Am J Transl Res 2023;15(12):6858-6866
www.ajtr.org /ISSN:1943-8141/AJTR0150351

Brief Communication

Bioinformatic analysis reveals the clinical value of SASH3 in survival prognosis and immune infiltration of acute myelocytic leukemia (AML)

Yufei Li[1], Lin Wang[1], Xueyuan Jia[1], Yanru Yang[2], Zhengqi Qiu[1,2]

[1]Faculty of Medicine, Macau University of Science and Technology, Room PP-R203, Est. Seak Pai Van Praia Park, Rés-Do-Chão R, Coloane, Macau 999078, China; [2]Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Health Science Center, Shenzhen University, Shenzhen 518060, Guangdong, China

Received March 25, 2023; Accepted November 21, 2023; Epub December 15, 2023; Published December 30, 2023

Abstract: Acute myeloid leukemia (AML), a malignant clonal disease, is the most prevalent form of leukemia, and it is associated with a poor prognosis and unfavorable treatment outcomes in both pediatric and adult populations. Accordingly, enhancing anti-tumor responses using immunomodulators is a promising therapeutic strategy and a new avenue for treating AML. In this study, we used publicly available data from The Cancer Genome Atlas and Genotype-Tissue Expression databases to investigate the correlation between SAM and SH3 domain-containing 3 (SASH3) and AML, and we performed Cox regression and Kaplan-Meier analyses to assess the clinical characteristics associated with overall survival among patients with AML. Additionally, we analyzed the relationship between immune infiltration and SASH3. Compared with that in the normal group, patients with AML were characterized by significantly higher levels of SASH3 expression (P = 3.05e-34), which was strongly associated with survival outcomes. We observed a significant correlation between SASH3 expression and the expression of cancer-related genes (*HCK*, *SYK*, *FYN*, *ITGB2*, *PIK3CD*, *FGR*, *PIK3R5*, *VAV1*, *LCP2*, and *GRB2*) and pathways. Our findings in this study indicate that SASH3 plays a key role in AML development and survival outcomes and in the regulation of small GTPase-mediated signal transduction and immune-related pathways. Accordingly, targeting SASH3 may offer a promising approach for the treatment of AML and may potentially influence the progression of other cancers via multiple immune pathways.

Keywords: SASH3, acute myeloid leukemia, TCGA, bioinformatics, prognosis

Introduction

Although external environmental factors play a prominent role in the development of acute myeloid leukemia (AML), the disease is characterized by significant genetic heterogeneity [1]. AML is the most prevalent type of leukemia, arising from hematopoietic stem and progenitor cells. By inducing similar changes in peripheral blood, the malignant cells of AML can suppress normal hematopoiesis. Regardless of the treatment method employed, AML management entails a high economic burden and necessitates a complex treatment process. In the United States alone, there are approximately 19,000 new cases of AML annually, with 10,000 deaths resulting from AML-associated complications [2]. Additionally, AML has a high recurrence rate, which further compounds the disease burden [3]. Although over the past three decades, the prognosis for children with AML has undergone a significant improvement; however, up to 30% of patients in pediatric outpatient clinics remain unhealed. Moreover, the cure rate for patients with AML above 60 years of age is between only 5% and 15%, and infectious events due to weakened immunity are difficult to prevent [4, 5]. Although the etiology of AML, as a specific type of cancer without paraneoplastic tissues, has yet to be sufficiently clarified, the development of new treatments for AML is imminent.

Sterile alpha motif (SAM) and Src homology-3 (SH3) domain-containing 3 (SASH3), referred to as SLY, is a putative adaptor protein postulated

to play an important role in the organization of signaling complexes and propagation of signal transduction cascades in lymphocytes [6]. SASH3 was initially identified based on an adhesion assay screen utilizing a T-cell lymphoma complementary DNA library [7], and an independent study subsequently identified SASH3 as a protein phosphorylated at the Ser27 residue following T-cell receptor ligation and activation of protein kinase C. More recently, it has been discovered that mutations in SASH3 can lead to a novel X-linked congenital immunodeficiency disorder [8]. The occurrence of this disease may be closely associated with the effects of SASH3 on T cell receptor (TCR) signaling. Furthermore, it has been established that T-cell progenitors derived from bone marrow CD34 cells with SASH3 deficiencies are less likely to survive [6]. In different cancer studies, such as those focusing on lung and breast cancer, SASH3 has been observed to show differences in expression and to be closely linked with disease diagnosis and prognosis. However, the expression of SASH3 is higher in the breast tissues of patient with breast cancer than in controls and paraneoplastic tissues. In contrast, the opposite has been observed in patients with lung cancer, in whom SASH3 expression is significantly higher in both normal and paracancerous tissues [9, 10]. Although much of the research on SASH3 has focused on cancer and immunodeficiency diseases, the extent of this research is limited, and there are few studies specifically investigating the overexpression of SASH3.

