H2A Histone Family Member Z (H2AFZ) Serves as a Prognostic Biomarker in Lung Adenocarcinoma: Bioinformatic Analysis and Experimental Validation

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

AE 1,2 **Zongkuo Li**
A 1,2 **Menglong Hu**
C 1,2 **Jinhuan Qiu**
C 1,2 **Junkai Feng**
C 1 **Ruizhen Zhang**
D 1 **Huifang Wu**
D 1 **Guiming Hu**
G 1 **Jingli Ren**

1 Department of Pathology, The Second Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, PR China
2 Academy of Medical Sciences, Zhengzhou University, Zhengzhou, Henan, PR China

Background: H2A histone family member Z (H2AFZ) is a special subtype in the H2A histone family, which participates in the regulation of gene transcription. Nevertheless, little is known about the role of H2AFZ in the tumor microenvironment and genetic factors associated with lung cancer.

Material/Methods: The expression of H2AFZ in LUAD was analyzed via Tumor Immune Estimation Resource (TIMER), the Cancer Genome Atlas (TCGA), and Gene Expression Omnibus (GEO) databases at the mRNA level. To detect the protein expression level of H2AFZ, immunohistochemistry (IHC) was performed using LUAD tissues and non-tumor lung tissues. Kaplan-Meier survival analysis and Cox regression analysis were conducted to identify the effect of H2AFZ expression on overall survival (OS) based on TCGA-LUAD and the GEO dataset GSE68465 cohorts, and our LUAD patient cohort was used for validation. Identification of signaling pathways associated with the expression of H2AFZ was performed using Gene Set Enrichment Analysis (GSEA). The influences of expression of H2AFZ on tumor immune-infiltrating cell (TIICs) were assessed via TIMER and CIBERSORT.

Results: The expression of H2AFZ was increased in LUAD tissues at both mRNA and protein levels. In addition, high expression of H2AFZ predicted poor OS and might be an independent prognostic predictor in LUAD patients. Moreover, H2AFZ affected the relative proportion of TIICs and was positively associated with Myeloid-derived suppressor cells (MDSC) infiltration level in LUAD.

Conclusions: H2AFZ was upregulated in LUAD and related to poor prognosis of LUAD patients; thus, it could be an underlying prognostic biomarker correlated with immune infiltration in LUAD.

Keywords: Adenocarcinoma of Lung • Biomarkers • H2A.Z Protein, Arabidopsis

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/933447

📄 3716   🗂 6   📊 8   📖 39

Background

Based on the analysis of GLOBOCAN 2018, the most common cause of death among all malignancies worldwide is lung cancer [1]. The proportion of non-small cell lung cancer (NSCLC) is highest in all lung cancer cases [2], with LUAD occupying a large part of NSCLC [3]. Although recent progress in therapeutic methods has greatly improved clinical outcomes, the prognosis of LUAD patients is still far from satisfactory [4]. Previously, studies have reported widespread use of immunotherapy as an effective treatment strategy in patients with LUAD [4, 5]. Nonetheless, due to specific differences in patients, only a tiny proportion benefit from immunotherapy. In this regard, there have been efforts by scientists to search for more meaningful immune-related biomarkers as targeted agents and prognostic predictors in patients with LUAD.

H2A histone family member Z (H2AFZ) is the most expressed histone variant in the H2A family. The variant is significantly involved in key cell growth and division processes, including chromosome segregation [6], cell cycle progression [7], and maintenance of heterochromatin [8]. The involvement of H2AFZ in these processes has been linked to the main mechanisms of its carcinogenic effects [9]. Recently, many studies have concentrated on the relationships between H2AFZ expression and cancers, including breast cancer [10-13], prostate cancer [14-16], bladder cancer [17], colorectal cancer [18], and its role in tumor progression. Elsewhere, the overexpression of H2AFZ has been reported to be more pronounced during the metastatic stage of cancer [18,19]. A study based on meta-analysis identified H2AFZ as a commonly dysregulated gene in lung cancer and an essential gene associated with molecular pathohistological subtypes of lung cancer [20]. However, few studies have explored the role and mechanism of H2AFZ in LUAD. Recent advances in the study of H2AFZ have opened up new insights for the study of LUAD.

Herein, using abundant TCGA, Gene Expression Omnibus (GEO), and TIMER databases, we attempted to investigate the effect of H2AFZ on tumor-infiltrating immune cells (TIICs) and its prognostic value in lung adenocarcinoma (LUAD). We also attempted to verify differential expression via immunohistochemistry, and analyze the correlation between expression level of H2AFZ and clinical features, as well as the OS of LUAD patients. The function enrichment analyses were carried out to identify the potential molecular mechanism of H2AFZ in LUAD. Furthermore, we studied the correlations between H2AFZ expression and immune cell subtypes to identify whether H2AFZ could be a new molecular target associated with immune infiltration for the treatment of LUAD patients.

Material and Methods

Data Obtained from Different Databases

The mRNA sequencing information as well as corresponding clinical parameters for samples were obtained from The Cancer Genome Atlas (TCGA) (https://tcga-data.nci.nih.gov/tcga/) and GEO database (https://www.ncbi.nlm.nih.gov/geo/). The TCGA-LUAD dataset comprised 594 samples, including 535 LUAD samples and 59 non-tumor samples. Due to the lack of M classification in most samples, we excluded the M classification from our study. Therefore, a total of 439 tumor cases with complete clinical data were obtained for further analyses. Three datasets, namely GSE68465, GSE10072, and GSE43458, were selected from the GEO database. The GSE10072 dataset comprised 107 samples containing mRNA sequencing information of 49 non-tumor cases and 58 LUAD cases. The GSE43458 dataset consisted of 30 non-tumor samples and 80 LUAD samples for a total of 110 samples. Similarly, there were 439 tumor samples in the GSE68465 dataset after excluding samples with incomplete information. The mRNA sequencing data from the GEO database were normalized by Bioinformatics Array Research Tool (BART) [21], which can comprehensively analyze the extensive microarray datasets on GEO and correct the data to make the results more authentic and reliable.

Differential Analysis of H2AFZ mRNA Expression

TIMER 2.0 (http://timer.cistrome.org) is an open-access data mining database. It includes multiple oncology datasets, samples and corresponding sequencing data. The differential expression of H2AFZ in multiple tumors was identified using the *GENE_DE* module in the TIMER 2.0 database. The downloaded mRNA expression profile and clinical data from the TCGA-LUAD cohort were sorted out and merged using Perl programming language (https://www.perl.org/). After that, the mRNA expression profile of H2AFZ was extracted from the TCGA-LUAD dataset and visualized with "limma" and "beeswarm" packages of R version 4.1.0. Importantly, GSE10072 and GSE43458 cohorts were only used to verify the difference in H2AFZ mRNA expression between normal lung and LUAD sample, owing to lack of corresponding clinical information.

Patient Samples

We randomly selected 109 LUAD samples and 55 para-cancerous samples in the Department of Pathology, the Second Affiliated Hospital of Zhengzhou University from 2015 to 2018. None of the patients had received preoperative chemoradiotherapy, and their clinical diagnosis was confirmed by the pathology department of our hospital. All samples were subjected to immunohistochemistry experiments, and clinical information of the samples were collected for survival analysis. The

use of human specimens in this study was approved by the Ethics Committee of our hospital (No. 2021311).

Immunohistochemistry

About 3-um-thick tissues were cut from formalin-fixed paraffin-embedded (FFPE) specimens and were placed on the slides. The sections were deparaffinized and hydrated, then heat-mediated antigen retrieval was carried out with sodium citrate buffer (pH 6.0). The sections were treated with 3% catalase solution and incubated for 15 min to inactivate endogenous peroxidase activity. The primary antibody of H2AFZ (Abcam, Cat# ab150402, RRID: AB_2891240, US) at 1: 2500 dilution was added to the sections and incubated for 1 h. The sections were then rinsed with 0.3% phosphate-buffered saline 3 times and treated with secondary goat anti-rabbit antibody at room temperature. After staining, washing, and dehydrating, the sections were covered with glass coverslips.

Comprehensive judgments were made according to the percentage of positive cells combined with staining intensity. The staining intensity score standard was: colorlessness of 0; low intensity of 1; medium intensity of 2; strong intensity of 3. Under the microscope 100× field of vision, each section was observed randomly in 5 fields and 100 or 200 cells were counted. The scores ranged from 1 to 4 depending on the percentage of positive cells: 1; 0-25% of positive cells, 2; 26-50%, 3; 51-75%, and 4; 75-100%. The 2 scores were multiplied to then yield a final IHC score [22]. The IHC results were confirmed by the independent observation of at least 2 pathologists. In accordance with the median value of the IHC score, LUAD samples were divided into 2 groups: a score less than 7 was considered as low H2AFZ expression, while ≥7 was considered as high H2AFZ expression.

Functional Enrichment Analyses of H2AFZ and Co-Expression Genes

Following LUAD data in the TCGA database, H2AFZ and the top 100 relevant genes were identified using the LinkedOmics web server (http://www.linkedomics.org) [23]. To determine the potential role of H2AFZ in LUAD, Gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment analyses and visualization were carried out on H2AFZ co-expressed genes using R packages "clusterProfiler, org.Hs.eg.db, enrichplot and ggplot2" [24]. A q-value less than 0.05 showed that GO terms and KEGG pathways were significantly enriched.

