OLFML2A expression in acute myeloid leukemia, with exploratory findings in acute lymphoblastic leukemia: Prognostic significance and re-analysis of public single-cell transcriptomic data

XUAN LU[1*], JINLIAN LI[2*], XINRU ZHAO[1], WANCHUAN ZHUANG[3] and XINGXING CHAI[3]

[1]Anhui Province Green Food Collaborative Technology Service Center for Rural Revitalization, Hefei Normal University, Hefei, Anhui 230601, P.R. China; [2]Department of Hematology, The People's Hospital of The Qiandongnan Miao and Dong Autonomous Prefecture, Qiandongnan, Guizhou 556099, P.R. China; [3]Department of Hematology, The Second People's Hospital of Lianyungang, Lianyungang, Jiangsu 222000, P.R. China

Received December 8, 2025; Accepted May 18, 2026

DOI: 10.3892/ol.2026.15821

Abstract. This retrospective study analyzed peripheral blood mononuclear cell samples from 34 patients with acute leukemia [26 with acute myeloid leukemia (AML) and 8 with acute lymphoblastic leukemia (ALL)] and 15 healthy controls to investigate olfactomedin-like 2A (OLFML2A) expression and its potential clinical significance. OLFML2A mRNA expression was significantly higher in the patient cohort than in the control group, and expression was higher in AML than in ALL. In the institutional cohort, elevated OLFML2A expression in AML was associated with fusion-gene positivity, chromosomal abnormalities and adverse-risk disease. Findings in ALL were based on a small sample (n=8) and should therefore be regarded as exploratory. Although a possible subtype-dependent difference between AML and ALL was considered, the ALL subgroup was too small to support a definitive directional conclusion. This observation should therefore be regarded as hypothesis-generating and may further support disease-specific OLFML2A biology. Receiver operating characteristic analysis in the local cohort suggested preliminary diagnostic discrimination, particularly for AML, although no independent external diagnostic validation cohort was available. In the combined local cohort, higher OLFML2A expression was associated with shorter progression-free survival, whereas overall survival did not differ significantly. Because the TCGA validation cohort consisted of AML cases only, all public-database survival analyses were restricted to AML and supported OLFML2A as an adverse prognostic factor in that disease. Re-analysis of public single-cell transcriptomic datasets detected OLFML2A expression in natural killer cells, monocytes/macrophages, endothelial cells, hematopoietic progenitor-related populations and malignant cells, providing supportive transcriptomic context rather than direct mechanistic validation. Overall, these findings support OLFML2A as a candidate biomarker in AML, whereas the observations in ALL remain preliminary and require validation in larger cohorts.

Introduction

Acute leukemia (AL) comprises biologically distinct hematologic malignancies, primarily acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL), which differ in cellular origin, molecular drivers, treatment strategies and clinical behavior (1-3). AML arises from myeloid progenitors and is increasingly classified according to integrated morphological, immunophenotypical, cytogenetic and molecular features, whereas ALL originates from lymphoid progenitors and is further divided into B-cell and T-cell lineages (2,3). The global age-standardized incidence and mortality rates of leukemia in 2022 were 5.3 and 3.1 per 100,000 individuals, respectively. In China, a nationwide study estimated 43,275 new AL cases and 27,049 deaths in 2019, with an age-standardized incidence of 2.83 per 100,000 individuals. AML accounts for ~70% of AL cases in China, representing the predominant disease burden (4,5). These epidemiological data further support the clinical need for subtype-specific biomarker evaluation, particularly in AML. Although both diseases may present with increased blast burden in peripheral blood (PB) and bone marrow, biomarkers identified in one subtype cannot

Correspondence to: Dr Xingxing Chai or Dr Wanchuan Zhuang, Department of Hematology, The Second People's Hospital of Lianyungang, 41 Hailian East Road, Lianyungang, Jiangsu 222000, P.R. China
E-mail: C18261303999@163.com
E-mail: wanchuanzhuang@163.com

*Contributed equally

Key words: OLFML2A, acute myeloid leukemia, acute lymphoblastic leukemia, prognosis, single-cell transcriptome

be assumed to apply equally to the other. Accordingly, studies evaluating candidate biomarkers across AL subtypes should clearly distinguish AML-specific findings from exploratory observations in ALL.

The olfactomedin-like 2A (OLFML2A) gene encodes an extracellular matrix-associated glycoprotein that belongs to the olfactomedin family (6-11). Olfactomedin-domain-containing proteins have been implicated in extracellular matrix organization, cell adhesion, tissue development and tumor-associated signaling. Previous studies have suggested that OLFML2A may contribute to tumor progression through pathways related to cell proliferation, migration and Wnt/β-catenin signaling (12-15). Of note, our group previously reported that OLFML2A overexpression was associated with unfavorable prognosis in AML, based mainly on bioinformatic analyses (16). However, PBMC-based expression validation, exploratory comparison with ALL and cell-type-level transcriptomic context had not been comprehensively described. This is relevant as OLFML2A expression measured in PBMCs may reflect not only leukemic blasts but also non-malignant immune-cell and progenitor-related components. The present study was therefore designed not to introduce a completely new AML biomarker, but to extend that earlier observation by adding an independent institutional PB mononuclear cell (PBMC) quantitative (q)PCR cohort, comparing AML with a small exploratory ALL subgroup, and re-analyzing public single-cell datasets to describe the cellular distribution of OLFML2A transcripts. Public single-cell transcriptomic resources, including TISCH2, provide useful platforms for examining gene-expression patterns across annotated cell populations and may help contextualize PBMC-based findings (17).

Given the absence of a comparable The Cancer Genome Atlas (TCGA) cohort for ALL and the limited size of the present ALL subgroup, the main inferential focus of the current study is AML. The ALL-related findings are presented as preliminary observations only. By integrating local PBMC expression data, AML-focused TCGA validation and re-analysis of public single-cell transcriptomic datasets, the present study sought to refine the clinical context of OLFML2A expression and to determine whether OLFML2A warrants further investigation as a candidate biomarker, particularly in AML.

Patients and methods

Study population. This retrospective cohort study enrolled 34 patients with AL who were admitted to the Second People's Hospital of Lianyungang City (Lianyungang, China) between January 2019 and March 2023. AML and ALL were diagnosed on the basis of bone marrow morphology, histochemical examination, immunophenotyping and routine clinical diagnostic work-up in accordance with the relevant Chinese guidelines for adult AML and ALL (18,19), and all patients received guideline-based chemotherapy. PB samples for qPCR were collected at admission, before treatment was initiated for the current disease episode. Fusion-gene status was determined using the institutional leukemia fusion-gene assay panel performed during routine diagnosis, and cytogenetic status was assessed by conventional chromosome analysis with adjunct fluorescence *in situ* hybridization when clinically indicated (20,21). Because PBMC-based expression can be influenced by leukemic burden and cellular composition, the institutional findings were interpreted cautiously. The inclusion criteria were as follows: i) Confirmed AML or ALL; ii) complete clinical data; and iii) signed informed consent from patients or their families. The exclusion criteria were as follows: i) Concomitant myelodysplastic syndromes; ii) other active malignancies; iii) systemic infectious diseases; iv) inability to complete follow-up; v) autoimmune diseases; vi) acute promyelocytic leukemia; and vii) prior antitumor treatment before admission. In addition, 15 healthy individuals who underwent a routine physical examination at the same hospital during the same period were consecutively recruited as controls. The healthy controls had no known hematological malignancy, active infection, autoimmune disease or other active malignancy and were not individually matched to the leukemia cohort. Among the AML cases, 26 patients were included, comprising 18 males (69.2%) and 8 females (30.8%); 5 patients (19.2%) had relapsed disease and the median age was 58.8 years (range, 32-80 years). AML risk stratification was performed according to the National Comprehensive Cancer Network AML Guidelines, version 3.2019 (22). Among the ALL cases, 8 patients were included, comprising 4 males (50.0%) and 4 females (50.0%); 1 patient (12.5%) had relapsed disease and the median age was 37.8 years (range, 14-78 years). Given the small size of the ALL subgroup, clinicopathologic and prognostic analyses in ALL were considered exploratory. The clinicopathological characteristics of the AML and ALL cohorts are summarized in Table I. The age and sex distributions of the AML, ALL and healthy control groups are summarized in Table II. This study was approved by the Ethics Committee of the Second People's Hospital of Lianyungang City (approval no. 2023-064). Written informed consent for the research use of de-identified clinical information and biological samples was obtained at the time of sample collection or hospital admission. For participants younger than the legal age of consent, written informed consent was obtained from a parent or legal guardian.