In this study, we sought to evaluate the diagnostic and prognostic values of SASH3 in human AML using data obtained from The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) databases. To better understand the potential pathophysiological pathways associated with AML pathogenesis and their associations with SASH3, we performed Gene Ontology (GO) functional and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses. In addition, we analyzed immune cell infiltration. The results of this study provide new insights into the underlying mechanisms of AML and indicate that SASH3 may have potential applications as a diagnostic and prognostic biomarker.

Materials and methods

Bioinformatic mining to determine the expression of SASH3 in AML

The Gene Expression Profiling Interactive Analysis (GEPIA) data referenced in this study are available in a public repository at the GEPIA website (http://gepia2.cancer-pku.cn). GEPIA is a powerful web server for analyzing and visualizing RNA sequence expression data. In this study, normal (GTEx) and tumor (TCGA) tissue samples were compared using the resources provided in GEPIA, as no paracancerous tissues were available.

Genes co-expressed with SASH3

In this study, we evaluated the genes co-expressed with SASH3 in AML using data obtained from the UALCAN (http://ualcan.path.uab.edu/) and GEPIA websites. The web tool (http://bioinformatics.psb.ugent.be/webtools/Venn/) extracted the top 500 genes most closely co-expressed with SASH3 from each database. We obtained a final set of 380 commonly co-expressed genes based on the intersection of these gene sets.

GO functional and KEGG pathway enrichment analyses

To identify gene functions and pathways pertinent to this study, we performed GO and KEGG enrichment analyses conducted using WebGestalt (http://www.webgestalt.org/). This website provides a range of powerful analytical tools for studying biological functions. GO annotations can be categorized into three functional categories based on common genes, namely, biological processes (BP), cellular components (CC), and molecular functions (MF). KEGG pathway enrichment results were obtained using the clusterProfiler package in R version 4.2.1. We used the default WebGestalt settings to calculate enrichment and corrected p-values using the Benjamini-Hochberg False Discovery Rate (FDR) method, with an FDR-corrected p-value < 0.05 employed as the threshold for defining significance.

Immune infiltration analysis

Single-sample gene set enrichment analysis is a commonly used method for assessing the

extent of immune cell infiltration in tumors [11]. In this study, we used the GSVA package to analyze the relationship between SASH3 and 24 different types of immune-infiltrating cells. Spearman correlation analysis was used to examine the relationship between the levels of SASH3 expression and immune cell enrichment scores.

Construction of protein-protein interaction networks and identification of hub genes

To examine the roles of associated proteins for different molecular functions and biological pathways, protein-protein interactions (PPIs) can be represented in terms of networks. In this study, we performed protein interaction analysis for SASH3 by searching for protein-protein interaction networks in the STRING database (https://string-db.org/). The resulting genes were displayed using Cytoscape (version 3.8.0), and we used the Maximal Clique Centrality method to rank genes and identify the top 10 genes as hub genes.

Nomogram model analysis

Cox proportional hazards regression was used to perform both univariate and multivariate analyses. We used the Wilcoxon signed-rank test and logistic regression to analyze the relationships between clinical pathological features and SASH3 expression. We identified independent prognostic predictors using a Cox proportional hazards regression model based on the "Enter" method. A nomogram was then designed to predict the 1-, 3-, and 5-year prognosis based on the multivariable Cox regression model results. Finally, we estimated the probability of event occurrence based on an alignment chart and compared this with the actual percentage of event occurrence. In this regard, a curve that plots close to the reference line indicates a high degree of calibration.

Statistical analysis

For all statistical analyses, we used the R statistical package. The Wilcoxon rank-sum and paired sample *t*-test were applied to assess the statistical significance of SASH3 expression in the non-paired and paired tissues, respectively. Correlations between clinical characteristics and SASH3 expression were assessed using the Wilcoxon rank-sum test and logistic regression, and Spearman correlation analysis was performed to analyze the immunological data. For this study, statistical significance was set at the $P < 0.05$ level.

Results

Relationships between SASH3 expression and clinicopathological variables

Analysis of the Timer 2.0 database (http://timer.comp-genomics.org/) revealed that SASH3 is differentially expressed in various cancers and is particularly evident in AML (**Figure 1A**). Our combined analysis of the TCGA (https://cancergenome.nih.gov/) and GTEx (https://www.gtexportal.org/) databases revealed that the expression of SASH3 was significantly higher in patients with AML (P = 3.05e-34) (**Figure 1B**). The Kaplan-Meier survival curve shown in **Figure 1C** indicates that patients with AML characterized by high SASH3 expression had a worse prognosis and higher mortality rates than those with low SASH3 expression (P = 0.002). Moreover, when used as the sole determinant, SASH3 was observed to have predictive value regarding the survival of patients with AML at the 1-, 3-, and 5-year stages (**Figure 1D**).

