Integrated Analysis and Validation Reveal CYTH4 as a Potential Prognostic Biomarker in Acute Myeloid Leukemia

Hong Wang  
Shenzhen University General Hospital

Yishu Xiao  
Shenzhen University General Hospital

Wei Zhou  
Shenzhen University Medical School

Yonghui Li (✉ yonghuililab@163.com)  
Shenzhen University General Hospital

Research Article

Keywords: CYTH4, acute myeloid leukemia, biomarker, clinical implications, functional analysis

Posted Date: May 30th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2986303/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Background

Acute myeloid leukemia (AML) is a clonal hematological malignancy with high mortality rates, and the clinical and genomic heterogeneity of AML has complicated therapy development. Identifying novel markers is urgently in need for AML. Cytohesins are a subfamily of guanine nucleotide exchange factors (GEFs) activating ADP-ribosylation factor (ARF) family GTPases. While previous studies have reported the important roles of cytohesins in various cancers, their function in AML remains unclear. Therefore, we performed this study to explore the prognostic impact of cytohesin-4 (CYTH4) and investigate the underlying molecular functions.

Methods

We obtained RNA sequencing data and AML clinical data from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) datasets to investigate gene expression and survival. Using R software, we identified differentially expressed genes between the high-CYTH4 group and the low-CYTH4 group. We conducted functional enrichment analysis by performing GO, KEGG, and GSEA analyses. CIBERSORTx tool was used to explore the proportions of different immune cell types. We also evaluated the molecular function of CYTH4 by examining cell growth, cell cycle, apoptosis, and colony-forming ability using CYTH4-knockdown AML cell lines.

Results

CYTH4 was significantly overexpressed in AML when compared with other cancers and normal tissues. High CYTH4 expression was associated with old age (p = 0.014), complex karyotype (p = 0.048), and higher risk status (p = 0.001). Patients with high CYTH4 expression had poor overall survival (OS) (high vs. low, HR = 1.58, 95%CI 1.04–2.45, p = 0.032) and event-free survival (EFS) (high vs. low, HR = 1.84, 95%CI 1.13–2.94, p = 0.013), and these patients could benefit from transplantation (transplantation vs. chemotherapy, HR = 0.35, 95%CI 0.20–0.60, p = 0.0001). Multivariate analysis showed high CYTH4 expression was independently associated with inferior OS (HR = 1.01, 95%CI 1.00-1.03, p = 0.017) and EFS (HR = 1.02, 95%CI 1.00-1.03, p = 0.034). Functional analysis showed that CYTH4 was involved in immunoregulation. In vitro validation showed knockdown of CYTH4 adversely affected cell growth and induced cell apoptosis.

Conclusions

CYTH4 is highly expressed in AML and can potentially function as a prognostic biomarker.
Background

Acute myeloid leukemia (AML) is a heterogenous clonal hematopoietic malignancy characterized by the arrest of myeloid cell maturation and disorder of differentiation, resulting in the abnormal accumulation of immature malignant cells in the bone marrow and the disruption of the normal hematopoietic process[1]. Consequently, AML patients exhibit a range of clinical symptoms, and the mortality rate is high[2]. The clinical and genomic diversity exacerbates the difficulty of therapy, necessitating the development of new biomarkers and treatment strategies. Over the past decades, risk assessment and treatment selection for AML have relied on morphology, immunophenotype, cytogenetics, and molecular features. Molecular testing is crucial because molecular changes often precede morphological abnormalities. However, clinical molecular testing mainly focused on gene mutation and fusion genes[3]. Recent studies have highlighted the critical roles of transcriptional dysregulation in AML leukemogenesis [4–6], and genome sequencing might serve as an alternative[7] or complement[8] to traditional testing in the diagnosis and prognosis of the disease.

Cytohesins, including cytohesin 1–4 (CYTH1-4), are a subfamily of guanine nucleotide exchange factors (GEFs). Cytohesins activate ADP-ribosylation factor (ARF) family GTPases which are involved in several essential biological functions, such as cytoskeletal organization[9], cell migration[10, 11], and cell signaling[12]. CYTH1-4 share a similar structural organization with an N-terminal coiled-coil motif, a central Sec7 domain, and a C-terminal pleckstrin homology (PH) domain[13]. Data from several sources have demonstrated the effects of cytohesins in carcinogenesis and cancer progression. Surveys conducted by Lee et al.[14] showed that CYTH2 was upregulated in malignant melanoma and contributed to tumor growth. CYTH2 was also reported to be upregulated in colorectal cancers[15] and hepatocellular carcinoma[16] and related with poor prognosis[15, 16]. This might be explained because CYTH2 enhanced the epidermal growth factor (EGF) pathway[17]. The study by Fu et al.[18] demonstrated that CYTH3 was upregulated in hepatocellular carcinoma and associated with tumor progression. Inhibiting cytohesins could inhibit the proliferation of gefitinib-resistant lung cancer cells, as Bill et al. reported[19]. Moreover, Zhang et al. comprehensively analyzed public datasets and revealed that high CYTH4 expression portends worse survival in ovarian cancer [20].

