

Original Research

# Expression Pattern and Prognostic Analysis of Branched-Chain Amino Acid Catabolism-Related Genes in Non-Small Cell Lung Cancer

Xiaojun Yao<sup>1,\*</sup>, Yulan Deng<sup>2</sup>, Jian Zhou<sup>2</sup>, Liangshuang Jiang<sup>1</sup>, Yijie Song<sup>1</sup>

<sup>1</sup>Department of Thoracic Surgery, The Public Health Clinical Center of Chengdu, 610061 Chengdu, Sichuan, China

<sup>2</sup>Department of Thoracic Surgery, West China Hospital, Sichuan University, 610041 Chengdu, Sichuan, China

\*Correspondence: [xiaojunyao\\_work@163.com](mailto:xiaojunyao_work@163.com) (Xiaojun Yao)

Academic Editor: Taeg Kyu Kwon

Submitted: 21 October 2022 Revised: 6 February 2023 Accepted: 10 February 2023 Published: 8 June 2023

## Abstract

**Background:** The purpose of our study is to analyze the expression pattern and prognostic value of catabolism-related enzymes of branched-chain amino acids (BCAAs) in non-small cell lung cancer (NSCLC). **Methods:** Differential expression analysis, mutation, copy number variation (CNV), methylation analysis, and survival analysis of BCAAs catabolism-related enzymes in NSCLC were performed using the Cancer Genome Atlas (TCGA) database. **Results:** Six and seven differentially expressed genes were obtained in lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), respectively. IL4I1 was located at the core regulatory nodes in the gene co-expression networks of both LUAD and LUSC. The AOX1 mutation rate was the highest in both LUAD and LUSC. For CNV, IL4I1 was up-regulated in both LUAD and LUSC with an increase in copy number, whereas AOX1 and ALDH2 were differentially regulated in the two subtypes of lung cancer. In patients with NSCLC, high expression of IL4I1 was associated with lower overall survival (OS), and low expression of ALDH2 predicted shorter disease-free survival (DFS). ALDH2 expression was related with LUSC survival. **Conclusions:** This study explored the biomarkers of BCAAs catabolism related to the prognosis of NSCLC, which provided a theoretical foundation to guide the clinical diagnosis and treatment of NSCLC.

**Keywords:** catabolism-related enzyme of BCAAs; NSCLC; expression pattern; survival analysis; TCGA

## 1. Introduction

Lung cancer is one of the malignant tumors with the highest morbidity and mortality worldwide and has become a major public health concern [1,2]. Despite considerable progress in therapeutic strategies, the 5-year survival rate of lung cancer in China has remained between 10% and 20% over the past decade [3]. Therefore, the key to improving the survival rate for lung cancer is not only improving the treatment but also improving the level of screening and using more abundant analytical methods to find biomarkers that are closely related to the development and prognosis of lung cancer.

Branched-chain amino acids (BCAAs) include leucine, isoleucine, and valine. Plasma levels of BCAAs and their metabolic enzymes are expressed to varying degrees in multiple cancers and have a very close relationship with tumor occurrence and development. They are considered important markers for early tumor screening and prognosis, and provide a very meaningful research prospect for the development of novel therapeutic drugs in the direction of targeted treatment of amino acid metabolism enzymes [4–10]. However, there have been no systematic studies on the expression pattern of the BCAAs catabolic enzyme in non-small cell lung cancer (NSCLC) and its correlation with prognosis. Therefore, it is crucial to screen for key BCAAs catabolic enzymes to identify new biomarkers for the prognosis of NSCLC.

In this study, sets of catabolic enzyme genes related to BCAAs were established using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment. Transcriptome and clinical data of lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) were obtained from the Cancer Genome Atlas (TCGA) database. Based on the multidimensional bioinformatic analysis, the expression pattern of the BCAAs catabolic enzyme in NSCLC and its correlation with prognosis were explored to identify novel biomarkers for the prognosis of NSCLC and to provide references for the future active exploration and development of new therapeutic targets for NSCLC.

## 2. Materials and Methods

### 2.1 Data Collection

The KEGG database (<https://www.kegg.jp/kegg/pathway.html>) was used to search the catabolism pathways of human BCAAs. The valine, leucine and isoleucine degradation pathway is displayed in **Supplementary Fig. 1**. In total, 44 gene sets of related metabolic enzymes were identified as the main study objects (**Supplementary Fig. 2**). The transcriptome profiles and corresponding clinical information of LUAD, LUSC, and adjacent normal tissues were downloaded from the TCGA (<http://tcga-data.nci.nih.gov/>) dataset using the RTCGAToolbox 2.28.0 package in R 4.2 software.



## 2.2 Expression Pattern of BCAAs Catabolism-Related Enzymes in NSCLC

The data were transformed from fragments per kilobase of exon per million reads mapped (FPKM) value, and  $\log_2$  of its FPKM value was used as the measure of the gene expression level. Differentially expressed genes in NSCLC tissues compared to those in adjacent normal tissues were calculated using the edgeR package. The Benjamini and Hochberg multiple testing methods were applied to determine the false discovery rate (FDR).  $FDR < 0.05$  and  $|\log_2 \text{fold change}| > 1$  were selected as the cutoff criteria. The heatmap.2 function of the R gplots 3.1.3 package was utilized to generate a hierarchical cluster analysis of differentially expressed genes. The Pearson correlation coefficient was used to obtain gene co-expression network pairs in LUAD and LUSC versus normal controls. The Cytoscape software was applied to visualize the gene co-expression network.

## 2.3 Mutation, Copy Number Variation (CNV) and Methylation Analysis of BCAAs Catabolism-Related Enzymes

The mutation data were processed and visualized using the maftools 2.14.0 R package (<https://github.com/PoissonAlien/maftools>). For CNV, the loss and gain of copy numbers have been identified using the Genomic Identification of Significant Targets in Cancer (GISTIC) algorithm. The 5-valued spectrum (-2, -1, 0, 1, 2) was used to indicate changes in CNV. -2, -1, 0, 1, and 2 represent homozygous deletion of copy number, heterozygous deletion of copy number, no variation in copy number, amplification of low-dose copy numbers and amplification of high-dose copy numbers, respectively. Due to the noise of low-dose amplification or deletion, we mainly referred to copy number changes of 2 and -2 in the analysis, taking 2 as amplification and -2 as deletion, while others considered copy number unchanged. For methylation analysis, ChAMP 3.8 in the R package was used to filter the data, fill in the missing values, and calculate the differential methylation probe and the differential methylation region. Differential methylation sites between lung cancer and normal lung tissues were obtained using limma 3.38.2 in the R package. The Benjamini and Hochberg multiple testing methods were applied to acquire the FDR. The FDR of  $< 0.05$  was considered statistically significant. Hierarchical cluster analysis and waterfall plots of differential methylation sites were visualized through gplots in the R package.

## 2.4 Survival Analysis of BCAA Catabolism-Related Enzymes

SPSS 22.0 statistical software (IBM Corp., Armonk, NY, USA) was used for the survival analysis. The Kaplan-Meier curve and the logarithmic rank test were used to evaluate high and low gene expression in overall survival (OS) and disease-free survival (DFS). The Cox proportional haz-

ard regression model was used to perform univariate and multivariate analysis of independent risk factors related to postoperative OS in lung cancer patients to calculate the risk ratio (HR) and 95% confidence interval (CI). Statistical results with  $FDR < 0.05$  were considered significant.

