

The Prognostic Value of RASGEF1A RNA Expression and DNA Methylation in Cytogenetically Normal Acute Myeloid Leukemia

Xue He

Capital Medical University

Weilong Zhang

Peking University Third Hospital

Wei Fu

Peking University Third Hospital

Xiaoni Liu

The First Affiliated Hospital of Gannan Medical University

Ping Yang

Peking University Third Hospital

Jing Wang

Peking University Third Hospital

Shaoliang Li

Capital Medical University

Xiuru Zhang

Capital Medical University

Gehong Dong

Capital Medical University

Changjian Yan

Gannan Medical University

Yali Zhao

Gannan Medical University

Hongmei Jing (✉ hongmeijing@bjmu.edu.cn)

Peking University Third Hospital

Research Article

Keywords: cytogenetically normal acute myeloid leukemia, RASGEF1A, expression, DNA methylation, prognosis

Posted Date: July 29th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-754409/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 significantly heterogeneous malignancy of the blood. Cytogenetic abnormalities are crucial for the prognosis of AML. However, since more than half of patients with AML are cytogenetically normal AML (CN-AML), new markers are badly required for predicting prognosis. In recent years, gene abnormalities are considered to be strong prognostic factors of CN-AML, already having clinical significance for treatment. In addition, the relationship of methylation in some genes and AML prognosis predicting has been discovered. RASGEF1A is a guanine nucleotide exchange factors of Ras, and has been reported to relate to malignancy. However, there is no research about the relationship of RASGEF1A gene and CN-AML.

Methods

By integrating the Cancer Genome Atlas (TCGA) database 75 patients with CN-AML and 240 Gene Expression Omnibus (GEO) database CN-AML samples, we examined the association between RASGEF1A's RNA expression level and DNA methylation of and AML patients' prognosis. Then, we investigated the RASGEF1A RNA expression and DNA methylation's prognostic value in 77 patients with AML after allogeneic hematopoietic stem cell transplantation (Allo-HSCT) as well as 101 AML patients after chemotherapy respectively. We investigated the association between sensitivity to Crenolanib and expression level of RASGEF1A in patients by integrating 191 CN-AML patients from BeatAML dataset. Finally, we integrated the expression and methylation of RASGEF1A to predict the CN-AML patients' prognosis.

Results

We found that RASGEF1A gene high expression group predicted poorer event-free survival (EFS) ($P < 0.0001$) as well as overall survival (OS) ($P < 0.0001$) in CN-AML samples, and the identical results were found in AML patients receiving chemotherapy ($P < 0.0001$) and Allo-HSCT ($P < 0.0001$). RASGEF1A RNA expression level is an CN-AML patients' independent prognostic factor (EFS: HR = 5.5534, 95% CI: 1.2982–23.756, $P = 0.0208$; OS: HR = 5.3615, 95% CI: 1.1014–26.099, $P = 0.0376$). The IC₅₀ (half maximal inhibitory concentration) of Crenolanib of CN-AML samples with RASGEF1A high expression level is lower. In addition, patients with high RASGEF1A methylation level had significant favorable prognosis (EFS: $P < 0.0001$, OS: $P < 0.0001$). Furthermore, the integrative analysis of expression and methylation of RASGEF1A could classify CN-AML patients into subgroups with different prognosis (EFS: $P < 0.0001$, OS: $P < 0.0001$).

Conclusion

Higher RASGEF1A RNA expression and lower DNA methylation predicts CN-AML patients' poorer prognosis. The RASGEF1A high expression level from patients with CN-AML have better sensitivity to Crenolanib. The integrative analysis of RASGEF1A RNA expression and DNA methylation can provide a more accurate classification for prognosis. Lower RASGEF1A expression is a favorable prognostic factor for AML patients receiving chemotherapy or Allo-HSCT.

