International Journal of Biology and Life Sciences
ISSN: 2957-9511 | Vol. 7, No. 2, 2024

Analysis of ALDOC Expression and Prognostic Value in Acute Myeloid Leukemiahe

Jianwei Huang [1, 2, *]

[1] Jinan University, Guangzhou, Guangdong, China
[2] The First Affiliated Hospital of Jinan University, Guangzhou, Guangdong, China
[*] **Corresponding author Email**: jyhlhjw@126.com

Abstract: Objective: This study aimed to investigate the expression pattern of ALDOC in patients with acute myeloid leukemia (AML) and to assess its potential as a prognostic marker. Methods: Data for this study were derived from the public databases TCGA and GTEx. Based on the expression levels of ALDOC in different age groups, the impact of ALDOC expression on the prognosis of AML patients was evaluated using univariate and multivariate Cox regression analyses, in conjunction with Kaplan-Meier survival curves. Bioinformatics methods were also employed to explore the possible mechanisms of ALDOC action in AML. Results: A significant difference in the expression of the metabolic gene ALDOC was found between AML patients and normal individuals ($P < 0.05$). ALDOC was identified as a prognostic marker for patients under the age of 60. Bioinformatics analysis suggested that high expression of ALDOC could upregulate pathways involved in allograft rejection, response to IFN-α, response to IFN-γ, and phagocytosis, particularly indicating a poorer prognosis in patients under the age of 60. Conclusion: High expression of ALDOC is associated with poor prognosis in AML patients under the age of 60, and ALDOC may serve as a marker for AML prognosis and a promising potential therapeutic target for further research.

Keywords: Acute Myeloid Leukemia; Bioinformatics Analysis; ALDOC.

1. Introduction

Acute myeloid leukemia (AML) is the most common and highly malignant type of leukemia in adults to date. The global incidence of AML is 3-5 cases per 100,000 people, with a median age at initial diagnosis of approximately 67.1 years [1, 2]. The malignant proliferation of poorly differentiated cells in AML is associated with genetic triggers, including amplification, deletion, rearrangement, and point mutations of genetic material, and exhibits a high degree of heterogeneity[3, 4]. Due to the heterogeneity of AML, leukemic cells exhibit a variety of metabolic characteristics involving multiple metabolic pathways[5]. Clinically, about 45% of AML patients have a normal karyotype[6]. Certain important biological processes, such as metabolic targeting analysis, may assist in patient stratification and management, and reveal new therapeutic targets[7], thereby improving patient prognosis. We identified metabolic genes with significant associations with clinical prognosis based on our analysis and found that ALDOC is the encoding gene for fructose-bisphosphate aldolase C[8]. ALDOC has been reported in various solid tumors such as colorectal, gallbladder, and breast cancer, and is highly correlated with prognosis[9-12]. However, the role of ALDOC in hematological malignancies is currently unknown, and whether its expression level can serve as a predictive factor for prognosis in AML patients remains unclear. This study primarily explores its prognostic value in AML based on database analysis, laying the foundation for further investigation of the role and mechanism of metabolic genes like ALDOC in AML and providing potential new therapeutic targets for AML treatment.

2. Materials and Methodsrganization

2.1. Selection of Metabolic Genes

In this study, we obtained gene sequencing data and clinical information from 173 AML tumor patients through the TCGA database and downloaded 337 gene expression samples from healthy humans from the GTEx database as controls. Metabolic genes were sourced from the Gene Set Enrichment Analysis website (c2.cp.kegg.v7.2.symbols.gmt). The Wilcox-test was utilized to analyze differences between the normal and tumor groups. Genes with a $|\log2FC| > 2$ (fold change > 4) and $FDR < 0.05$ were selected as differentially expressed genes. Genes associated with prognosis were identified based on overall survival time and survival status. The R heatmap package was used to plot gene heatmaps and volcano plots. In the merged gene expression database, survival time and survival status for both control and tumor groups were included. Data were processed using the R survival package ($P < 0.01$), and survival curves associated with prognosis were plotted ($P < 0.01$). Descriptive statistics were used to report the basic clinical characteristics of patients. The Mann-Whitney U test and the chi-squared test (Pearson's chi-squared test) were used to compare numerical and categorical data, respectively, with continuity corrections or Fisher's exact test applied when necessary. All statistical analyses were performed using R commands.

2.2. Clinical and Molecular Characteristic Analysis Based on ALDOC Levels

To further analyze the clinical and molecular characteristics associated with different ALDOC expression levels across various age groups, we divided the patients into four groups based on the median ALDOC expression level and whether the age was greater than 60 years (including 60 years). We explored the associations with clinical factors such

as platelet count, bone marrow blast percentage, peripheral blood blast percentage, transplant status, FAB classification, karyotype, and common genetic mutation phenotypes. FAB classification includes M0 to M7; karyotype includes normal karyotype, complex karyotype, del (5), etc.; common genetic mutations included genes such as CEBPA, RUNX1, AXSL1, TP53, DNMT3A, etc. Statistical difference analysis was performed using R language functions, with the Mann-Whitney U test for numerical data and Person's chi-squared test or Fisher's exact test for categorical data. A P-value < 0.05 was considered statistically significant.

2.3. Univariate and Multivariate Cox Proportional Hazards Analysis Based on Age Grouping

This study assessed the prognostic significance of clinical and molecular characteristics in two age groups (<60 years and ≥60 years) using univariate and multivariate Cox proportional hazards regression models. The dependent variable was survival outcome and overall survival time. Covariates in the univariate Cox model included ALDOC expression level, prognostic stratification, bone marrow and peripheral blood precursor cell proportions, transplant status, and genetic conditions of CEBPA, RUNX1, AXSL1, TP53, DNMT3A, etc. Factors with a P-value < 0.2 in the univariate Cox model were included as covariates in the multivariate Cox regression model. Statistical tools were the R language survival package, with a significance level set at $P < 0.05$.

2.4. Survival Comparison Based on ALDOC Expression Level and Age Grouping

Among the 151 AML patients with precise survival time data, we used the Kaplan-Meier method to calculate the overall survival rates for patients with different ALDOC expression levels (divided by the median ALDOC expression level) in different age groups (whether age is ≥60 years) and plotted the curves. The Log-rank test was used to compare whether the curves were different, with a test threshold set at $P < 0.05$.