Detection of OLFML2A mRNA. A total of 3 ml of peripheral venous blood were collected from patients with AML or ALL at admission and from healthy controls during routine physical examination. PBMCs were isolated by density-gradient centrifugation. Total RNA was extracted from PBMCs using the Total RNA Extraction Kit (cat. no. DP419; Tiangen Biotech Co., Ltd.) with a TRIzol-based protocol. according to the manufacturer's instructions. Reverse transcription and real-time qPCR were performed using the Toyobo ReverTra Ace qPCR Master Mix cDNA First-Strand Synthesis Kit (cat. no. FSQ-101; Toyobo Life Science) according to the manufacturer's instructions. Briefly, cDNA was synthesized from total RNA by reverse transcription at 37˚C for 15 min, 50˚C for 5 min and 98˚C for 5 min. qPCR was performed on an ABI 7500 qPCR instrument (Thermo Fisher Scientific, Inc.) using 2X TSINGKE Master qPCR Mix (Tsingke Biotechnology Co., Ltd.) according to the manufacturer's instructions. cDNA was synthesized from total RNA by reverse transcription. The forward primer for OLFML2A was 5'-CACGCCTACGTC CACAAGG-3' and the reverse primer was 5'-TCATAGTGC CTCAACTGCTCA-3'. The forward primer for the internal

Table I. Association of olfactomedin-like 2A expression with clinicopathological features in AML and exploratory ALL.

Clinicopathological characteristic	Number of cases (n=26)	AML group	F/Z	P-value	Number of cases (n=8)	ALL group	F/Z	P-value
Sex			0.50	0.617			0.29	0.773
Male	18	4.92 (8.50)			4	2.11 (3.90)		
Female	8	5.57 (34.33)			4	0.69 (4.89)		
Age, years			0.10	0.918			1.50	0.184
<60	12	4.92 (8.89)			6	1.54 (2.03)		
≥60	14	4.41 (13.02)			2	4.27 (3.06)		
White blood cell count, x10^9/l			1.27	0.205			1.05	0.335
<20	17	7.44 (10.10)			7	5.45 (4.69)		
≥20	9	2.98 (9.33)			1	0.19		
Hemoglobin, g/l			0.39	0.697			0.30	0.773
<75	15	5.00 (8.13)			4	3.69 (3.27)		
≥75	11	3.70 (15.25)			4	0.69 (4.89)		
Platelet count, x10^9/l			1.27	0.205			0.66	0.531
<50	9	4.84 (8.50)			2	1.19 (1.30)		
≥50	17	8.13 (28.92)			6	2.56 (2.71)		
Proportion of primitive cells, %			1.18	0.237			0.02	0.982
<60	12	6.28 (7.06)			3	2.25 (2.74)		
≥60	14	3.53 (13.56)			5	2.21 (2.55)		
AML FAB classification			0.05	0.959			-	-
M1-M2	14	4.98 (7.99)				-	-	
M4-M6	12	4.35 (13.70)				-	-	
ALL immune typing			-	-			0.74	0.488
T cell	-	-			2	3.36 (4.36)		
B cell	-	-			6	1.84 (1.93)		
Prognostic risk stratification			3.45	0.005			8.04	0.007
Good	9	2.87 (7.48)			3	0.86 (1.09)		
Secondary	10	3.53 (4.71)			3	1.13 (0.96)		
Poor quality	7	14.64 (29.63)			2	5.90 (0.75)		
Extramedullary lesions			0.82	0.411			0.03	0.974
Yes	12	2.92 (13.75)			4	2.25 (2.93)		
No	14	6.28 (7.06)			4	2.19 (2.27)		
Fusion gene			2.67	0.008			6.65	0.001
Positive	15	9.57 (12.70)			2	5.19 (0.75)		
Negative	11	2.15 (4.79)			6	0.99 (0.93)		
Chromosome abnormality			2.37	0.018			3.60	0.011
Yes	10	9.83 (9.81)			3	4.64 (2.25)		
No	16	2.92 (6.35)			5	0.77 (0.85)		
Blood transfusion dependence, times			1.64	0.102			0.25	0.808
<27	15	3.70 (7.94)			4	2.46 (2.77)		
≥27	11	8.13 (20.90)			4	1.99 (2.43)		

Table I. Continued.

Clinicopathological characteristic	Number of cases (n=26)	AML group	F/Z	P-value	Number of cases (n=8)	ALL group	F/Z	P-value
Hospitalization time, days			0.26	0.792			1.03	0.343
<130	16	4.27 (13.70)			6	1.71 (2.43)		
≥130	10	5.06 (6.46)			2	3.74 (2.31)		
Daily average hospitalization cost, yuan			0.22	0.031			1.25	0.257
<2,500	14	2.92 (7.24)			5	1.42 (2.24)		
≥2,500	12	9.21 (18.42)			3	3.55 (2.50)		

Values are presented as the median (interquartile range) OLFML2A expression values. AML, acute myeloid leukemia; ALL, acute lymphoblastic leukemia; FAB, French-American-British; PBMCs, peripheral blood mononuclear cells; RT-qPCR, reverse transcription quantitative PCR.

Table II. Age and sex distribution of patients with AML (n=26) or ALL (n=8) and healthy controls (n=15).

Characteristic	AML group	ALL group	Healthy controls	P-value
Sex, n				0.585
Male	18	4	9	
Female	8	4	6	
Mean age ± SD, years	58.8±13.5	37.8±24.0	53.13±14.25	0.008

Values are presented as n or mean ± standard deviation. Sex distribution was compared using the χ^2 test, and age was compared using one-way analysis of variance. AML, acute myeloid leukemia; ALL, acute lymphoblastic leukemia.

control GAPDH was 5'-CATCGCTCAGACACCATGGGG AAGG-3' and the reverse primer was 5'-ATCTCGGGCGGG AGTAG-3'. The total reaction volume was 20 μl and consisted of 2 μl cDNA, 0.8 μl each of the forward and reverse primers, 10 μl 2X TSINGKE Master qPCR Mix (Tsingke Biotechnology Co., Ltd.), 6 μl RNase-free ddH_2O and 0.4 μl 50X ROX Reference Dye II (Low). The thermal cycling conditions were 95˚C for 1 min, followed by 40 cycles of 95˚C for 15 sec and 60˚C for 40 sec. Each qPCR assay was performed in triplicate for each sample. Relative OLFML2A mRNA expression was calculated using the $2^{-\Delta\Delta Cq}$ method (23).

Follow-up and prognostic indicators. Patients were followed through outpatient visits or telephone interviews from the date of diagnosis until June 30, 2023. Survival analyses in the local cohort were considered exploratory because of the small sample size and the inclusion of both AML and ALL. Patients with AL were divided into high- and low-OLFML2A expression groups using the median relative expression value as the cutoff (high expression, ≥ median; low expression, < median). Progression-free survival (PFS) was defined as the interval from the start of treatment to disease recurrence, progression or last follow-up. Overall survival (OS) was defined as the interval from the start of treatment to death from any cause or last follow-up.

Data acquisition. RNA-sequencing (RNA-seq) data from The Cancer Genome Atlas-Acute Myeloid Leukemia (TCGA-LAML) project, processed through the STAR pipeline, were downloaded from the Genomic Data Commons portal (https://portal.gdc.cancer.gov; accessed January 2024) and organized for analysis. Because this public cohort represents AML rather than ALL, all TCGA-based survival analyses, Cox regression analyses, Lasso modeling and clinicopathologic validation analyses were restricted to AML and were not used to validate findings in ALL. Transcripts per million (TPM)-formatted expression data and corresponding clinical data were extracted and transformed using log2(TPM+1) normalization. After samples lacking either RNA-seq or clinical information were excluded, 147 samples were retained for subsequent analyses. The R package survival (version 3.5-5) was used to test proportional-hazards assumptions and fit survival models, and the results were visualized with the survminer (version 0.4.9) and ggplot2 (version 3.4.4) packages (24-26). Survival differences between groups in Kaplan-Meier analyses were compared with the log-rank test. For Kaplan-Meier curves, P-values and hazard ratios (HRs) with 95% confidence intervals (CIs) were obtained from log-rank tests and univariate Cox regression.