Survival Analysis and Cox Regression Analyses

The extracted sequencing data with corresponding clinical parameters in the TCGA-LUAD cohort and GSE68465 dataset were merged. Based on the median expression level of H2AFZ, deemed as the cut-off value, all LUAD cases in the TCGA-LUAD, GSE68465, and our hospital cohorts were respectively divided into low and high expression groups. This was followed by Kaplan-Meier analysis of OS between the 2 groups using the log-rank test. Then, the Cox regression analysis was carried out on the TCGA-LUAD and GSE68465 cohorts to assess the predictive ability of H2AFZ expression on OS.

Gene Set Enrichment Analysis (GSEA)

To explore the molecular role of H2AFZ and its oncogenic mechanism in LUAD, tumor-associated signaling pathways related to H2AFZ expression in the TCGA-LUAD dataset were identified through GSEA 3.0 software [25]. We selected the hallmark gene set (h.all.v7.2.symbols.gmt) as the internal reference. A phenotype label was determined based on the H2AFZ expression level. Every gene set was permuted 1000 times to calculate the normalized enrichment score (NES). Pathways related to gene sets were screened using the nominal (NOM) P value <0.05, and false discovery rate (FDR) q-value <0.05 as the criteria. Then, visualization of those results was performed using R packages "ggplot2, plyr, grid and gridExtra" [26].

Correlation Analysis of H2AFZ with Immune-Infiltrating Cells

TIMER 2.0 can systematically analyze the relationships between target genes and multiple immune cells in different cancer types following gene expression sequencing profiling in the TCGA database. The association between TIIC infiltration level and the expression level of H2AFZ was assessed using the *Gene* module of TIMER 2.0. In addition, we also investigated whether H2AFZ affected the OS of patients with LUAD by affecting the infiltration level of MDSC using the "*OUTCOME*" module. Finally, under the median H2AFZ expression value, the TCGA-LUAD cases were divided into 2 subgroups: high- and low-H2AFZ expression groups. The difference in immune cells composition between the 2 groups was then analyzed using CIBERSORT.

Statistical Analysis

The difference in the expression of H2AFZ between LUAD tissues and non-tumor tissues was determined using the Wilcoxon test. The Kruskal-Wallis test was applied to observe changes in the expression of H2AFZ among different groups based on pathological stages and T and N classifications. To assess the influence of H2AFZ on the prognosis of patients, Kaplan-Meier analysis was performed on all LUAD samples. The log-rank test was carried out to determine the difference in the survival curves between the 2 groups. The expression of H2AFZ that correlated with the clinic-pathological parameters was
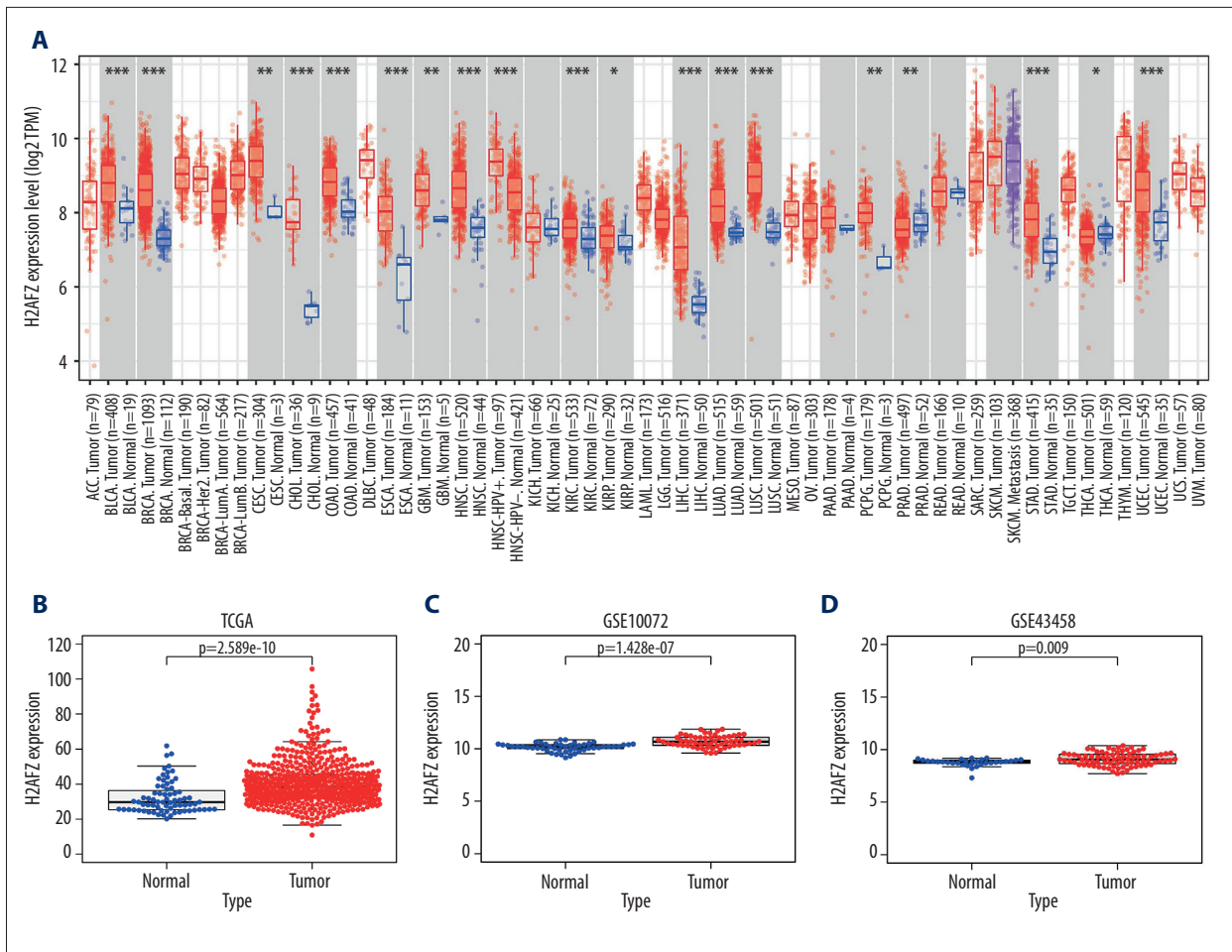
Figure 1. H2AFZ expression in LUAD tissue at the mRNA level. (**A**) Differential expression of H2AFZ in different cancer tissues compared with normal tissues in the Timer 2.0 database (* p<0.05, ** p<0.01, *** p<0.001). (**B-D**) Comparisons of H2AFZ expression levels between LUAD tissues and normal lung tissues in different datasets including TCGA, GSE10072 and GSE43458 datasets. R 4.1.0 software, (R Development Core Team, Vienna).

evaluated using the chi-square (χ^2) test. SPSS 22.0 software, R 4.1.0 software, and GraphPad Prism 7.0 software were used for statistical analyses, and *p*<0.05 was considered as a statistically significant.

Results

The Expression Level of H2AFZ mRNA was Upregulated in LUAD Tissues

Analyses of the TIMER 2.0 database indicated that H2AFZ was overexpressed in many cancer types compared with adjacent normal tissues, such as bladder, breast, cervical, colorectal, gastric, kidney, liver, and lung cancers (**Figure 1A**). Comparison of the expression level of H2AFZ in LUAD tissues with their adjacent non-tumor lung tissues via analysis of RNA sequencing data of TCGA-LUAD samples showed that the expression

level of H2AFZ mRNA was upregulated in LUAD samples (**Figure 1B**). Besides, the expression level of H2AFZ was verified in GSE10072 and GSE43458 datasets (**Figure 1C, 1D**). Therefore, H2AFZ might promote the occurrence of a variety of cancers, especially LUAD.