2.5. Selection of Differentially Expressed Genes Between High and Low ALDOC Expression Groups

Raw analysis data was sourced from the TCGA database. AML patients were divided into high and low ALDOC expression groups based on the median ALDOC expression level. Conditions were set as $|\log2FC| > 1$ (fold change in expression > 2) and adj. $P < 0.05$ as thresholds for gene selection. The R software's "DESeq2" package was used for the differential gene expression analysis between the two groups. The "pheatmap" package and "ggplot2" package were used to create heatmaps and volcano plots, respectively, to visually display the results.

2.6. GO Functional and KEGG Enrichment Analysis of Differentially Expressed Genes Between High and Low ALDOC Expression Groups

To investigate the functions of differentially expressed genes between the high and low ALDOC expression groups in AML patients, pathway enrichment analysis was performed using KEGG. A P-value < 0.05 was considered statistically significant. The R software's "clusterProfiler" package was used to visualize the enrichment results.

2.7. GSEA Enrichment Analysis of Differentially Expressed Genes Between High and Low ALDOC Expression Groups

Considering that GO and KEGG analyses mainly utilize differentially expressed genes, ignoring the contributions of the remaining genes in these pathways, gene set enrichment analysis (GSEA) was used to comprehensively calculate the enrichment of predefined functional gene sets across all ranked genes in AML patients, observing the main cellular pathways involved and changes in these pathways (upregulation or downregulation). The R software's "clusterProfiler" package was used for GSEA enrichment analysis of the high and low ALDOC expression groups, using the HallMark.gmt from the MSig DB Collections as a reference gene set.

3. Results

3.1. Selection Results of Differentially Expressed Metabolic Genes

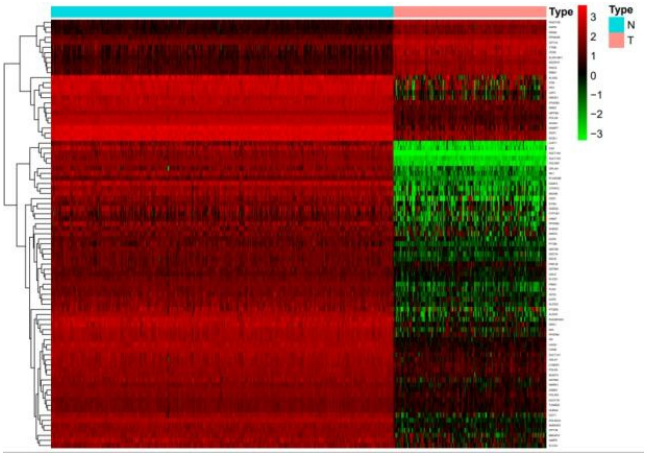


Figure 1. Heatmap of Differentially Expressed Metabolic Genes in Acute Myeloid Leukemia

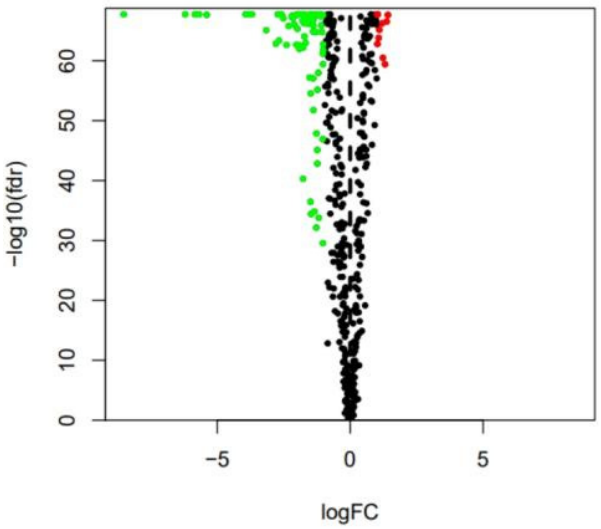


Figure 2. Volcano Plot of Differentially Expressed Metabolic Genes in Acute Myeloid Leukemia

To identify differentially expressed metabolic genes between the leukemia tumor group and the normal control

group, a total of 106 genes with differential expression were obtained based on the criteria of |log2FC| > 2 and FDR < 0.05, and heatmaps and volcano plots were constructed. Among these, there were 26 upregulated metabolic genes and 80 downregulated metabolic genes. Figure 1 displays the heatmap of differentially expressed metabolic genes in acute myeloid leukemia, and Figure 2 illustrates the volcano plot of differentially expressed metabolic genes in acute myeloid leukemia.

3.2. Analysis of Prognostic Correlation with Differentially Expressed Genes

To further determine which members among the differentially expressed metabolic genes are highly correlated with the clinical prognosis of acute myeloid leukemia (AML), we obtained complete survival times and survival statuses from AML patients in the TCGA database. Since there were no missing visits among the 151 patients, no data exclusion was necessary for this analysis. We processed gene expression data and clinical survival information using the R survival package, and based on the median level of gene expression, AML patients were divided into high and low expression groups. Kaplan-Meier estimation was applied to estimate survival curves, and the Log Rank test was conducted ($P < 0.01$). Six differential metabolic genes were found to be closely related to prognosis, namely ALDOC, RDH5, ARG1, PMM1, AK1, and DGKA. The survival curves for these genes are illustrated in Figure 3. The Kaplan-Meier curves indicated that high expression of the aforementioned genes was associated with poorer overall survival rates, with P-values of 0.0002 (Figure 3A), 0.0005 (Figure 3B), 0.007 (Figure 3D), 0.007 (Figure 3E), and 0.009 (Figure 3F), respectively. In contrast, low expression of ARG1 was associated with poorer overall survival, with a P-value of 0.006 (Figure 3C). Kaplan-Meier Survival Curves for Differentially Expressed Metabolic Genes in AML:(A) ALDOC, (B) RDH5, (C) ARG1, (D) PMM1, (E) AK1, (F) DGKA; High expression of these genes correlates with poorer overall survival, except for ARG1, where low expression is associated with poorer survival.