GO and KEGG pathway enrichment analyses

Based on a comparison of the two co-expression datasets, we identified 380 commonly co-expressed genes (**Figure 2A**), for which we performed functional and pathway enrichment analyses using the WebGestalt web tool based on three functional categories: Molecular Function (MF), Cellular Component (CC), and Biological Process (BP). **Figure 2B** displays the results of the BP analysis, which revealed significant enrichment in the biological regulation, response to stimulus, and metabolic process categories. For CC, we identified membrane, cytosol, and vesicle as the top three enriched categories. In contrast, MF analysis revealed the importance of protein binding, ion binding, and transferase activity. KEGG analysis revealed that the top three enriched pathways were mast cell-mediated immunity, tolerance induction, and response to lipoprotein particles (**Figure 2C**).

Prognostic performance of SASH3 in AML

Cox univariate regression analysis revealed a significant correlation between SASH3 expres-

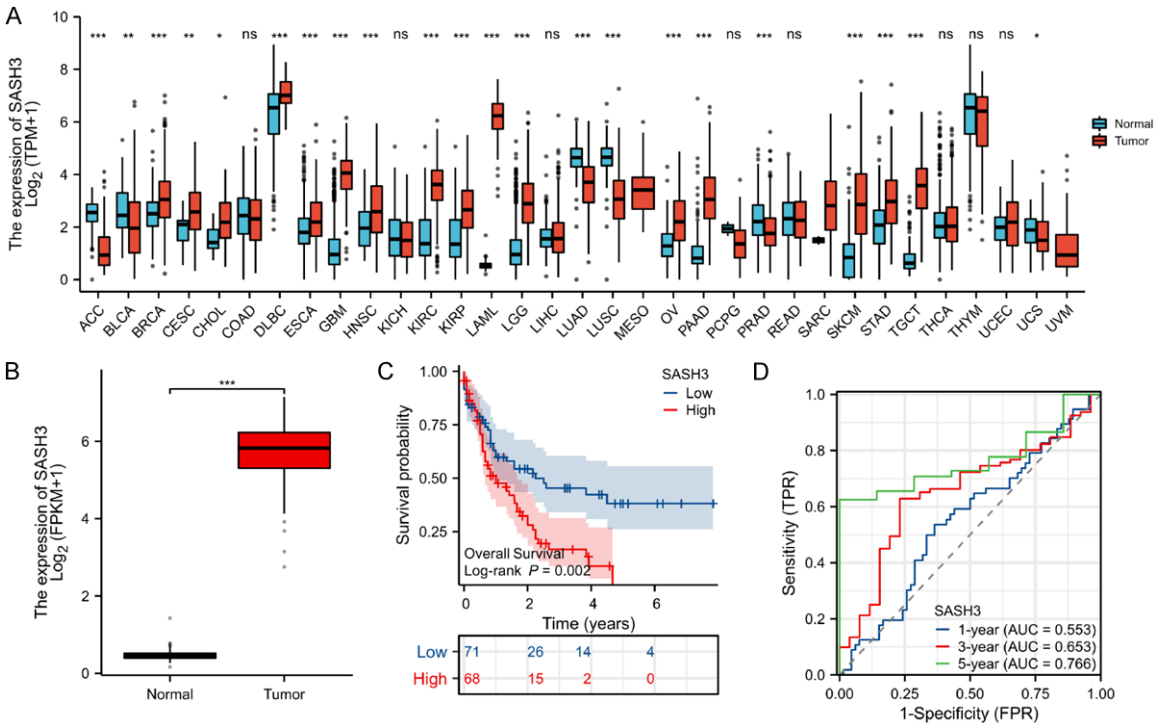

Figure 1. Expression of SASH3 is strongly associated with acute myelocytic leukemia (AML). A. SASH3 mRNA expression levels in different tumor types from the TIMER database. B. SASH3 is differentially expressed in normal and tumor patients. C. KM (Kaplan-Meier) survival curves based on SASH3 expression levels. D. Time-dependent ROC curves for survival prediction.

sion and the overall survival of patients with AML (hazard ratio [HR]: 2.033; 95% confidence interval (CI): 1.314-3.145; P = 0.001) (**Figure 3A**). Cox multifactorial regression analysis yielded similar results, identifying significant associations between overall survival and age, cytogenetic risk, and SASH3 (HR: 1.671; 95% CI: 1.028-2.715; P = 0.038) (**Figure 3B**). Additionally, we constructed a survival-validated nomogram plot based on this information (**Figure 3C**), which can contribute to an accurate prediction of survival among patients with AML (**Figure 3D**). Furthermore, classical univariate receiver operating characteristic (ROC) curve analysis revealed that this scoring criterion had a high area under the curve (AUC) of 0.807 (**Figure 3E**).

