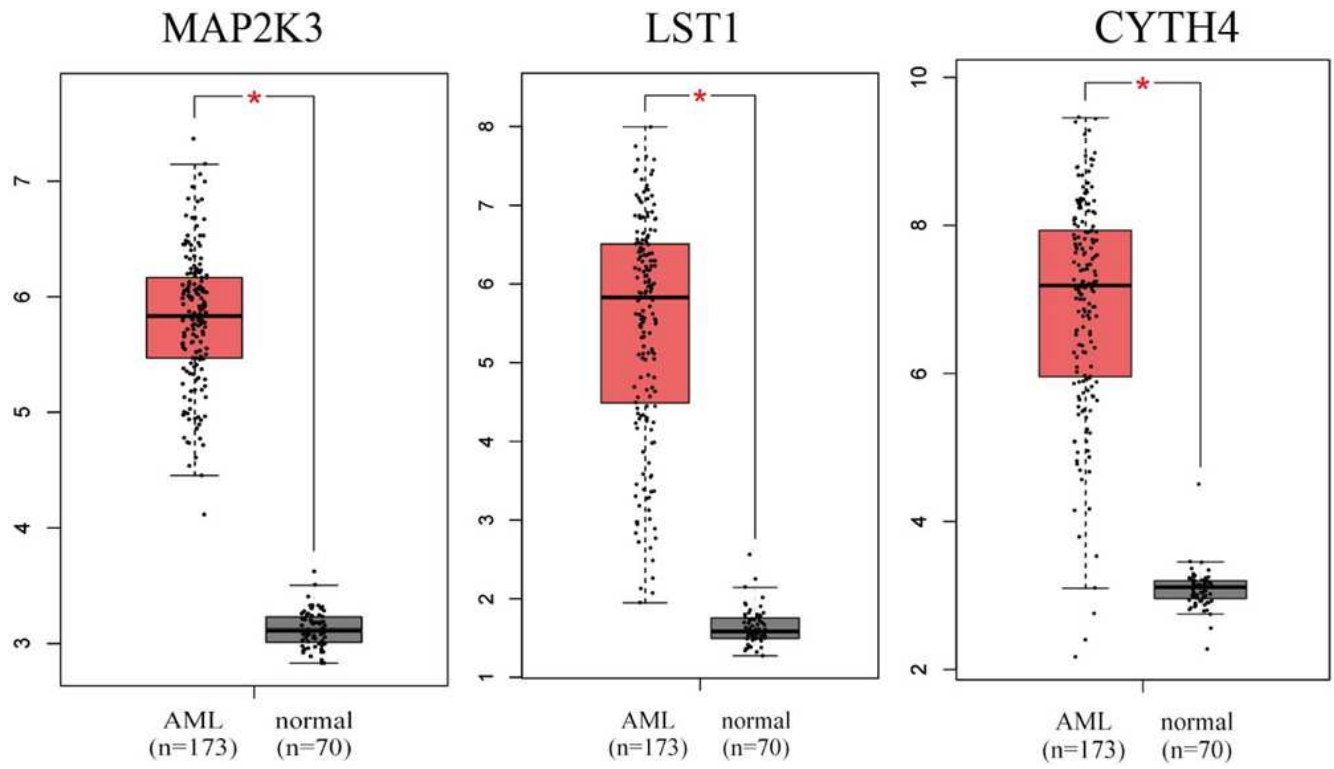


Figure 7. *The expression analysis of MAP2K3, LST1, and CYTH4 in AML patients. *P<0.05.*