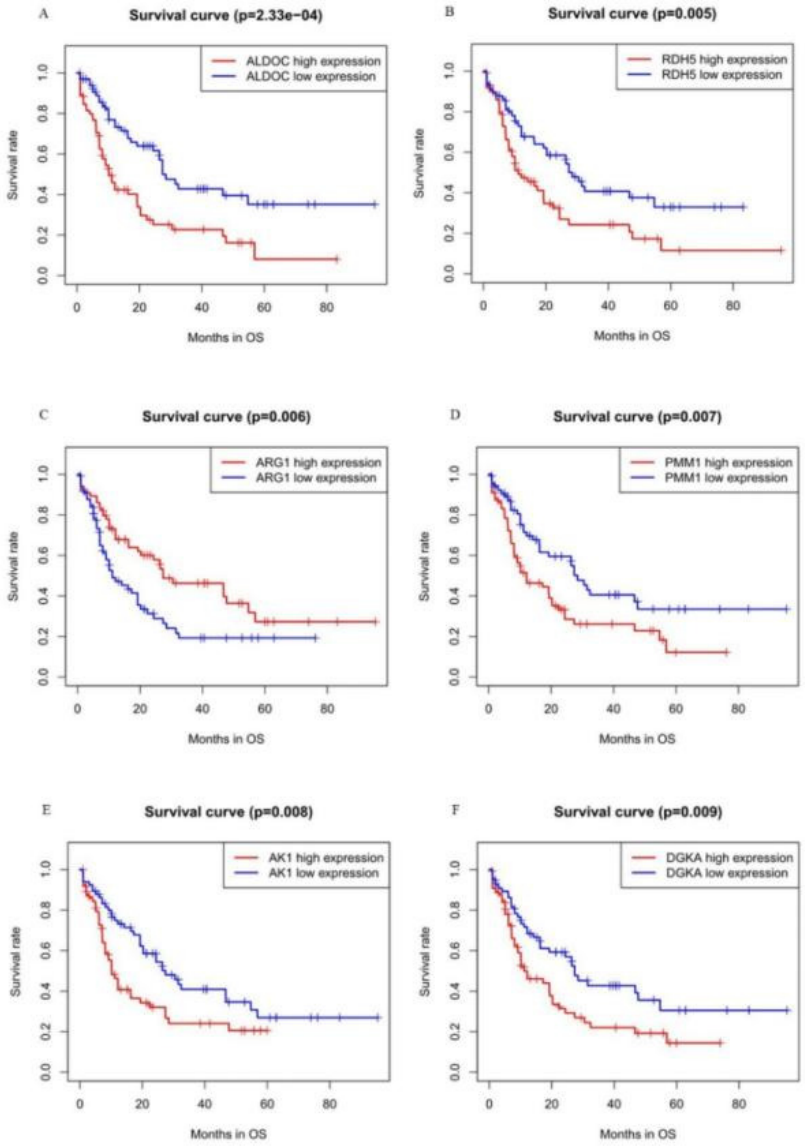


Figure 3. Correlation of Six Differentially Expressed Metabolic Genes with Clinical Prognosis

3.3. Basic Clinical Characteristics Correlated with Prognostic Differential Genes

To further investigate the correlation between the expression levels of differential metabolic genes and the basic clinical characteristics of acute myeloid leukemia (AML) patients, AML patients were defined as high and low expression groups based on the median expression level, with

76 cases in the high expression group and 75 in the low expression group. Table 1 shows the relationship between the expression levels of the six differential metabolic genes correlated with prognosis and the clinical information characteristics of AML patients. The results indicate that four prognostic-related differential metabolic genes (ALDOC, PMM1, AK1, DGKA) have a significant correlation with age (with 60 years as the boundary); four prognostic-related differential metabolic genes (ALDOC, PMM1, AK1, RDH5) have a significant correlation with FAB classification (M3 group and non-M3 group); three prognostic-related differential metabolic genes (ALDOC, PMM1, AK1) suggest a clinically relevant possibility of correlation with genetic molecular risk stratification. Consistent with large-scale epidemiological incidence statistics and clinical phenomena, none of the above genes were found to be correlated with gender in the TCGA clinical data. In summary, ALDOC, AK1, and PMM1 are most closely related to basic clinical characteristics, including age, FAB classification, and risk stratification. This study selected ALDOC as the target gene for subsequent analysis.

Table 1. Relationship between the Expression Levels of Six Prognostic-Related Differential Metabolic Genes and Clinical Information Characteristics in AML Patients

Characters	Age		FAB Typing		Risk		Sex	
	<60	≥60	M3	Non M3	Non-Poor	Poor	Female	Male
ALDOC								
High expression	32	44	1	75	47	29	30	46
Low expression	52	23	14	61	68	7	38	37
P		0.0013*		0.0009*		7.34E-05*		0.2229
PMM1								
High expression	32	44	2	74	50	26	31	45
Low expression	52	23	13	62	65	10	37	38
P		0.0013*		0.0060*		0.0048*		0.3726
AK1								
High expression	35	41	3	73	49	27	32	44
Low expression	49	26	12	63	66	9	36	39
P		0.0263*		0.0275*		0.0013*		0.5725
DGKA								
High expression	33	43	6	70	55	21	36	40
Low expression	51	24	9	66	60	15	32	43
P		0.0040*		0.5678		0.3631		0.6766
RDH5								
High expression	37	39	2	74	62	14	31	45
Low expression	47	28	13	62	53	22	37	38
P		0.1175		0.0060*		0.1668		0.3726
ARG1								
High expression	44	32	7	69	55	21	36	40
Low expression	40	35	8	67	60	15	32	43
P		0.6889		0.9784		0.3631		0.6766

3.4. Correlation of ALDOC with Clinical and Pathological Features in Different Age Groups

To ascertain the clinical significance of ALDOC expression in patients across various age brackets, patients were categorized into an elderly group (age ≥ 60 years, n=67) and a non-elderly group (age < 60 years, n=84), with the median ALDOC expression level serving as the cutoff within each group. The correlation of ALDOC with clinical and pathological features in elderly AML patients is detailed in Table 2. In the elderly group, the Mann-Whitney U test identified significant correlations between ALDOC expression levels and platelet count (P = 0.0823), bone marrow blasts (P = 0.0275), and peripheral blood blasts (P = 0.0353). Additionally, using the Person's chi-squared test or Fisher's exact test, significant correlations were observed between ALDOC expression levels and genetic abnormalities such as del (5) (P = 0.0378), RUNX1 (P = 0.0212), and TP53 (P = 0.0212). The correlation of ALDOC with clinical and pathological features in non-elderly AML patients is presented in Table 3. Employing the same methodology, it was found that in the non-elderly group, ALDOC expression levels were correlated with FAB classification (with P values for M2, M3, and M5 being 0.0852, 0.0024, and 0.0066, respectively), and significantly associated with CEBPA (P = 0.0047), ASXL1 (P = 0.0809), and DNMT3A (P = 0.0106).

3.5. Univariate/Multivariate Cox Regression Analysis of Prognostic Factors in AML Patients

We further constructed univariate and multivariate Cox proportional hazards regression models for AML patients of different age groups. Figure 4 shows the results of the univariate regression analysis for overall survival (OS) in AML patients aged 60 years or older. The results indicate that receiving a transplant (P = 0.0007) and TP53 (P = 0.0019) are independent prognostic factors affecting the overall survival of AML patients.