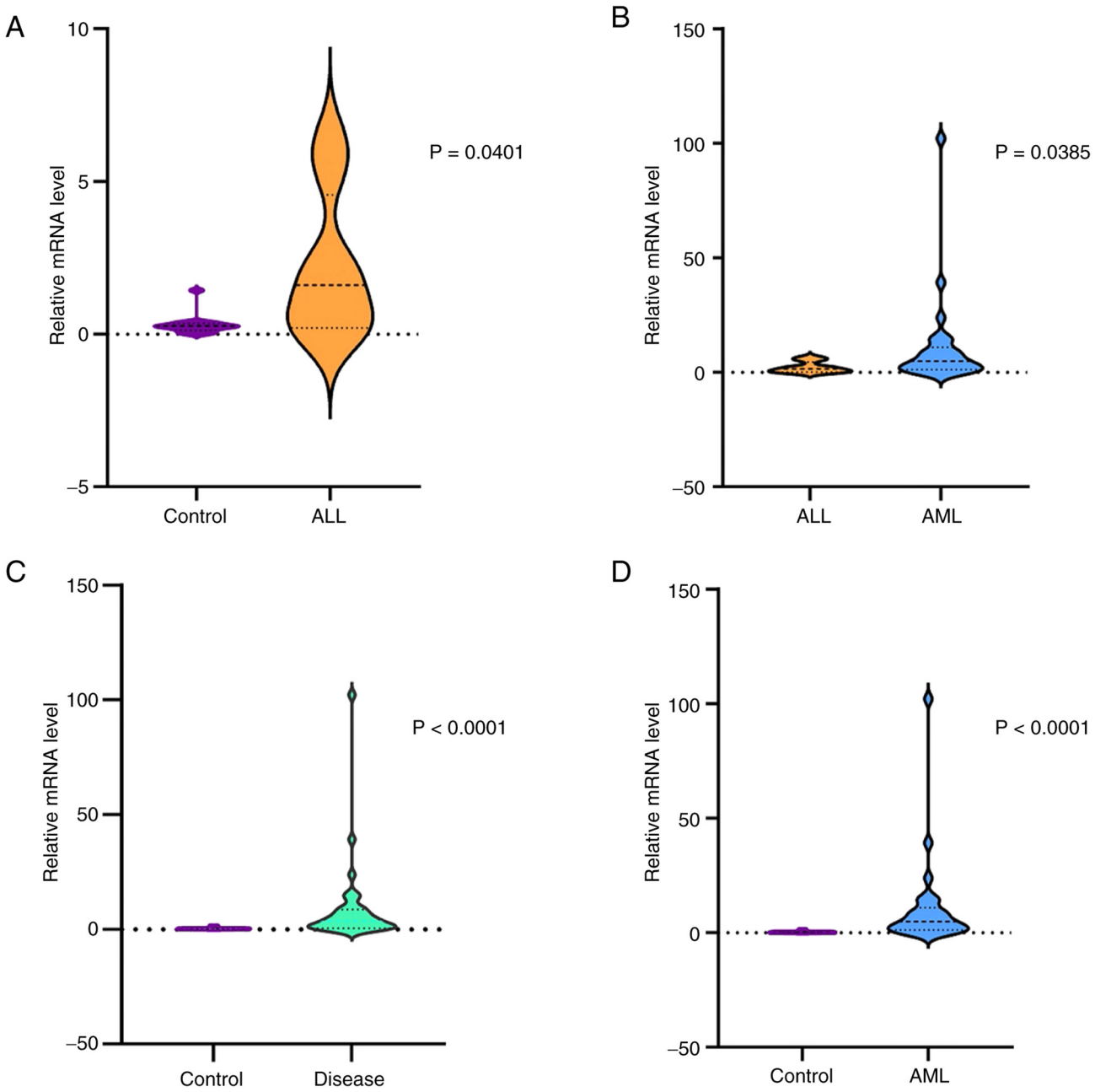

Figure 1. Comparison of OLFML2A expression levels in AML, exploratory ALL and healthy controls. OLFML2A mRNA expression in peripheral blood mononuclear cells was measured by reverse transcription quantitative PCR. Relative expression values are presented as 2-ΔΔCq values, normalized to GAPDH and calibrated to the control sample. Data are shown as violin plots, with dashed lines indicating the median and interquartile range. (A) Comparison between healthy controls and patients with ALL. (B) Comparison patients with ALL and AML. (C) Comparison between healthy controls and all AL cases. (D) Comparison between healthy controls and patients with AML. AL, acute leukemia; AML, acute myeloid leukemia; ALL, acute lymphoblastic leukemia; OLFML2A, olfactomedin-like 2A.

Single-cell data analysis. To provide transcriptomic context rather than direct sequencing data from the study cohort, publicly available single-cell datasets with detailed annotations from the Tumor Immune Single-cell Hub 2 (TISCH https://tisch.compbio.cn; accessed January 2024) database were re-analyzed. Data processing and downstream analyses were performed in R using the MAESTRO (version 1.5.1) and Seurat (version 4.3.0) packages (27,28). The preprocessing workflow included quality control, SCTransform normalization, a variance-stabilizing normalization method implemented in Seurat, followed by feature selection and scaling (29). Cells were then clustered using the dimensionality-reduction and clustering functions implemented in Seurat. Finally, Seurat visualization tools were used to generate plots showing cell relationships and OLFML2A expression patterns. These analyses were intended to provide supportive transcriptomic evidence and did not constitute direct mechanistic validation in patient-derived samples.

Statistical analysis. Bioinformatic analyses were performed with R software (version 4.0.3; R Foundation for Statistical Computing). Statistical analyses of clinical data were

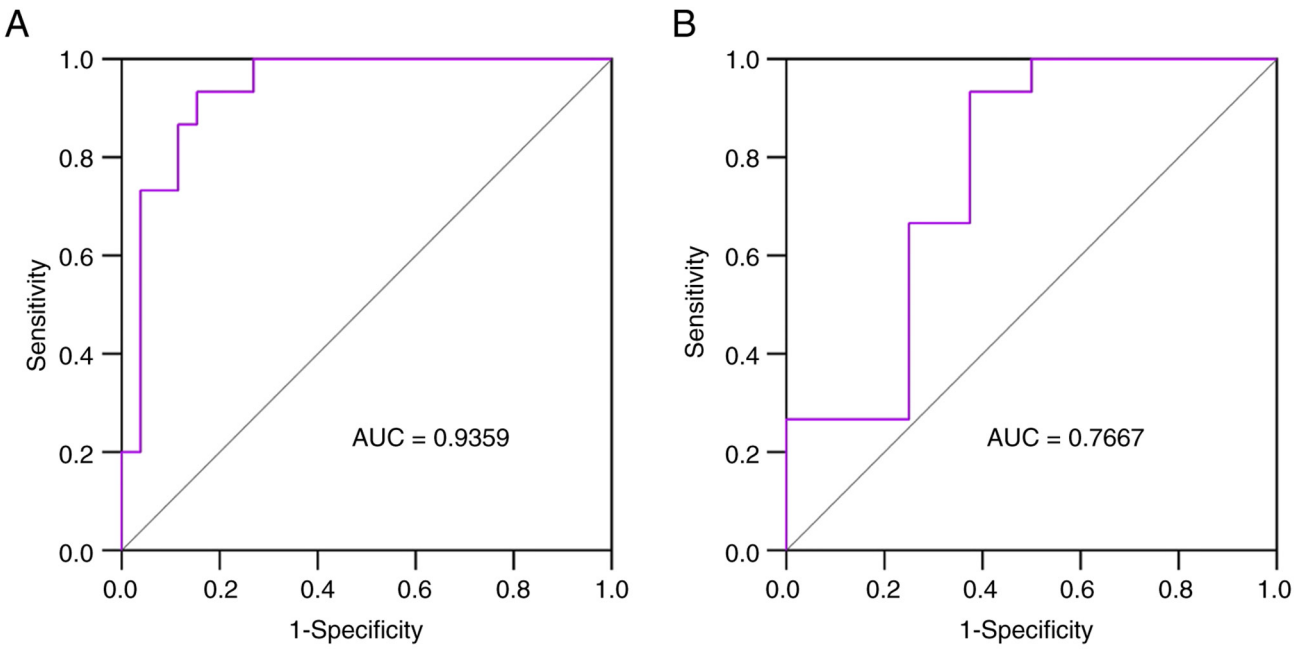

Figure 2. ROC curves for the diagnosis of AML and ALL based on relative OLFML2A mRNA expression levels in the peripheral blood. ROC curves for the local cohort. ROC curves showing the diagnostic performance of relative olfactomedin-like 2A mRNA expression in the local institutional cohort. (A) ROC curve for distinguishing patients with AML from healthy controls. (B) ROC curve for distinguishing patients with ALL from healthy controls. ROC, receiver operating characteristic; AUC, area under the ROC curve.

conducted using SPSS 28.0 (IBM Corp.), and GraphPad Prism Version 9 (Dotmatics) was used for visualization. Continuous variables with skewed distributions are presented as the median (interquartile range) and were compared using the Mann-Whitney U-test. Ranked variables were compared using the Wilcoxon rank-sum test. PFS and OS curves were generated using the Kaplan-Meier method and compared with the log-rank test. Bone marrow smear cytology or bone marrow biopsy pathology served as the reference standard, and receiver operating characteristic (ROC) curve analysis was used to assess the diagnostic performance of relative OLFML2A mRNA expression for AML and ALL. Given the small and heterogeneous local cohort, especially the ALL subgroup (n=8), ALL subgroup analyses were considered exploratory and descriptive. ROC analyses were performed only within the local cohort and should therefore be interpreted as preliminary because no external diagnostic validation cohort or disease-control cohort was available. A two-sided $P<0.05$ was considered to indicate statistical significance.

Results

Comparison of OLFML2A expression levels in AML, exploratory ALL and healthy controls. In the subgroup comparisons, OLFML2A mRNA expression was higher in ALL than in healthy controls ($Z=2.07$, $P=0.040$; Fig. 1A), and the expression in AML was significantly higher than that in ALL ($Z=2.07$, $P=0.039$; Fig. 1B). When all AL cases were compared with healthy controls, the median relative OLFML2A mRNA expression level was significantly higher in the leukemia group ($Z=4.38$, $P<0.001$; Fig. 1C). OLFML2A mRNA expression was also significantly higher in AML than in healthy controls ($Z=4.66$, $P<0.001$; Fig. 1D). Among these comparisons, the AML vs. healthy control result was considered the most reliable as it was based on a larger AML subgroup and showed the strongest statistical signal. As PBMC-based expression is influenced by sample composition and the ALL subgroup was small, the ALL-related findings should be regarded as preliminary.