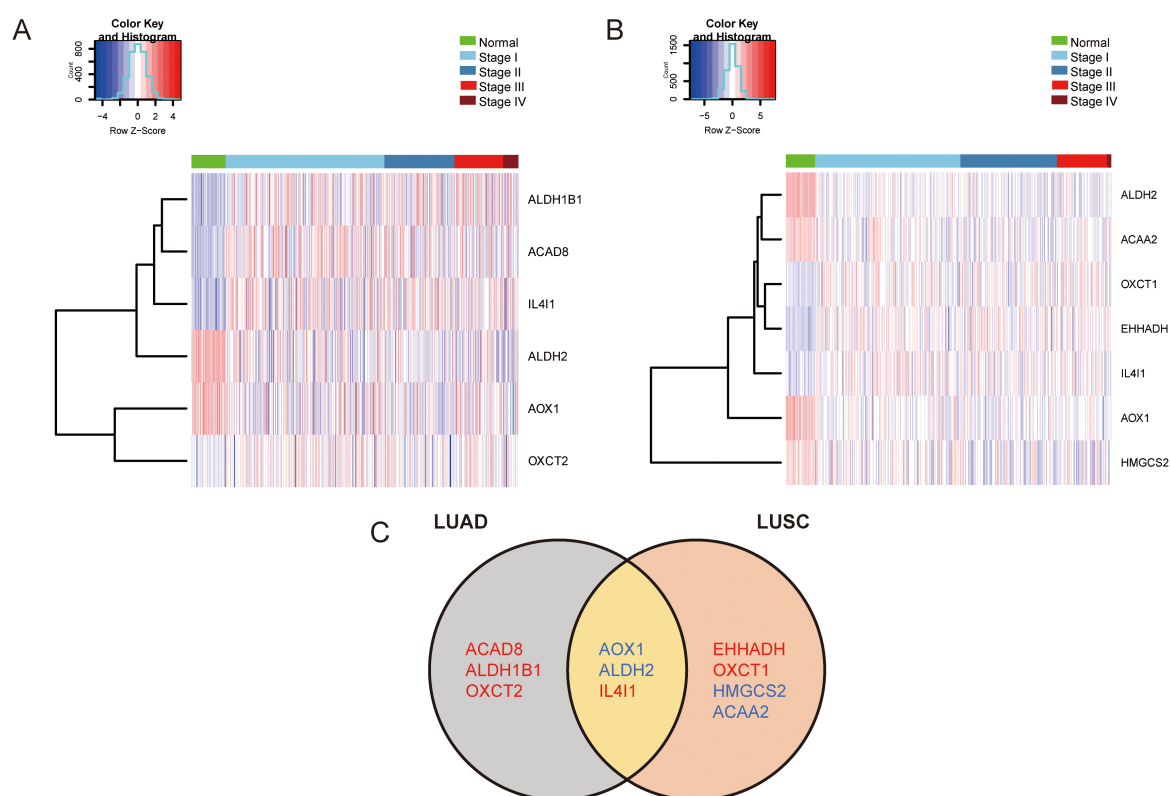
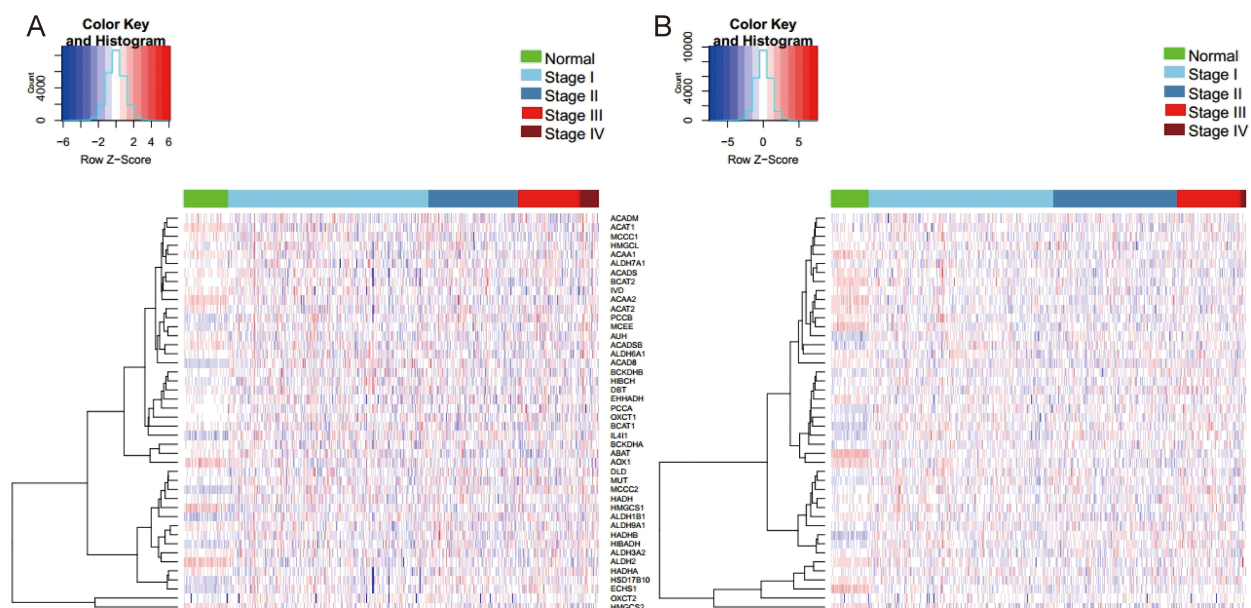
## 3. Results

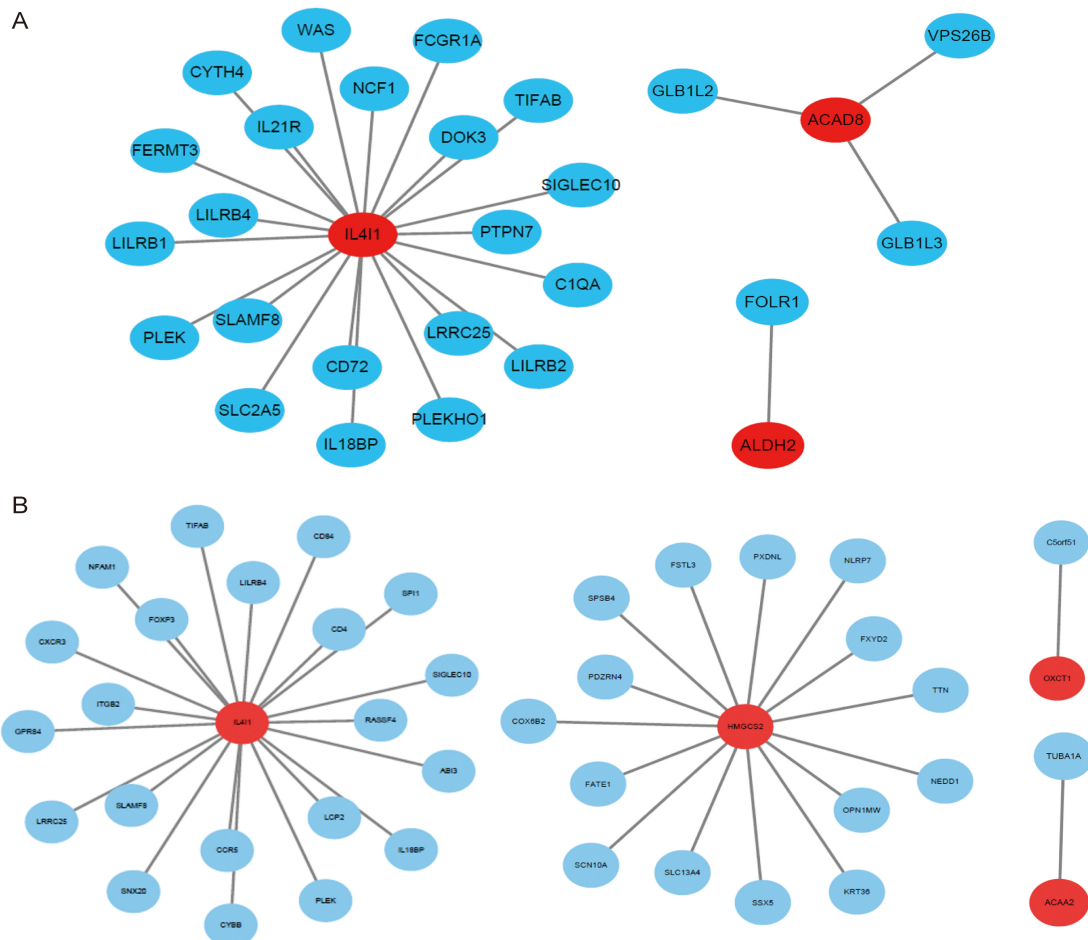
### 3.1 Expression Pattern of BCAAs Catabolism-Related Enzymes in NSCLC