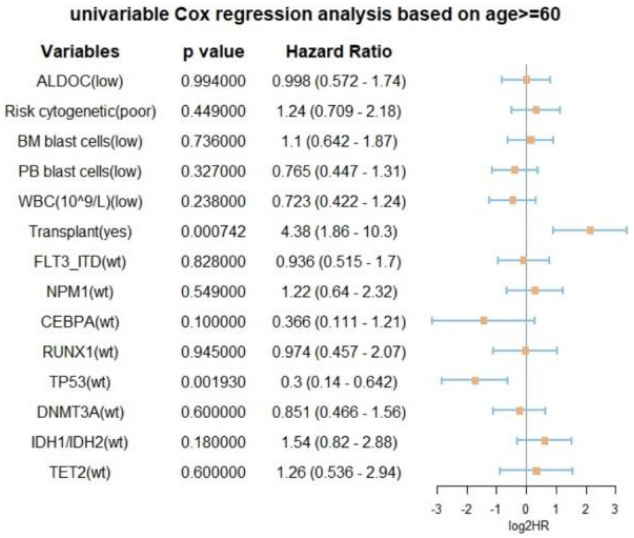


Figure 4. Univariate Cox Regression Analysis for Overall Survival in AML Patients Aged 60 and Above

In the Cox regression model, variables with P < 0.2 in the univariate Cox regression were included in the multivariate Cox regression analysis. The covariates that met the inclusion threshold were transplant receipt (P = 0.0007), CEBPA (P = 0.1), TP53 (P = 0.0019), and IDH1/IDH2 (P = 0.18). Figure 5 presents the results of the multivariate regression analysis for overall survival (OS) in AML patients aged 60 or above. The results show that transplant receipt (HR 3.87 [1.62-9.26], P = 0.002) and TP53 (HR 0.42 [0.19-0.93], P = 0.033) are independent prognostic factors affecting the overall survival of AML patients. In AML patients aged 60 or above, the death rate of patients who received a transplant after chemotherapy is 3.87 times the death rate of patients who only received chemotherapy.

Table 2. Clinical and Molecular Characteristics of Patients Aged ≥60 Years Based on ALDOC Expression Levels

Characters	elderly group (age≥60) high (n = 42)	low (n = 25)	*P*
Gender/n (%)			
Female	16 (0.24)	12 (0.18)	
Male	26 (0.39)	13 (0.19)	0.5899
HB/×g/l, median (range)	9 (7-13)	10 (8-13)	0.2627[&]
WBC/$\times10^9$/l,median(range)	8.35 (0.6-98.2)	30.8 (0.9-131.5)	0.1182[&]
PLT/$\times10^9$/l, median (range)	62.5 (11-174)	41 (11-168)	0.0823[&]
BM blast/%, median (range)	62 (0-98)	78 (11-99)	0.0275[&]
PB blast/%, median (range)	20 (0-84)	49 (0-97)	0.0353[&]
FAB subtypes/n (%)			
M0	6 (0.09)	3 (0.04)	1.0000[*]
M1	9 (0.13)	5 (0.07)	1.0000
M2	10 (0.15)	6 (0.09)	1.0000
M3	1 (0.01)	3 (0.04)	0.1425[*]
M4	8 (0.12)	6 (0.09)	0.8638
M5	5 (0.07)	2 (0.03)	0.7040[*]
M6	1 (0.01)	0 (0)	1.0000[*]
M7	6 (0.09)	3 (0.04)	1.0000[*]
Karyotype/n (%)			
Normal Karyotype	17 (0.25)	9 (0.13)	0.9168
Complex Karyotype	9 (0.13)	2 (0.03)	0.1891[*]
del(7)	0 (0)	0 (0)	1.0000[*]
inv(16)	0 (0)	2 (0.03)	0.1357[*]
del(5)	0 (0)	0 (0)	0.0378
t(8;21)	1 (0.01)	0 (0)	1.0000[*]
t(9;11)	1 (0.01)	0 (0)	1.0000[*]
others	15 (0.22)	12 (0.18)	0.4629
FLT3-ITD/n (%)			
Wild type	31 (0.46)	16 (0.24)	
Mutation	11 (0.16)	9 (0.13)	0.5669
NPM1/n (%)			
Wild type	35 (0.52)	16 (0.24)	
Mutation	7 (0.1)	9 (0.13)	0.1339
CEBPA/n (%)			
Wild type	41 (0.61)	23 (0.34)	
Mutation	1 (0.01)	2 (0.03)	0.5507[*]
RUNX1/n (%)			
Wild type	33 (0.49)	25 (0.37)	
Mutation	9 (0.13)	0 (0)	0.0212[*]
ASXL1/n (%)			
Wild type	40 (0.6)	25 (0.37)	
Mutation	2 (0.03)	0 (0)	0.5251[*]
TP53/n (%)			
Wild type	33 (0.49)	25 (0.37)	
Mutation	9 (0.13)	0 (0)	0.0212[*]
DNMT3A/n (%)			
Wild type	29 (0.43)	20 (0.3)	
Mutation	13 (0.19)	5 (0.07)	0.4882
IDH1/IDH2/n (%)			
Wild type	34 (0.51)	16 (0.24)	
Mutation	8 (0.12)	9 (0.13)	0.2106
TET2/n (%)			

Characters	elderly group (age≥60) high (n = 42)	low (n = 25)	*P*
Wild type	37 (0.55)	23 (0.34)	
Mutation	5 (0.07)	2 (0.03)	0.7040*
Therpy/n (%)			
Chemotherapy-alone	32 (0.48)	21 (0.31)	
Transplant	10 (0.15)	4 (0.06)	0.6529

In the fourth column of the table, "&" indicates the use of the Mann-Whitney U test, "*" indicates the use of Fisher's exact test, and no marking indicates the use of the Person's chi-squared test.