Although many studies have reported the clinical and pathological implications of cytohesins in cancer, their roles in leukemia remain unexplored. In the present study, we aimed to investigate the expression of CYTH4 in AML and explore its potential clinical implications. We also sought to identify genes associated with CYTH4 in AML, hoping to provide a promising prognostic biomarker for AML.

Methods

Gene expression analysis of CYTH4

In this study, gene expression analysis of CYTH4 was conducted using various public datasets and online platforms. The Human Protein Atlas (HPA) dataset (https://www.proteinatlas.org/, accessed on 14
October 2022) was utilized to analyze the expression of CYTH4 in different healthy human tissues. The CCLE dataset (https://sites.broadinstitute.org/ccle/, accessed on 15 October 2022)[21] was used to compare the expression of CYTH4 among various cancer types. The HPA dataset was also used to display the expression of CYTH4 in different cancer cell lines (accessed on 14 October 2022).

The Cancer Genome Atlas dataset (TCGA, https://portal.gdc.cancer.gov/) is a large public dataset containing both genome and clinical information spanning 33 cancer types[22]. The Genotype-Tissue Expression (GTEx, https://gtexportal.org/home/) dataset contains extensive RNA sequencing data and complex signatures from 54 types of non-diseased tissue sites across nearly 1000 individuals[23]. In this study, both TCGA and GTEx datasets were used. TIMER2.0 platform (http://timer.cistrome.org/, accessed on 20 October 2022)[24] was used in this study to compare the expression of CYTH4 between tumor and normal tissues based on TCGA. GEPIA website (http://geopia.cancer-pku.cn/, accessed on 21 October 2022)[25] was adopted to compare the expression of CYTH4 between AML patients (TCGA dataset) and normal cases (GTEx dataset). ULCAN platform (http://ualcan.path.uab.edu, accessed on 24 October 2022)[26] was used for visualizing the expression data among different AML FAB subtypes in the TCGA dataset.

**Survival analysis**

TCGA dataset was employed to investigate the survival significance of CYTH4 (accessed on 26 October 2022) [22]. A total of 151 AML samples with intact RNA sequencing data and survival status were included in our study. The cohort was divided into two groups based on the median expression of CYTH4: the low-CYTH4 group and the high-CYTH4 group. Overall survival (OS) and Event-free survival (EFS) were calculated with the Kaplan–Meier method and compared using the log-rank test. Univariate and multivariate survival analysis were performed with the Cox regression model and described with hazard ratios (HRs). A stepwise forward procedure was used in the multivariate analysis. Data from Gene Expression Omnibus (GEO) project (GSE10358, GSE14468) was also used in the survival analysis.

**Differential gene expression analysis and gene association analysis**

The DESeq2 package in R software was used to screen differentially expressed genes between the low-CYTH4 group and the high-CYTH4 group in AML. Significantly differentially expressed genes were defined using an adjusted p < 0.05 and |FC|>1.5. Gene association analysis was carried out using LinkedOmics (http://www.linkedomics.org/login.php, accessed on 1 November 2022)[26], a publicly available portal for analyzing multi-omics data based on the TCGA dataset. Pearson's correlation test was used to search for CYTH4-associated genes. Significantly associated genes were determined based on the criteria of FDR < 0.05 and |Pearson correlation coefficient|>0.5.

**GO analysis, GSEA analysis, STRING interaction analysis and CIBERSORT analysis**

Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were done with DAVID online tool (https://david.ncifcrf.gov/, accessed on 11 November 2022)
[27]. Gene Set Enrichment Analysis (GSEA) was performed with MSigDB (https://www.gsea-msigdb.org/gsea/index.jsp, accessed on 13 November 2022)[28]. Genes interacting with CYTH4 were also investigated with STRING (https://string-db.org/, accessed on 19 November 2022). The CIBERSORTx platform (https://cibersortx.stanford.edu/, accessed on 1 December 2022)[29] was used to compare the difference of immune cell infiltration between low-CYTH4 and high-CYTH4 groups.

**Cell culture**

MOLM-13, NOMO-1, and THP-1 cell lines were maintained in RPMI-1640 medium supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin. HEK293T was maintained in DMEM medium supplemented with 10% FBS and 1% penicillin-streptomycin. All cell lines were authenticated and regularly tested for mycoplasma contamination.