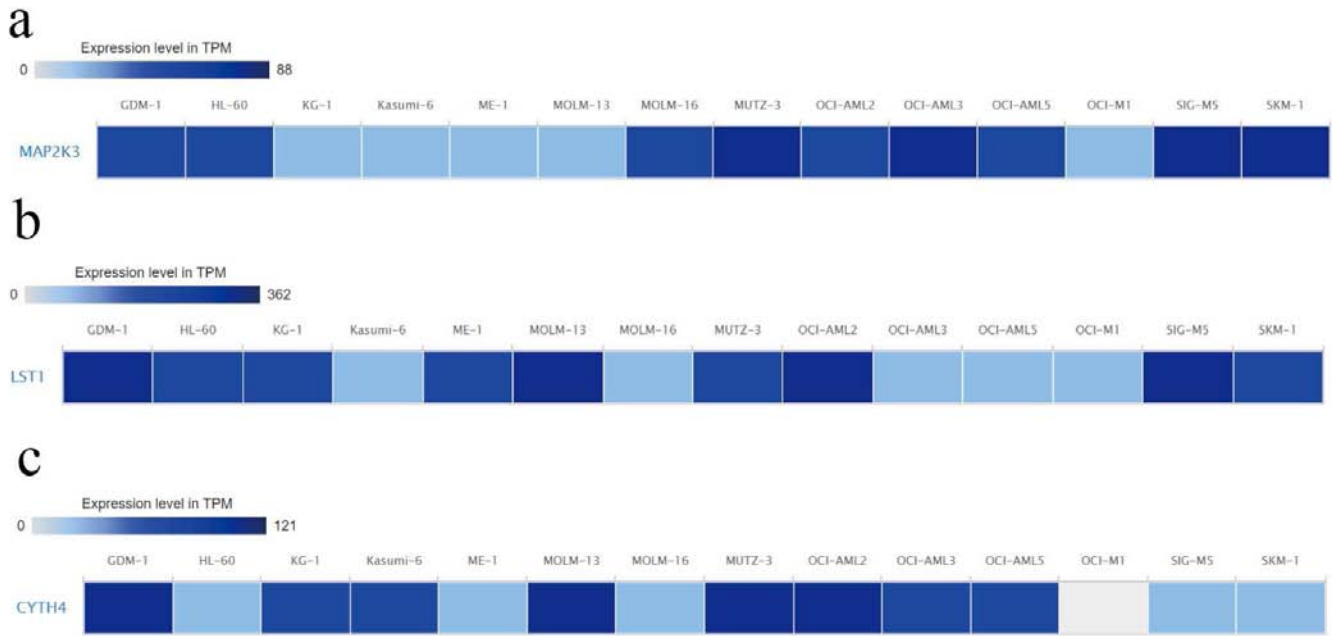


Figure 8. *The expression of MAP2K3, LST1, and CYTH4 in AML cell lines.*

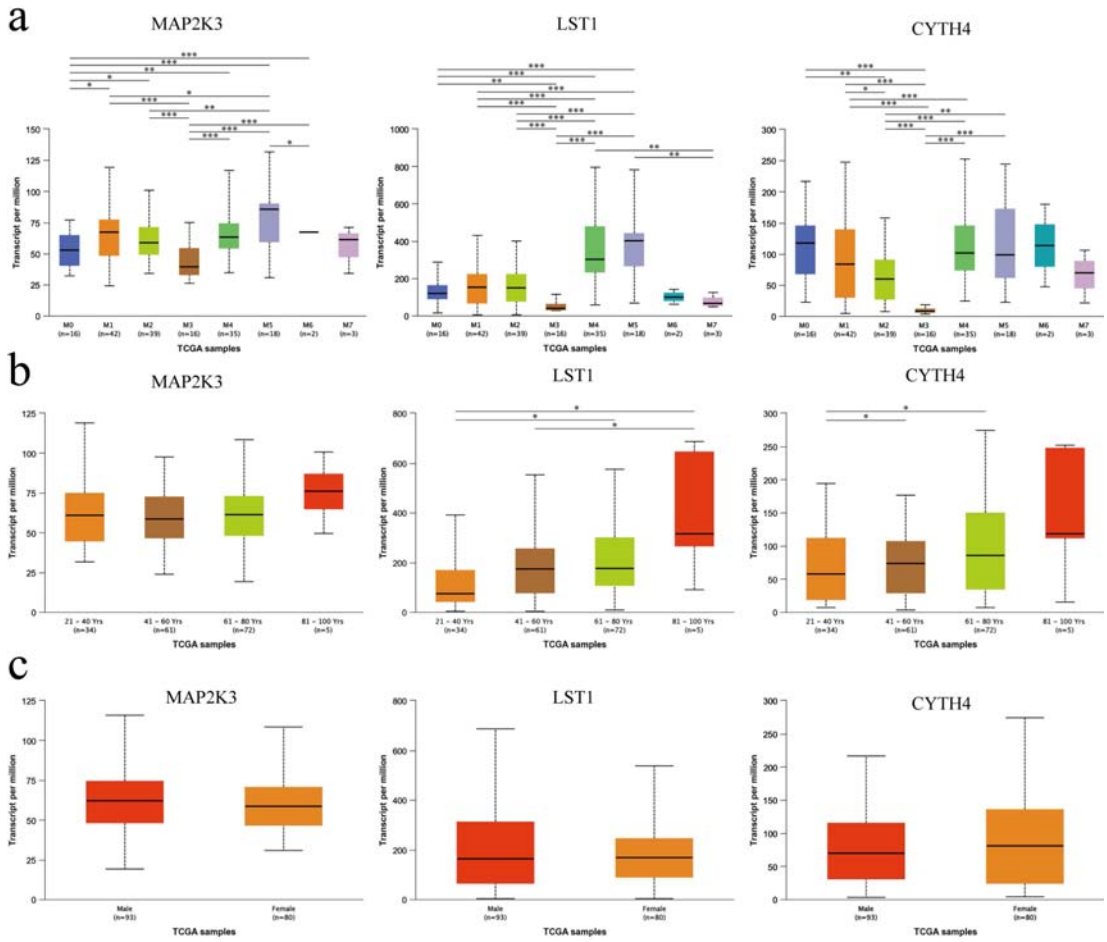


Figure 9. *The expression differences of the genes in AML subtypes and their relationship with AML patient's age, and gender. (a) Gene expression in AML patients of different subtypes. (b) Gene expression in AML patients of different ages. (c) Gene expression in AML patients of different genders. *P<0.05; **P<0.01; ***P<0.001.*