SASH3 and related hub genes may regulate AML via immune pathways

We used the 380 identified co-expressed genes from patients with AML to construct a protein-protein interaction network based on reference to the STRING database. To identify the most abundant interacting genes, we used Maximal Clique Centrality scores generated using the CytoHubba plugin and selected the top 10 genes, as shown in **Figure 4A**. These 10 genes (*VAV1, SYK, FYN, ITGB2, PIK3CD, FGR, PIK3R5, HCK, LCP2*, and *GRB2*) were identified as potential key regulators of the SASH3 network. We further analyzed the KEGG pathway of the central gene using WebGestalt, focusing mostly on immune and cellular material transport pathways, and the results were consistent with the findings of previous analyses (**Figure 4B**). To determine the levels of immune cell infiltration, we performed single-sample gene set enrichment analysis and Spearman analysis to assess the association between SASH3 and immune cell infiltration. SASH3 expression was accordingly observed to be positively correlated with the infiltration of iDC (R = 0.387, P < 0.001), TReg (R = 0.418, P < 0.001), and Th17 (R = 0.388, P < 0.001) cells, totaling 16 types of immune cells (**Figure 4C**).

Discussion

To the best of our knowledge, this study represents the first investigation of SASH3 mRNA

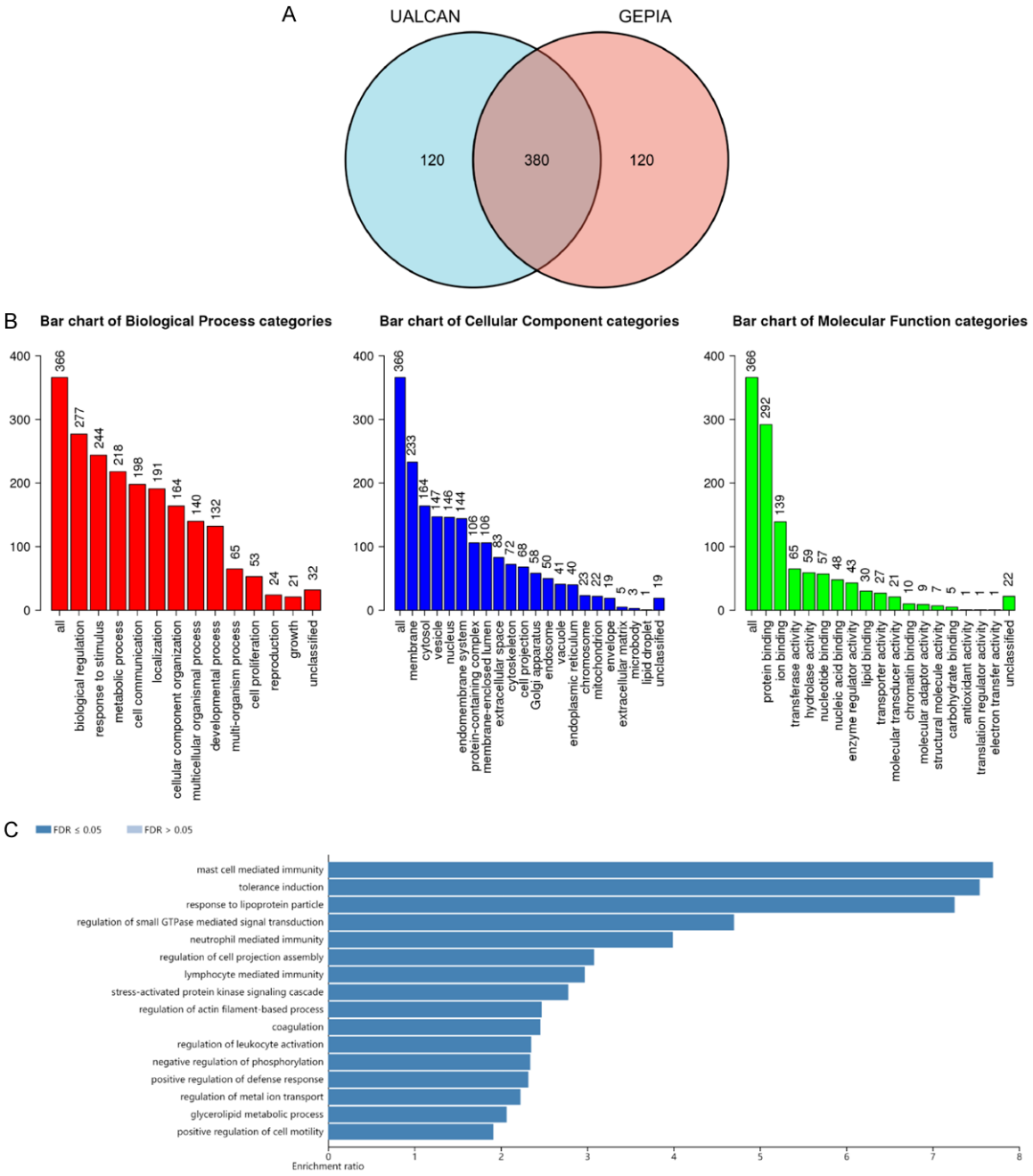

Figure 2. Enrichment analysis of KEGG and GO for genes co-expressed with SASH3 in acute myelocytic leukemia (AML). A. The Venn diagram represents the intersection of the top 500 positively corrected genes between the GEPIA and UALCAN databases. B. GO enrichment of co-expressed genes in biological processes, cellular components, and molecular functions. C. KEGG analysis based on SASH3 co-expressed genes.

expression and prognosis in AML. Previous studies have reported the associations between SASH3 and proliferative diseases and cancer in humans [12]. Emerging evidence indicates that SASH3, a key bridging protein, is a positive regulator of tumor immune responses [13]. SASH3 is involved in immune pathways in lung cancer, such as those associated with natural killer cell-mediated cytotoxicity, PD-L1 expression, and the PD-1 checkpoint pathway [9, 14, 15]. SASH3 promotes the immune response against lung tumor invasion, as evidenced by its low expression in normal and paracancerous tissues. Conversely, in patients