Table 3. Clinical and Molecular Characteristics of Patients Aged <60 Years Based on ALDOC Expression Levels

Characters	non-elderly group (age<60) high (n = 34)	low (n = 50)	*P*
Gender/n (%)			
Female	16 (0.19)	24 (0.29)	
Male	18 (0.21)	26 (0.31)	1.0000
HB/×g/l, median (range)	9 (6-12)	9 (6-13)	0.6502 &
WBC/$\times 10^9$/l, median(range)	30.4 (1.5-171.9)	22.55 (0.4-223.8)	0.1077&
PLT/$\times 10^9$/l, median (range)	53 (9-351)	39 (9-232)	0.2185&
BM blast/%, median (range)	76 (0-95)	67 (0-100)	0.3407&
PB blast/%, median (range)	41 (0-91)	45 (0-97)	0.7185&
FAB subtypes/n (%)			
M0	4 (0.05)	2 (0.02)	0.2164*
M1	12 (0.14)	9 (0.11)	0.1235
M2	5 (0.06)	17 (0.2)	0.0852
M3	0 (0)	11 (0.13)	0.0024*
M4	5 (0.06)	10 (0.12)	0.7401
M5	7 (0.08)	1 (0.01)	0.0066*
M6	1 (0.01)	0 (0)	0.4048*
M7	4 (0.05)	2 (0.02)	0.2164*
Karyotype/n (%)			
Normal Karyotype	14 (0.17)	20 (0.24)	
Complex Karyotype	5 (0.06)	2 (0.02)	0.1133*
del(7)	1 (0.01)	1 (0.01)	1.0000*
inv(16)	3 (0.04)	5 (0.06)	1.0000*
del(5)	0 (0)	1 (0.01)	1.0000*
t(8;21)	2 (0.02)	3 (0.04)	1.0000*
t(9;11)	1 (0.01)	0 (0)	0.4048*
others	10 (0.12)	20 (0.24)	0.4460
FLT3-ITD/n (%)			
Wild type	23 (0.27)	38 (0.45)	
Mutation	11 (0.13)	12 (0.14)	0.5529
NPM1/n (%)			
Wild type	21 (0.25)	41 (0.49)	
Mutation	13 (0.15)	9 (0.11)	0.0691
CEBPA/n (%)			
Wild type	34 (0.4)	40 (0.48)	
Mutation	0 (0)	10 (0.12)	0.0047*
RUNX1/n (%)			
Wild type	32 (0.38)	48 (0.57)	
Mutation	2 (0.02)	2 (0.02)	1.0000*
ASXL1/n (%)			
Wild type	34 (0.4)	50 (0.6)	
Mutation	0 (0)	0 (0)	0.0809

Characters	non-elderly group (age<60) high (n = 34)	low (n = 50)	*P*
TP53/n (%)			
Wild type	32 (0.38)	50 (0.6)	
Mutation	2 (0.02)	0 (0)	0.1609 *
DNMT3A/n (%)			
Wild type	22 (0.26)	45 (0.54)	
Mutation	12 (0.14)	5 (0.06)	0.0106
IDH1/IDH2/n (%)			
Wild type	27 (0.32)	45 (0.54)	
Mutation	7 (0.08)	5 (0.06)	0.2118 *
TET2/n (%)			
Wild type	33 (0.39)	46 (0.55)	
Mutation	1 (0.01)	4 (0.05)	0.6438 *
Therpy/n (%)			
Chemotherapy-alone	11 (0.13)	20 (0.24)	
Transplant	23 (0.27)	30 (0.36)	0.6294

In the fourth column of the table, "&" indicates the use of the Mann-Whitney U test, "*" indicates the use of Fisher's exact test, and no marking indicates the use of the Person's chi-squared test.

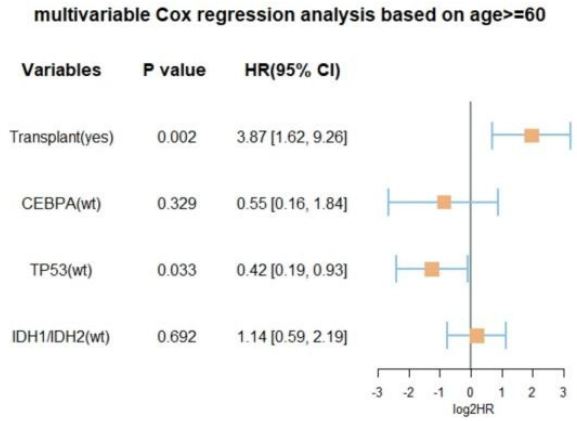


Figure 5. Multivariate Cox Regression Analysis for Overall Survival in AML Patients Aged 60 or Above

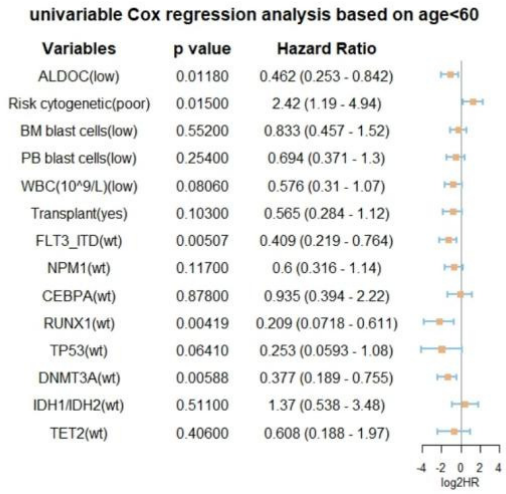


Figure 6. Univariate Cox Regression Analysis for Overall Survival in AML Patients Aged Under 60

It represents the results of the univariate Cox regression analysis for overall survival (OS) in AML patients aged under 60 at figure 6. The results of the univariate Cox regression analysis show that ALDOC expression level (P = 0.0118), prognostic stratification (P = 0.0150), FLT3-ITD (P = 0.005), RUNX1 (P = 0.0042), and DNMT3A (P = 0.0059) are independent prognostic factors that affect the survival of AML patients aged under 60.

It represents the results of the multivariate Cox regression analysis for overall survival (OS) in AML patients aged under 60 at figure 7. The multivariate Cox regression analysis shows that FLT3-ITD (P = 0.007), RUNX1 (P = 0.001), and DNMT3A (P = 0.022) are independent prognostic factors affecting the survival of AML patients in this age group.