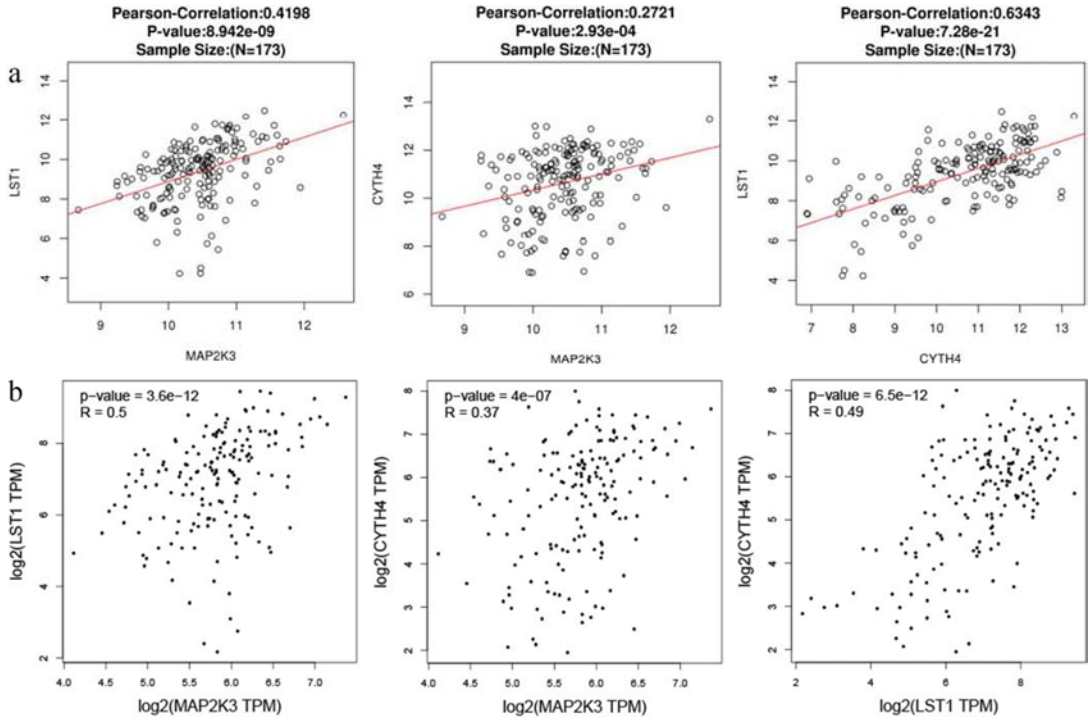


Figure 10. *(a) The relationship between the three genes in AML, analyzed by LinkedOmics (p <0.05). (b) The relationship between the three genes in AML, analyzed by GEPIA (p <0.05).*