Expression of H2AFZ Was Upregulated at the Protein Level in LUAD Tissues

Immunohistochemical staining performed on specimens including 109 LUAD tissues and 55 non-tumor tissues revealed that H2AFZ was mainly located in the nucleus of LUAD cells, but H2AFZ protein expression was not recognized in adjacent normal tissues. Results of immunohistochemical staining are presented in **Figure 2A and 2B**. After independent scoring by 2 pathologists, the differences between LUAD and non-tumor tissues were statistically significant at *P*<0.05 (**Figure 2C**). The expression of H2AFZ in LUAD tissues was increased at
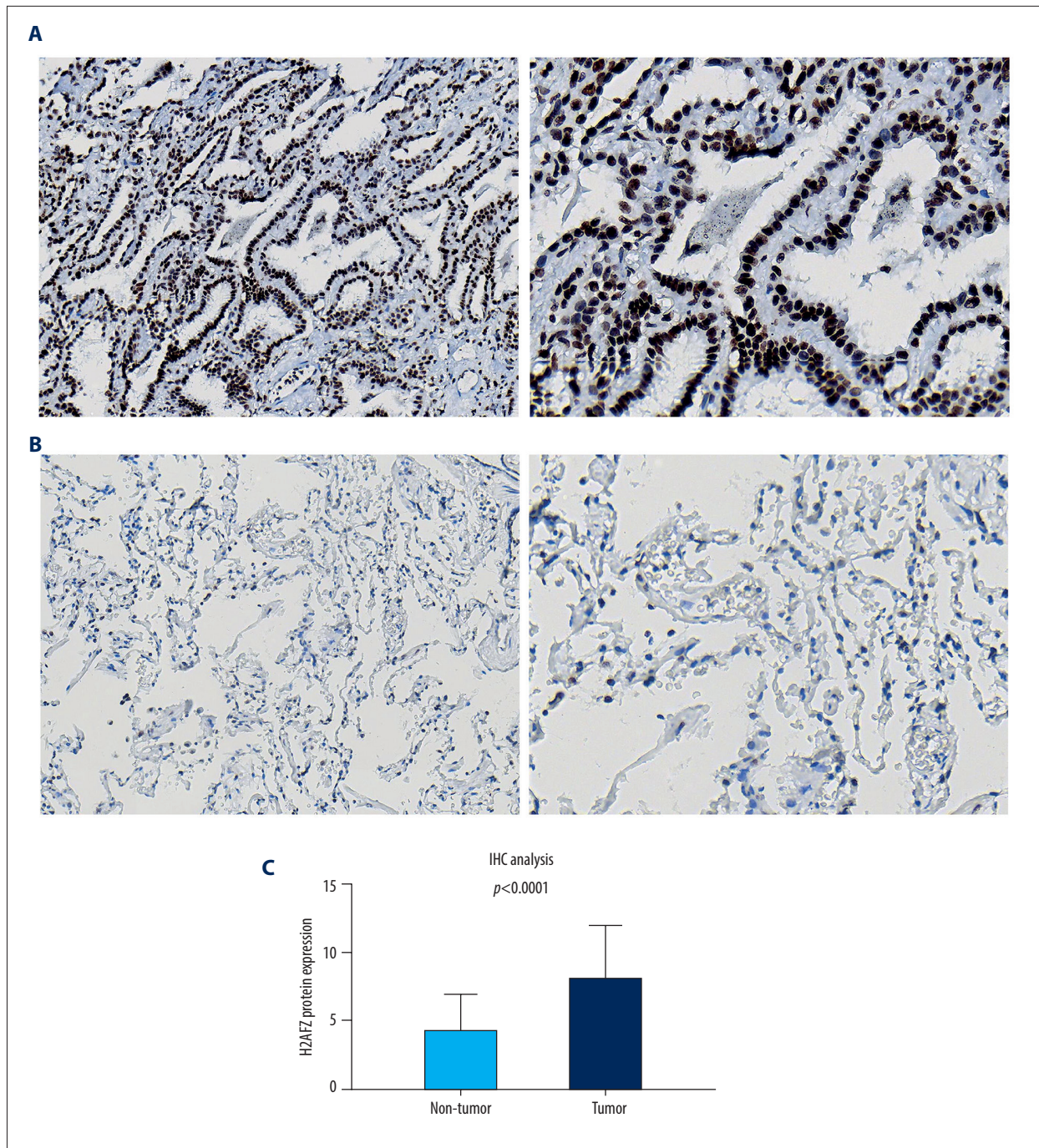
Figure 2. Immunohistochemistry of H2AFZ in the LUAD tissues and normal lung tissues. (**A**) Nuclear staining in LUAD tissues (**left panel:** 20×, **right panel:** 40×). (**B**) Cells in adjacent normal tissues are not stained (**left panel:** 20×, **right panel:** 40×). (**C**) Differential expression of H2AFZ protein in LUAD tissues and normal lung tissues. GraphPad Prism 7.0 software (GraphPad Software Inc., La Jolla, CA, USA).

A

Positively correlated significantly genes
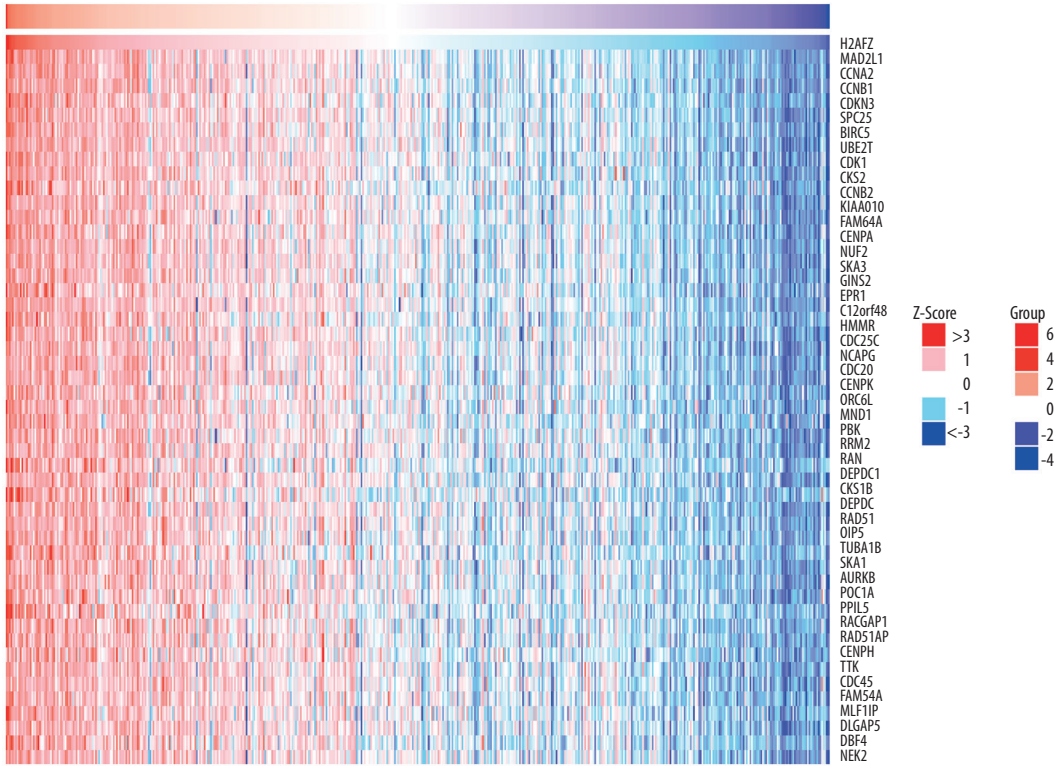
Negatively correlated significantly genes
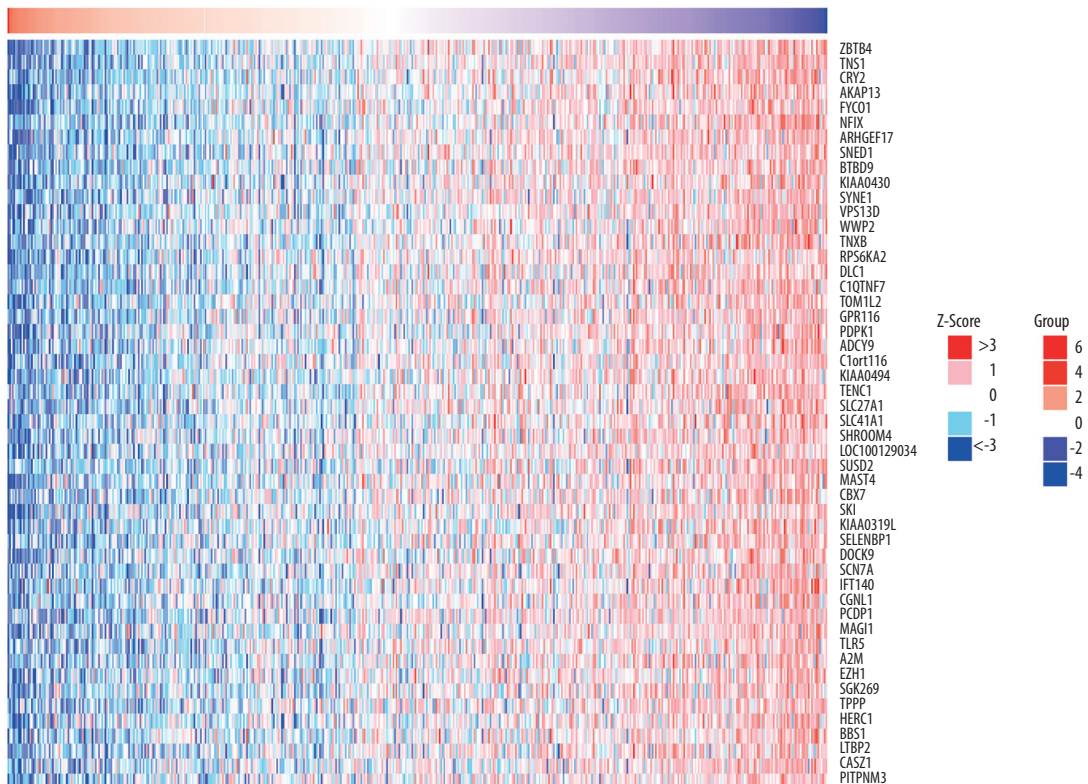
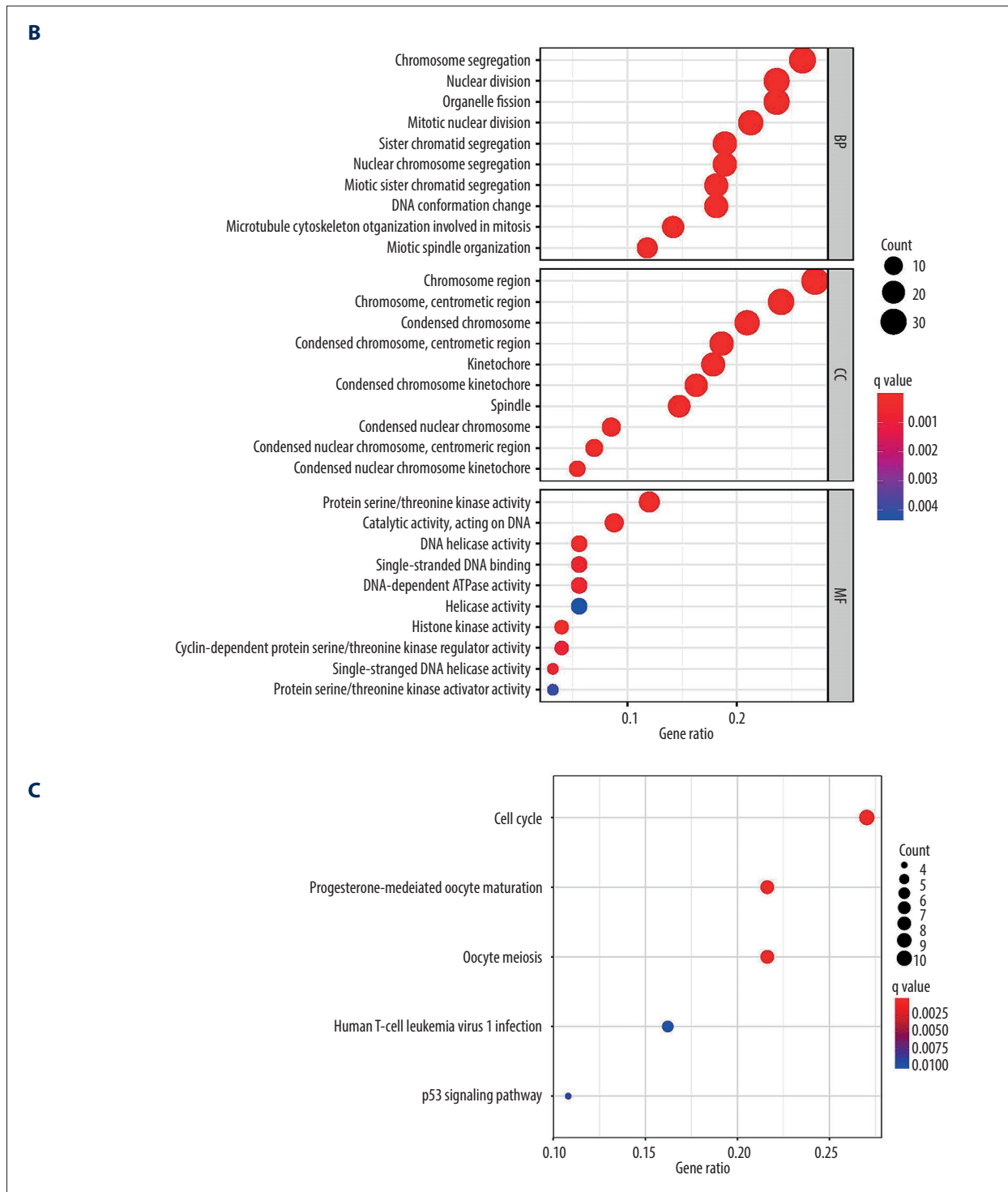**Figure 3.** Gene enrichment analysis of H2AFZ co-expressed genes based on the TCGA-LUAD data. (**A**) Heatmaps showing genes positively and negatively correlated with H2AFZ in LUAD (top50). (**B, C**) Enriched GO terms and enriched KEGG pathways of H2AFZ-associated genes. R 4.1.0 software, (R Development Core Team, Vienna).
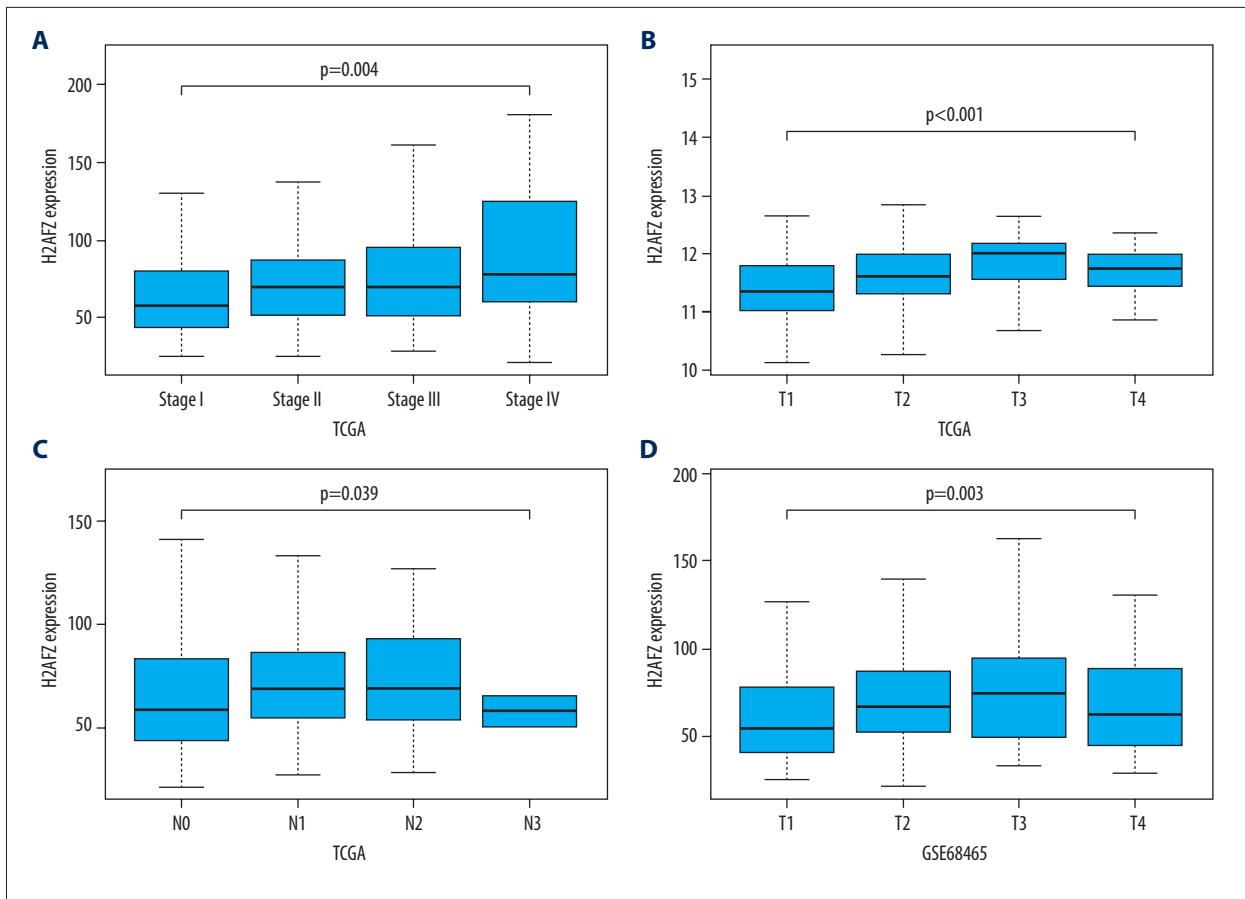
Figure 4. The expression level of H2AFZ in subgroups of patients with LUAD. (**A-C**) The expression of H2AFZ is grouped by pathological stage, T, and N classification based on TCGA-LUAD data. (**D**) Differential mRNA expression of H2AFZ in LUAD with different T classifications in GSE68465. R 4.1.0 software (R Development Core Team, Vienna).