Gene expression profiles and corresponding clinical data for NSCLC were obtained from the TCGA database. In this study, 505 LUAD tissues and 59 normal adjacent samples from patients with LUAD, and 501 LUSC tissues and 51 normal adjacent samples from patients with LUSC were included. We analyzed the expression of 44 BCAAs catabolic enzymes between NSCLC and normal lung tissues. Hierarchical clustering analysis of 44 BCAAs catabolism-related enzymes in LUAD and LUSC is shown in Fig. 1A,B, respectively. As shown in Fig. 2A,B, compared to normal lung tissues, there were six differentially expressed genes (*ALDH1B1*, *ACAD8*, *IL4I1*, *OXCT2*, *ALDH2*, and *AOX1*) and seven differentially expressed genes (*OXCT1*, *EHHADH*, *IL4I1*, *ALDH2*, *ACAA2*, *AOX1* and *HMGCS2*) in LUAD and LUSC, respectively, with statistical significance ( $FDR < 0.05$  and  $|\log_2 \text{fold change}| > 1$ ). *ALDH1B1*, *ACAD8*, *IL4I1*, and *OXCT2* were up-regulated, while *ALDH2* and *AOX1* were down-regulated in LUAD. *OXCT1*, *EHHADH*, and *IL4I1* were up-regulated, whereas *ALDH2*, *ACAA2*, *AOX1*, and *HMGCS2* were down-regulated in LUSC. Among these, *IL4I1*, *ALDH2*, and *AOX1* were differentially expressed in both LUAD and LUSC (Fig. 2C). We also built the gene co-expression network for the differentially expressed genes. Three genes (*IL4I1*, *ACAD8*, and *ALDH2*) and four genes (*IL4I1*, *HMGCS2*, *ACAA2*, and *OXCT1*) were in the core regulatory nodes of the gene co-expression network in LUAD and LUSC, respectively (Fig. 3).

### 3.2 Mutation, CNV, and Methylation Analysis of BCAAs Catabolism-Related Enzymes

The mutation information for each gene in each sample is displayed in a waterfall plot, where various colors with annotations at the bottom represent the different types of mutations (Fig. 4). Somatic mutations were found in 66 (28.7%) of 230 LUAD samples, and somatic mutations occurred in 36 (81.8%) of 44 enzymes related to the catabolism of BCAAs in 66 patients with somatic mutations (Fig. 4A). Among these, the *AOX1* mutation rate was the highest (3%, 7/230). As showed in Fig. 4B, somatic mutations were found in 59 (33.15%) of 178 LUSC samples, and somatic mutations occurred in 37 (84.1%) of the 44 BCAAs catabolism-related enzymes in 59 patients with somatic mutations. Among these, the *AOX1* mutation rate was the highest (6.7%, 12/178).





**Fig. 3. Gene co-expression network of significantly differentially expressed BCAAs catabolism-related enzymes in NSCLC. (A) LUAD. (B) LUSC. Ellipses were used to represent nodes, and lines were used to represent edges.**

We then evaluated the CNV patterns of the differentially expressed genes in LUAD and LUSC. We found that *OXCT2*, *AOX1*, *ACAD8*, *ALDH2*, and *IL4I1* were mainly copy number amplifications in LUAD, whereas *ALDH1B1* was mainly a copy number deletion (Table 1). In LUSC, *EHHADH*, *OXCT1*, *AOX1*, *IL4I1*, and *HMGCS2* showed copy number amplification, whereas *ACAA2* and *ALDH2* displayed copy number deletions (Table 2). According to the correlation analysis of CNV and gene expression levels, *IL4I1*, *ACAD8*, and *OXCT2* were up-regulated in LUAD with an increase in copy number, whereas *ALDH2*, *ALDH1B1*, and *AOX1* were oppositely regulated (Fig. 5A). *IL4I1*, *OXCT1*, and *EHHADH* were up-regulated in LUSC with increasing copy number; *ACAA2* and *ALDH2* were down-regulated with deletion of copy number; and *HMGCS2* and *AOX1* were down-regulated with increased copy number (Fig. 5B).

Methylation analysis of BCAAs catabolism-related enzymes was performed in LUAD and LUSC. Hierarchical cluster analysis of the methylation pattern in LUAD and LUSC is shown in Fig. 6A,B, respectively. Among them, *IL4I1* and *ALDH2* were hypomethylated in LUAD (FDR

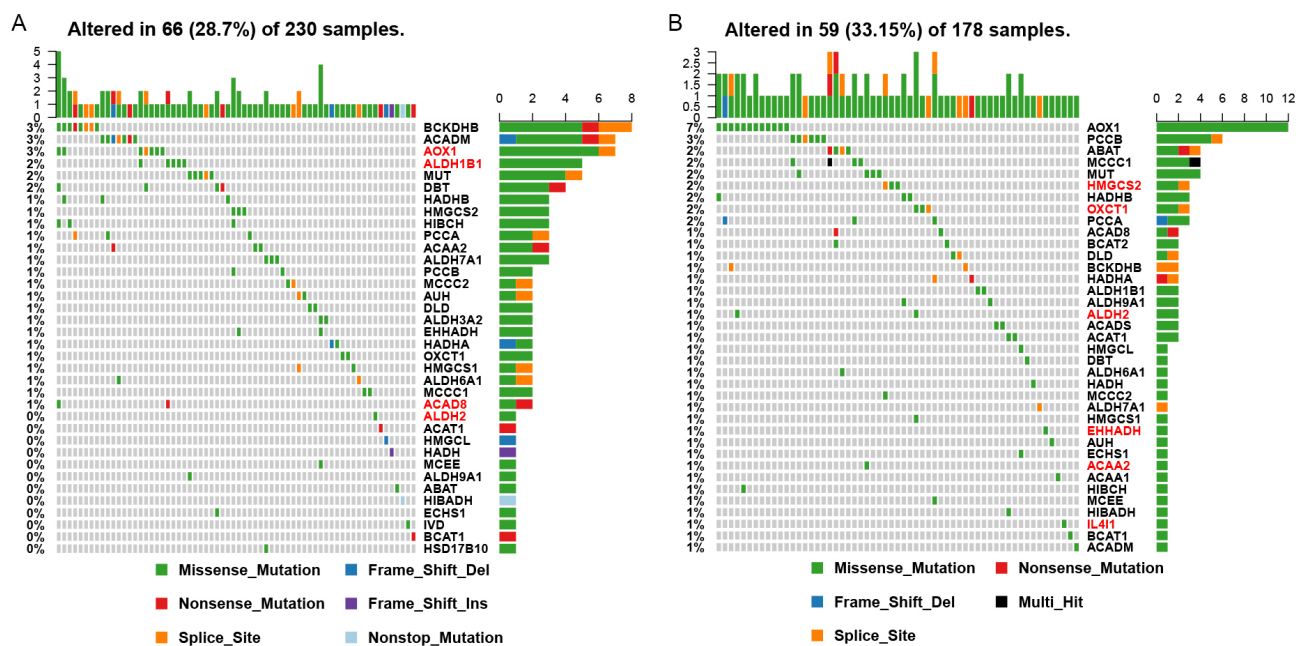
**Table 1. CNV of differentially expressed BCAAs catabolism enzymes in LUAD.**

Gene	Copy number (n = 494)					Variability rate
	-2	-1	0	1	2	
<i>OXCT2</i>	0	88	278	118	10*	4.94%*
<i>ACAD8</i>	7	103	272	104	8*	1.6%*
<i>AOX1</i>	0	39	324	127	4*	0.8%*
<i>ALDH2</i>	2	111	264	114	3*	0.6%*
<i>IL4I1</i>	1	176	245	70	2*	0.4%*
<i>ALDH1B1</i>	7#	217	210	58	2	1.4%#