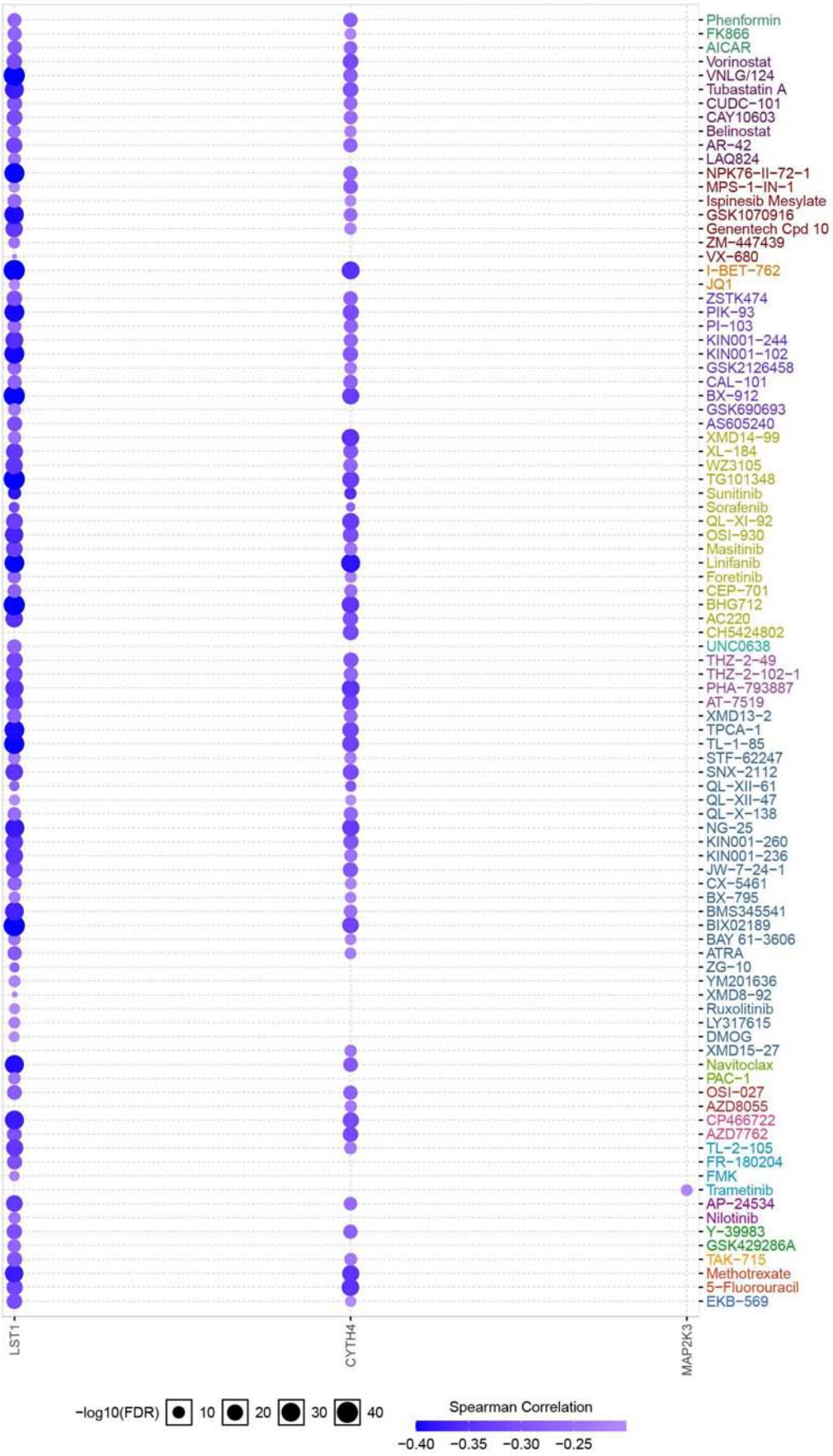


Figure 11. *Drug sensitivity analysis of the candidate genes in AML (GSCALite). The relation between gene expression and drugs was detected by Spearman correlation. The negative correlation suggests that the gene high expression is sensitive to the drug, vise verse.*

3.7. The Relationship Between Genes Expression and Drug Sensitivity in AML Cells

We next analyzed the correlation between the three genes expression (MAP2K3, LST1, and CYTH4) and small-molecule drugs sensitivity in AML cells using the GDSC IC50 drug data from the GSCALite database (Figure 11). Drug sensitivity analysis showed that AML cells with over-expression of MAP2K3, LST1, and CYTH4 were sensitive to certain small molecule drugs (including targeted drugs and non-targeted drugs). For example, we found that the sensitivity to the anti-metabolite methotrexate was associated with the over-expression of LST1 and CYTH4 in AML cells. Similarly, AML cell lines with overexpressing LST1 and CYTH4 also showed sensitivity to sorafenib (Nexavar), which drug targets were PDGFR, KIT, VEGFR, RAF. However, AML cells with high MAP2K3 expression showed sensitivity only to trametinib, which mainly acting on ERK MAPK signaling pathway. Detailed information about clinical small-molecule drugs were presented in Table 2.

4. Discussion

Despite advances in understanding the molecular mechanism of the disease initiation and progression, AML is still the hematopoietic tumor with high morbidity and mortality [24]. Therefore, investigating the AML pathogenesis and exploring new biomarkers have emerged as a hot research field. Microarray technology, as a newly developed molecular biology technology in recent years, could enable us to discover the nature of diseases, and it has proven to be a reliable way to research for potential human tumor biomarkers, including AML.

In the current study, the gene expression profiles of GSE65409 and GSE67936 were acquired from the GEO dataset to screen the DEGs between AML patients and healthy donor samples. Ultimately, 181 up-regulated genes were selected. Then, the interactions among the top 100 up-regulated genes were analyzed through DAVID database. GO enrichment analyses manifested that the up-regulated genes mainly participated in immune response, cell adhesion, and cell proliferation. Similarly, it has been reported tumor cells could accelerate cancer progression by down-regulating the expression of immunogenic molecules to avoid immune response [25, 26]. KEGG pathway enrichment results revealed that the overlapped DEGs participated in multiple pathways, such as the T cell receptor signaling pathway, TNF signaling pathway, and chemokine signaling pathway. Numerous researches disclosed that these pathways play a critical role in human cancer progression [27-29]. Therefore, the results suggested the up-regulated DEGs might be closely related to tumorigenesis and progression of AML, which was in agreement with the previous studies.