the protein level, which was consistent with its expression at the mRNA level.

GO Annotation and KEGG Pathway Enrichment of H2AFZ Co-Expressed Genes

The LinkedOmics database was applied to seek the genes associated with H2AFZ in LAUD. As presented in the heatmap (**Figure 3A**), 50 genes positively and negatively were correlated with H2AFZ, respectively. To identify the molecular functions of H2AFZ, GO and KEGG enrichment analyses were conducted on its co-expressed genes. It was found that the most significant GO biological process (BP) terms were "chromosome segregation" and "nuclear division". Furthermore, H2AFZ co-expressed genes were found to be prevalently enriched in "chromosomal region" in terms of the cellular compartment (CC). Results of molecular function (MF) analysis identified that H2AFZ might participate in regulating "protein serine/threonine kinase activity" (**Figure 3B**). Several KEGG signaling pathways suggested that H2AFZ and associated genes have the potential to participate in processes like "cell cycle", "progesterone-mediated

oocyte maturation", and "p53 signaling pathway" (**Figure 3C**). These results demonstrated that H2AFZ might promote tumor cell proliferation by regulating certain critical processes in cell division.

The Relationship Between the Expression of H2AFZ and Clinical Features in LUAD

On extracting the expression value of H2AFZ from clinical data of patients from the TCGA-LUAD cohort, the effect of tumor pathological stage on H2AFZ expression (*P*<0.05) was the most significant clinical feature observed (**Figure 4A**). T and N classifications also influenced the expression of H2AFZ, with the highest expression occurring in T3 and N2 classifications LUAD patients (**Figure 4B, 4C**). The above analyses were then performed in GSE48465, and similar results were obtained (**Figure 4D**).

By setting median expression value as the cut-off, TCGA-LUAD samples and GSE68465 were separated into 2 groups. As shown in **Table 1**, the increased expression of H2AFZ was remarkably associated with pathological stage, T classification, N

Table 1. Association between H2AFZ expression and clinicopathologic characteristics in patients with LUAD in TCGA-LUAD cohort.

Characteristics	H2AFZ expression		Total	p Value
	Low (n=220)	High (n=219)		
Age				
<65	93	103	196	0.316
≥65	127	116	243	
Gender				
Male	88	113	201	**0.015**
Female	132	106	238	
Pathological stage				
I	139	100	239	**0.002**
II	44	61	105	
III	31	43	74	
IV	6	15	21	
T classification				
T1	92	61	153	**0.022**
T2	102	128	230	
T3	17	21	38	
T4	9	9	18	
N classification				
N0	160	130	290	**0.032**
N1	33	49	82	
N2	26	39	65	
N3	1	1	2	

Bold values indicate *p*<0.05.

Table 2. Association between H2AFZ expression and clinicopathologic characteristics in patients with LUAD in GSE68465.