**shRNA construction, lentivirus preparation, and infection**

We constructed shRNA targeting CYTH4 and non-targeting control using synthesized shRNA-encoded DNA oligos and cloned them into the pLKO.1-puro vector (Addgene). The designed target sequences were as follows: scramble (TGAGGAAATTGCGGCTTATTT), shCYTH4 #1 (TRCN0000242587, CCGCCAAGGGTATCCAGTATT), shCYTH4 #2 (TRCN0000242586, TTGCACGGTTCCTGTATAAAG). The lentivirus was produced in HEK293T cells by transfecting the shRNA vector together with the packing vectors pLP/VSVG and psPAX2 (Addgene). Cells were subsequently infected with lentiviral particles via two rounds of “spinoculation” with polybrene (8ug/ml).

**RNA isolation, cDNA preparation, and qPCR**

Total RNA was isolated from cells using TRIzol reagent (Invitrogen) following the manufacturer’s instructions. Isolated RNA was converted into complementary DNA (cDNA) using a reverse transcription system (Transgene, #AT341). The expression of CYTH4 was detected by quantitative real-time PCR (qPCR) analysis using SYBR Fast Universal (Sigma-Aldrich, #KK4601) on the ABI Prism 7500 sequence detection system (Applied Biosystems). GAPDH was used as the reference gene for normalization. The following CYTH4 primer sequences were used: Forward ATTGGGCGCAAGAAGTTCAAC, Reverse TTTATACAGGAACCGTGCAATGT. The following GAPDH primer sequences were used: Forward CTCTGCTCCTCCTGTTCGAC, Reverse GCCCAATACGACCAAATCC.

**Western blot analysis**

Cells were lysed using RIPA buffer supplemented with 1 mM phenylmethane sulfonl fluoride. The concentration of total protein was measured using a BCA assay. Equal amounts of protein (~ 30ug) were loaded and separated by 12% SDS-PAGE. The proteins were then transferred onto polyvinylidene fluoride membranes and detected by immunoblotting with the ECL Western Blotting Substrate (Millipore). CYTH4 (NOVUS, H00027128-B01P) and β-Actin (Cell Signaling Technology, 5125) antibodies were used in this study.
Cell proliferation, cell cycle, apoptosis, and in vitro colony formation assay

Cell proliferation was assessed by cell number counting using Trypan blue staining. Cell cycle analysis was performed by fixing cells in 75% ice-cold ethanol at 4°C overnight, staining with propidium iodide (PI) for 30 minutes, and then analyzing them with a FACS flow cytometer (Beckman Coulter). Apoptosis was detected by using Annexin V-fluorescein isothiocyanate (FITC)/PI apoptosis detection kit (DOJINDO) according to the manufacturer’s instructions. Briefly, cells were stained with Annexin V-FITC and PI at room temperature for 15 minutes and analyzed with a FACS flow cytometer (Beckman Coulter). The colony formation assay was performed using MethoCult™ H4434 Classic (Stem Cell Technologies) culture system. Colonies (≥ 50 cells) were counted after 10 days.

Statistical analysis

SPSS (IBM, version 22.0) and GraphPad Prism 7 were used for the statistical analyses. Clinical features between the two groups were compared using the chi-square test for categorical variables and Fisher’s exact test for the expected frequency of an event (< 5 in any cell of 2×2 tables). Continuous variables were compared with the nonparametric Mann–Whitney U test. One-way ANOVA was used to compare scramble cells and CYTH4-knockdown cells, with scramble cells serving as the control. Dunnett’s multiple comparison method was used for testing. Results were presented as mean ± standard deviation. A two-sided p value < 0.05 was considered statistically significant.

Results

CYTH4 is upregulated in AML cell lines.

We focused on CYTH4 because its expression was much higher than other cytohesins in AML (Additional file 1, Fig. S1). We first explored the HPA dataset to examine the RNA tissue specificity of CYTH4 in healthy human. Results showed that the expression of CYTH4 was enhanced in the bone marrow and lymphoid tissues while being expressed at low levels in other measured tissues (Fig. 1A).

We then analyzed the level of CYTH4 expression in cell lines based on the latest next-generation sequencing data from the CCLE dataset (Additional file 2, Table S1). The result showed that CYTH4 was particularly highly expressed in lymphoma and leukemia cell lines, followed by thyroid cancer (Fig. 1B). The analysis from the HPA dataset also presented that CYTH4 was highly expressed in myeloid cancer cells than in other cancer cell lines including the brain, liver, kidney, et al. (Fig. S2). Further, we used the data from CCLE leukemia cell lines to compare the expression level of CYTH4 in different types of leukemia and found that CYTH4 was highly expressed in AML when compared with acute lymphoblastic leukemia (ALL) and chronic myeloid leukemia (CML) (Fig. 1C). RT-PCR detection of some commonly used leukemia cell lines also showed that CYTH4 was highly expressed in AML cell lines (Fig. 1D)

CYTH4 is upregulated in patients with acute myeloid leukemia.
To investigate CYTH4 expression in human cancers, the TCGA dataset was analyzed. Figure 2A displayed an overview of the different expression of CYTH4 between tumors and normal tissues across the TCGA dataset, suggesting that CYTH4 expression was higher in AML patients than in other tumors. CYTH4 expression was also significantly upregulated in AML when compared to healthy donors (Fig. 2B). These results for the AML samples corresponded with those in the cell lines (Fig. 1B). Furthermore, differences in CYTH4 expression were observed among AML subtypes, with M3-AML patients demonstrating the lowest expression of CYTH4 (Fig. 2C).