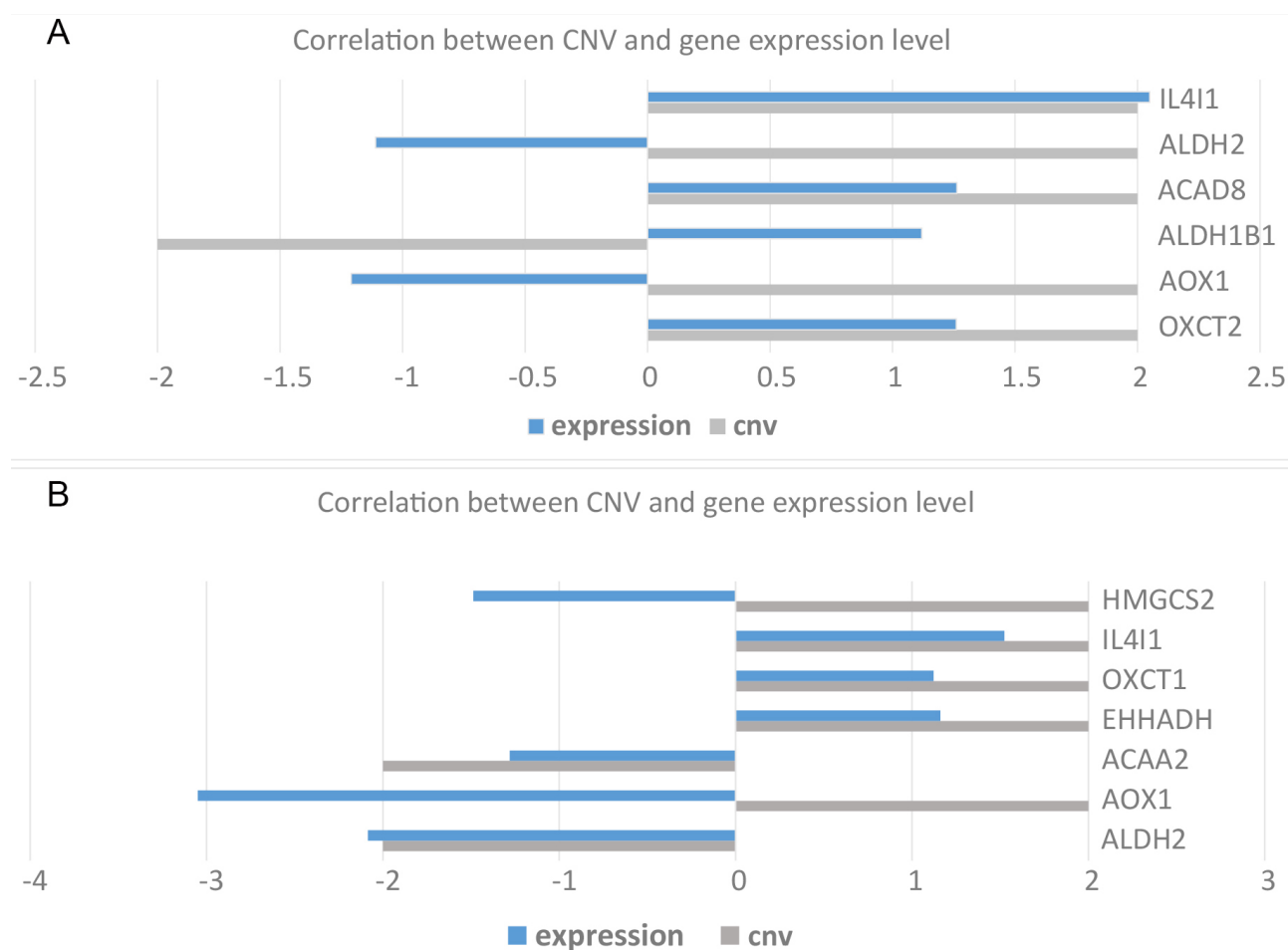
\* represents copy number amplification; # indicates missing copy number.

<0.05; Table 3), and *ALDH2*, *ACAA2*, *OXCT1*, *EHHADH*, and *IL4I1* were hypomethylated and *AOX1* was hypermethylated in LUSC (FDR <0.05; Table 4). According to a correlation analysis of methylation patterns and gene expression levels, *IL4I1* showed hypermethylation corresponding to its up-regulation in LUAD (Fig. 7A), and *OXCT1*, *EHHADH*, and *IL4I1* showed hypermethylation corresponding to their up-regulation in LUSC (Fig. 7B).

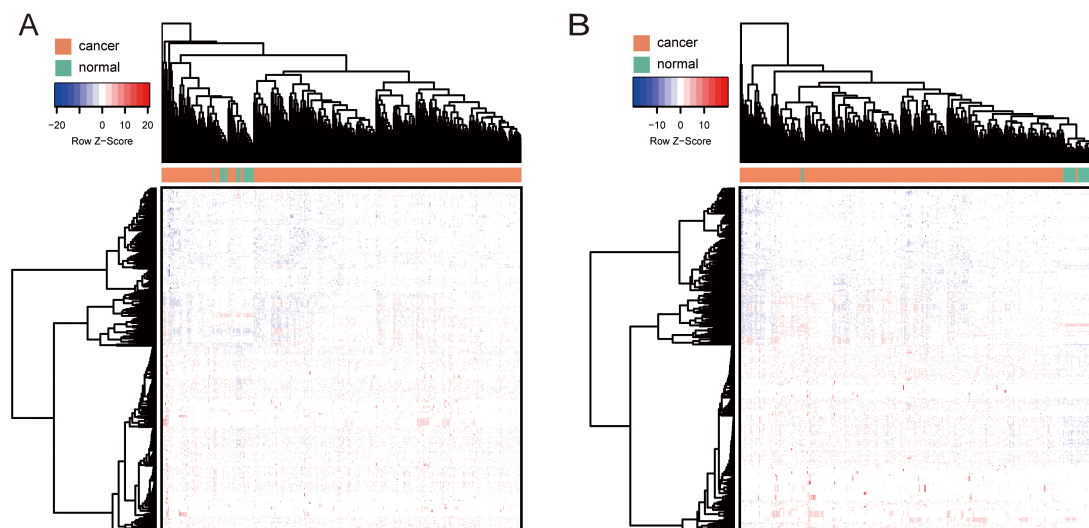




**Fig. 4. Mutation of significantly differentially expressed BCAAs catabolism-related enzymes in NSCLC. (A) LUAD. (B) LUSC.** The red font shows significantly differentially expressed BCAAs catabolism-related enzymes.



**Fig. 5. Correlation diagram of CNV and expression levels in NSCLC with significant difference in BCAAs catabolism. (A) LUAD. (B) LUSC.**



**Fig. 6. Hierarchical cluster analysis of the methylation pattern in NSCLC. (A) LUAD. (B) LUSC.** The row and column represented the methylation level and samples, respectively. The color scale indicated the level of methylation.

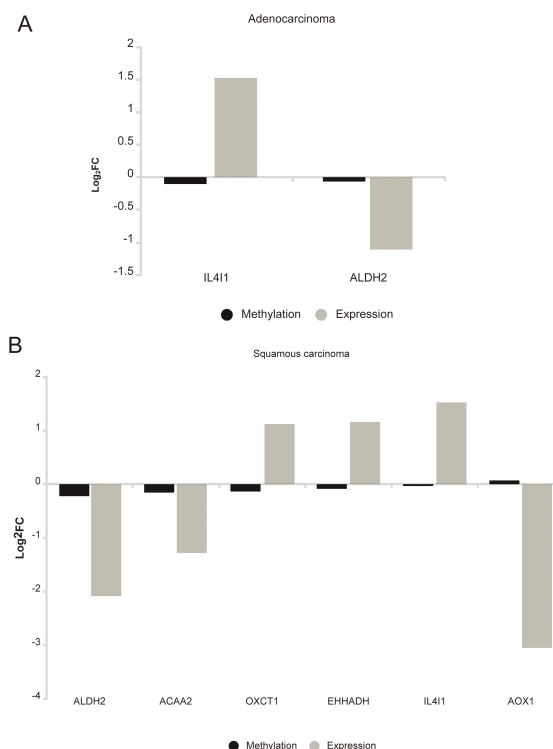
**Table 2. CNV of differentially expressed BCAAs catabolism enzymes in LUSC.**

Gene	Copy number (n = 490)					Variability rate
	-2	-1	0	1	2	
<i>EHHADH</i>	0	4	48	233	205*	41.8%
<i>OXCT1</i>	1	26	110	303	50*	10.2%
<i>AOX1</i>	1	74	278	123	14*	2.9%
<i>HMGCS2</i>	13	165	215	83	14*	2.9%
<i>IL4I1</i>	0	132	197	158	3*	0.6%
<i>ACAA2</i>	10 <sup>#</sup>	193	197	85	5	2% <sup>#</sup>
<i>ALDH2</i>	2 <sup>#</sup>	77	291	119	1	0.4% <sup>#</sup>

\* represents copy number amplification; # indicates missing copy number.