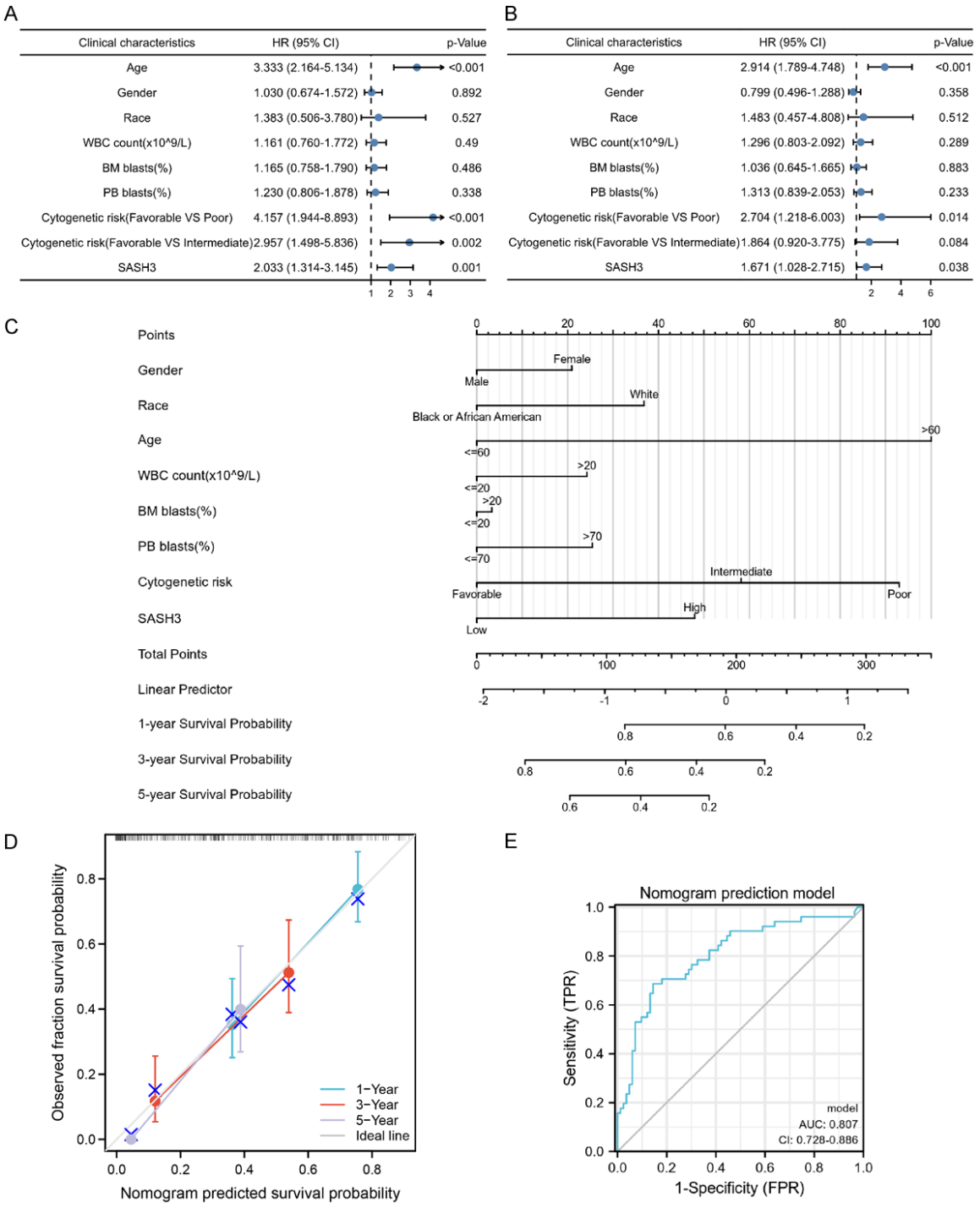

Figure 3. SASH3 could be an independent prognostic marker for acute myelocytic leukemia (AML). A. Univariate Cox regression analysis of SASH3 expression and clinicopathologic characteristics. B. Multivariate Cox regression analysis of SASH3 expression and clinicopathologic characteristics. C. Nomogram of SASH3 expression and clinicopathologic characteristics. D. Prognostic Calibration: the horizontal coordinate is the probability of survival predicted by the model; the vertical coordinate is the actual observed probability of survival. E. Receiver Operating Characteristic (ROC) Curve for Prognostic Scoring Incorporating important Factors.

with breast cancer, the expression of SASH3 is higher in cancerous tissues and normal and paraneoplastic tissues [16]. These findings indicate that SASH3 may be activated and play