Diagnostic performance of OLFML2A expression in the local cohort. ROC curve analysis in the local institutional cohort showed that the area under the curve (AUC) for diagnosing AML on the basis of relative OLFML2A mRNA expression was 0.936 (95% CI, 0.8615-1.000), with an optimal cutoff value of 0.780. For ALL, the AUC was 0.767 (95% CI, 0.5352-0.9982), with an optimal cutoff value of 0.558 (Fig. 2). These estimates suggest preliminary diagnostic discrimination, particularly for AML; however, they should be interpreted with caution because the analysis was derived from the same small cohort, lacked external validation and compared patients only with healthy controls rather than with other hematologic disorders encountered in routine practice.

Exploratory correlation between relative OLFML2A expression and prognosis in the local cohort. The median follow-up duration for the 34 patients was 15 months (range, 4-67 months), and no patients were lost to follow-up. During follow-up, 17 patients experienced disease progression and 12 died. Kaplan-Meier analysis showed that the 3-year PFS rate in the overall cohort was 26.7% (Fig. 3A), whereas the 3-year OS rate was 35.8% (Fig. 3B). Using the median relative OLFML2A mRNA expression level as the cutoff, the cohort was divided into high-expression and low-expression groups, with 17 patients in each group. Kaplan-Meier analysis showed

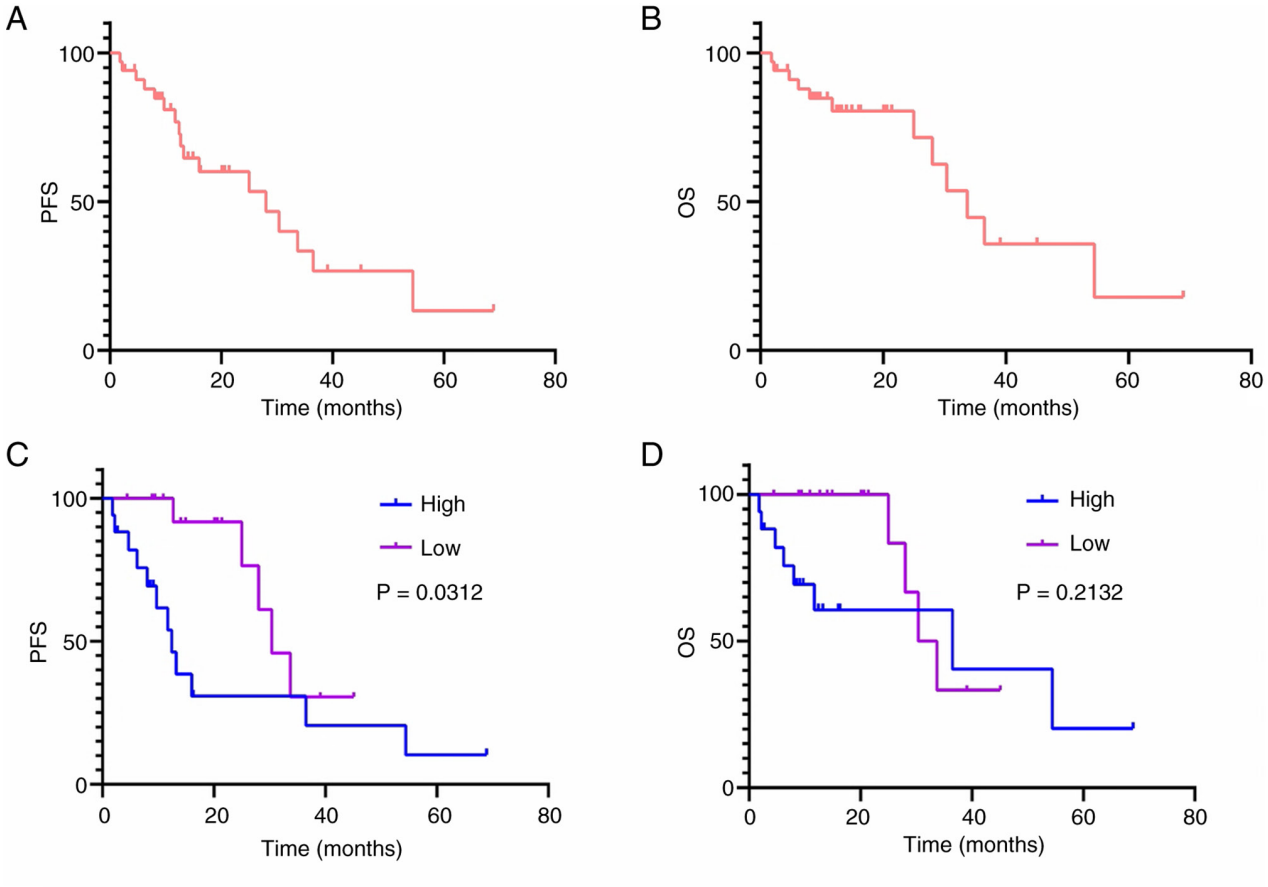

Figure 3. Kaplan-Meier survival curves for the combined local AL cohort according to OLFML2A expression. The local cohort included both patients with AML and patients with ALL. (A) PFS curve for the entire local AL cohort. (B) OS curve for the entire local AL cohort. (C) Comparison of PFS between patients with high and low OLFML2A expression. (D) Comparison of OS between patients with high and low OLFML2A expression. OLFML2A expression in PBMCs was measured by reverse transcription quantitative PCR and evaluation with the $2^{-\Delta\Delta Cq}$ method; the median relative expression level was 3.53 and was used as the cutoff. PFS, progression-free survival; OS, overall survival; AL, acute leukemia; AML, acute myeloid leukemia; ALL, acute lymphoblastic leukemia; OLFML2A, olfactomedin-like 2A.

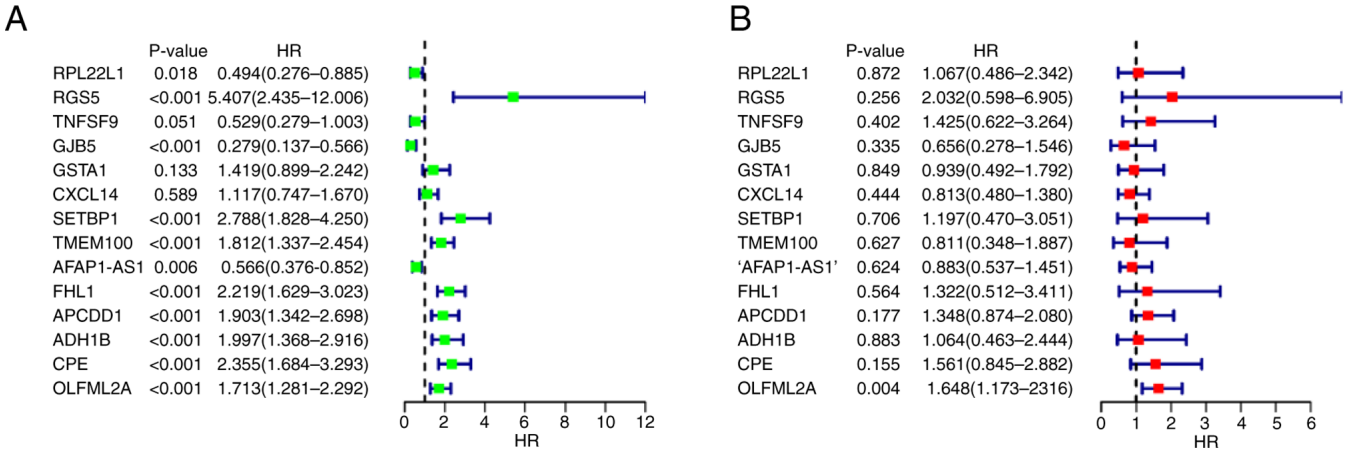

Figure 4. Independent prognostic analysis of related genes. (A) Univariate Cox regression analysis. (B) Multivariate Cox regression analysis. Each line segment represents an HR and its 95% confidence interval; longer segments indicate wider confidence intervals. HR, hazard ratio.

significantly shorter PFS in the high-expression group than in the low-expression group (χ^2=4.64, P=0.031, HR=2.840; Fig. 3C). By contrast, OS did not differ significantly between the two groups (χ^2=1.55, P=0.213, HR=2.023; Fig. 3D). Because this analysis combined biologically distinct diseases and included only 8 ALL cases, these findings should be viewed as exploratory and hypothesis-generating rather than definitive.