Characteristics	H2AFZ expression		Total	p Value
	Low (n=220)	High (n=219)		
Age				
<65	100	112	212	0.233
≥65	120	107	227	
Gender				
Male	97	124	221	**0.010**
Female	123	95	218	
T classification				
T1	97	52	149	**<0.001**
T2	112	139	251	
T3	7	21	28	
T4	4	7	11	
N classification				
N0	159	140	299	0.132
N1	40	47	87	
N2	21	32	53	
N3				

Bold values indicate *p*<0.05.

Table 3. Association between H2AFZ expression and clinicopathological characteristics in patients with LUAD in our cohort.

Characteristics	H2AFZ expression		Total	p Value
	Low (n=220)	High (n=219)		
Age				
<65	29	51	80	0.865
≥65	10	19	29	
Gender				
Male	14	40	54	**0.033**
Female	25	30	55	
Pathological stage				
I	25	36	61	**0.039**
II	5	23	28	
III	8	6	14	
IV	1	5	6	
T classification				
T1	23	45	68	0.801
T2	10	18	28	
T3	5	5	10	
T4	1	2	3	
N classification				
N0	23	41	64	0.679
N1	4	11	15	
N2	12	18	30	
N3				

Bold values indicate *p*<0.05.

classification, and sex in the TCGA-LUAD cohort, all at *P*<0.05. Elsewhere, it was found that H2AFZ expression was related to T classification and sex in GSE68465, all at *P*<0.05 (**Table 2**). In our cohort, H2AFZ overexpression was associated with tumor pathological stage and sex (**Table 3**). The above results revealed that H2AFZ overexpression might be an important element of tumor progression.

H2AFZ Overexpression Predicted Poor Overall Survival in LUAD Patients

K-M survival analysis and log-rank test done on the low and high H2AFZ expression groups in the TCGA-LUAD cohort and GSE68465 dataset to analyze survival differences between the groups in TCGA-LUAD and GSE68465 samples showed that high H2AFZ expression was markedly correlated with decreased OS in LUAD patients, at *P*<0.05 (**Figure 5A, 5B**). The above analysis carried out on LUAD samples from our hospital revealed similar results (**Figure 5C**), at *P*<0.05.

H2AFZ Serves as an Independent Prognostic Marker in LUAD Patients

Cox regression analyses were conducted in TCGA-LUAD and GSE68465 cohorts to identify whether H2AFZ expression could independently influence the OS of LUAD patients. In the TCGA-LUAD cohort, univariate analysis demonstrated that advanced pathological stage, T and N classifications, and increased H2AFZ expression were closely related to shorter OS period (all *P*<0.001). Elsewhere, results of the multivariate analysis indicated that both H2AFZ expression and pathological stage are potential indicators that predict survival of LUAD patients (**Table 4, Figure 6A**). Similarly, as represented in **Table 5** and **Figure 6B**, H2AFZ mRNA expression, age, and T and N classifications were independent predictors for OS of LUAD patients in the GSE68465 dataset (all at *P*<0.001). Taken together, our results suggest that H2AFZ could be a new independent factor predicting poor prognosis of LUAD patients.
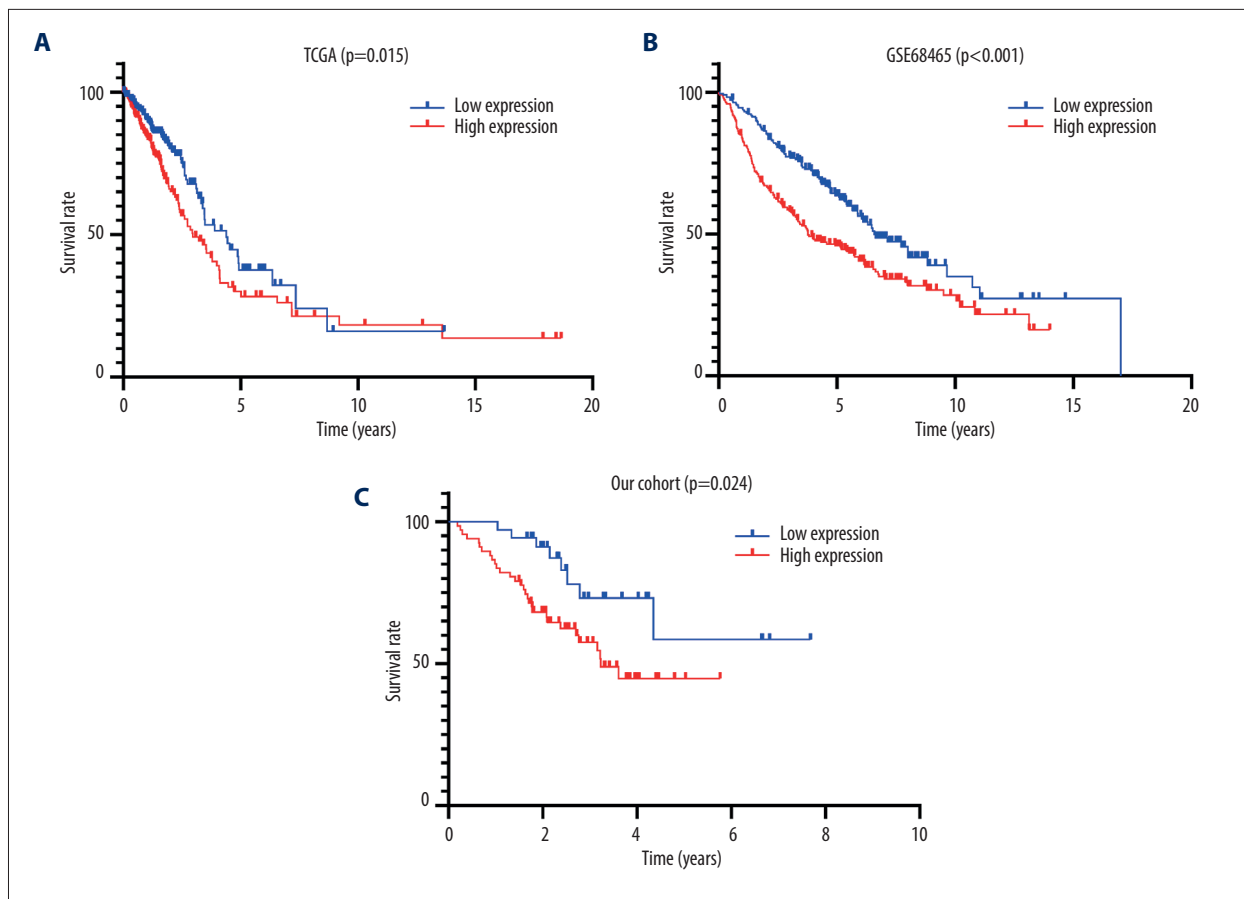
Figure 5. Kaplan-Meier survival curves comparing the high and low expression of H2AFZ in TCGA-LUAD, GSE68465, and our cohorts. (**A**) Kaplan-Meier (KM) survival analysis of H2AFZ in LUAD based on the TCGA-LUAD cohort. (**B**) Kaplan-Meier (KM) survival analysis of H2AFZ in LUAD based on the GSE68465. (**C**) Kaplan-Meier (KM) survival analysis of H2AFZ in LUAD based on our cohort. GraphPad Prism 7.0 software (GraphPad Software Inc., La Jolla, CA, USA).

Enrichment Analyses Associated with H2AFZ Using GSEA

To determine the potential functions of H2AFZ and its effect on LUAD carcinogenesis, the H2AFZ expression data obtained from the TCGA-LUAD dataset were used to determine cancer hallmarks significantly related to H2AFZ through GSEA. According to NES, FDR *q*-value, and nominal *P* value, 10 significant signaling pathways were enriched in the high expression group of H2AFZ: DNA repair, MYC targets V2, G2M checkpoint, E2F targets, oxidative phosphorylation, mitotic spindle, MYC targets V1, unfolded protein response, and mTORC1 signaling pathways, as well as the PI3K/AKT/mTOR signaling pathway (**Table 6, Figure 7**).

H2AFZ Expression Might Affect Immune Infiltration Levels in LUAD

The study further sought to understand whether the expression of H2AFZ was associated with the level of immune infiltration in LUAD patients. First, we sought to investigate immune cell subtypes affected by the expression of H2AFZ in

LUAD. Comparisons of the relative proportions of immune cells between the low and high H2AFZ expression groups showed that there were significant differences in some immune cell subtypes ($P<0.05$), including memory B cells, resting mast cells, CD8+ T cells, activated dendritic cells, resting memory CD4+ T cells, monocytes, activated memory CD4+ T cells, eosinophils, follicular helper T cells, M1 macrophages, resting dendritic cells, and T cells gamma delta (**Figure 8A**).

Results of the correlation analyses between expression of H2AFZ and immune cell subtypes using the TIMER 2.0 database indicated that the expression level of H2AFZ were negatively correlated with the infiltration levels of MDC cells (Rho=-0.257, *P*=7.54e-09), NK cells (Rho=-0.275, *P*=5.55e-10), T cell regulatory (Tregs) (Rho=-0.224, *P*=4.78e-07), and macrophage M2 (Rho=-0.334, *P*=2.57e-14). On the contrary, the expression level of H2AFZ was positively correlated with the levels of MDSC (Rho=0.606, *P*=7.55e-51) and macrophage M1 (Rho=0.327, *P*=9.22e-14) (**Figure 8B**). Finally, the influence of expression of H2AFZ and the infiltration level of MDSC on the OS of LUAD

Table 4. Univariate analysis and multivariate analysis of the correlation of H2AFZ expression with OS among LUAD patients in TCGA cohort.