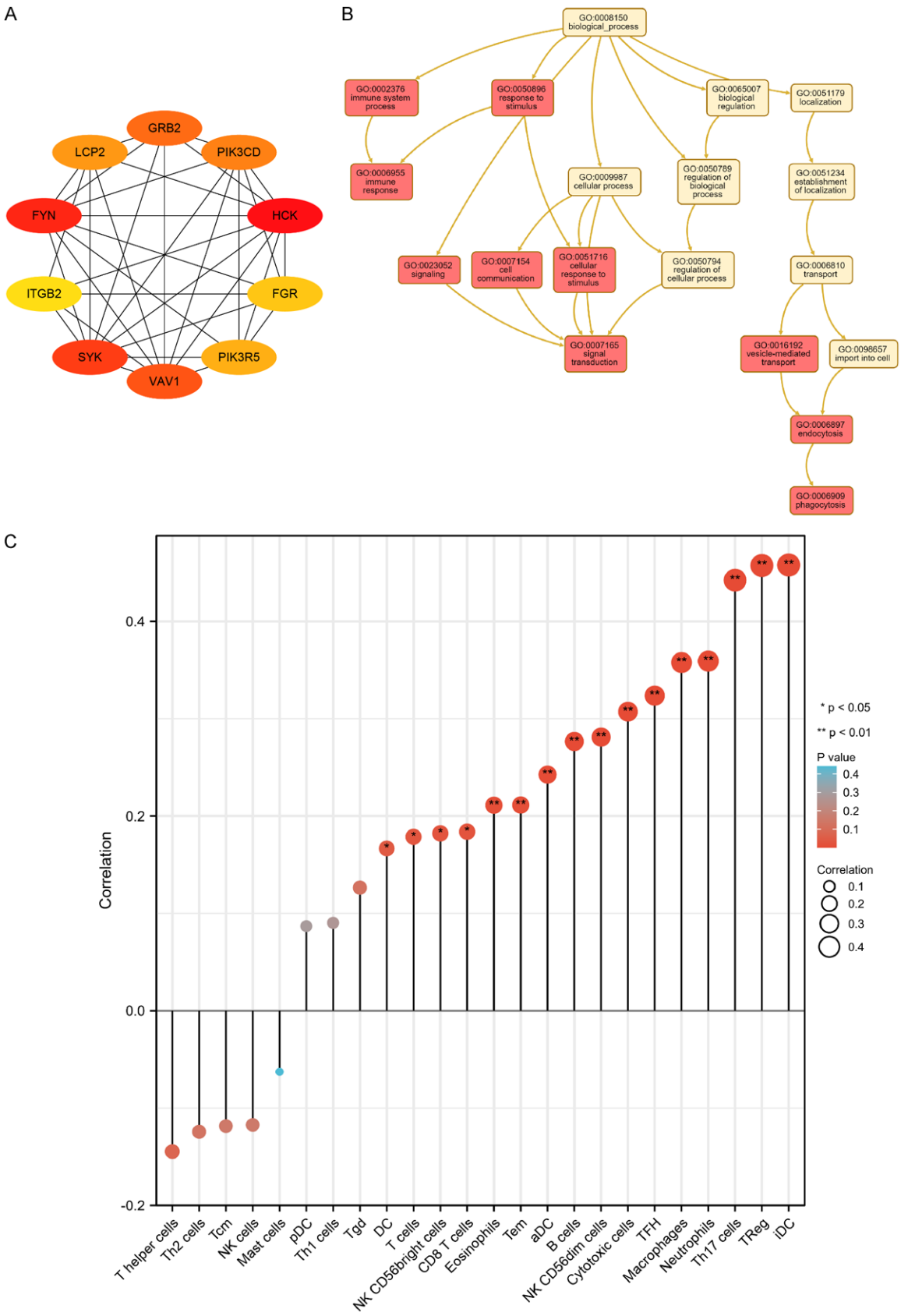

Figure 4. Protein-protein interaction (PPI) network of related genes reveals the correlation between immune cell infiltration and SASH3 expression in Acute Myelocytic Leukemia (AML) patients. A. The CytoHubba toolbox in Cytoscape

was used to identify the top 10 hub genes. B. Analysis of the hub gene enriched KEGG pathway using WebGestalt. C. Spearman's analysis of the relationships between infiltration levels of 24 immune cell types and SASH3 expression profiles.

a pivotal immune role in defending against invasive breast tumors. The findings of relevant prognostic studies indicate that higher SASH3 expression may be associated with more efficient clearance by the immune system and longer survival periods.

Our findings in this study have revealed that SASH3 is overexpressed in AML and that high SASH3 expression is strongly associated with a poor prognosis, thereby indicating its potential application value as a prognostic marker in AML, a disease characterized by abnormal proliferation, clonal expansion, abnormal differentiation, and reduced apoptosis of myeloid hematopoietic progenitors [17, 18]. Although a robust immune response is typically anticipated when aberrantly differentiated, long-lived myeloid hematopoietic progenitors undergo malignant transformation and uncontrollable proliferation, our findings indicate that despite the increased degree of immune cell infiltration, which SASH3 probably induces, the immune response fails to prevent tumorigenesis and spread, and instead contributes to an elevated cytogenetic risk. Given that high peripheral blood leukocyte count, and neutropenia are key characteristics of AML, it would be worth investigating whether SASH3 is associated with excess immunity or immune disorders and whether it can trigger the excessive elimination of defective neutrophils and normal cells by immune cells. However, few studies have examined the potential role of SASH3 in the development of the immune system and the signaling events involved in immune responses.