**Clinical features differ between low-CYTH4 group and high-CYTH4 group.**

To investigate the clinical significance of CYTH4 in AML, we divided patients into low-CYTH4 and high-CYTH4 groups based on the median expression of CYTH4 (Additional file 2, Table S2). We then compared the clinical characteristics between the two groups and demonstrated the results in Table 1. We discovered that patient age varied significantly between the two groups, with patients in the high-CYTH4 group being older than those in the low-CYTH4 group (median age, 60 vs. 53, p = 0.014). Moreover, in the M4-AML subtype, there were more patients with high CYTH4 expression than those with low expression (p = 0.008), while all M3-AML patients were in the low expression group (p < 0.0001). This discovery is in line with the previous result that M3-AML patients had the lowest CYTH4 expression (Fig. 2C). The high-CYTH4 group had more cases with complex karyotypes than the low expression group (p = 0.048). As for risk status, low expression of CYTH4 was significantly associated with good-risk status (p = 0.001). No significant difference was found between the low-CYTH4 and high-CYTH4 groups concerning sex, bone marrow (BM) blasts, white blood cell (WBC) count, and peripheral blood (PB) blasts.
Table 1: Characteristics of AML patients (n = 151) in TCGA database grouped by CYTH4 expression.

<table>
<thead>
<tr>
<th>Clinicopathological characteristic</th>
<th>Low-CYTH4 (n = 75)</th>
<th>High-CYTH4 (n = 76)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, male/female</td>
<td>43/32</td>
<td>39/37</td>
<td>0.458</td>
</tr>
<tr>
<td>Median age, years (range)</td>
<td>53(22–76)</td>
<td>60(21–88)</td>
<td>0.014</td>
</tr>
<tr>
<td>Median BM blasts, % (range)</td>
<td>74(30–100)</td>
<td>70.5(30–98)</td>
<td>0.064</td>
</tr>
<tr>
<td>Median WBC, (\times 10^9/L) (range)</td>
<td>12.6(0.4-171.9)</td>
<td>24.5(0.7-223.8)</td>
<td>0.093</td>
</tr>
<tr>
<td>Median PB blasts, % (range)</td>
<td>40(0–97)</td>
<td>37(0–96)</td>
<td>0.933</td>
</tr>
<tr>
<td>FAB classifications</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M0</td>
<td>5</td>
<td>10</td>
<td>0.182</td>
</tr>
<tr>
<td>M1</td>
<td>18</td>
<td>18</td>
<td>0.964</td>
</tr>
<tr>
<td>M2</td>
<td>20</td>
<td>17</td>
<td>0.539</td>
</tr>
<tr>
<td>M3</td>
<td>15</td>
<td>0</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>M4</td>
<td>8</td>
<td>21</td>
<td>0.008</td>
</tr>
<tr>
<td>M5</td>
<td>6</td>
<td>9</td>
<td>0.430</td>
</tr>
<tr>
<td>M6</td>
<td>1</td>
<td>1</td>
<td>0.992</td>
</tr>
<tr>
<td>M7</td>
<td>1</td>
<td>0</td>
<td>0.312</td>
</tr>
<tr>
<td>NA</td>
<td>1</td>
<td>0</td>
<td>0.312</td>
</tr>
<tr>
<td>Fusion gene, no.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal karyotype</td>
<td>25</td>
<td>37</td>
<td>0.081</td>
</tr>
<tr>
<td>RUNX1-RUNX1T1</td>
<td>7</td>
<td>0</td>
<td>0.006</td>
</tr>
<tr>
<td>PML-RARA</td>
<td>15</td>
<td>0</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>MYH11-CBFB</td>
<td>3</td>
<td>7</td>
<td>0.327</td>
</tr>
<tr>
<td>BCR-ABL1</td>
<td>2</td>
<td>1</td>
<td>0.620</td>
</tr>
<tr>
<td>Complex karyotype</td>
<td>5</td>
<td>13</td>
<td>0.048</td>
</tr>
<tr>
<td>Risk level (Molecular)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>24</td>
<td>7</td>
<td>0.001</td>
</tr>
<tr>
<td>Intermediate</td>
<td>33</td>
<td>48</td>
<td>0.018</td>
</tr>
<tr>
<td>Poor</td>
<td>17</td>
<td>19</td>
<td>0.737</td>
</tr>
</tbody>
</table>
Clinicopathological characteristic | Low-CYTH4 (n = 75) | High-CYTH4 (n = 76) | p value
--- | --- | --- | ---
NA | 1 | 2 | 0.568
Gene mutation
TET2 | 6 | 6 | 0.981
DNMT3A | 20 | 16 | 0.418
CEBPA | 3 | 10 | 0.045
RUNX1 | 3 | 11 | 0.027
NPM1 | 19 | 19 | 0.962
TP53 | 4 | 7 | 0.359
WT1 | 6 | 4 | 0.499
FLT3 | 25 | 18 | 0.189
KIT | 2 | 5 | 0.253
IDH2 | 5 | 11 | 0.119
IDH1 | 9 | 5 | 0.251
NRAS | 3 | 4 | 0.712
PTPN11 | 3 | 3 | 0.987
KRAS | 3 | 4 | 0.712