Independent prognostic analysis of OLFML2A in the TCGA AML cohort using Cox regression. Using the TCGA-LAML

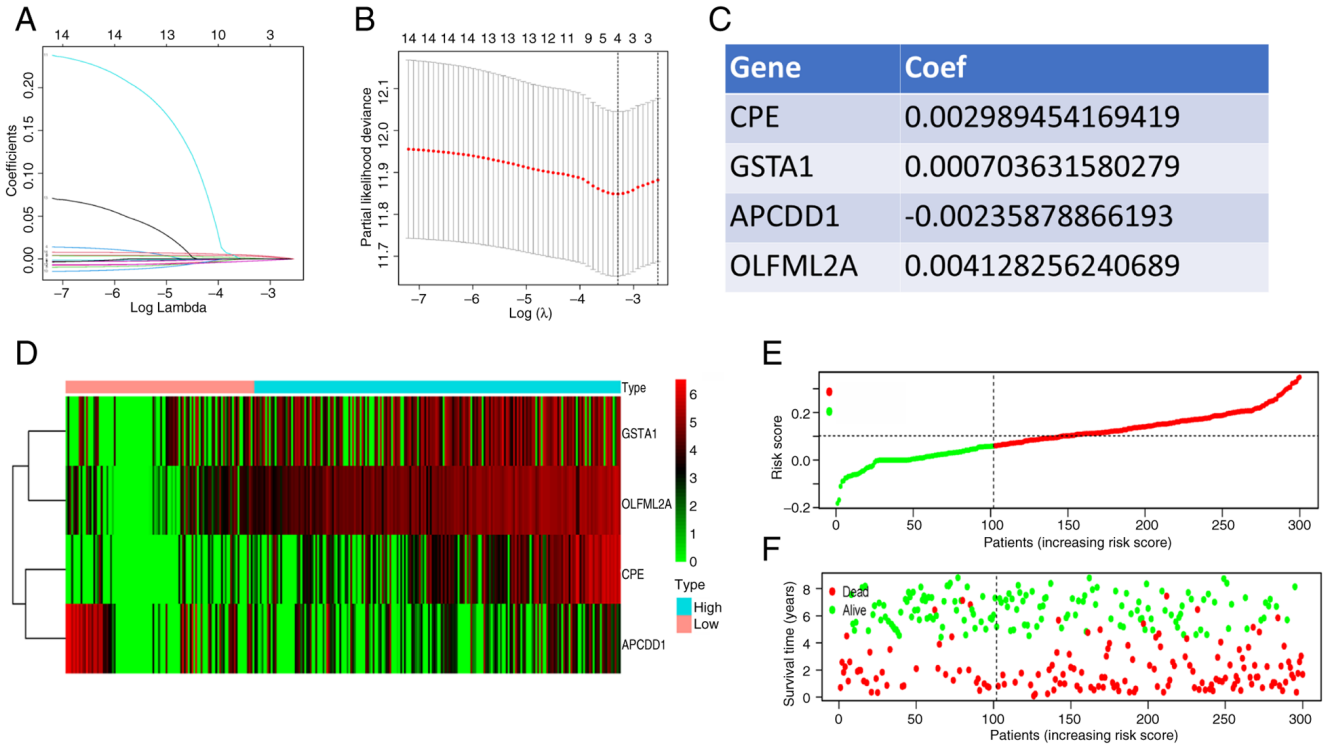

Figure 5. Screening of AML-related prognostic genes in The Cancer Genome Atlas acute myeloid leukemia cohort using Lasso regression. (A) Trajectories of Lasso coefficients. (B) Cross-validation plot for selection of the optimal log(λ) value. (C) Four genes retained by the LASSO Cox regression model; the coefficients shown are the regression coefficients used to calculate the risk score. (D) Heatmap of the expression levels of the four genes. (E) Distribution of risk scores in the high-risk and low-risk groups. (F) Distribution of survival status according to increasing risk score. AML, acute myeloid leukemia; LASSO, least absolute shrinkage and selection operator; Coef, coefficient.

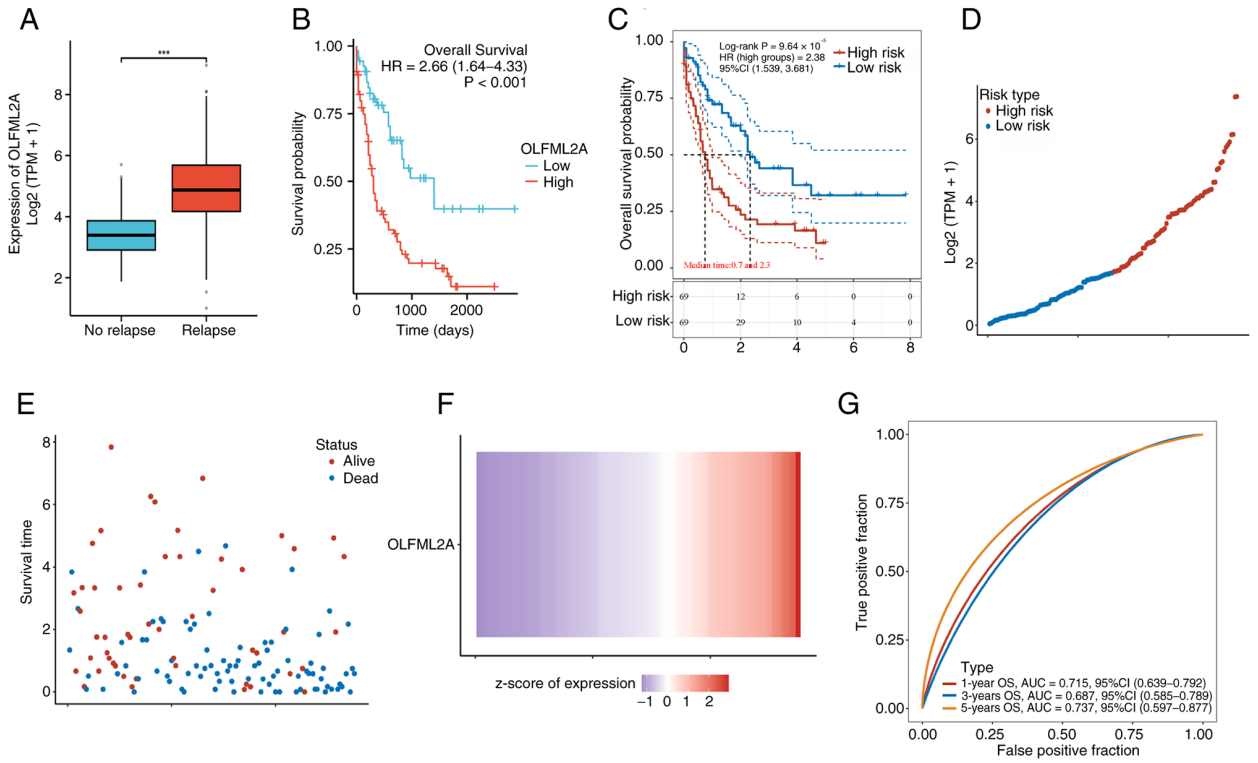

Figure 6. Relationship between OLFML2A expression and prognosis in The Cancer Genome Atlas AML cohort. (A) Differential OLFML2A expression between relapsed and non-relapsed AML cases. ***P<0.0001. (B) Kaplan-Meier survival analysis of OS according to high and low OLFML2A expression. (C) Kaplan-Meier analysis of OS according to high-risk and low-risk groups defined by the prognostic risk model. (D) Distribution of risk type according to increasing risk score. (E) Distribution of survival time and survival status according to increasing risk score. (F) Heatmap illustrating the relationship between OLFML2A expression and risk score. (G) Time-dependent ROC curves showing the predictive performance of the risk model for 1-, 3- and 5-year OS. ROC, receiver operating characteristic; OS, overall survival; AML, acute myeloid leukemia; AUC, area under the ROC curve; HR, hazard ratio; OLFML2A, olfactomedin-like 2A.

Table III. Clinicopathological characteristics of the low- and high-OLFML2A expression groups in The Cancer Genome Atlas acute myeloid leukemia cohort.

Characteristic	Low OLFML2A expression (n=74)	High OLFML2A expression (n=73)	P-value
Sex			0.795
Female	33 (22.4)	31 (21.1)	
Male	41 (27.9)	42 (28.6)	
Age, years			0.082
≥60	48 (32.7)	37 (25.2)	
>60	26 (17.7)	36 (24.5)	
WBC count, x10^9/l			0.217
≥20	42 (28.6)	34 (23.1)	
>20	32 (21.8)	39 (26.5)	
BM blasts, %			0.281
≥20	26 (17.7)	32 (21.8)	
>20	48 (32.7)	41 (27.9)	
PB blasts, %			0.280
≥70	38 (25.9)	31 (21.1)	
>70	36 (24.5)	42 (28.6)	
Cytogenetic risk			0.040
Favorable	20 (13.6)	9 (6.1)	
Intermediate/normal	40 (27.2)	42 (28.6)	
Poor	14 (9.5)	22 (15)	
OS status			<0.001
Alive	37 (25.2)	15 (10.2)	
Dead	37 (25.2)	58 (39.5)	

Values are expressed as n (%). WBC, white blood cell; BM, bone marrow; PB, peripheral blood; OS, overall survival; OLFML2A, olfactomedin-like 2A.