Introduction

Acute myeloid leukemia (AML) is a highly heterogeneous malignancy of the blood in genetic basis and prognosis. The main feature of AML is the clonal proliferation of myeloid blasts in bone marrow, peripheral blood and/or other tissues, which may cause impaired hematopoiesis and bone marrow failure, with symptoms like frequent infections, easy bruising or bleeding, anemia, bone or abdominal pain, etc.[1–3] The incidence of AML is approximately 4.3 per 100,000 persons in the United States and 3.7 per 100,000 persons in Europe.[4] As a disease commonly incident in elderly people, the median diagnostic age of AML is 67 years old, and nearly one-third of AML patients are diagnosed at the age of 75 or more.[1] For a long time, the common chemotherapy of AML has been “3 + 7”, which are 3 days of daunorubicin or idarubicin combining 7 days of cytarabine to acquire complete remission (CR).[2, 5] 70–80% of patients under 60 years old can enter CR after receiving several anthracycline-based chemotherapy combinations. The situation is worse in older patients. Only approximately 40–65% of older patients will achieve CR, but 85% will relapse in 2 to 3 years.[2] In the Medical Research Council (MRC) AML11 trial, the AML patients receiving DAT (thioguanine, cytarabine, and daunorubicin) have a 5-year overall survival (OS) of 12%, the patients with AML receiving ADE (etoposide, cytarabine, and daunorubicin) have a 5-year OS of 8% and the patients with AML receiving MAC (mitoxantrone-cytarabine) have a 5-year OS of 10%.[6] In the Leukemia Research fund (LRF) AML14 trial, the 1-year survival and 3-years survival of 1137 intensively treated patients were 47% and 19%, and the rates of 275 non-intensively treated patients were 19%, 2%.[7] Another study showed that only 40 ~ 50% of patients over 65 years old can achieve CR. The 5-year survival rate for young AML patients is 40% and is only 5% in elderly patients.[4] Since common treatments are not effective for some AML patients, precision medicine is crucial for AML, which means more prognostic factors are badly required.

Cytogenetic abnormalities have long been considered one of the most significant factors in the prognosis of AML, which have been affirmed in clinical practice, for instance patients with inv(16), t(15; 17) as well as t(8; 21) have better estimate.[8, 9] However, nearly one-half of the AML patients are cytogenetically normal (CN), cannot predicting prognosis by cytogenetic markers.[9] In recent year, several genetic mutations have been described, providing more possibility for predicting prognosis.[9, 10] The 4-years survival rate of AML patients with CEBPA mutations is nearly 60%, and AML patients with NPM1 mutation have a 50% 4-year survival rate approximately. Patients with mutations of the CEBPA or NPM1 genes get a favorable prognosis, and BAALC, MLL-PTD and CEBPA can help to predict the response of patients obtaining allogeneic hematopoietic stem cell transplantation (ALLO-HSCT). FLT3–Internal Tandem Duplication (FLT3-ITD) patients own an adverse prognosis, whose 4-years survival rate is only 20–25%.[11–13] Mutations such as RUNX1, DNMT3A and ASXL1 are cytogenetically normal AML (CN-

AML) patients' adverse prognostic factors.[14–16] Above findings illustrated that gene abnormalities are extreme valuable prognostic factors for CN-AML, providing research directions for us to find precise risk stratifications.

DNA methylation is a direct chemical modification of DNA that form a 5-methylcytosine through transferring a methyl group to the cytosine's C5 position. DNA methylation controls gene expression mostly through recruiting proteins can inhibit expression of gene or impairing the transcriptional activators' binding, mainly to reduce gene expression.[17] In addition, methylation also exerts its effects by other mechanisms, including splicing variants, down regulating microRNA, fostering DNA rearrangements, etc.[18–20] Previous studies have demonstrated that DNA methylation levels are closely related to the AML patients' prognosis. For example, high GADD45A methylation level can predict poor survival of AML patients, and the decrease of SCIN expression due to promoter methylation results in adverse prognosis of AML patients, while high HOXA5 methylation levels are valuable biomarkers for prognosis in AML patients.[21–23] Methylation further affects the prognosis of disease through several mechanisms, indicating that gene methylation also has important research significance in AML patients.