BM, bone marrow; WBC, white blood cell; PB, peripheral blood. A two-sided p-value < 0.05 was considered statistically significant.

**High expression of CYTH4 is associated with poor survival in AML.**

To investigate the prognostic significance of CYTH4 expression in AML, we compared survival between high and low expression groups in different datasets. As shown in Fig. 3A and 3B, high CYTH4 expression was significantly associated with unfavorable OS (high vs. low, HR = 1.58, 95%CI 1.04–2.45, p = 0.032) and EFS (high vs. low, HR = 1.84, 95%CI 1.13–2.94, p = 0.013). This conclusion was validated in GSE10358 (Fig. 3C) and GSE14468 (Fig. 3D) datasets. Next, with the TCGA dataset, we compared the survival between low and high CYTH4 expression grouped by treatment. The results showed that high CYTH4 expression was associated with poor OS (Fig. 3E; high vs. low, HR = 2.01, 95%CI 1.19–3.41, p = 0.009) in patients treated with chemotherapy alone. However, in cases who received both chemotherapy and transplantation, no significant difference was found (Fig. 3F, p = 0.370). We then compared the survival between patients treated with chemotherapy alone and those with chemotherapy plus
transplantation grouped by CYTH4 expression level. Interestingly, transplantation did not show a significant difference in the low-CYTH4 group (Fig. 3G, p = 0.419), but significantly improved survival in the high-CYTH4 expression group compared to chemotherapy alone (Fig. 3H; transplantation vs. chemotherapy, HR = 0.35, 95%CI 0.20–0.60, p = 0.0001). These findings suggest that transplantation may attenuate the adverse effect of high CYTH4 expression on patient survival in AML.

To further explore the prognostic effect of CYTH4 in AML, we did univariate and multivariate survival analysis using Cox regression model. When combining age, WBC, FLT3 mutation, TP53 mutation, transplantation in the multivariate Cox analysis, results confirmed that high expression of CYTH4 was independently associated with inferior OS (HR = 1.01, 95%CI 1.00-1.03, p = 0.017) and EFS (HR = 1.02, 95%CI 1.00-1.03, p = 0.034) (Table 2).

Table 2
Univariate and multivariate analysis of OS and EFS in AML patients in TCGA dataset (n = 151).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>OS HR (95% CI), p value</th>
<th>EFS HR (95% CI), p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Univariate</td>
<td>Multivariate</td>
</tr>
<tr>
<td>CYTH4</td>
<td>1.02 (1.01–1.03), 0.001</td>
<td>1.01 (1.00-1.03), 0.017</td>
</tr>
<tr>
<td>Sex</td>
<td>0.98 (0.66–1.46), 0.924</td>
<td>1.07 (0.66–1.72), 0.796</td>
</tr>
<tr>
<td>Age</td>
<td>2.07 (1.33–3.20), 0.001</td>
<td>1.74 (1.10–2.75), 0.018</td>
</tr>
<tr>
<td>WBC</td>
<td>1.00 (1.00-1.01), 0.020</td>
<td>1.01 (1.00-1.01), 0.018</td>
</tr>
<tr>
<td>BM Blast</td>
<td>1.00 (0.99–1.01), 0.977</td>
<td>1.00 (0.98–1.01), 0.529</td>
</tr>
<tr>
<td>PB blast</td>
<td>1.00 (0.99-1.00), 0.485</td>
<td>1.01 (1.00-1.02), 0.007</td>
</tr>
<tr>
<td>Karyotype</td>
<td>1.04 (0.97–1.12), 0.302</td>
<td>1.11 (1.01–1.23), 0.033</td>
</tr>
<tr>
<td>FLT3 mutation</td>
<td>1.54 (1.01–2.38), 0.045</td>
<td>1.75 (1.11–2.77), 0.016</td>
</tr>
<tr>
<td>NPM1 mutation</td>
<td>0.87 (0.56–1.37), 0.554</td>
<td>0.71 (0.43–1.20), 0.200</td>
</tr>
<tr>
<td>TP53 mutation</td>
<td>5.09 (2.64–9.85), &lt;0.001</td>
<td>6.12 (2.97–12.64), &lt;0.001</td>
</tr>
<tr>
<td>Transplantation</td>
<td>0.53 (0.36–0.81), 0.003</td>
<td>0.55 (0.36–0.84), 0.006</td>
</tr>
</tbody>
</table>
OS, overall survival; EFS, event-free survival; HR, hazard ratio; CI, confidence interval; WBC, white blood count; BM, bone marrow; PB, peripheral blood. A two-sided p-value < 0.05 was considered statistically significant.