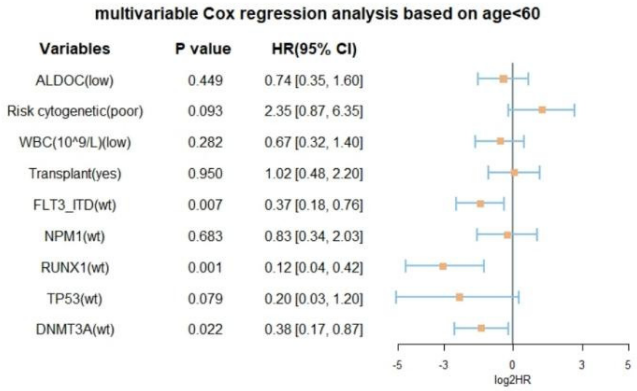


Figure 7. Multivariate Cox Regression Analysis for Overall Survival in AML Patients Under 60 Years of Age

3.6. Prognostic Value of ALDOC Expression Levels in AML Patients

To further assess the prognostic value of ALDOC in elderly and non-elderly patients in terms of overall survival (OS), we utilized the survminer package to plot Kaplan-Meier curves and performed the Log-rank test to compare the differences between the curves. Figure 8 illustrates the prognostic situation of ALDOC expression levels in AML patients aged 60 or above, while Figure 9 shows the prognostic situation in AML patients under 60 years of age. The results are consistent with our Cox proportional hazards regression model, indicating that in elderly patients, there is no difference in

overall survival between those with high ALDOC expression and those with low ALDOC expression (P = 0.99). In contrast, in non-elderly patients, those with high ALDOC expression have a shorter overall survival time compared to those with low ALDOC expression (P = 0.0099).

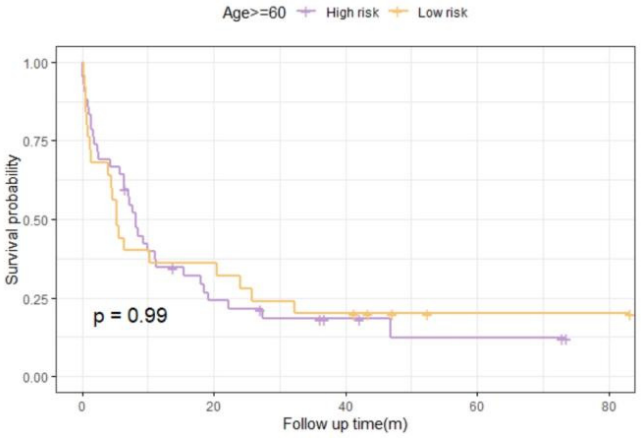


Figure 8. Prognostic Situation of ALDOC Expression Levels in AML Patients Aged 60 or Above

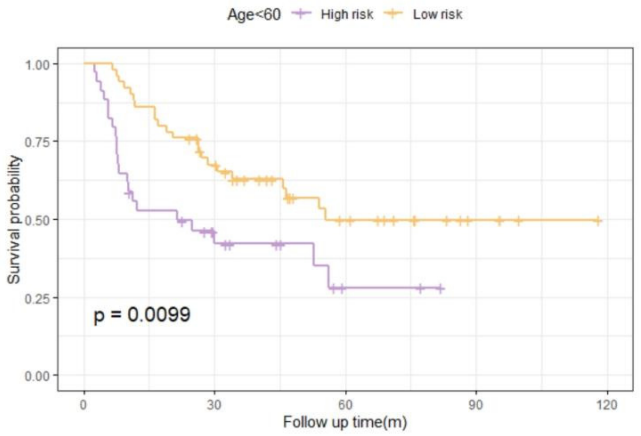


Figure 9. Prognostic Situation of ALDOC Expression Levels in AML Patients Under 60 Years of Age

3.7. Selection of Differentially Expressed Genes Between High and Low ALDOC Expression Groups

Referencing the TCGA database and using |log2FC| > 1 and adj. P < 0.05 as the selection criteria, a total of 120 differentially expressed genes were identified between the high and low ALDOC expression groups. Of these, 106 genes were upregulated, and 14 genes were downregulated. Based on these results, a heatmap of differentially expressed genes (Figure 10) and a volcano plot (Figure 11) were constructed. The top 5 upregulated genes (by P.adjust value) are: ECHDC3, BCL9, FHL1, EPB41L2, TNS3, as shown in Table 4, which indicates their upregulation in the high expression group and downregulation in the low expression group. The top 5 downregulated genes are: LPO, UNCX, SIX3, MPO, NTNG2, as demonstrated in Table 5, which shows their downregulation in the high expression group and upregulation in the low ALDOC expression group.

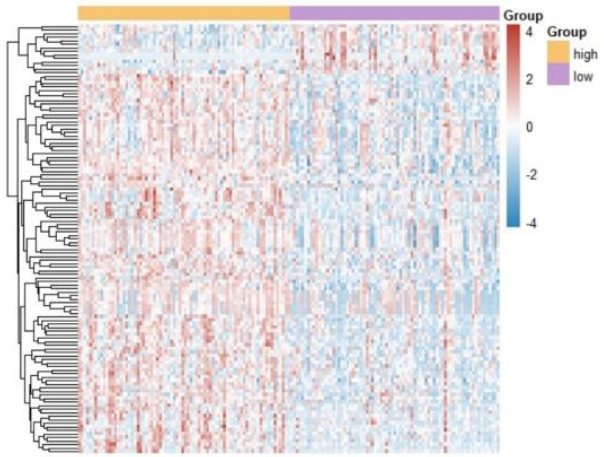


Figure 10. Heatmap of Differentially Expressed Genes Between High and Low ALDOC Expression Groups

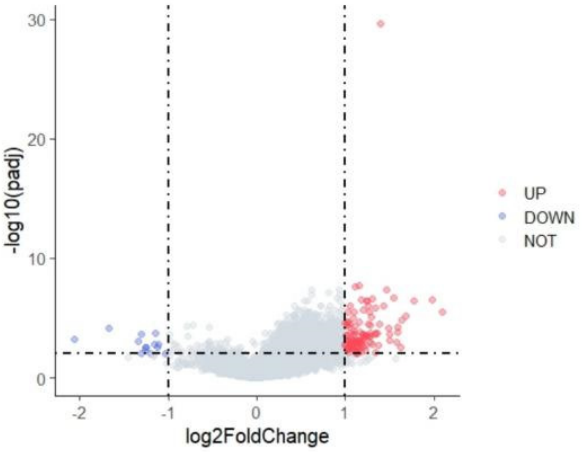


Figure 11. Volcano Plot of Differentially Expressed Genes Between High and Low ALDOC Expression Groups

Table 4. Top 5 Upregulated Genes When ALDOC is Highly Expressed (by P.adjust Value)

Gene	log_2FC	*P*.adjust
ECHDC3	1.161353	2.05E-08
BCL9	1.111203	2.35E-08
FHL1	1.461722	4.06E-08
EPB41L2	1.550612	2.00E-07
TNS3	1.311653	2.41E-07

Table 5. Top 5 Downregulated Genes When ALDOC is Highly Expressed (by P.adjust Value)

Gene	log2FC	*P*.adjust
LPO	-1.66915	7.03E-05
UNCX	-1.14146	0.000187
SIX3	-1.30797	0.000206
MPO	-2.06572	0.000574
NTNG2	-1.33617	0.000806