After a sequence of screening, analysis, and confirmation, a total of 3 candidate genes (MAP2K3, LST1, and CYTH4) were screened in connection with the poor outcome of AML patients, which was not reported in previous studies. BloodSpot database analysis showed that over-expression of the three candidate genes in AML patients may indicate high-risk groups at the level of genetic stratification. Meanwhile, the GEPIA and EMBL-EBI datasets confirmed the high expression of MAP2K3, LST1, and CYTH4 in AML patients and AML cell lines. Subsequently, in the UALCAN database, the expression differences of the three genes in the AML different subtypes were presented. It is worth noting that the three genes showed relatively low expression levels in M3 type AML, which may be due to the pathogenesis of M3 being inconsistent with other subtypes, and it was related to the formation of PML/RARA fusion genes. Moreover, we found that expression levels of the three genes was highest in elderly patients (especially patients over 60 years old), but the difference between genders had not statistically significant. To comprehensively analyze the biological role of MAP2K3, LST1, and CYTH4 in AML, the correlation analysis of the genes in AML were carried out by the LinkedOmics and GEPIA database. The results suggested that there was a positive correlation between MAP2K3, LST1, and CYTH4 in AML. The treatment of traditional drugs leads to adaptive resistance of patients, which is the main reason for the poor outcome of AML patients. Drug sensitivity analysis showed that AML patients with high expression of LST1 and CYTH4 were sensitive to most small-molecule drugs, while AML cells with high MAP2K3 expression were only sensitive to Trametinib. All of the evidence indicated that the three candidate genes might have potential application values in prognostic prediction and targeted drug therapy of AML.

Mitogen-activated protein kinase kinase 3 (MAP2K3, MKK3) is the main member of the bi-specific protein kinase [30]. Increasing evidence emphasizes MAP2K3 was involved in the progression and invasion of human tumor cells. Some studies demonstrated that MAP2K3, as a transcriptional target of up-regulated mutant (mut) p53, could maintain the proliferation and existence of human tumor cells [31]. In wild-type (wt) and mutp53-carrying cells, MAP2K3 deficiency induced endoplasmic reticulum stress and autophagy, which was conducive to stabilizing WTP53 and degrading mutp53, respectively [30]. In general, MAP2K3 is expected to be a promising anticancer therapeutic target in mutp53- and wtp53-carrying tumors. Leukocyte-specific transcript 1 (also termed LST1, B144, SLC21A6, OATP2, or OATP-C), is a myeloid leukocyte-specific membrane-anchored proteins encoded in the histocompatibility complex, with comprehensive selective splicing and immunomodulatory functions [32, 33]. It is mainly expressed in myeloid cells and as a negative regulator to participated in myeloid cell signaling [32]. Furthermore, it is also a key regulator of self-renewal, acts as a carcinogen to promotes oncogenesis [34]. Currently, the carcinogenic effect of LST1 aberrant expression was reported in hepatocellular carcinoma, breast cancer, and bladder cancer [33, 35, 36]. The carcinogenic mechanism of cytohesin family proteins CYTH4 (Cytohesin-4) is still poorly understood. Zhang et al. revealed [37] that CYTH4 was closely relevant to

multiple immune cells (such as $CD8^+$T cells, $CD4^+$T cells, and neutrophils) and crucial immune checkpoints (such as CTLA4, CD274, and PDCD1). Meanwhile, previous researches revealed the protein participated in tumorigenesis and progression as a carcinogenic factor in ovarian cancer and breast cancer [37, 38].

To the best of our knowledge, this is the first discovery of the possible carcinogenic role of MAP2K3, LST1, and CYTH4 in AML, which may be the potential prognostic markers and drug treatment targets for AML. However, there are some limitations in our study: Firstly, the sample size of our expression profile analysis was small, further studies with larger sample sizes are necessary. Secondly, the DEGs were obtained from clinical case samples, and verification of pre-clinical trials is necessary. Lastly, drug sensitivity data was just derived from AML cells, and these predictions are worthy of animal experiment verification in future studies. Overall, we look forward to this research will lay a sufficient theoretical foundation for the following experimental verification and provide reliable guidance for the clinical drug treatment of AML patients in the future.