The findings of our pathway analysis revealed that the molecular pathways linked to SASH3 encompass immune, cellular bio-transport, and the regulation of small GTPase-mediated signal transduction. GTPase binding is implicated in numerous intracellular processes, including cytoskeletal dynamics, signal transduction, and membrane trafficking [19, 20]. Dysregulated expression of SASH3 is hypothesized to be associated with the pathogenesis of various cancer types, and this could be a plausible mechanism, whereas SASH3 influences the progression of AML.

In conclusion, our findings in this study provide preliminary insights into the potential role of SASH3 as a prognostic marker and therapeutic target in AML. However, it is important to note that these findings are based primarily on bioinformatics analyses and have yet to be experimentally validated through biochemical and animal studies. Accordingly, further research is required to verify our findings, elucidate the underlying mechanisms, and evaluate the therapeutic potential of targeting SASH3 in AML treatment.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (81900264). Many thanks to Dr. Jia Liu for supporting this study.

Disclosure of conflict of interest

None.

Address correspondence to: Zhengqi Qiu, Faculty of Medicine, Macau University of Science and Technology, Room PP-R203, Est. Seak Pai Van Praia Park, Rés-Do-Chão R, Coloane, Macau 999078, China. E-mail: caspiansev@163.com; 1800243014@email.szu.edu.cn

References

[1] Shallis RM, Wang R, Davidoff A, Ma X and Zeidan AM. Epidemiology of acute myeloid leukemia: recent progress and enduring challenges. Blood Rev 2019; 36: 70-87.

[2] Wang ES. Treating acute myeloid leukemia in older adults. Hematology Am Soc Hematol Educ Program 2014; 2014: 14-20.

[3] Yi M, Li A, Zhou L, Chu Q, Song Y and Wu K. The global burden and attributable risk factor analysis of acute myeloid leukemia in 195 countries and territories from 1990 to 2017: estimates based on the global burden of disease study 2017. J Hematol Oncol 2020; 13: 72.

[4] Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, Jemal A, Kramer JL and Siegel RL. Cancer treatment and survivorship statistics, 2019. CA Cancer J Clin 2019; 69: 363-385.