Parameters	Univariate analysis			Multivariate analysis		
	HR	95% CI	*p* Value	HR	95% CI	*p* Value
Age	1.012	0.995-1.029	0.145	1.020	1.004-1.037	0.017
Gender	1.084	0.787-1.493	0.619	0.900	0.649-1.248	0.526
Pathological stage	1.655	1.426-1.921	**<0.001**	1.500	1.204-1.868	**<0.001**
T classification	1.535	1.259-1.872	**<0.001**	1.176	0.950-1.455	0.136
N classification	1.699	1.419-2.033	**<0.001**	1.162	0.907-1.488	0.236
H2AFZ	1.499	1.189-1.891	**<0.001**	1.494	1.185-1.884	**<0.001**

Bold values indicate *p*<0.05. HR – hazard ratio; CI – confidence interval.

patients was also analyzed. As presented in **Figure 8C**, patients with high H2AFZ expression and high level of MDSC infiltration were likely to have poorer prognosis, suggesting that H2AFZ influences the prognosis of patients with LUAD by affecting the MDSC to some degree.

Discussion

Lung adenocarcinoma (LUAD) is a serious threat to human health and it is an important cause of cancer death worldwide. With the continuous progress of scientific research, new insights into the diagnosis and treatment of LUAD continue to develop. However, the development of metastasis and recurrence after treatment makes the prognosis of LUAD patients poor. Therefore, the search for new biomarkers and therapeutic targets has become the main direction of research on the treatment of LUAD.

Genome and epigenetic changes are essential drivers of the development of cancer. Moreover, histone variants and their post-translational modifications also play an essential role in the occurrence and development of cancer. H2AFZ serves as a histone H2A variant and cell cycle-related gene. The variant participates in the activation, inhibition, and extension of transcription during gene transcription. It also plays a vital role in the histology and stability of chromatin [27]. Its expression is upregulated in a variety of tumor tissues, affecting the prognosis of patients [28]. A report on breast cancer suggested that H2AFZ induced the expression of Myc protein by enriching the promoter region of the Myc gene [29]. In malignant melanoma, high expression of H2AFZ was necessary for cell proliferation and was related to decreased OS [30].

In our study, we attempted to characterize the role of H2AFZ upregulation in the development of LUAD and explore how it participates in this process. First, we observed that H2AFZ expression in LUAD tissues was remarkably increased at the mRNA level based on TIMER 2.0, TCGA, and GEO databases. To verify the expression difference of H2AFZ at the protein level, we performed IHC on clinical LUAD tissues and non-tumor tissues. The results of immunohistochemistry were consistent with those of the above databases. We found that H2AFZ might be related to the occurrence of LUAD. We also found that the expression level of H2AFZ in the advanced stage of LUAD was obviously higher than that in the early stage and was related to the T classification. This implies that the higher the expression level of H2AFZ, the higher the malignant degree of LUAD. In addition, a higher expression level of H2AFZ indicated poor prognosis, and the results of univariate and multivariate analyses suggested that H2AFZ could independently predict OS of LUAD patients. We also collected detailed information on patients in our hospital and carried out survival analysis on them, and the results were consistent with bioinformatic analysis.

Some studies have proposed that metabolic abnormalities are not only associated with the occurrence and progression of tumors, but also the abnormal expression of genes correlated with metabolic pathways. This may be realized through direct involvement in cell regulation which promotes the abnormal proliferation of tumor cells [17,31]. It has been reported that H2AFZ might be involved in the occurrence of liver cancer by accelerating the cell cycle process and regulating related proteins [13]. Regarding LUAD, these studies were consistent with the signaling pathways (cell cycle, pyrimidine metabolism, p53 signaling pathway, and DNA replication) as determined in our enrichment analysis. In addition, we also screened out several cancer-related signal pathways closely associated with H2AFZ high expression through GSEA. Many studies have shown that certain signaling pathways, such as DNA repair, G2M checkpoint, mitotic spindle, E2F targets, and Myc targets V1/2, were involved in cell growth and cell proliferation, leading to cancer invasion and metastasis and low patient survival [32,33]. Some scientists also believe that PI3K/AKT/mTOR and unfolded protein response
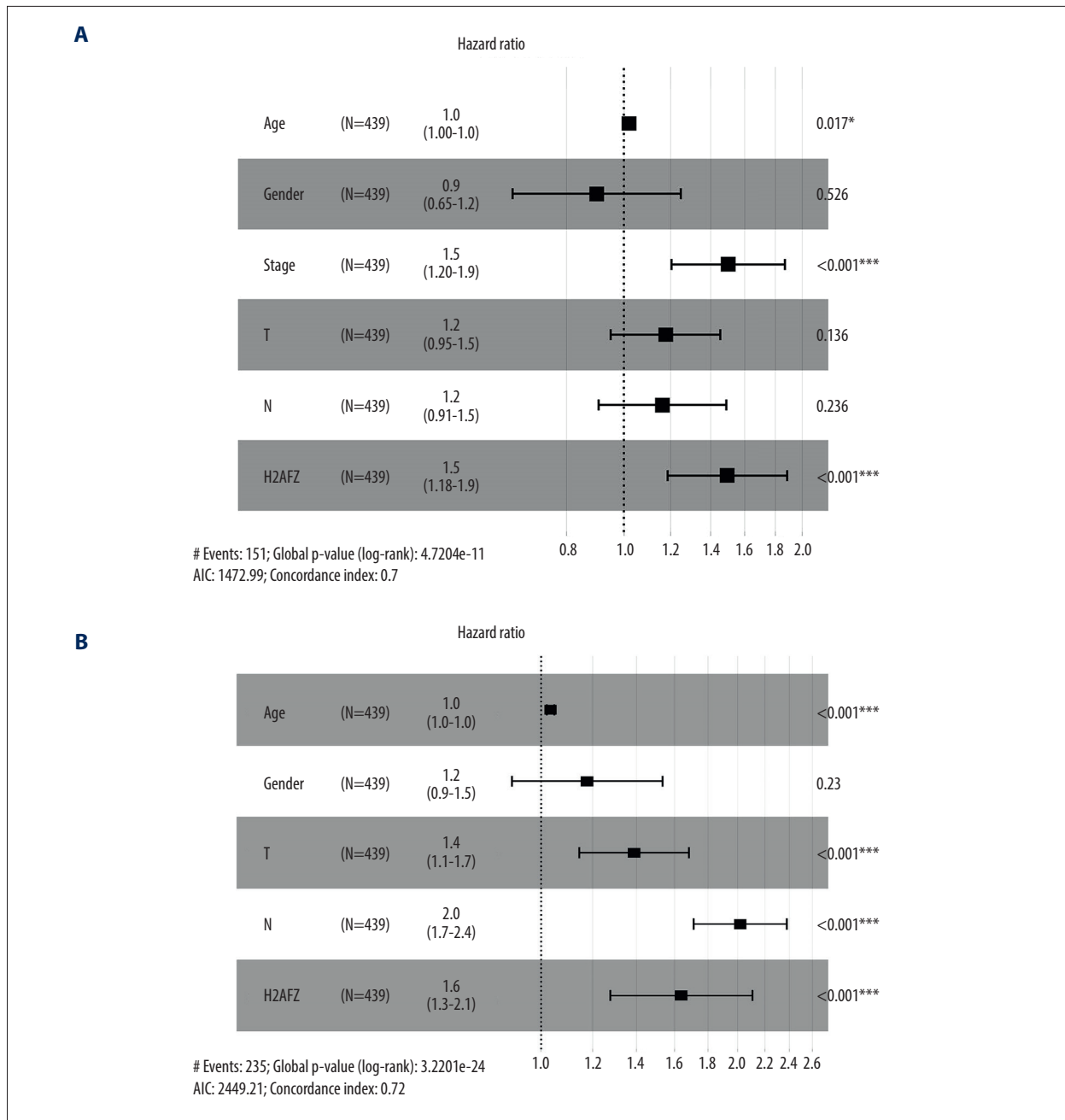
Figure 6. Forest plot for the multivariate Cox proportional hazard regression model in 2 datasets. (**A**) TCGA-LUAD cohort. (**B**) GSE68465. HR – hazard ratio; CI – confidence interval. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. R 4.1.0 software, (R Development Core Team, Vienna).

pathways are closely correlated with the occurrence and progression of various tumors, especially non-small cell lung cancer [34,35]. These analyses provided strong shreds of evidence for us to explore the pathogenesis of LUAD. However, these results may need to be further validated by in vitro experiments.

Recently, it was reported that the tumor microenvironment has a great influence on the occurrence, development, and metastasis

of tumors [36,37]. In accordance with the results of GSEA enrichment, H2AFZ is related to the mTORC1 signaling pathway in LUAD, and the mTORC1 signaling pathway is an integral part in the regulation of the tumor microenvironment [38], which allowed us to study the correlation between H2AFZ and immune cells. According to TCGA data, the difference in proportions of partial immune cells in the high and low H2AFZ expression groups was statistically significant. Immune cells infiltration in LUAD varied

Table 5. Univariate analysis and multivariate analysis of the correlation of H2AFZ expression with OS among LUAD patients in GSE68465.

Parameters	Univariate analysis			Multivariate analysis		
	HR	95% CI	*p* Value	HR	95% CI	*p* Value
Age	1.026	1.013-1.040	**<0.001**	1.034	1.020-1.048	**<0.001**
Gender	1.426	1.101-1.849	**0.007**	1.176	0.903-1.534	0.230
T classification	1.664	1.387-1.988	**<0.001**	1.387	1.144-1.683	**0.001**
N classification	2.012	1.712-2.365	**<0.001**	2.017	1.710-2.378	**<0.001**
H2AFZ	1.652	1.307-2.088	**<0.001**	1.639	1.276-2.105	**<0.001**

Bold values indicate *p*<0.05. HR – hazard ratio; CI – confidence interval.