### 3.3 Survival Analysis of BCAAs Catabolism-Related Enzymes

To determine the prognostic value of enzymes related to BCAAs catabolism, we evaluated the effects of differentially expressed genes on the OS and DFS of patients with LUAD and LUSC. In LUAD patients, the expression of five of the six differentially expressed genes, including *ALDH1B1* ( $p = 0.958$ ), *ALDH2* ( $p = 0.077$ ), *OXCT2* ( $p = 0.617$ ), *IL4I1* ( $p = 0.492$ ), and *AOX1* ( $p = 0.288$ ), was not significantly correlated with DFS. The DFS of patients with LUAD in the high expression group of *ACAD8* was significantly longer than that of patients with low *ACAD8* expression ( $p < 0.001$ , Fig. 8A). Furthermore, as shown in Fig. 8B–E, high expression of *ALDH1B1* ( $p = 0.029$ ) and low expression of *ACAD8* ( $p < 0.001$ ), *ALDH2* ( $p = 0.011$ ) and *OXCT2* ( $p = 0.017$ ) were associated with poor OS in patients with LUAD. However, *IL4I1* ( $p = 0.149$ ) and *AOX1* ( $p = 0.378$ ) did not correlate with OS in patients with LUAD.



**Fig. 7. Correlation diagram of the methylation level and expression level of BCAAs catabolism enzymes in NSCLC. (A) LUAD. (B) LUSC.**

In patients with LUSC, the expression of *ALDH2* ( $p = 0.837$ ), *EHHADH* ( $p = 0.359$ ), *AOX1* ( $p = 0.059$ ), *ACAA2* ( $p = 0.183$ ) and *OXCT1* ( $p = 0.779$ ) was not significantly associated with DFS, whereas high expression of *IL4I1* ( $p = 0.012$ ) and *HMGCS2* ( $p = 0.010$ ) was associated with poor DFS (Fig. 9A–B). Furthermore, as shown in Fig. 9C–E, in addition to the expression of *ACAA2* ( $p = 0.805$ ), *OXCT1*

**Table 3. Methylation analysis of differentially expressed BCAAs catabolism-related enzymes in LUAD.**

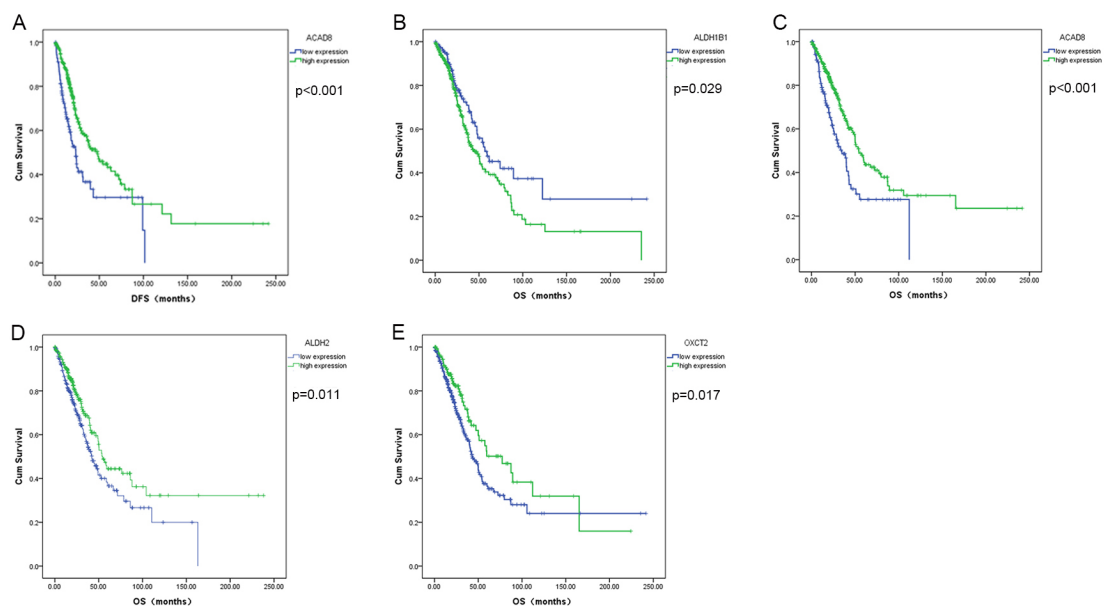
Gene	Probe	Log <sub>2</sub> FC	<i>p</i> value	FDR	Cancer AVG	Normal AVG	Delta beta
<i>IL4I1</i> *	cg10805880*	−0.10435*	$6.07 \times 10^{-6}$	$3.03 \times 10^{-5}$	0.2921	0.39645	−0.10435*
<i>IL4I1</i>	cg0638809	−0.06198	$2.46 \times 10^{-5}$	0.00011	0.17722	0.2392	−0.06198
<i>ALDH2</i>	cg10887937	−0.06601	0.00104	0.00313	0.77884	0.84485	0.06601

Log<sub>2</sub>FC indicates the differential expression multiple after transformation. Cancer AVG represents the methylation value of the probe in cancer samples; Normal AVG indicates the methylation value of the probe in normal samples; Delta beta represents the absolute differential methylation value. \* represents the maximum Beta value of absolute value of multiple probes corresponding to the same gene and the corresponding probe and Log<sub>2</sub>FC value.

**Table 4. Methylation analysis of differentially expressed BCAAs catabolism-related enzymes in LUSC.**

Gene	Probe	log <sub>2</sub> FC	<i>p</i> value	FDR	Cancer AVG	Normal AVG	Delta beta
<i>ALDH2</i> *	cg10887937*	−0.22205*	$1.76 \times 10^{-14}$	$1.57 \times 10^{-13}$	0.66984	0.89189	−0.22205*
<i>ALDH2</i>	cg22158248	−0.14350	$1.44 \times 10^{-7}$	$4.86 \times 10^{-7}$	0.47053	0.61403	−0.1435
<i>ACAA2</i> *	cg03345145*	−0.15567*	$1.45 \times 10^{-9}$	$6.65 \times 10^{-9}$	0.75873	0.9144	−0.15567*
<i>ACAA2</i>	cg23825830	−0.02635	0.00855	0.013038	0.078	0.10434	−0.02634
<i>OXCT1</i> *	cg06537708*	−0.13572*	$2.74 \times 10^{-10}$	$1.40 \times 10^{-9}$	0.73252	0.86823	−0.13571*
<i>OXCT1</i>	cg17317280	−0.06634	0.00255	0.004269	0.83218	0.89853	−0.06634
<i>EHHADH</i> *	cg11935638*	−0.08541*	$4.28 \times 10^{-8}$	$1.57 \times 10^{-7}$	0.20847	0.29388	−0.08541*
<i>EHHADH</i>	cg08095700	−0.05792	$4.54 \times 10^{-5}$	0.000103	0.1981	0.25602	−0.05792
<i>IL4I1</i>	cg06388099	−0.03433	0.018178	0.026101	0.14216	0.17649	−0.03433
<i>AOX1</i> *	cg13000082*	0.06938*	0.004132	0.00667	0.13969	0.07031	0.06936*
<i>AOX1</i>	cg04380340	0.06444	0.003021	0.004993	0.12397	0.05953	0.06444