RasGEF1A belongs to RasGEF family and is capable of regulating the transformation and activation of Rap2A, a member of Ras family.[24] Studies showed that Rap2a is a direct target of p53, and the overexpression of Rap2a may promote the invasion and migration of many cancers, including lung cancer, renal cell carcinoma, osteosarcoma, nasopharyngeal carcinoma, and gliomas of the central nervous system.[25–28] A study exhibited that the expression of RASGEF1A is significant for intrahepatic cholangiocarcinoma cells' survival and migration, while the suppression of RASGEF1A inhibits the growth of ICC cells.[29] It has been reported that RASGEF1A is highly expressed in brain, moderately expressed in bone marrow, spleen, lymph node, testis and gall bladder, weakly expressed in other 21 examined tissues. [30] But the prognostic significance of the expression and methylation of RASGEF1A genes in CN-AML has not been reported so far. Our study integrated CN-AML patients' two independent groups to studied the prognosis value of RASGEF1A RNA expression level and DNA methylation level in CN-AML patients.

Methods

Data source

We integrated RNA expression profiles and DNA methylation profiles of 75 the Cancer Genome Atlas (TCGA) CN-AML samples and 240 Gene Expression Omnibus (GEO) CN-AML patients.[31, 32] We analyzed the relation between survival and RASGED1A RNA expression through 75 TCGA patients CN-AML and 240 GSE12417 CN-AML. We analyzed the relation between post-treatment survival rates and RASGED1A RNA expression of 67 patients with CN-AML receiving Allo-HSCT and 92 CN-AML patients with only chemotherapy from TCGA database. We analyzed the relation between the prognosis and the level of RASGED1A methylation by analyzing 85 CN-AML patients, 77 AML patients with Allo-HSCT and 101 AML patients with only chemotherapy from TCGA database. We investigated the association between sensitivity to Crenolanib and expression level of RASGED1A in patients by integrating 191 CN-

AML patients from BeatAML dataset.[33] This research was in consistent with the Declaration of Helsinki.

AML patients were selected established on the following criteria. The 75 TCGA AML patients were normal karyotype and received chemotherapy or transplant treatment. Their gene expression was measured using RNA-seq. The 92 AML patients from TCGA database were normal karyotype or abnormal karyotype and received chemotherapy treatment. Their gene expression was measured using RNA-seq. The 67 AML patients from TCGA database were normal karyotype or abnormal karyotype and received Allo-HSCT treatment. Their gene expression was measured using RNA-seq. The 162 AML of GSE12417 U133B dataset and the 78 AML of GSE12417 U133 Plus dataset were normal karyotype and received chemotherapy or transplant treatment. Their gene expression was measured using microarray. The 191 CN-AML patients from BeatAML were normal karyotype received chemotherapy or transplant treatment. Their gene expression was measured using RNA-seq.

RNA expression profiles and DNA methylation profiles analysis

RNA expression microarray of each CN-AML sample from GEO database was calculated through the method of robust multiarray averaging method (RMA). The expression levels of each probe were transformed with \log_2 . RNA expression data (RNA-seq) and DNA methylation data (HumanMethylation450 chip) were obtained from TCGA database. RNA expression levels of each gene were displayed with RPKM (Reads Per Kilobase per Million mapped reads). It was transformed with $\log_2(\text{FPKM} + 1)$. The methylation of each gene was calculated from the probes related to this gene and was showed from 0 to 1 (0 for no methylation and 1 from 100% of DNA methylation). The P-value of event-free survival (EFS) and OS for each gene were screened from the RNA expression profiles or DNA methylation profiles of each dataset. Patients of each dataset were classified into RASGEF1A-high expression and RASGEF1A-low expression group with the RASGEF1A RNA expression or RASGEF1A DNA methylation with survminer package.