Table 6. Gene sets enriched in the high H2AFZ expression phenotype.

Gene set name	NES	NOM *p*-value	FDR *q*-value
HALLMARK_MTORC1_SIGNALING	2.66	0.000	0.000
HALLMARK_MYC_TARGETS_V1	2.50	0.000	0.000
HALLMARK_DNA_REPAIR	2.40	0.000	0.000
HALLMARK_G2M_CHECKPOINT	2.39	0.000	0.000
HALLMARK_E2F_TARGETS	2.37	0.000	0.000
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	2.29	0.000	0.000
HALLMARK_MYC_TARGETS_V2	2.24	0.000	0.000
HALLMARK_OXIDATIVE_PHOSPHORYLATION	2.18	0.000	0.001
HALLMARK_PI3K_AKT_MTOR_SIGNALING	2.16	0.000	0.001
HALLMARK_MITOTIC_SPINDLE	1.88	0.012	0.015

NES – normalized enrichment score; NOM – nominal; FDR – false discovery rate.
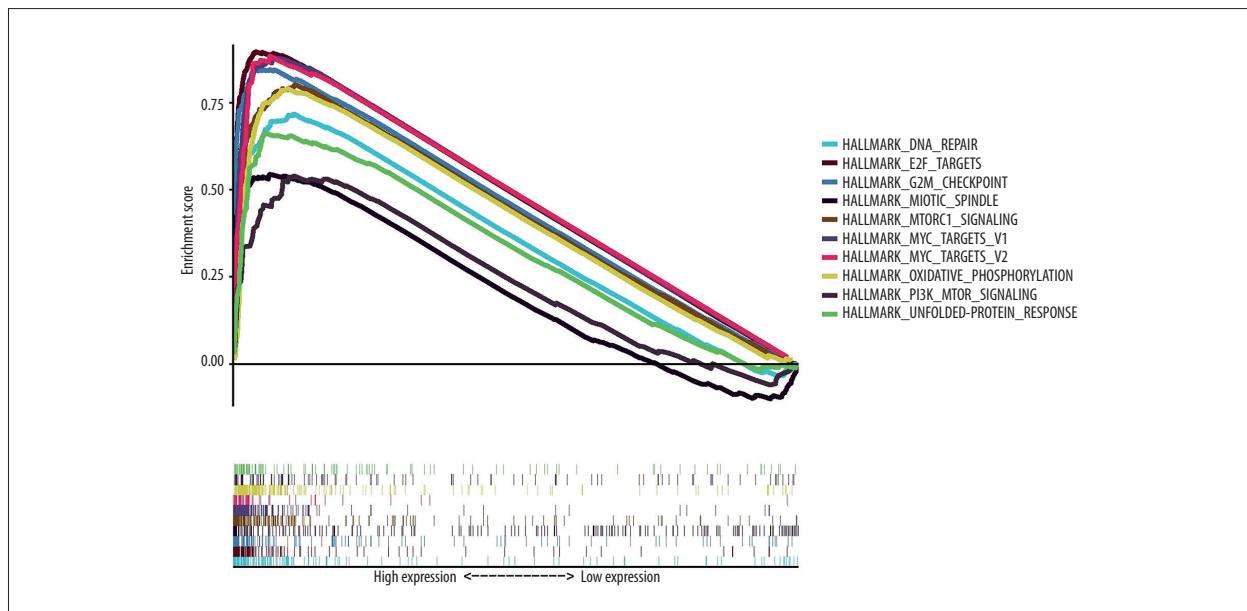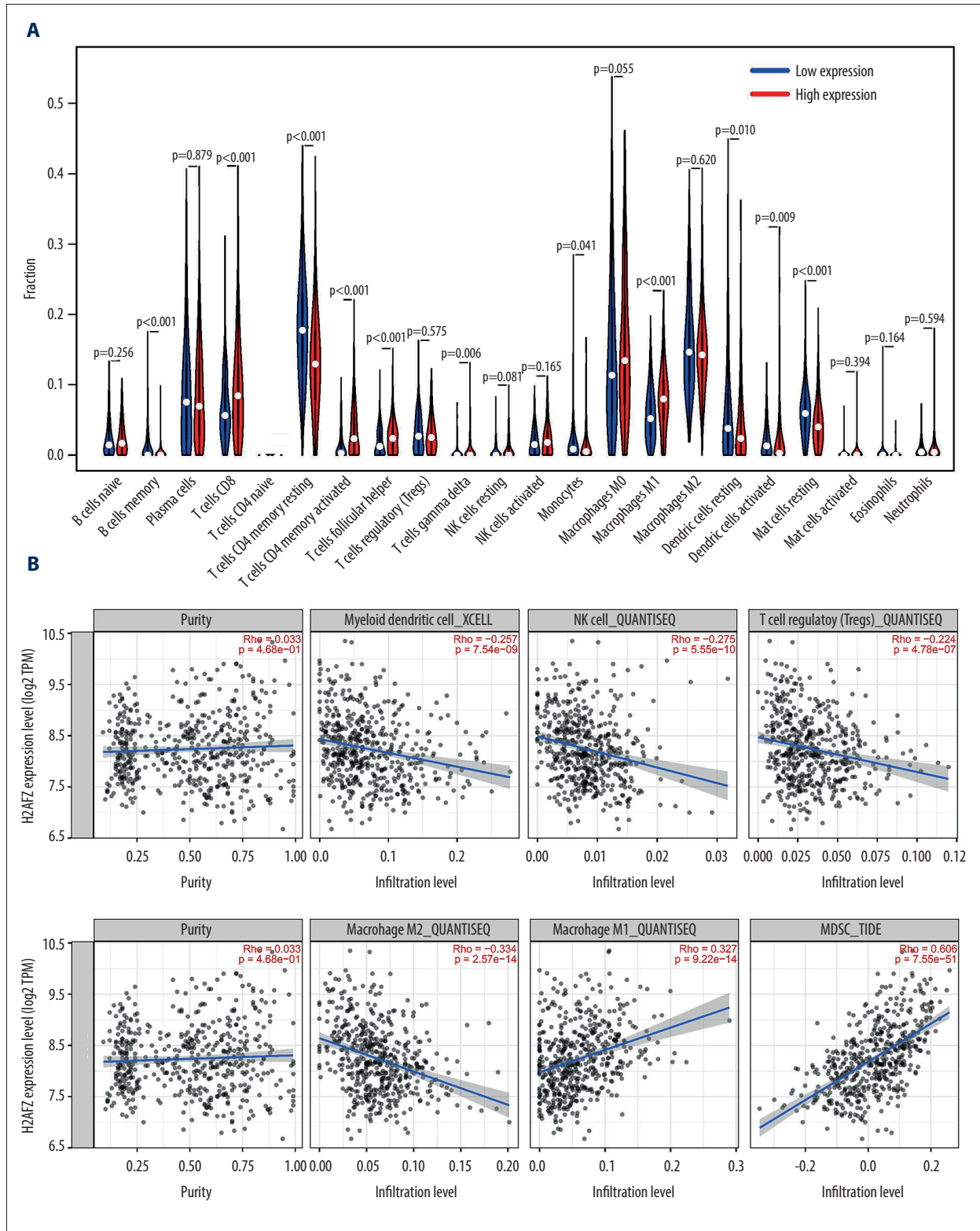

Figure 7. The significantly enriched signaling pathways in the high-expression phenotypes of H2AFZ in LUAD. GSEA 3.0 software (Massachusetts Institute of Technology, Cambridge, MA, USA).
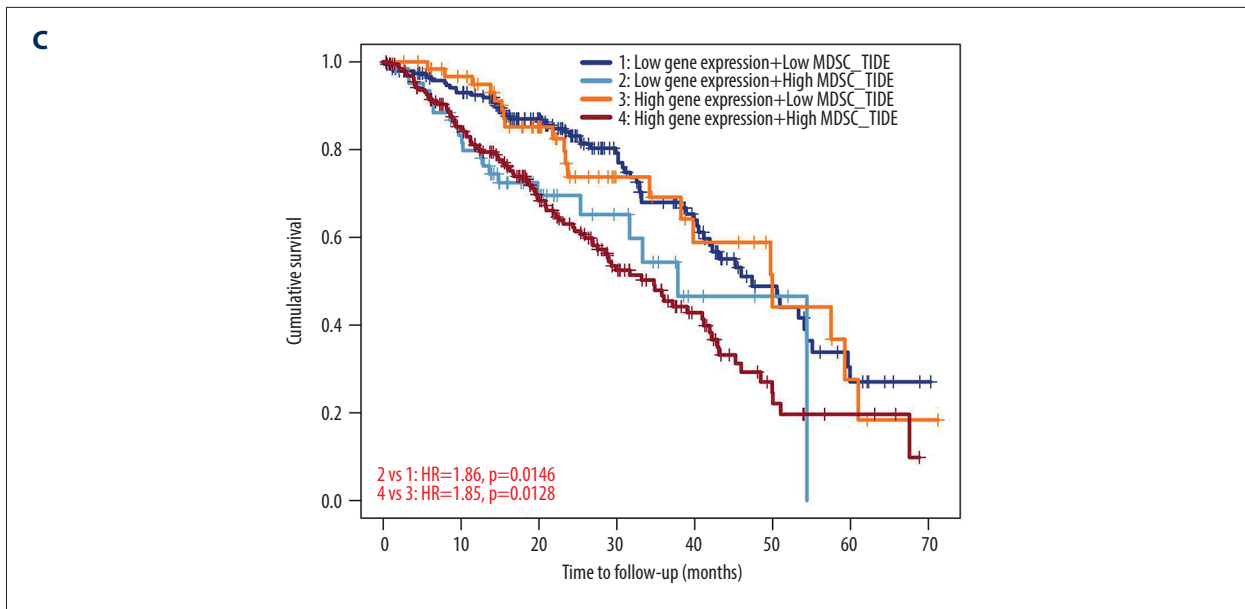
C