5. Conclusions

In summary, this study systematically analyzed the dysfunctional genes related to the tumorigenesis and prognosis of AML. Eventually, three up-regulated genes (MAP2K3, LST1, and CYTH4) were identified in connection with the poor outcome of AML patients and might be regarded as novel biomarkers for drug screening. The over-expression of the three candidate genes in AML patients may suggest high-risk population at the level of genetic stratification, which will provide powerful guidance for clinical therapeutics.

Conflicts of Interest

All the authors do not have any possible conflicts of interest.

Acknowledgements

The authors would like to thank the Overseas Chinese Affairs Office of the State Council Key Discipline Construction Fund (51205002).

References

[1] Rowe JM, Tallman MS. How I treat acute myeloid leukemia [J]. Blood. 2010; 116 (17): 3147-56.

[2] Weinberg OK, Sohani AR, Bhargava P, et al. Diagnostic work-up of acute myeloid leukemia [J]. Am J Hematol. 2017; 92 (3): 317-21.

[3] Cai SF, Levine RL. Genetic and epigenetic determinants of AML pathogenesis [J]. Seminars in hematology. 2019; 56 (2): 84-9.

[4] Li S, Mason CE, Melnick A. Genetic and epigenetic heterogeneity in acute myeloid leukemia [J]. Current opinion in genetics & development. 2016; 36: 100-6.

[5] Burnett A, Wetzler M, Löwenberg B. Therapeutic advances in acute myeloid leukemia [J]. J Clin Oncol. 2011; 29 (5): 487-94.

[6] Stein EM, Tallman MS. Emerging therapeutic drugs for AML [J]. Blood. 2016; 127 (1): 71-8.

[7] Dillon R, Hills R, Freeman S, et al. Molecular MRD status and outcome after transplantation in NPM1-mutated AML [J]. Blood. 2020; 135 (9): 680-8.

[8] Roboz GJ. Current treatment of acute myeloid leukemia [J]. Current opinion in oncology. 2012; 24 (6): 711-9.

[9] Yin F, Shu L, Liu X, et al. Microarray-based identification of genes associated with cancer progression and prognosis in hepatocellular carcinoma [J]. J Exp Clin Cancer Res. 2016; 35 (1): 127.

[10] Cai D, Yang Y, Jiang Z, et al. Identification of 11 Differentially Expressed Hub Genes and Their Upstream microRNAs in Nasal-type NK/T-cell Lymphoma Based on Clinical Sample Analysis %J Cancer Research Journal [J]. 2020; 8 (4).

[11] Rung J, Brazma A. Reuse of public genome-wide gene expression data [J]. Nat Rev Genet. 2013; 14 (2): 89-99.

[12] Xu Y, Li X, Han Y, et al. A New Prognostic Risk Model Based on PPAR Pathway-Related Genes in Kidney Renal Clear Cell Carcinoma [J]. PPAR Res. 2020; 2020: 6937475.

[13] Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses [J]. Nucleic Acids Res. 2017; 45 (W1): W98-W102.

[14] Liu CJ, Hu FF, Xia MX, et al. GSCALite: a web server for gene set cancer analysis [J]. Bioinformatics. 2018; 34 (21): 3771-2.

[15] McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation [J]. Nucleic Acids Res. 2012; 40 (10): 4288-97.

[16] Franceschini A, Szklarczyk D, Frankild S, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration [J]. Nucleic Acids Res. 2013; 41 (Database issue): D808-15.

[17] Chin CH, Chen SH, Wu HH, et al. cytoHubba: identifying hub objects and sub-networks from complex i nteractome [J]. BMC Syst Biol. 2014; 8 Suppl 4: S11.

[18] Huang DW, Sherman BT, Tan Q, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists [J]. Genome Biol. 2007; 8 (9): R183.

[19] Bagger FO, Kinalis S, Rapin N. BloodSpot: a database of healthy and malignant haematopoiesis updated with purified and single cell mRNA sequencing profiles [J]. Nucleic Acids Res. 2019; 47 (D1): D881-d5.

[20] Li W, Cowley A, Uludag M, et al. The EMBL-EBI bioinformatics web and programmatic tools framework [J]. Nucleic Acids Res. 2015; 43 (W1): W580-4.