[5] Sekeres MA, Guyatt G, Abel G, Alibhai S, Altman JK, Buckstein R, Choe H, Desai P, Erba H,

Hourigan CS, LeBlanc TW, Litzow M, MacEachern J, Michaelis LC, Mukherjee S, O'Dwyer K, Rosko A, Stone R, Agarwal A, Colunga-Lozano LE, Chang Y, Hao Q and Brignardello-Petersen R. American Society of Hematology 2020 guidelines for treating newly diagnosed acute myeloid leukemia in older adults. Blood Adv 2020; 4: 3528-3549.

[6] Delmonte OM, Bergerson JRE, Kawai T, Kuehn HS, McDermott DH, Cortese I, Zimmermann MT, Dobbs AK, Bosticardo M, Fink D, Majumdar S, Palterer B, Pala F, Dsouza NR, Pouzolles M, Taylor N, Calvo KR, Daley SR, Velez D, Agharahimi A, Myint-Hpu K, Dropulic LK, Lyons JJ, Holland SM, Freeman AF, Ghosh R, Similuk MB, Niemela JE, Stoddard J, Kuhns DB, Urrutia R, Rosenzweig SD, Walkiewicz MA, Murphy PM and Notarangelo LD. SASH3 variants cause a novel form of X-linked combined immunodeficiency with immune dysregulation. Blood 2021; 138: 1019-1033.

[7] Beer S, Simins AB, Schuster A and Holzmann B. Molecular cloning and characterization of a novel SH3 protein (SLY) preferentially expressed in lymphoid cells. Biochim Biophys Acta 2001; 1520: 89-93.

[8] Labrador-Horrillo M, Franco-Jarava C, Garcia-Prat M, Parra-Martínez A, Antolín M, Salgado-Perandrés S, Aguiló-Cucurull A, Martinez-Gallo M and Colobran R. Case report: X-linked SASH3 deficiency presenting as a common variable immunodeficiency. Front Immunol 2022; 13: 881206.

[9] Chen X, Yuan Y, Ren W, Zhou F, Huang X, Pu J, Niu X and Jiang X. Pan-cancer integrated analysis identification of SASH3, a potential biomarker that inhibits lung adenocarcinoma progression. Front Oncol 2022; 12: 927988.

[10] Bai M, Pan Q and Sun C. Tumor purity coexpressed genes related to immune microenvironment and clinical outcomes of lung adenocarcinoma. J Oncol 2021; 2021: 9548648.

[11] Guan M, Jiao Y and Zhou L. Immune infiltration analysis with the CIBERSORT method in lung cancer. Dis Markers 2022; 2022: 3186427.

[12] Priebe V, Sartori G, Napoli S, Chung EYL, Cascione L, Kwee I, Arribas AJ, Mensah AA, Rinaldi A, Ponzoni M, Zucca E, Rossi D, Efremov D, Lenz G, Thome M and Bertoni F. Role of ETS1 in the transcriptional network of diffuse large B cell lymphoma of the activated B cell-like type. Cancers (Basel) 2020; 12: 1912.

[13] Jaufmann J, Franke FC, Sperlich A, Blumendeller C, Kloos I, Schneider B, Sasaki D, Janssen KP and Beer-Hammer S. The emerging and diverse roles of the SLy/SASH1-protein family in health and disease-overview of three multifunctional proteins. FASEB J 2021; 35: e21470.

[14] Zhang TJ, Cao JS, Li Q, An GS, Ni JH, Jia HT and Li SY. A systematic analysis of dysregulated long non-coding RNAs/microRNAs/mRNAs in lung squamous cell carcinoma. Am J Med Sci 2020; 360: 701-710.

[15] Li J, Du Q, Sun J, Xiang L and Wang S. Identification and validation of a novel phagocytosis regulators-related signature with potential prognostic and immunotherapeutic value in patients with lung adenocarcinoma. Front Oncol 2022; 12: 988332.

[16] Shi Y, Chen S, Xing H, Jiang G, Wu N, Liu Q, Sakamoto N, Kuno T, Sugiura R, Xiao Q, Jin F, Fang Y and Yao F. Comprehensive analysis of prognostic microenvironment-related genes in invasive breast cancer. Front Oncol 2022; 11: 576911.

[17] Aul C, Bowen DT and Yoshida Y. Pathogenesis, etiology and epidemiology of myelodysplastic syndromes. Haematologica 1998; 83: 71-86.

[18] Doucette K, Karp J and Lai C. Advances in therapeutic options for newly diagnosed, high-risk AML patients. Ther Adv Hematol 2021; 12: 20406207211001138.

[19] Aslan JE. Platelet Rho GTPase regulation in physiology and disease. Platelets 2019; 30: 17-22.

[20] Prieto-Dominguez N, Parnell C and Teng Y. Drugging the small GTPase pathways in cancer treatment: promises and challenges. Cells 2019; 8: 255.