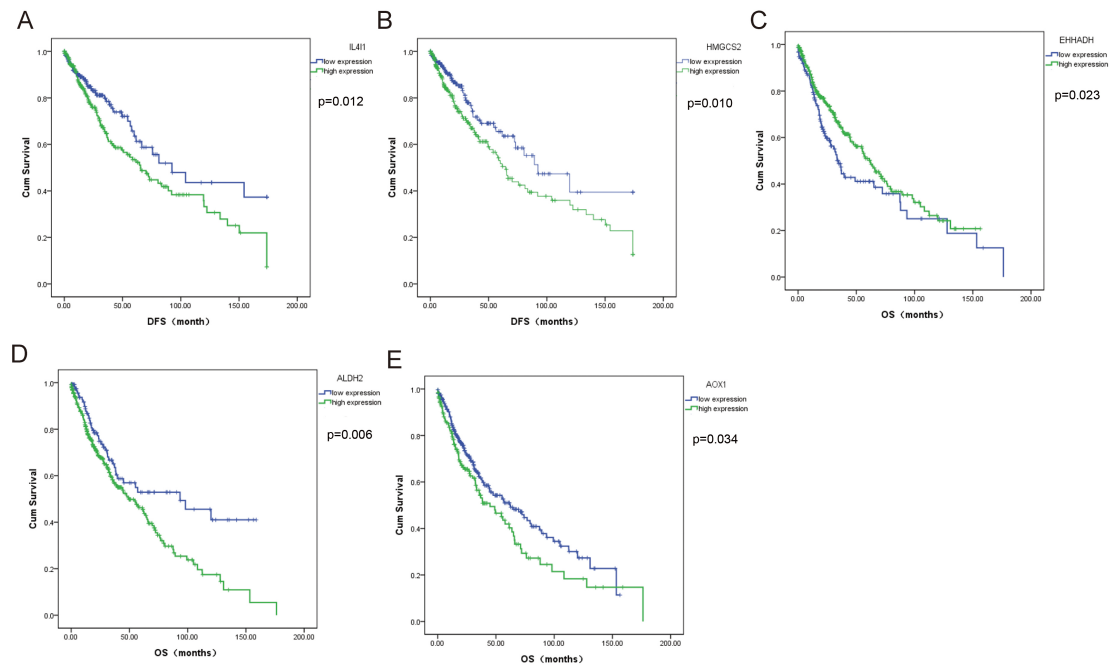
Log<sub>2</sub>FC indicates the differential expression multiple after transformation. Cancer AVG represents the methylation value of the probe in cancer samples; Normal AVG indicates the methylation value of the probe in normal samples; Delta beta represents the absolute differential methylation value. \* represents the maximum Beta value of absolute value of multiple probes corresponding to the same gene and the corresponding probe and Log<sub>2</sub>FC value.



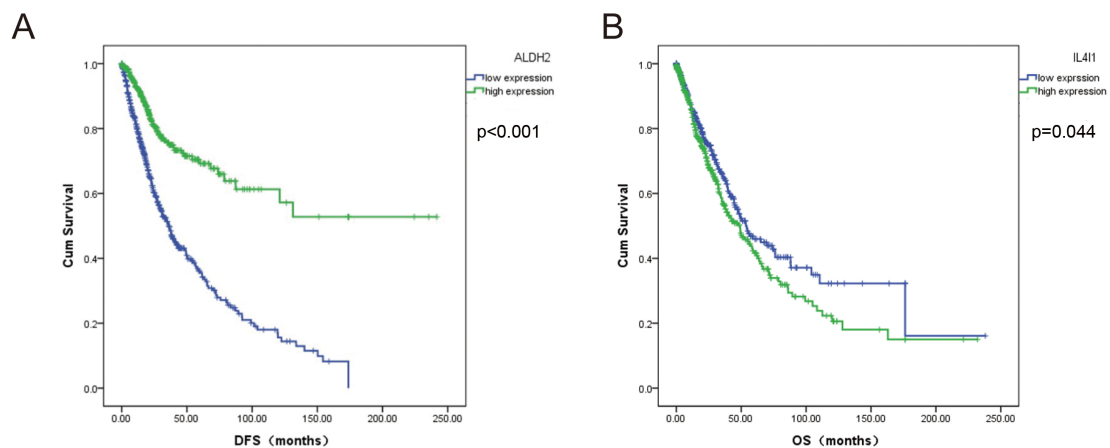
**Fig. 8. Survival analysis of significantly differentially expressed BCAA catabolism-related enzymes in LUAD. (A) *ACAD8*, (B) *ALDH1B1*, (C) *ACAD8*, (D) *ALDH2*, (E) *OXCT2*.**

( $p = 0.258$ ), *IL4I1* ( $p = 0.130$ ), *HMGCS2* ( $p = 0.190$ ), low expression of *EHHADH* ( $p = 0.023$ ), high expression of *ALDH2* ( $p = 0.006$ ) and *AOX1* ( $p = 0.034$ ) in patients with LUSC were significantly associated with worse OS (Fig. 9C–E).

In addition, we integrated the survival data of LUAD and LUSC and analyzed the effect of the expression of the catabolic enzyme gene of BCAAs, which were differentially expressed jointly in both LUAD and LUSC, on the OS and DFS of patients with NSCLC. The DFS of patients with



**Fig. 9.** Survival analysis of significantly differentially expressed BCAAs catabolism-related enzymes in LUSC. (A) *IL4I1*, (B) *HMGS2*, (C) *EHHADH*, (D) *ALDH2*, (E) *AOX1*.



**Fig. 10.** Survival analysis of significantly differentially expressed BCAAs catabolism-related enzymes in NSCLC. (A) *ALDH2*, (B) *IL4I1*.

NSCLC and low *ALDH2* expression was relatively poor ( $p < 0.001$ , Fig. 10A). However, *IL4I1* ( $p = 0.059$ ) and *AOX1* ( $p = 0.322$ ) were not significantly associated with DFS in NSCLC patients. *ALDH2* ( $p = 0.077$ ) and *AOX1* ( $p = 0.228$ ) were not significantly correlated with OS, whereas the high expression group of *IL4I1* had a worse prognosis than the low expression group (Fig. 10B).

Finally, Cox regression analysis was performed for genes with significant effects on the OS of LUAD and LUSC. Univariate Cox regression analysis showed that *ACAD8* and *OXCT2* expression, lymph node metastasis, and Tumor Node Metastasis (TNM) stage were predictors of poor prognosis in patients with LUAD (Table 5). In multivariate Cox regression analysis, *ACAD8* expression,

lymph node metastasis, and TNM stage were independent predictors of prognosis in patients with LUAD (Table 5). The expression of *ALDH2* and the stage were related to LUSC survival in both the univariate and multivariate Cox regression analysis (Table 6).

#### 4. Discussion

The occurrence and development of tumors are complex processes. Amino acid catabolic enzymes are over-expressed in a variety of cancers, providing not only cellular energy and metabolites for the anabolic process but also serving as a mechanism for cancer cells to escape immunity [6,11]. Multiple studies have indicated that branched-chain aminotransferase (BCAT), an enzyme that catalyzes



**Table 5. Univariate and multivariate OS analysis in patients with LUAD using the Cox proportional hazard regression model (n = 500).**

Variables	Univariate analysis			Multivariate analysis		
	Hazard ratio	95% CI	p value	Hazard ratio	95% CI	p value
<i>ALDH1B1</i> expression High vs Low	1.349	0.978–1.861	0.068			
<i>ACAD8</i> expression High vs Low	0.601	0.442–0.818	0.001*	0.553	0.405–0.756	<0.001*
<i>ALDH</i> expression High vs Low	0.839	0.625–1.127	0.244			
<i>OXCT2</i> expression High vs Low	0.677	0.482–0.950	0.024*	0.722	0.511–1.021	0.066
Gender Male vs female	1.090	0.812–1.462	0.567			
Age (year) >60 vs ≤60	1.154	0.838–1.590	0.380			
Lymph node metastasis Positive vs Negative	2.437	1.815–3.272	<0.001*	1.831	1.287–2.605	0.001*
TNM stage III–IV vs I–II	2.518	1.841–3.443	<0.001*	1.884	1.298–2.735	0.001*
Smoking History Yes vs No	0.822	0.602–1.123	0.218			

\* represents  $p < 0.05$ .