AML cohort, clinical variables and gene-expression profiles were analyzed to evaluate the prognostic relevance of OLFML2A. In the univariate Cox regression analysis, OLFML2A expression was significantly associated with prognosis, suggesting that OLFML2A may represent an important prognostic gene in AML (Fig. 4A). Multivariate Cox regression analysis further supported OLFML2A as an independent prognostic factor within this AML cohort (Fig. 4B). These public-database findings apply to AML and should not be generalized to ALL.

Screening of AML-related prognostic genes using Lasso regression. To further identify independent prognostic genes, Lasso regression analysis was performed in the TCGA-LAML AML cohort. Based on the coefficient trajectories and the corresponding Lasso model selection results, four genes were retained as independent prognostic genes in AML (Fig. 5A and B). CPE, GSTA1 and OLFML2A were associated with adverse disease biology, whereas APCDD appeared to be protective (Fig. 5C). A prognostic model based on these four genes showed that OLFML2A was more highly expressed in the high-risk group than in the low-risk group (Fig. 5D). As the risk score increased, both disease risk and the number of deaths also increased (Fig. 5E and F).

Relationship between OLFML2A expression and prognosis in the TCGA AML cohort. In the TCGA-LAML AML cohort, OLFML2A expression was higher in relapsed patients than in non-relapsed patients (Fig. 6A). Additional analyses showed that the OLFML2A high-expression group had worse OS than the low-expression group (Fig. 6B). Kaplan-Meier analysis based on the prognostic risk model also showed poorer OS in the high-risk group than in the low-risk group (Fig. 6C). When cases were ordered according to the model risk score, the distribution of risk type and survival events suggested an increase in high-risk classification and deaths, together with a shortening of survival time, as the risk score increased (Fig. 6C and D). Consistent with this pattern, the gene-expression heatmap supported OLFML2A as a risk factor within the TCGA-LAML AML cohort (Fig. 6E), and the time-dependent ROC curve indicated acceptable predictive performance for 1-, 3- and 5-year OS (Fig. 6F). Comparisons of clinical characteristics between the high- and low-expression groups in the TCGA cohort further showed significant differences in cytogenetic risk and OS status (Table III).

OLFML2A expression patterns across cell types in public single-cell datasets. Re-analysis of public single-cell

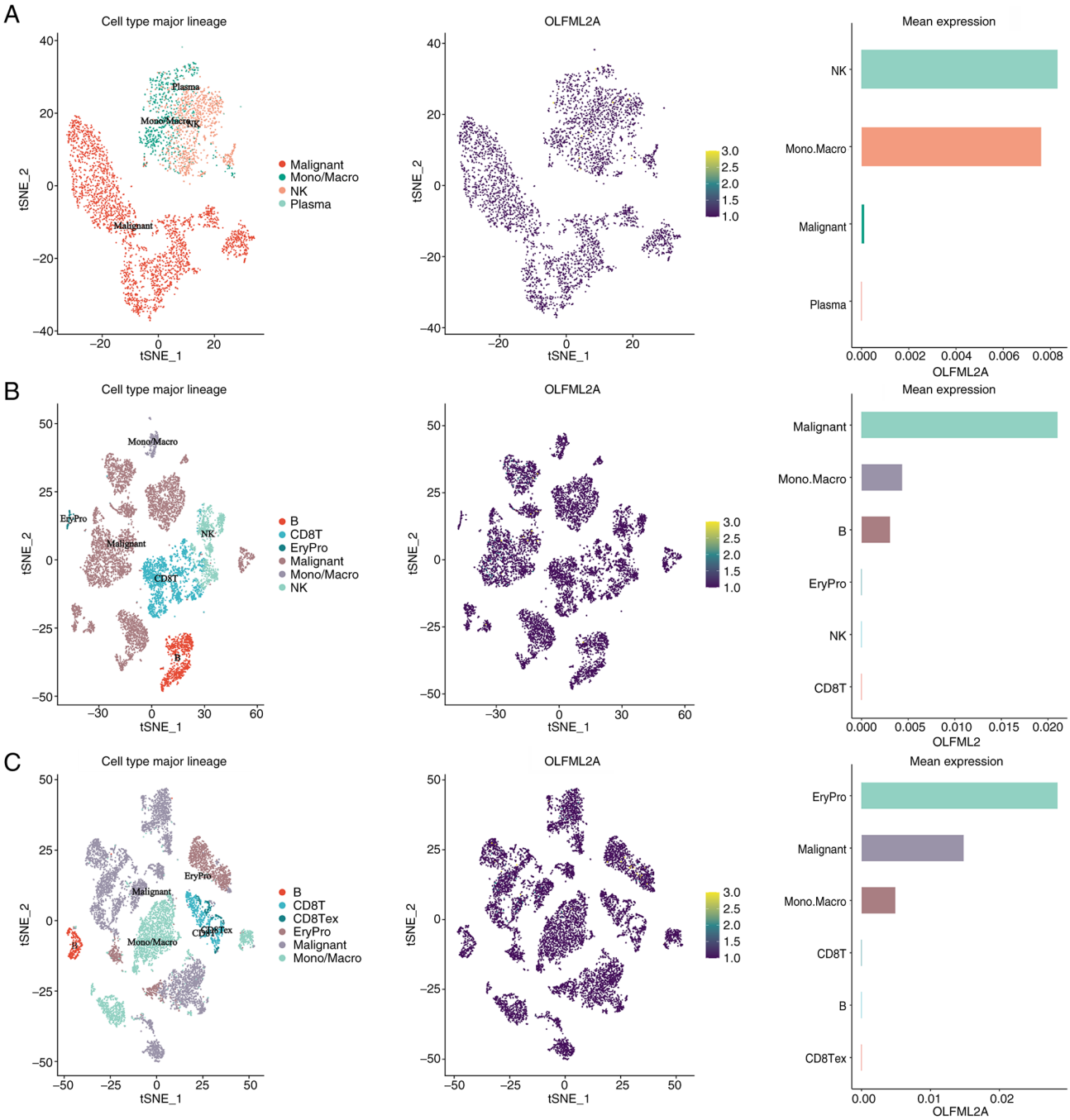

Figure 7. (A) Representative t-SNE plots showing cell-type major-lineage annotation, OLFML2A expression distribution and mean OLFML2A expression across malignant cells, monocytes/macrophages, NK cells and plasma cells. (B) Representative t-SNE plots showing cell-type major-lineage annotation, OLFML2A expression distribution and mean OLFML2A expression across B cells, CD8T cells, erythroid progenitor cells, malignant cells, monocytes/macrophages and NK cells. (C) Representative t-SNE plots showing cell-type major-lineage annotation, OLFML2A expression distribution and mean OLFML2A expression across B cells, CD8T cells, exhausted CD8T cells, erythroid progenitor cells, malignant cells and monocytes/macrophages. In each row, the left panel shows major cell-lineage clustering, the middle panel shows OLFML2A expression intensity and the right panel shows mean OLFML2A expression by annotated cell type. The tSNE_1 and tSNE_2 axes represent unitless t-SNE embedding coordinates. Different colors indicate distinct cell types or relative expression levels, as indicated. B, B cell; CD8T, $CD8^+$ T cell; CD8Tex, exhausted $CD8^+$ T cell; EryPro, erythroid progenitor cell; Mono/Macro, monocyte/macrophage; NK, natural killer cell; OLFML2A, olfactomedin-like 2A; t-SNE, t-distributed stochastic neighbor embedding.

datasets related to AL showed that OLFML2A transcripts were detectable across multiple cell types (Figs. 7 and 8). Relatively higher expression was observed in natural killer cells, monocytes/macrophages, malignant cells, endothelial cells, hematopoietic stem/progenitor-related populations and erythroid progenitor cells (Figs. 7A-C, and 8A and B). In addition, correlation analysis showed that OLFML2A expression was positively correlated with several immune-cell-related populations, including Th1 cells, NK CD56dim cells, TFH cells and macrophages, whereas a negative correlation was observed for mast cells (Fig. 8C). These findings provide supportive transcriptomic context for the cellular distribution of OLFML2A, but they do not constitute direct mechanistic validation in the present study cohort.