**CYTH4-associated gene analysis**

To better understand the role of CYTH4 in AML, we compared the transcriptomes between the high-CYTH4 group and the low-CYTH4 group in the TCGA dataset. A total of 793 genes showed significantly different expression (adjusted p < 0.05, |FC|>1.5) including 572 genes significantly upregulated in the high-CYTH4 group and 221 genes downregulated in the high-CYTH4 group (Fig. 4A; Additional file 2, Table S3). We then used LinkedOmics tools to perform correlation analysis and identified a total of 451 significantly co-expressed genes with a cut-off value of FDR < 0.05 and |Pearson correlation coefficient|>0.5 (Fig. 4B, Additional file 2, Table S4). Of these co-expressed genes, 326 genes were positively correlated with CYTH4 expression, and 125 were negatively correlated. By integrating the results of these two analyses, we identified 214 genes that were upregulated in the high-CYTH4 group and positively correlated with CYTH4 expression (Fig. 4C). Conversely, only 27 genes were found to be both downregulated in the high-CYTH4 group and negatively correlated with CYTH4 expression (Fig. 4D). The overlapping genes were further analyzed for their biological functions.

**Functional enrichment analysis of overlapping genes.**

We then sought to explore the possible biological function of CYTH4. GO analysis and KEGG pathway enrichment were performed and the results were shown in Fig. 5A and 5B. These overlapping genes were significantly associated with immune response, regulation of T cell proliferation, cytokine production (IFN-γ, IL-6), and signaling transduction. We also conducted gene set enrichment analysis (GSEA) and the enrichment plot showed that these overlapping CYTH4-associated genes were significantly enriched in the gene set related to the immune response (Fig. 5C). Additionally, we employed the STRING website to investigate the genes that interacted with CYTH4 and the results were presented in Fig. 5D. Then, we estimated the fractions of 22 distinct immune cell types using the CIBERSORTx algorithm. The result showed that high CYTH4 expression was significantly correlated with memory B cells, activated CD4 memory T cells, and monocytes (Fig. 5E).

**In vitro validation of the function of CYTH4**

To further investigate the function of CYTH4 in AML, we did in vitro validation using MOLM-13, NOMO-1, and THP-1 AML cell lines. Lentivirus expressing shRNA significantly reduced the mRNA and protein expression levels of CYTH4 (Fig. 6A; Additional file 1, Fig. S3). Cell number counting showed that CYTH4 knockdown significantly suppressed the cell growth of AML cells (Fig. 6B). Cell cycle assays revealed a significant G0/G1 phase arrest in all three AML cell lines (Fig. 6C). Moreover, a reduced S phase population was recorded in MOLM-13 and THP-1 cells, while a drop in the G2/M phase population was observed in NOMO-1 cells (Fig. 6C) The results of colony-forming assays demonstrated that silencing of CYTH4 significantly impaired the clonogenic potential of leukemia cell lines (Fig. 6D). Increased
apoptosis was also recorded in leukemia cell lines following transfection with CYTH4 shRNA (Fig. 6E). Taken together, these results indicated that CYTH4 played an oncogenic role in AML cells.

Discussion

AML is a clonal hematological malignancy characterized by abnormally rapid proliferation, maturation arrest, and differentiation block of myeloid precursors[30]. Patients diagnosed with AML usually have poor outcomes and high mortality rates[2]. Nowadays, the diagnosis of AML mainly relies on the analysis of bone marrow morphology, immunophenotype, cytogenetics, and molecular features, which also form the basis for risk stratification[1, 8]. The complexity and heterogeneity of AML shed light on the importance of precision medicine and the detection of robust biomarkers. Recent studies highlight the feasibility of gene expression assay in AML management, with an improvement in risk stratification efficiency and prognostic capacity[7, 8]. In the present study, we found that high expression of CYTH4 was associated with poor survival in AML and it might be used as a prognostic biomarker.