**Table 6. Univariate and multivariate OS analysis in patients with LUSC using the Cox proportional hazard regression model (n = 494).**

Variables	Univariate analysis			Multivariate analysis		
	Hazard ratio	95% CI	p value	Hazard ratio	95% CI	p value
<i>ALDH2</i> expression High vs Low	1.467	1.045–2.058	0.027*	1.480	1.054–2.077	0.023*
<i>EHHADH</i> expression High vs Low	0.917	0.658–1.277	0.607			
<i>AOX1</i> expression High vs Low	1.157	0.772–1.735	0.480			
Gender Male vs female	0.968	0.674–1.390	0.858			
Age (year) >60 vs ≤60	1.179	0.852–1.633	0.566			
Lymph node metastasis Positive vs Negative	1.057	0.764–1.462	0.738			
TNM stage III–IV vs I–II	1.433	1.001–2.050	0.049*	1.450	1.014–2.075	0.042*
Smoking History Yes vs No	0.914	0.576–1.450	0.702			

\* represents  $p < 0.05$ .

the first step of BCAA catabolism, is over-expressed in many malignant tumors [7,10,12,13]. BCAT is highly expressed in lung cancer and promotes the proliferation of lung cancer cells [10]. Therefore, it is necessary to systematically study the expression patterns of BCAAs catabolic enzymes in NSCLC and their correlation with disease prognosis.

Through multidimensional bioinformatic analysis, we found that the expression of BCAAs' metabolic enzymes *IL4I1*, *ALDH2*, and *AOX1* was specific to NSCLC and correlated with prognosis. *IL4I1* is a secreted L-amino acid oxidase that is induced by interleukin 4. *IL4I1* is highly expressed in lymphomas and associated with the prognosis of lymphomas [14]. *IL4I1* is a novel immunomodulatory

enzyme produced by mature dendritic cells, that inhibits the proliferation of effector T lymphocytes and promotes the development of regulatory T cells [15,16]. Local secretion of *IL4I1* in the immune synaptic cleft and its binding to CD3<sup>+</sup> lymphocytes may be important for the immunosuppressive mechanism of *IL4I1* [15]. In the current study, *IL4I1* was highly expressed in both LUAD and LUSC, with an increase in copy number. In the co-expression analysis network, *IL4I1* was in a key regulatory position in both LUAD and LUSC, indicating that *IL4I1* may play a key role in the development of lung cancer. In survival analysis, high expression of *L4I1* in LUSC was associated with poor DFS. We integrated survival data from LUAD and LUSC and found that *IL4I1* expression had a significant effect on OS, and its high expression indicated a lower OS. Our study revealed that *IL4I1* is closely related to the occurrence, development, and prognosis of NSCLC and is expected to become an important biomarker in the field of NSCLC immunotherapy and an effective predictor of the prognosis of NSCLC.

*AOX1* is a protein in the molybdoflavin family and an important enzyme involved in purine catabolism. More and more studies show that *AOX1* is involved in the pathophysiology of many clinical diseases [17,18]. *AOX1* promotes liver cell damage and fibrosis by increasing reactive oxygen species, which in turn may affect the metabolism and activity of drugs in the liver [19]. *AOX1* expression is reduced in hepatocellular carcinoma and correlates with a higher tumor stage, distant metastasis, or lymph node positive status [20]. The beneficial role of *Nrf2* in cancer prevention is essentially dependent on strict control of its activity, and relaxation of *Nrf2* is a key determinant of tumorigenesis and is found in many types of cancer [21]. Previous studies have shown that *AOX1* plays a critical role in the occurrence and development of tumors by regulating the *Nrf2* pathway [22]. In the present study, *AOX1* was poorly expressed in both LUAD and LUSC, indicating that *AOX1* may play an inhibitory role in the development and progression of NSCLC. DNA mutations of *AOX1* were highest in both LUAD and LUSC, especially LUSC, which has not been reported in previous studies. *AOX1* was mainly amplified by the copy numbers of LUAD and LUSC, but the expression of *AOX1* was negatively regulated with an increasing copy number. *AOX1* may play an important role in the occurrence and progression of NSCLC, but the underlying mechanisms need to be determined in further clinical and basic research experiments. *ALDH2* catalyzes the transformation of toxic methylmalonate semialdehyde into non-toxic methylmalonate via the valine catabolic pathway. *ALDH2* is mainly involved in liver metabolism and has been reported to play a role in liver diseases, especially alcoholic liver disease [23–25]. Among the different subtypes of acetaldehyde dehydrogenase, only *ALDH2* has better basic functions than other subtypes in the detoxification of acetaldehyde dehydrogenase [26]. Acetaldehyde, a substrate

catalyzed by *ALDH2* in the metabolic process, is closely related to a variety of tumors, and low expression of *ALDH2* in lung and liver cancer is associated with a poor prognosis [26,27]. Furthermore, aldehyde dehydrogenase is differentially expressed in lung cancer, and *ALDH2* was poorly expressed in lung cancer, while *ALDH1A1* and *ALDH3A1* were highly expressed in NSCLC [28]. Increasing evidence indicates that lung cancer may originate from tumor stem cells and that aldehyde dehydrogenase is a functional marker of lung cancer stem cells [29]. Several studies have reported that *ALDH2* is also a functional marker of lung cancer stem cells [29–31]. In the present study, *ALDH2* expression was down-regulated in both LUAD and LUSC. Low *ALDH2* expression was associated with poor OS in LUAD patients, which is consistent with previous reports [26]. However, high expression of *ALDH2* is associated with a poor prognosis for LUSC. *ALDH2* expression was related to LUSC survival in both the univariate and multivariate Cox regression analyses. There were some differences in the level of *ALDH2* expression in the survival of patients with different pathological subtypes of NSCLC, and there may be many unknown mechanisms and complex interference factors, which need to be confirmed in further studies.

## 5. Conclusions

Our study revealed the expression pattern and prognosis of differentially expressed BCAAs catabolism-related enzymes in NSCLC at multiple levels based on the TCGA database. First, we analyzed the expression of 44 BCAAs catabolic enzymes in NSCLC and normal lung tissues. A total of six differentially expressed genes (*ALDH1B1*, *ACAD8*, *IL4I1*, *OXCT2*, *ALDH2* and *AOX1*) and seven differentially expressed genes (*OXCT1*, *EHHADH*, *IL4I1*, *ALDH2*, *ACAA2*, *AOX1*, and *HMGCS2*) were identified in LUAD and LUSC, respectively. Among them, *IL4I1*, *ALDH2*, and *AOX1* were differentially expressed in both LUAD and LUSC. *IL4I1* participated in the first step of the catabolic process of L-isoleucine, metabolizing L-isoleucine to (S)-3-Methyl-2-oxopentanoate, producing Ammonia and Hydrogen peroxide. In the process of valine catabolism, *AOX1* mainly worked with aldehyde dehydrogenase protein family members *ALDH2* and *ALDH1B1* to oxidize Methylmalonate semialdehyde to Methylmalonate. Mutation, CNV, and methylation analysis of differentially expressed BCAAs catabolism-related enzymes were performed. Finally, a survival analysis of the differentially expressed BCAAs catabolism-related enzymes was performed. This study has some limitations. Our results were not verified in clinical NSCLC samples. This is a pilot study, and more experiments are needed to uncover the pathogenesis of differentially expressed BCAAs catabolism-related enzymes in NSCLC.

## Availability of Data and Materials

The data sets used and analyzed during the present study are available from the Cancer Genome Atlas (TCGA) public database.

## Author Contributions

XY contributed to the conception of the study. YD, JZ, LJ, and YS performed the data analyses. XY contributed significantly to the writing of the manuscript. All authors read and approved the final manuscript.

## Ethics Approval and Consent to Participate

Not applicable.

## Acknowledgment

Not applicable.

## Funding

This research received no external funding.

## Conflict of Interest

The authors declare no conflict of interest.

## Supplementary Material

Supplementary material associated with this article can be found, in the online version, at <https://doi.org/10.31083/j.fbl2806107>.