3.8. GO and KEGG Enrichment Analysis of Differentially Expressed Genes in High and Low ALDOC Expression Groups

We conducted a GO database enrichment analysis on the differentially expressed genes identified by comparing AML patients with normal individuals. A significant number of GO annotation items were obtained, with P < 0.05 considered

statistically significant. For the biological process (BP), cellular component (CC), and molecular function (MF) categories, we obtained 131, 47, and 22 items, respectively. The top 10 items from each category (by P.adjust value) were selected and visualized in a bar chart (Figure 12) with the number of genes on the x-axis and the gene function items on the y-axis. Figure 12 indicates that the top 10 enriched items in the BP category are: phagocytosis, response to interferon-γ, cell response to interferon-γ, and antigen processing and presentation, among others. The top 10 enriched items in the CC category are: endosomal vesicle, endosomal vesicle membrane, clathrin-coated vesicle membrane, and MHC class II protein complex, among others. The top 10 enriched items in the MF category (by P.adjust value) are: amide binding, peptide binding, growth factor binding, and immune receptor activity, among others.

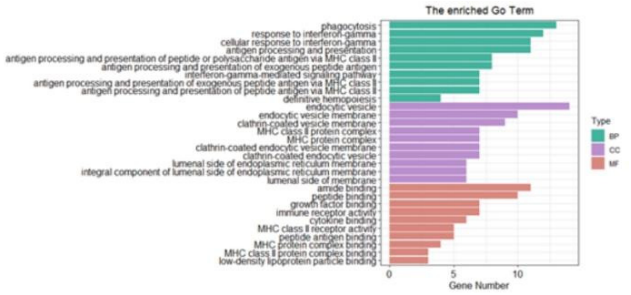


Figure 12. Top 10 GO Functional Analysis Results of Differentially Expressed Genes in AML Patients Compared to Normal Individuals

To further explore the main enrichment pathways of the genes identified in the GO functional analysis, we conducted a KEGG enrichment analysis and created a bubble chart. The top 10 pathways obtained from KEGG enrichment (by P.adjust value) were visualized in Figure 13, with the x-axis representing the number of enriched genes and the y-axis representing the significance of enrichment. The figure indicates that genes associated with ALDOC expression are primarily involved in pathways such as phagosomes, toxoplasmosis, hematopoietic cell lineage, and antigen processing and presentation.

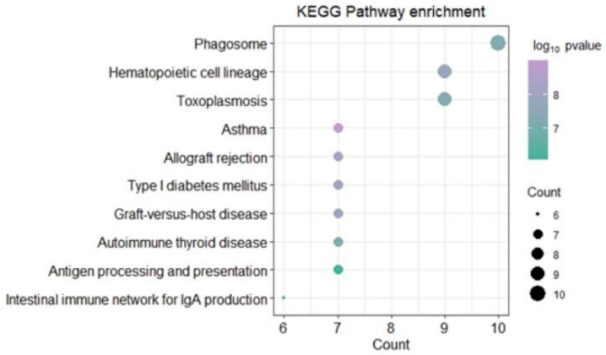


Figure 13. KEGG Enrichment Analysis for the ALDOC Gene

3.9. GSEA Enrichment Analysis of Differentially Expressed Genes Between High and Low ALDOC Expression Groups

To understand the relationship between other genes in AML patients with ALDOC and certain predefined gene sets, we conducted GSEA analysis for both the high and low ALDOC expression groups. The standard of P.adjust < 0.05 was set for statistical significance, and a total of 30 gene sets were obtained. The top 4 gene sets by P.adjust value are presented in Figure 14. As shown in the figure, all four relevant gene sets are enriched at the top of the ALDOC high expression group (represented by the red section), with all ES (Enrichment Score) values being positive. This indicates that the allogeneic rejection, IL2 STAT5 signaling, response to interferon-alpha, and response to interferon-gamma gene sets are upregulated in the ALDOC high expression group and downregulated in the low expression group.

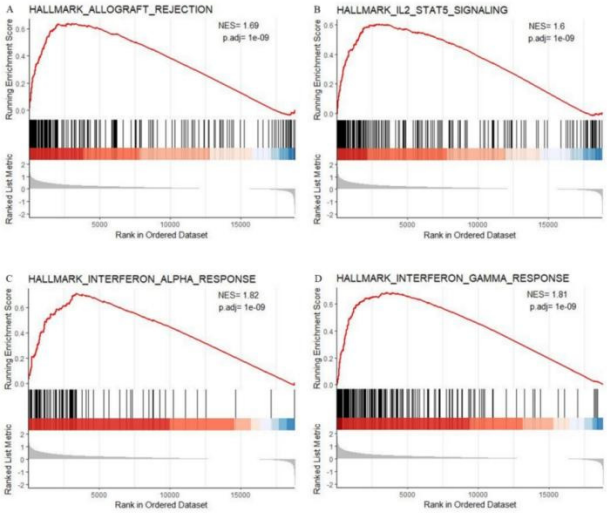


Figure 14. GSEA Enrichment Analysis of ALDOC High and Low Expression Groups

4. Discussion

Metabolic studies of cancer are among the most classic research areas in cancer biology, predating the discovery of oncogenes and tumor suppressor genes by about half a century[13]. In cancer cells, metabolic activity patterns are distinctly altered compared to normal cells, and these metabolic changes have the potential to promote the maintenance and exacerbation of their malignant characteristics, a concept known as metabolic reprogramming in cancer. In recent years, given the profound heterogeneity of cancer, many cancer metabolism studies have focused on the cell-autonomous effects of oncogenes in revealing the interplay between changes in gene expression and signal regulation and metabolic networks [14].

Acute myeloid leukemia (AML) is a disease characterized by the uncontrolled proliferation of myeloid precursor cells at various stages of maturation, with high malignancy and poor prognosis. The precise mechanisms by which metabolic reprogramming contributes to the initiation and progression of AML and alters prognosis are not yet fully understood, necessitating further in-depth research into the pathogenesis of AML. We combined healthy human peripheral blood samples from the GTEx database as a control group and analyzed changes in gene expression in known metabolic pathways. We rigorously selected differentially expressed genes based on strict criteria and identified six metabolic genes significantly associated with clinical prognosis through comprehensive clinical survival information. Further observation of the correlation between these six genes and basic clinical factors in AML patients (gender, age, FAB classification, prognostic stratification) revealed that, in particular, ALDOC, PMM1, and AK1, were correlated with all basic characteristics except gender. The prognostic significance of AK1 in AML has been reported in the literature[15], while PMM1 may not have a clear correlation

with clinical disease[16]. Given the important biological significance of ALDOC in promoting oncogenesis and leading to poor prognosis through glycolytic metabolism, and our preliminary results indicating the potential of ALDOC to predict the prognosis of AML patients, we chose ALDOC as the final target gene for subsequent analysis.