[21] Chandrashekar DS, Bashel B, Balasubramanya SAH, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses [J]. Neoplasia. 2017; 19 (8): 649-58.

[22] Vasaikar SV, Straub P, Wang J, et al. LinkedOmics: analyzing multi-omics data within and across 32 cancer types [J]. Nucleic Acids Res. 2018; 46 (D1): D956-D63.

[23] Carbon S, Ireland A, Mungall CJ, et al. AmiGO: online access to ontology and annotation data [J]. Bioinformatics. 2009; 25 (2): 288-9.

[24] Lowenberg B, Downing JR, Burnett A. Acute Myeloid Leukemia. [J]. N Engl J Med. 1999; 341 (14): 1051-62.

[25] Austin R, Smyth MJ, Lane SW. Harnessing the immune system in acute myeloid leukaemia [J]. Crit Rev Oncol Hematol. 2016; 103: 62-77.

[26] Davidson-Moncada J, Viboch E, Church SE, et al. Dissecting the Immune Landscape of Acute Myeloid Leukemia [J]. Biomedicines. 2018; 6 (4).

[27] Gasparini C, Celeghini C, Monasta L, et al. NF-κB pathways in hematological malignancies [J]. Cellular and Molecular Life Sciences. 2014; 71 (11): 2083-102.

[28] Wang M, Zhang C, Tian T, et al. Increased Regulatory T Cells in Peripheral Blood of Acute Myeloid Leukemia Patients Rely on Tumor Necrosis Factor (TNF)-alpha-TNF Receptor-2 Pathway [J]. Front Immunol. 2018; 9: 1274.

[29] Scala S. Molecular Pathways: Targeting the CXCR4-CXCL12 Axis--Untapped Potential in the Tumor Microenvironment [J]. Clin Cancer Res. 2015; 21 (19): 4278-85.

[30] Baldari S, Ubertini V, Garufi A, et al. Targeting MKK3 as a novel anticancer strategy: molecular mechanisms and therapeutical implications [J]. Cell Death Dis. 2015; 6: e1621.

[31] Gurtner A, Starace G, Norelli G, et al. Mutant p53-induced up-regulation of mitogen-activated protein kinase kinase 3 contributes to gain of function [J]. J Biol Chem. 2010; 285 (19): 14160-9.

[32] Draber P, Stepanek O, Hrdinka M, et al. LST1/A is a myeloid leukocyte-specific transmembrane adaptor protein recruiting protein tyrosine phosphatases SHP-1 and SHP-2 to the plasma membrane [J]. J Biol Chem. 2012; 287 (27): 22812-21.

[33] Cui Y, Konig J, Nies AT, et al. Detection of the human organic anion transporters SLC21A6 (OATP2) and SLC21A8 (OATP8) in liver and hepatocellular carcinoma [J]. Lab Invest. 2003; 83 (4): 527-38.

[34] Haupt KA, Enright AL, Ferdous AS, et al. The molecular basis of LST-1 self-renewal activity and its control of stem cell pool size [J]. Development. 2019; 146 (20).

[35] Hsu HM, Chu CM, Chang YJ, et al. Six novel immunoglobulin genes as biomarkers for better prognosis in triple-negative breast cancer by gene co-expression network analysis [J]. Sci Rep. 2019; 9 (1): 4484.

[36] D'Aloia A, Berruti G, Costa B, et al. RalGPS2 is involved in tunneling nanotubes formation in 5637 bladder cancer cells [J]. Exp Cell Res. 2018; 362 (2): 349-61.

[37] Zhang Q, Wang Q, Wu S, et al. Clinical implication and immunological characterisation of the ARF-GEF family member CYTH4 in ovarian cancer [J]. Autoimmunity. 2020; 53 (8): 434-42.

[38] Kennedy BM, Harris RE. Cyclooxygenase and lipoxygenase gene expression in the inflammogenesis of breast cancer [J]. Inflammopharmacology. 2018.