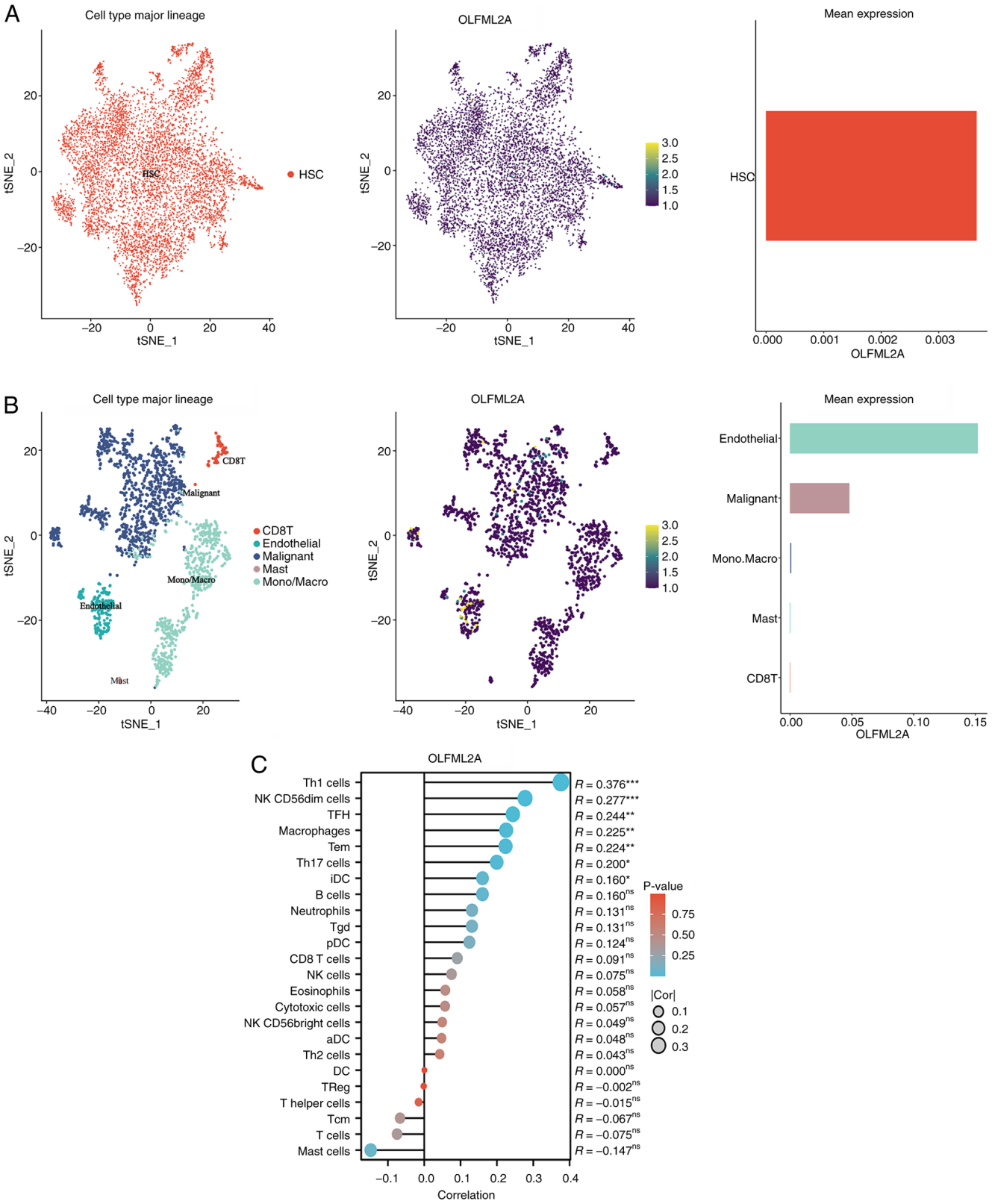

Figure 8. Additional public single-cell transcriptomic re-analysis and correlation analysis of OLFML2A expression in acute leukemia-related datasets. (A) Representative t-SNE plots from a public dataset showing hematopoietic stem cell annotation, OLFML2A expression distribution and mean OLFML2A expression in hematopoietic stem cells. (B) Representative t-SNE plots from another public dataset showing cell-type major-lineage annotation, OLFML2A expression distribution and mean OLFML2A expression across CD8T cells, endothelial cells, malignant cells, mast cells and monocytes/macrophages. (C) Correlation analysis showing associations between OLFML2A expression and immune-cell-related populations. In panels A and B, the left panel shows major cell-lineage clustering, the middle panel shows OLFML2A expression intensity and the right panel shows mean OLFML2A expression by annotated cell type. The tSNE_1 and tSNE_2 axes represent unitless t-SNE embedding coordinates. In panel C, the x-axis represents the correlation coefficient. Dot size indicates the absolute correlation coefficient, and color indicates the P-value. *P<0.05, **P<0.01 and ***P<0.001. aDC, activated dendritic cell; CD8T, CD8+ T cell; Cor, correlation coefficient; DC, dendritic cell; HSC, hematopoietic stem cell; iDC, immature dendritic cell; Mono/Macro, monocyte/macrophage; NK, natural killer cell; pDC, plasmacytoid dendritic cell; Tem, effector memory T cell; TFH, follicular helper T cell; Tgd, γδ T cell; Th, T helper cell; TReg, regulatory T cell; OLFML2A, olfactomedin-like 2A; t-SNE, t-distributed stochastic neighbor embedding..

Discussion

The present study was designed to extend, rather than duplicate, previous work by our group showing that OLFML2A overexpression is associated with unfavorable prognosis in AML (16). That previous study, which was based mainly on public transcriptomic datasets and bioinformatic analyses, suggested that OLFML2A may have diagnostic, prognostic and immune-related relevance in AML. The current institutional PBMC-qPCR cohort provides additional clinical-sample-based support for this AML-focused observation, although the small sample size and PBMC-based design require cautious interpretation. The principal added value of the current manuscript lies in three aspects: The inclusion of an independent institutional PBMC qPCR cohort, the reporting of preliminary observations in a small ALL subgroup and the re-analysis of public single-cell datasets to describe the cellular distribution of OLFML2A transcripts. This manuscript is therefore intentionally positioned as a primarily AML-focused study with exploratory ALL findings, rather than as evidence for a universal biomarker across all forms of AL.

In the local cohort, OLFML2A expression was higher in both AML and exploratory ALL samples than in healthy controls, with the most prominent signal observed in AML. However, AML and ALL are biologically distinct diseases and should not be interpreted as a single entity for biomarker inference (1-3,18,19). This distinction is also supported the data of the present study: Higher OLFML2A expression was associated with adverse clinicopathologic features in AML, whereas the limited ALL observations were insufficient to support a similar interpretation. As the TCGA validation dataset was restricted to AML, the strongest evidence from the present study supports a role for OLFML2A in AML rather than in AL broadly defined (30). In the broader cancer context, OLFML2A has been implicated in malignant phenotypes in several solid tumors (12-15). For example, OLFML2A downregulation was reported to inhibit glioma proliferation through suppression of Wnt/β-catenin signaling, and studies in triple-negative breast cancer suggested that OLFML2A may be involved in tumor-cell proliferation, migration, invasion and treatment-related biological responses (12-14). These findings are generally consistent with the concept that OLFML2A may participate in cancer-related processes, but they also emphasize that its biological significance may be context-dependent and should not be extrapolated directly across tumor types or leukemia subtypes.

The ALL subgroup in the institutional series of the present study was small (n=8) and included few outcome events. Accordingly, any clinicopathologic associations or prognostic inferences in ALL should be regarded as preliminary. Rather than supporting OLFML2A as a universal AL biomarker, the current data may point to subtype-specific biology that requires validation in larger, well-defined cohorts. In the present institutional cohort, the subtype-specific patterns were further examined and it was found that OLFML2A showed a more consistent adverse association in AML, whereas the ALL subgroup was too small to support a reliable directional conclusion; therefore, the ALL findings should not be interpreted as evidence of a favorable association, but rather as hypothesis-generating observations that further support the possibility that OLFML2A biology may be disease-specific. Similarly, the diagnostic performance reported here is likely optimistic because ROC analysis was performed in the same small institutional cohort and used healthy controls instead of clinically relevant disease controls.

Another important limitation is that OLFML2A expression was measured in PBMCs rather than in purified leukemic cells. Although the samples were collected at admission before treatment for the current disease episode, PBMC-based measurements can still be influenced by blast percentage, leukocyte subset composition and overall disease burden. Because the cohort size was limited, it was not possible to perform a robust adjustment for cellular composition. The PBMC findings should therefore be interpreted as clinically associated expression signals rather than as proof of leukemia cell-intrinsic upregulation.

The prognostic results also require cautious interpretation. In the local cohort, higher OLFML2A expression was associated with shorter PFS, whereas OS did not differ significantly between expression groups. These findings do not justify describing OLFML2A as a strong prognostic marker for all AL. The AML-focused TCGA analyses provide additional support for an adverse association in AML, but these analyses remain observational and cannot establish causality. Likewise, the public single-cell component provides supportive localization of OLFML2A transcripts across several cell populations, but it should not be interpreted as a direct mechanistic validation.

In summary, the current findings support OLFML2A as a candidate biomarker worthy of further investigation in AML. Larger multicenter studies, subtype-specific validation cohorts, adjustment for disease burden and cellular composition, and functional experiments will be necessary to determine whether OLFML2A has true diagnostic, prognostic or therapeutic relevance.

Several limitations should be emphasized. First, this was a single-center retrospective study with a small and heterogeneous cohort, especially in the ALL subgroup (n=8), which limited statistical power and precluded robust subtype-specific prognostic inference. Second, in the present cohort, OLFML2A expression was measured in PBMCs rather than in purified leukemic cells; therefore, variation in blast percentage, leukocyte subset composition and disease burden may have influenced the measured signal. Third, the TCGA validation dataset corresponded to AML only and could not validate the ALL findings. Fourth, the ROC analysis was based on the same small cohort and compared patients only with healthy controls, which likely overestimated diagnostic performance in real-world clinical settings. Fifth, the public single-cell component consisted of re-analysis of external datasets and provided supportive transcriptomic localization rather than direct experimental or mechanistic validation. Finally, the absence of functional experiments limited causal inference.