Cytohesins have been reported to play a pivotal role in various cancers, including but not limited to hepatocellular carcinoma, colorectal cancer, lung cancer, and ovarian cancer [10, 14, 16−19, 31]. We focused on CYTH4 for its higher expression than other cytohesins in AML. Notably, the expression of CYTH4 was substantially higher in AML compared to other cancers. CYTH4 expression was higher in AML than in healthy people. The HPA dataset showed that CYTH4 expression was predominantly enhanced in the bone marrow and lymphoid tissues compared to other organs. All these supported that CYTH4 was a significantly overexpressed gene in AML. The high expression and bone marrow specificity provided prerequisites for CYTH4 to be a possible biomarker in AML. Furthermore, the limited tissue specificity made it reasonable to hypothesize that it might be used as a therapeutic target. However, it remains to be determined whether CYTH4 could be used as a therapeutic target in AML. Bill’s study found that inhibition of cytohesins improved the treatment of gefitinib-resistant lung cancers while there were no studies on AML[19].

As for clinical characteristics, high CYTH4 expression was significantly correlated with elder age, complex karyotype, higher risk status, and RUNX1 mutation. These unfavorable factors are known to adversely affect the prognosis of AML patients [32, 33], indicating that CYTH4 might act as a negative prognostic factor. The survival analysis provided ultimate evidence that high expression of CYTH4 was associated with poor survival, validated in 3 different datasets. The multivariate analysis also confirmed the adverse prognostic effect of CYTH4. Additionally, it was observed that in patients with high CYTH4 expression, those who received chemotherapy plus transplantation had better survival outcomes than those who received chemotherapy alone. It suggested that CYTH4 might serve as an indicator to guide therapy, and transplantation could potentially overcome the adverse effect of high CYTH4 expression. In vitro functional analysis further confirmed that CYTH4 exerted oncogenic effects on AML cell lines, thereby underscoring its potential value as a prognostic biomarker in AML. These findings were consistent with previous studies that reported cytohesins have a variety of biological activities and are involved in cell
proliferation[16], migration[18], and invasion[15] during carcinogenesis. This study, therefore, suggests that CYTH4 is upregulated in AML and may play a crucial role in AML leukemogenesis.

When we compared the gene expression profile between the high-CYTH4 and the low-CYTH4 groups, more genes were upregulated in the high-CYTH4 group than downregulated. Meanwhile, the correlation analysis showed that the number of genes positively correlated with CYTH4 was much more than the negatively correlated genes. This is consistent with the function of CYTH4 as a guanine nucleotide exchange factor (GEF) to activate GTPase. In addition, the GO analysis showed that CYTH4 associated genes were more involved in the positive regulation process. In AML, our data showed that CYTH4-associated genes were largely involved in the immune response such as antigen processing and presentation, positive regulation of T cell proliferation, and positive regulation of T cell activation. Besides, it was related to the production of many cytokines, like IFN-γ, IL-6, and tumor necrosis factor. This is consistent with the study by Wang et al.[10] that they demonstrated CYTH2 participated in immunoregulation. The KEGG and GSEA analysis also proved that CYTH4 was involved in immune response, but the exact mechanism needs to be further explored. Immunotherapy, including checkpoint inhibitors and chimeric antigen receptor (CAR)-T therapy, is playing an increasingly important role in the treatment of leukemia. Further investigation of the role of CYTH4 in immunoregulation might provide novel insight into improving therapeutic efficacy and overcoming obstacles encountered in immunotherapy. Apart from the immunoregulation effect, the KEGG result showed that CYTH4 was related to allograft rejection. It suggested that CYTH4 might be involved in graft-versus-host disease (GVHD) in AML patients undergoing transplantation. Further investigation of CYTH4 might help us to improve the success rate of transplantation. However, this hypothesis requires further study.

The current study has some limitations that need to be addressed. Firstly, the results were mainly based on bioinformatics analysis of public datasets, more experimental validations especially in vivo validations are required to confirm the findings. Secondly, our study suggested CYTH4 expression might be used as a prognostic biomarker in AML. However, compared with gene mutation and cytogenetic abnormalities, there is much more to consider before the gene expression profile could be used as a biomarker. For instance, the optimal cut-off value between high and low expression is difficult to determine, because the expression is a relative concept. The accuracy, feasibility, and clinical utility need to be well demonstrated in larger patient cohorts. Nonetheless, this study provides valuable insights into the potential role of CYTH4 as a prognostic biomarker in AML and prompts further investigations into its clinical relevance and therapeutic potential.

**Conclusions**

CYTH4 is upregulated in AML and the high expression of CYTH4 is associated with poor survival. CYTH4 can potentially be used as a prognostic marker in AML.