After preliminarily establishing the correlation between ALDOC and the prognosis of AML patients, another focus of this study was to analyze the prognostic role of ALDOC in subgroup analysis based on age. AML is a hematological malignancy highly associated with age. During the preliminary screening of target genes, we identified a significant correlation between ALDOC gene expression levels and age. To further explore the clinical value of ALDOC expression levels in AML, we grouped patients into elderly and non-elderly groups, incorporating more clinical and molecular characteristics. We then constructed univariate and multivariate Cox proportional hazards regression models for AML patients of different age groups. Finally, we used Kaplan-Meier survival curves to further assess the prognostic value of ALDOC in elderly and non-elderly patients. The conclusions were consistent with our Cox proportional hazards regression model results, indicating that in non-elderly AML patients, ALDOC expression level is a high-predictive-value indicator of prognosis, and high ALDOC expression leads to poorer prognosis.

To understand the potential biological significance of ALDOC in AML, we divided 151 AML patients from the TCGA database into high and low ALDOC expression groups based on the median ALDOC expression level. After screening, we identified a total of 120 differentially expressed genes between the two groups, including 106 upregulated genes and 14 downregulated genes. GO analysis revealed that these genes primarily function in biological processes such as phagocytosis, response to IFN-γ, antigen processing, and presentation via MHC Class II; in cellular components such as endosomes and MHC protein complexes; and in molecular functions such as binding to various cytokines. KEGG analysis revealed that the main enriched pathways include phagosomes, antigen processing and presentation, and allogeneic rejection reactions.

Considering that the remaining genes apart from the differentially expressed ones may also play a role in the disease, we used GSEA analysis, which showed that functional gene sets such as "allogeneic rejection, IL2 STAT5 signaling, interferon alpha response, interferon gamma response" were enriched at the top of the high expression group, showing an upward trend in expression. Therefore, we speculate that high ALDOC expression in AML is significantly associated with allogeneic rejection, IL2 STAT5 signaling, interferon alpha response, and interferon gamma response pathways, leading to relatively poorer prognosis results.

5. Conclusion

The combined TCGA and GTEx primary data analysis confirmed that the metabolic gene ALDOC is differentially expressed in AML and that high ALDOC expression levels can predict poor OS in AML patients, especially for non-elderly (under 60 years of age) AML patients, where ALDOC is a valuable prognostic molecular marker. ALDOC, involved in glycolytic pathways, is significantly associated with allogeneic rejection, response to IFN-alpha, response to IFN-gamma, and phagocytosis in AML, indicating that ALDOC is a promising potential therapeutic target for further research.

References

[1] KHWAJA A, BJORKHOLM M, GALE R E, et al. Acute myeloid leukaemia [J]. Nat Rev Dis Primers, 2016, 2(16010.

[2] O'DONNELL M R, ABBOUD C N, ALTMAN J, et al. NCCN Clinical Practice Guidelines Acute myeloid leukemia [J]. J Natl Compr Canc Netw, 2012, 10(8): 984-1021.

[3] CANAANI J, BEOHOU E, LABOPIN M, et al. Impact of FAB classification on predicting outcome in acute myeloid leukemia, not otherwise specified, patients undergoing allogeneic stem cell transplantation in CR1: An analysis of 1690 patients from the acute leukemia working party of EBMT [J]. American journal of hematology, 2017, 92(4): 344-50.

[4] YAMAMOTO J F, GOODMAN M T. Patterns of leukemia incidence in the United States by subtype and demographic characteristics, 1997-2002 [J]. Cancer causes & control : CCC, 2008, 19(4): 379-90.

[5] SHANG S Q, YANG Y, YANG X J, et al. Endothelial lipase promotes acute myeloid leukemia progression through metabolic reprogramming [J]. Neoplasma, 2022.

[6] DöHNER H, ESTEY E, GRIMWADE D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel [J]. Blood, 2017, 129(4): 424-47.

[7] NG S W, MITCHELL A, KENNEDY J A, et al. A 17-gene stemness score for rapid determination of risk in acute leukaemia [J]. Nature, 2016, 540(7633): 433-7.

[8] FUJITA H, AOKI H, AJIOKA I, et al. Detailed expression pattern of aldolase C (Aldoc) in the cerebellum, retina and other areas of the CNS studied in Aldoc-Venus knock-in mice [J]. PLoS One, 2014, 9(1): e86679.

[9] FAN K, WANG J, SUN W, et al. MUC16 C-terminal binding with ALDOC disrupts the ability of ALDOC to sense glucose and promotes gallbladder carcinoma growth [J]. Exp Cell Res, 2020, 394(1): 112118.

[10] MARUYAMA R, NAGAOKA Y, ISHIKAWA A, et al. Overexpression of aldolase, fructose-bisphosphate C and its association with spheroid formation in colorectal cancer [J]. Pathol Int, 2022, 72(3): 176-86.

[11] YUAN J, XING H, LI Y, et al. EPB41 suppresses the Wnt/beta-catenin signaling in non-small cell lung cancer by sponging ALDOC [J]. Cancer letters, 2021, 499(255-64.

[12] REINSBOROUGH C W, IPAS H, ABELL N S, et al. BCDIN3D RNA methyltransferase stimulates Aldolase C expression and glycolysis through let-7 microRNA in breast cancer cells [J]. Oncogene, 2021, 40(13): 2395-406.

[13] DEBERARDINIS R J, CHANDEL N S. Fundamentals of cancer metabolism [J]. Sci Adv, 2016, 2(5): e1600200.

[14] VANDER HEIDEN M G, DEBERARDINIS R J. Understanding the Intersections between Metabolism and Cancer Biology [J]. Cell, 2017, 168(4): 657-69.

[15] QIN T, ZHAO H, SHAO Y, et al. High expression of AK1 predicts inferior prognosis in acute myeloid leukemia patients undergoing chemotherapy [J]. Biosci Rep, 2020, 40(6).

[16] CROMPHOUT K, KELDERMANS L, SNELLINX A, et al. Tissue distribution of the murine phosphomannomutases Pmm1 and Pmm2 during brain development [J]. Eur J Neurosci, 2005, 22(4): 991-6.