In conclusion, the present data support increased OLFML2A expression in PBMCs from patients with AL, with the most consistent evidence indicating a potential adverse association in AML. In the local cohort, higher expression was associated with shorter PFS, whereas OS did not differ significantly. TCGA-LAML analyses further support the

prognostic relevance of OLFML2A in AML. By contrast, the findings in ALL are preliminary and should not be interpreted as evidence of a universal AL biomarker. Additional subtype-specific, multicenter and mechanistic studies are required before clinical application can be considered.

Acknowledgements

Not applicable.

Funding

This work was financially supported by the Foundation of Scientific Research Project of Jiangsu Provincial Health Commission (grant no. Z2024060), the XZHMU-QL Joint Research Fund (grant no. QL-YB023) and Hefei Normal University School-level research project, Hefei, Anhui, China (grant no. 2022rcjj50).

Availability of data and materials

The data generated in the present study may be requested from the corresponding author.

Authors' contributions

XL was responsible for research design, data analysis and obtaining funding support. JL performed manuscript writing and statistical analysis. XZ was involved in research implementation and statistical analysis. WZ contributed to study supervision, data interpretation, manuscript review and revision. XC was responsible for research design, data collection and obtaining funding support. XL and XC confirm the authenticity of all the raw data. All authors have read and approved the final manuscript.

Ethics approval and consent to participate

This study was approved by the Ethics Committee of the Second People's Hospital of Lianyungang City (ethical approval no. 2023-064). General written informed consent for the use of clinical data and biological samples for research purposes was obtained at the time of sample collection or hospital admission. For participants younger than the legal age of consent, written informed consent was obtained from a parent or legal guardian. All patient data were anonymized before analysis.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Kantarjian H, Kadia T, DiNardo C, Daver N, Borthakur G, Jabbour E, Garcia-Manero G, Konopleva M and Ravandi F: Acute myeloid leukemia: Current progress and future directions. Blood Cancer J 11: 41, 2021.
2. Döhner H, Wei AH, Appelbaum FR, Craddock C, DiNardo CD, Dombret H, Ebert BL, Fenaux P, Godley LA, Hasserjian RP, *et al*: Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. Blood 140: 1345-1377, 2022.
3. Arber DA, Orazi A, Hasserjian RP, Borowitz MJ, Calvo KR, Kvasnicka HM, Wang SA, Bagg A, Barbui T, Branford S, *et al*: International consensus classification of myeloid neoplasms and acute Leukemias: Integrating morphologic, clinical, and genomic data. Blood 140: 1200-1228, 2022.
4. Daltveit DS, Morgan E, Colombet M, Steliarova-Foucher E, Bendahhou K, Marcos-Gragera R, Rongshou Z, Smith A, Wei H and Soerjomataram I: Global patterns of leukemia by subtype, age, and sex in 185 countries in 2022. Leukemia 39: 412-419, 2025.
5. Yin W, Yan X, Cai J, Han B, Yin P, Lu G, Cheng W, Zhang J, Wu H, Ren J, *et al*: Incidence, mortality, and survival associated with acute leukaemia subtypes by age group in China: A population-based cancer registry analysis and cohort study. Lancet Haematol 12: e808-e822, 2025.
6. Zeng LC, Han ZG and Ma WJ: Elucidation of subfamily segregation and intramolecular coevolution of the olfactomedin-like proteins by comprehensive phylogenetic analysis and gene expression pattern assessment. FEBS Lett 579: 5443-5453, 2005.
7. Barembaum M, Moreno TA, LaBonne C, Sechrist J and Bronner-Fraser M: Noelin-1 is a secreted glycoprotein involved in generation of the neural crest. Nat Cell Biol 2: 219-225, 2000.
8. Tomarev SI and Nakaya N: Olfactomedin domain-containing proteins: Possible mechanisms of action and functions in normal development and pathology. Mol Neurobiol 40: 122-138, 2009.
9. Chen L, Li H, Liu W, Zhu J, Zhao X, Wright E, Cao L, Ding I and Rodgers GP: Olfactomedin 4 suppresses prostate cancer cell growth and metastasis via negative interaction with cathepsin D and SDF-1. Carcinogenesis 32: 986-994, 2011.
10. Grover PK, Hardingham JE and Cummins AG: Stem cell marker olfactomedin 4: Critical appraisal of its characteristics and role in tumorigenesis. Cancer Metastasis Rev 29: 761-775, 2010.
11. Hill SE, Kwon MS, Martin MD, Suntharalingam A, Hazel A, Dickey CA, Gumbart JC and Lieberman RL: Stable calcium-free myocilin olfactomedin domain variants reveal challenges in differentiating between benign and glaucoma-causing mutations. J Biol Chem 294: 12717-12728, 2019.
12. Ma S, Duan L, Dong H, Ma X, Guo X, Liu J, Li G, Yu Y, Xu Y, Yuan G, *et al*: OLFML2A downregulation inhibits glioma proliferation through suppression of Wnt/β-catenin signaling. Front Oncol 11: 717917, 2021.
13. Wu C, Chen M, Zhang Q, Yu L, Zhu J and Gao X: Genomic and GeneChip expression profiling reveals the inhibitory effects of Amorphophalli Rhizoma in TNBC cells. J Ethnopharmacol 235: 206-218, 2019.
14. Zhao Q, Zhang K, Li Y, Ren Y, Shi J, Gu Y, Qiu S, Liu S, Cheng Y, Qiao Y and Liu Y: OLFML2A is necessary for anti-triple negative breast cancer effect of selective activator protein-1 inhibitor T-5224. Transl Oncol 14: 101100, 2021.
15. Gao X, Yang Z, Xu C, Yu Q, Wang M, Song J, Wu C and Chen M: GeneChip expression profiling identified OLFML2A as a potential therapeutic target in TNBC cells. Ann Transl Med 10: 274, 2022.
16. Lu X, Li Y, Yang Y, Zhuang W, Chai X and Gong C: OLFML2A overexpression predicts an unfavorable prognosis in patients with AML. J Oncol 2023: 6017852, 2023.
17. Han Y, Wang Y, Dong X, Sun D, Liu Z, Yue J, Wang H, Li T and Wang C: TISCH2: Expanded datasets and new tools for single-cell transcriptome analyses of the tumor microenvironment. Nucleic Acids Res 51: D1425-D1431, 2023.
18. Chinese Society of Hematology, Chinese Medical Association, Leukemia & Lymphoma Group: Chinese guidelines for diagnosis and treatment of adult acute myeloid leukemia (non-acute promyelocytic leukemia) (2023 edition). Chin J Hematol 44: 705-712, 2023.
19. Professional Committee of Hematologic Oncology, China Anti-Cancer Association; Chinese Society of Hematology, Chinese Medical Association, Leukemia & Lymphoma Group: Chinese guidelines for diagnosis and treatment of adult acute lymphoblastic leukemia (2024 edition). Chin J Hematol 45: 417-429, 2024.

20. Liu H, Shi Y, Tao S, Li Y, Wang C and Yu L: Clinical and molecular characteristics of acute myeloid leukemia and the dismal prognosis of TP53 mutations in a real-world setting. Hematology 28: 2223866, 2023.
21. Soler G, Ouedraogo ZG, Goumy C, Lebecque B, Aspas Requena G, Ravinet A, Kanold J, Véronèse L and Tchirkov A: Optical genome mapping in routine cytogenetic diagnosis of acute leukemia. Cancers (Basel) 15: 2131, 2023.
22. Tallman MS, Wang ES, Altman JK, Appelbaum FR, Bhatt VR, Bixby D, Coutre SE, De Lima M, Fathi AT, Fiorella M, *et al*: Acute myeloid leukemia, version 3.2019, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 17: 721-749, 2019.
23. Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 25: 402-408, 2001.
24. Therneau TM: A package for survival analysis in R. R package version 3.5-5. 2023.
25. Kassambara A, Kosinski M and Biecek P: survminer: Drawing survival curves using ggplot2. R package version 0.4.9. 2021.
26. Wickham H: ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.
27. Wang C, Sun D, Huang X, Wan C, Li Z, Han Y, Qin Q, Fan J, Qiu X, Xie Y, *et al*: Integrative analyses of single-cell transcriptome and regulome using MAESTRO. Genome Biol 21: 198, 2020.
28. Hao Y, Hao S, Andersen-Nissen E, Mauck WM III, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, *et al*: Integrated analysis of multimodal single-cell data. Cell 184: 3573-3587.e29, 2021.
29. Hafemeister C and Satija R: Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol 20: 296, 2019.
30. Wu M, Li A, Zhang T, Ding W, Wei Y, Wan C, Ke B, Cheng H, Jin C and Kong C: The novel prognostic analysis of AML based on ferroptosis and cuproptosis related genes. J Trace Elem Med Biol 86: 127517, 2024.