**Abbreviations**
AML, acute myeloid leukemia; cytohesin-4, CYTH4; TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GSEA, Gene Set Enrichment Analysis; OS, overall survival; EFS, event-free survival; HR, hazard ratio; HPA: Human Protein Atlas; CCLE, Cancer Cell Line Encyclopedia; GTEx, Genotype-Tissue Expression; FDR, false discovery rate; FBS, fetal bovine serum; BM, bone marrow; WBC, white blood cell; PB, peripheral blood.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data used in this study are openly available (TCGA, GEO, CCLE) or included in the article. Further inquiries can be directed to the corresponding author.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by grants from the National Natural Science Foundation of China (82070161, 82030076, 81970151, 81870134, and 82100164), China Postdoctoral Science Foundation (2022TQ0217), Natural Science Foundation of Guangdong Province (2022A1515012465), Natural Foundation of Shenzhen Science and Technology Innovation Commission (JCYJ20220531103003007), Natural Science Foundation of Shenzhen University General Hospital (SUGH2020QD008), Shenzhen Key Laboratory Foundation (ZDSYS20200811143757022) and Sanming Project of Shenzhen (SZSM202111004).

Author contributions

HW and YL designed the research. HW, YX, and WZ acquired data, analyzed data and performed statistical analysis. HW and YX performed in vitro function validation. HW drafted the manuscript. YX,
WZ and YL revised the manuscript. All authors read and approved the final manuscript.

Acknowledgments

Not applicable.

References


Figures
Figure 1

The expression of CYTH4 in human normal tissues and pan-cancer. (A) The expression of CYTH4 in various human normal tissues in the HPA dataset. (B) The expression of CYTH4 in different categories of cancer cell lines, analyzed with the CCLE dataset. (C) The expression of CYTH4 in different types of leukemia cell lines, analyzed with the CCLE dataset. (D) RT-PCR showing the expression of CYTH4 in
different leukemia cell lines. HPA, Human Protein Atlas, CCLE, Cancer Cell Line Encyclopedia. ****, p<0.0001.

Figure 2

The expression of CYTH4 in patients with AML. (A) The expression of CYTH4 in tumor and adjacent normal tissues across all TCGA tumors, analyzed by TIMER 2.0. (B) The comparison of CYTH4 expression between AML patients (TCGA) and healthy people (GTEx), analyzed by the GEPIA website. (C) The expression of CYTH4 in different AML FAB subtypes, analyzed by the UALCAN website. TCGA, The Cancer Genome Atlas; FAB, French–American–British classification systems. *, p<0.05; **, p<0.01; ***, p<0.001.
Survival analysis based on the expression level of CYTH4. Comparison of OS (A) and EFS (B) between high-CYTH4 group and low-CYTH4 group in TCGA dataset. Comparison of OS between high-CYTH4 group and low-CYTH4 group in GSE10358 (C) and GSE14468 (D) datasets. Comparison of OS between the high-CYTH4 group and low-CYTH4 group in patients treated with chemotherapy alone (E) or treated with chemotherapy plus transplantation (F). Comparison of OS between chemotherapy and transplantation in patients with low CYTH4 expression (G) or with high CYTH4 expression (H). OS, overall survival; EFS, event-free survival. A two-sided p-value < 0.05 was considered statistically significant.
**Figure 4**

**Identification of CYTH4-associated genes.** (A) Volcano plot of differentially expressed genes between the low-CYTH4 group and the high-CYTH4 group. The red dots represent genes significantly upregulated in the high-CYTH4 group and the blue dots represent downregulated genes. Adjusted p<0.05 and |FC|>1.5 were considered significantly expressed. (B) CYTH4 association analysis result by Linkedomics. Red dots represent genes positively correlated with CYTH4 and green dots represent genes negatively correlated. (C) Venn diagram showing the overlap of upregulated genes and positively correlated genes. (D) Venn diagram showing the overlap of downregulated genes and negatively correlated genes. FC, fold change.
Figure 5

**Functional enrichment analysis of CYTH4-associated genes.** (A) GO analysis result. (B) KEGG pathway enrichment analysis. (C) GSEA enrichment plots. (D) Genes interacted with CYTH4 in STRING. (E) CIBERSORTx result showing the different fraction of immune cells between the low-CYTH4 and high-CYTH4 groups. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GSEA, Gene Set Enrichment Analysis.
Figure 6

In vitro validation of the function of CYTH4. (A) Western blot showing reduced CYTH4 protein expression level after shRNA-mediated CYTH4 knockdown. (B) Cell counting displaying the inhibition of cell growth in CYTH4-silenced cells compared with control cells. (C) Flow cytometry analysis of cell cycle distribution in CYTH4-silenced cells and control cells. (D) Colony formation assays showing growth inhibition in CYTH4-silenced cells compared with control cells. (E) Flow cytometry showing increased apoptosis in
CYTH4-silenced cells compared with control cells. Bars in (C), (D), and (E) showed the mean ± SD from three independent biological replicates. *, p<0.05; **, p<0.01; ***, p<0.001; ***, p<0.0001.