## References

- [1] Chen W, Sun K, Zheng R, Zeng H, Zhang S, Xia C, *et al.* Cancer incidence and mortality in China, 2014. *Chinese Journal of Cancer Research*. 2018; 30: 1–12.
- [2] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*. 2018; 68: 394–424.
- [3] Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, *et al.* Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet (London, England)*. 2018; 391: 1023–1075.
- [4] Mayers JR, Wu C, Clish CB, Kraft P, Torrence ME, Fiske BP, *et al.* Elevation of circulating branched-chain amino acids is an early event in human pancreatic adenocarcinoma development. *Nature Medicine*. 2014; 20: 1193–1198.
- [5] Ericksen RE, Lim SL, McDonnell E, Shuen WH, Vadiveloo M, White PJ, *et al.* Loss of BCAA Catabolism during Carcinogenesis Enhances mTORC1 Activity and Promotes Tumor Development and Progression. *Cell Metabolism*. 2019; 29: 1151–1165.e6.
- [6] Ananieva E. Targeting amino acid metabolism in cancer growth and anti-tumor immune response. *World Journal of Biological Chemistry*. 2015; 6: 281–289.
- [7] Tönjes M, Barbus S, Park YJ, Wang W, Schlotter M, Lindroth AM, *et al.* BCAT1 promotes cell proliferation through amino acid catabolism in gliomas carrying wild-type IDH1. *Nature Medicine*. 2013; 19: 901–908.
- [8] Hattori A, Tsunoda M, Konuma T, Kobayashi M, Nagy T, Glushka J, *et al.* Cancer progression by reprogrammed BCAA metabolism in myeloid leukaemia. *Nature*. 2017; 545: 500–504.
- [9] Wang ZQ, Faddaoui A, Bachvarova M, Plante M, Gregoire J, Renaud MC, *et al.* BCAT1 expression associates with ovarian cancer progression: possible implications in altered disease metabolism. *Oncotarget*. 2015; 6: 31522–31543.
- [10] Mayers JR, Torrence ME, Danai LV, Papagiannakopoulos T, Davidson SM, Bauer MR, *et al.* Tissue of origin dictates branched-chain amino acid metabolism in mutant Kras-driven cancers. *Science (New York, N.Y.)*. 2016; 353: 1161–1165.
- [11] DeBerardinis RJ, Chandel NS. Fundamentals of cancer metabolism. *Science Advances*. 2016; 2: e1600200.
- [12] Dey P, Baddour J, Muller F, Wu CC, Wang H, Liao WT, *et al.* Genomic deletion of malic enzyme 2 confers collateral lethality in pancreatic cancer. *Nature*. 2017; 542: 119–123.
- [13] Conway ME, Hull J, El Hindy M, Taylor SC, El Amraoui F, Paton-Thomas C, *et al.* Decreased expression of the mitochondrial BCAT protein correlates with improved patient survival in IDH-WT gliomas. *Brain Pathology (Zurich, Switzerland)*. 2016; 26: 789–791.
- [14] Carbonnelle-Puscian A, Copie-Bergman C, Baia M, Martin-Garcia N, Allory Y, Haioun C, *et al.* The novel immunosuppressive enzyme IL4I1 is expressed by neoplastic cells of several B-cell lymphomas and by tumor-associated macrophages. *Leukemia*. 2009; 23: 952–960.
- [15] Aubatin A, Sako N, Decrouy X, Donnadiou E, Molinier-Frenkel V, Castellano F. IL4-induced gene 1 is secreted at the immune synapse and modulates TCR activation independently of its enzymatic activity. *European Journal of Immunology*. 2018; 48: 106–119.
- [16] Boulland ML, Marquet J, Molinier-Frenkel V, Möller P, Guiter C, Lasoudris F, *et al.* Human IL4I1 is a secreted L-phenylalanine oxidase expressed by mature dendritic cells that inhibits T-lymphocyte proliferation. *Blood*. 2007; 110: 220–227.
- [17] Hille R, Hall J, Basu P. The mononuclear molybdenum enzymes. *Chemical Reviews*. 2014; 114: 3963–4038.
- [18] Garattini E, Mendel R, Romão MJ, Wright R, Terao M. Mammalian molybdo-flavoenzymes, an expanding family of proteins: structure, genetics, regulation, function and pathophysiology. *The Biochemical Journal*. 2003; 372: 15–32.
- [19] Neumeier M, Weigert J, Schäffler A, Weiss TS, Schmid C, Büttner R, *et al.* Aldehyde oxidase 1 is highly abundant in hepatic steatosis and is downregulated by adiponectin and fenofibric acid in hepatocytes in vitro. *Biochemical and Biophysical Research Communications*. 2006; 350: 731–735.
- [20] Sigrüener A, Buechler C, Orsó E, Hartmann A, Wild PJ, Terracciano L, *et al.* Human aldehyde oxidase 1 interacts with ATP-binding cassette transporter-1 and modulates its activity in hepatocytes. *Hormone and Metabolic Research*. 2007; 39: 781–789.
- [21] Geismann C, Arlt A, Sebens S, Schäfer H. Cytoprotection “gone astray”: Nrf2 and its role in cancer. *OncoTargets and Therapy*. 2014; 7: 1497–1518.
- [22] Maeda K, Ohno T, Igarashi S, Yoshimura T, Yamashiro K, Sakai M. Aldehyde oxidase 1 gene is regulated by Nrf2 pathway. *Gene*. 2012; 505: 374–378.
- [23] Wang RS, Nakajima T, Kawamoto T, Honma T. Effects of aldehyde dehydrogenase-2 genetic polymorphisms on metabolism of structurally different aldehydes in human liver. *Drug Metabolism and Disposition: the Biological Fate of Chemicals*. 2002; 30: 69–73.
- [24] Qiu L, Zhong W. Relationship between Human aldH2\*2 and Alcoholic Diseases. *Journal of Medical Molecular Biology*. 2007; 4: 164–167.
- [25] Borràs E, Coutelle C, Rosell A, Fernández-Muixi F, Broch M, Crosas B, *et al.* Genetic polymorphism of alcohol dehydrogenase in europeans: the ADH2\*2 allele decreases the risk for

- alcoholism and is associated with ADH3\*1. *Hepatology* (Baltimore, Md.). 2000; 31: 984–989.
- [26] Chen X, Legrand AJ, Cunniffe S, Hume S, Poletto M, Vaz B, *et al.* Interplay between base excision repair protein XRCC1 and ALDH2 predicts overall survival in lung and liver cancer patients. *Cellular Oncology* (Dordrecht). 2018; 41: 527–539.
  - [27] Lessel D, Vaz B, Halder S, Lockhart PJ, Marinovic-Terzic I, Lopez-Mosqueda J, *et al.* Mutations in SPRTN cause early onset hepatocellular carcinoma, genomic instability and progeroid features. *Nature Genetics*. 2014; 46: 1239–1244.
  - [28] Moreb JS, Baker HV, Chang LJ, Amaya M, Lopez MC, Ostmark B, *et al.* ALDH isozymes downregulation affects cell growth, cell motility and gene expression in lung cancer cells. *Molecular Cancer*. 2008; 7: 87.
  - [29] Liu X, Wang L, Cui W, Yuan X, Lin L, Cao Q, *et al.* Targeting ALDH1A1 by disulfiram/copper complex inhibits non-small cell lung cancer recurrence driven by ALDH-positive cancer stem cells. *Oncotarget*. 2016; 7: 58516–58530.
  - [30] Yu Y, Wang YY, Wang YQ, Wang X, Liu YY, Wang JT, *et al.* Antiangiogenic therapy using endostatin increases the number of ALDH+ lung cancer stem cells by generating intratumor hypoxia. *Scientific Reports*. 2016; 6: 34239.
  - [31] Corominas-Faja B, Oliveras-Ferraros C, Cuyàs E, Segura-Carretero A, Joven J, Martin-Castillo B, *et al.* Stem cell-like ALDH(bright) cellular states in EGFR-mutant non-small cell lung cancer: a novel mechanism of acquired resistance to erlotinib targetable with the natural polyphenol silibinin. *Cell Cycle* (Georgetown, Tex.). 2013; 12: 3390–3404.