Figure 8. Correlation of H2AFZ with TIICs in LUAD. (**A**) Proportions of the 22 tumor-infiltrating immune cell subtypes in high and low H2AFZ expression groups. (**B**) Correlation between H2AFZ expression and immune cells in LUAD. (**C**) The combined impact of H2AFZ expression and MDSC infiltration on the survival of LUAD patients. R 4.1.0 software (R Development Core Team, Vienna).

according to H2AFZ expression, including M1 macrophages, memory B cells, and CD8+ T cells. Therefore, increased H2AFZ might alter the composition of immune cells in the LUAD microenvironment. In addition, the expression of H2AFZ was positively correlated with the infiltration level of MDSC. Being immature myeloid cells, MDSC could promote growth and metastasis of tumors by inducing immunosuppression, and ultimately affect the therapeutic effect of patients [39]. In our research, high expression of H2AFZ combined with the high level of MDSC infiltration might be an important reason for the poor prognosis of LUAD patients. As such, H2AFZ might suppress the function of immune cells by increasing the infiltrating level of MDSC, thus being a potential immune therapy target.

Nonetheless, there are still some limitations in our research. First, the clinical data was incomplete and did not provide essential information, such as tumor grades. Moreover, there is a need for more clinical trials to be conducted to demonstrate the function and the prognostic value of H2AFZ.

Conclusions

Compared with adjacent normal tissues, H2AFZ was observably overexpressed in LUAD tissues, which participated in the occurrence and development of LUAD, leading to a poor prognosis of LUAD patients. Meanwhile, H2AFZ played a crucial role in the regulation and recruitment of LUAD immune-infiltrating cells. Therefore, H2AFZ is speculated to participate in tumor prevention, early diagnosis, and tumor immunotherapy in the future. Thus, H2AFZ may serve as a new biomarker and an important therapeutic target for LUAD.

Declaration of Figures' Authenticity

All figures submitted have been created by the authors, who confirm that the images are original with no duplication and have not been previously published in whole or in part.

References:

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. Cancer J Clin. 2018;68(6):394-424

2. Barlesi F, Mazieres J, Merlio JP, et al. Routine molecular profiling of patients with advanced non-small-cell lung cancer: Results of a 1-year nationwide programme of the French Cooperative Thoracic Intergroup (IFCT). Lancet. 2016;387(10026):1415-26

3. Bade BC, Dela Cruz CS. Lung cancer 2020: Epidemiology, etiology, and prevention. Clin Chest Med. 2020;41(1):1-24

4. Mantovani A, Marchesi F, Malesci A, et al. Tumour-associated macrophages as treatment targets in oncology. Nat Rev Clin Oncol. 2017;14(7):399-416

5. Saito M, Suzuki H, Kono K, et al. Treatment of lung adenocarcinoma by molecular-targeted therapy and immunotherapy. Surg Today. 2018;48(1):1-8

6. Rangasamy D, Greaves I, Tremethick DJ. RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. Nat Struct Mol Biol. 2004;11(7):650-55

7. Dhillon N, Oki M, Szyjka SJ, et al. H2A.Z functions to regulate progression through the cell cycle. Mol Cell Biol. 2006;26(2):489-501

8. Meneghini MD, Wu M, Madhani HD. Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. Cell. 2003;112(5):725-36

9. Yang HD, Kim PJ, Eun JW, et al. Oncogenic potential of histone-variant H2A.Z.1 and its regulatory role in cell cycle and epithelial-mesenchymal transition in liver cancer. Oncotarget. 2016;7(10):11412-23

10. Hua S, Kallen CB, Dhar R, et al. Genomic analysis of estrogen cascade reveals histone variant H2A.Z associated with breast cancer progression. Mol Syst Biol. 2008;4:188

11. Zucchi I, Mento E, Kuznetsov VA, et al. Gene expression profiles of epithelial cells microscopically isolated from a breast-invasive ductal carcinoma and a nodal metastasis. Proc Natl Acad Sci USA. 2004;101(52):18147-52

12. Brunelle M, Nordell Markovits A, Rodrigue S, et al. The histone variant H2A.Z is an important regulator of enhancer activity. Nucleic Acids Res. 2015;43(20):9742-56

13. Svotelis A, Gévry N, Grondin G, Gaudreau L. H2A.Z overexpression promotes cellular proliferation of breast cancer cells. Cell Cycle. 2010;9(2):364-70

14. Baptista T, Graça I, Sousa EJ, et al. Regulation of histone H2A.Z expression is mediated by sirtuin 1 in prostate cancer. Oncotarget. 2013;4(10):1673-85

15. Valdés-Mora F, Gould CM, Colino-Sanguino Y, et al. Acetylated histone variant H2A.Z is involved in the activation of neo-enhancers in prostate cancer. Nat Commun. 2017;8(1):1346

16. Dryhurst D, McMullen B, Fazli L, et al. Histone H2A.Z prepares the prostate specific antigen (PSA) gene for androgen receptor-mediated transcription and is upregulated in a model of prostate cancer progression. Cancer Lett. 2012;315(1):38-47

17. Kim K, Punj V, Choi J, et al. Gene dysregulation by histone variant H2A.Z in bladder cancer. Epigenetics Chromatin. 2013;6(1):34

18. Dunican DS, McWilliam P, Tighe O, et al. Gene expression differences between the microsatellite instability (MIN) and chromosomal instability (CIN) phenotypes in colorectal cancer revealed by high-density cDNA array hybridization. Oncogene. 2002;21(20):3253-57

19. Zlatanova J, Thakar A. H2A.Z: View from the top. Structure. 2008;16(2):166-79

20. Amelung JT, Bührens R, Beshay M, Reymond MA. Key genes in lung cancer translational research: A meta-analysis. Pathobiology. 2010;77(2):53-63

21. Amaral ML, Erikson GA, Shokhirev MN. BART: Bioinformatics array research tool. BMC Bioinformatics. 2018;19(1):296

22. Xiao H, Tong R, Cheng S, et al. BAG3 and HIF-1 α coexpression detected by immunohistochemistry correlated with prognosis in hepatocellular carcinoma after liver transplantation. Biomed Res Int. 2014;2014:516518

23. Vasaikar SV, Straub P, Wang J, Zhang B. LinkedOmics: Analyzing multi-omics data within and across 32 cancer types. Nucleic Acids Res. 2018;46(D1):D956-63

24. Yu G, Wang LG, Han Y, He QY. clusterProfiler: An R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284-87

25. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102(43):15545-50

26. Chen S, Wei Y, Liu H, et al. Analysis of collagen type X alpha 1 (COL10A1) expression and prognostic significance in gastric cancer based on bioinformatics. Bioengineered. 2021;12(1):127-37

27. Vange P, Bruland T, Beisvag V, et al. Genome-wide analysis of the oxyntic proliferative isthmus zone reveals ASPM as a possible gastric stem/progenitor cell marker over-expressed in cancer. J Pathol. 2015;237(4):447-59

28. Pai VC, Hsu CC, Chan TS, et al. ASPM promotes prostate cancer stemness and progression by augmenting Wnt-Dvl-3-β-catenin signaling. Oncogene. 2019;38(8):1340-53

29. Taty-Taty GC, Courilleau C, Quaranta M, et al. H2A.Z depletion impairs proliferation and viability but not DNA double-strand breaks repair in human immortalized and tumoral cell lines. Cell Cycle. 2014;13(3):399-407

30. Vardabasso C, Gaspar-Maia A, Hasson D, et al. Histone variant H2A.Z.2 mediates proliferation and drug sensitivity of malignant melanoma. Mol Cell. 2015;59(1):75-88

31. Horikoshi N, Sato K, Shimada K, et al. Structural polymorphism in the L1 loop regions of human H2A.Z.1 and H2A.Z.2. Acta Crystallogr D Biol Crystallogr. 2013;69(Pt 12):2431-39

32. Schulze A, Oshi M, Endo I, Takabe K. MYC targets scores are associated with cancer aggressiveness and poor survival in ER-positive primary and metastatic breast cancer. Int J Mol Sci. 2020;21(21):8127

33. Park SM, Choi EY, Bae DH, et al. The LncRNA EPEL promotes lung cancer cell proliferation through E2F target activation. Cell Physiol Biochem. 2018;45(3):1270-83

34. Tan AC. Targeting the PI3K/Akt/mTOR pathway in non-small cell lung cancer (NSCLC). Thorac Cancer. 2020;11(3):511-18

35. Madden E, Logue SE, Healy SJ, et al. The role of the unfolded protein response in cancer progression: From oncogenesis to chemoresistance. Biol Cell. 2019;111(1):1-17

36. Garner H, de Visser KE. Immune crosstalk in cancer progression and metastatic spread: A complex conversation. Nat Rev Immunol. 2020;20(8):483-97

37. Joyce JA, Fearon DT. T cell exclusion, immune privilege, and the tumor microenvironment. Science. 2015;348(6230):74-80

38. Kim LC, Cook RS, Chen J. mTORC1 and mTORC2 in cancer and the tumor microenvironment. Oncogene. 2017;36(16):2191-201

39. Kalathil SG, Thanavala Y. Importance of myeloid derived suppressor cells in cancer from a biomarker perspective. Cell Immunol. 2021;361:104280