Statistics

The statistical results were analyzed through ggplot2 and survminer packages of R software v3.1.3. Survival curves were designed by Kaplan-Meier estimation. Survival analysis was performed by the Log-rank test. The unpaired t test was employed to compare the average values of two groups. The Fisher's exact test was employed for enumeration data of two or more groups. Cox proportional hazards regression models of multivariate analysis were employed to make a judgement about the prognosis value of multiple biomarkers related to CN-AML.

Results

The 75 TCGA CN-AML samples' baseline characteristics

We examine the baseline characteristics of the 75 TCGA CN-AML including 10 CN-AML patients with RASGEF1A-low expression and 65 patients with RASGEF1A-high expression and 240 GSE12417 CN-AML patients including 162 from GSE12417 U133B and 78 from GSE12417 U133 Plus. In 75 TCGA CN-AML patients, we found that the baseline features of the two groups matched. Except for recurrence ($P = 0.034$, Fisher's Exact test), other indicators were not statistically significant. These indicators not only included gender, race, FAB type, nuclear type, risk, induction, transplantation, pre-transplantation, gene mutations such as DNMT3A, NPM1, NET2 ($P > 0.05$, Fisher's exact test), but also included age, bone marrow blast cell (BM_BLAST), peripheral blood WBC (WBC), peripheral blood blast cell (PB_BLAST, $P > 0.05$, unpaired t-test, Table 1). In GSE12417 U133B including 21 patients with low RASGEF1A expression and 141 patients with high RASGEF1A expression we found that the baseline features of two groups are match including French-American-British (FAB) type ($P > 0.05$, Fisher's Exact test) and age ($P > 0.05$, unpaired t-test). In 78 from GSE12417 U133 Plus including 12 patients with low RASGEF1A expression and 66 patients with high RASGEF1A expression, we found that the baseline features of two groups are match including FAB type ($P > 0.05$, Fisher's Exact test) and age ($P > 0.05$, unpaired t-test, Table S1 and S2).

Table 1
Baseline CN-AML patient characteristics in the RASGEF1A-high expression and RASGEF1A-low expression group from TCGA database.

Characteristics	Level	RASGEF1A-low	RASGEF1A-high	P-value
n		10	65	
Sex (%)	Female	6 (60.0)	32 (49.2)	0.736
	Male	4 (40.0)	33 (50.8)	
Race (%)	Black	0 (0.0)	3 (4.6)	0.82
	Others	0 (0.0)	1 (1.5)	
	Unknown	1 (10.0)	12 (18.5)	
	White	9 (90.0)	49 (75.4)	
FAB (%)	M0	0 (0.0)	3 (4.6)	0.968
	M1	4 (40.0)	20 (30.8)	
	M2	2 (20.0)	17 (26.2)	
	M4	2 (20.0)	13 (20.0)	
	M5	2 (20.0)	10 (15.4)	
	M7	0 (0.0)	1 (1.5)	
	Unknown	0 (0.0)	1 (1.5)	
Age (mean (sd))		56.60 (14.01)	54.25 (17.55)	0.687
BM_BLAST (mean (sd))		73.40 (16.74)	70.32 (18.68)	0.625
WBC (mean (sd)) × 10 ⁹ /L		20.79 (24.69)	55.77 (58.08)	0.065
PB_BLAST (mean (sd))		36.40 (32.12)	45.86 (33.59)	0.407
Karyotype (%)	Normal	10 (100.0)	65 (100.0)	NA
Risk (%)	Intermediate	10 (100.0)	65 (100.0)	NA
Induction (%)	7 + 3	7 (70.0)	24 (36.9)	0.807
	7 + 3 + 3	2 (20.0)	17 (26.2)	
	7 + 3 + 3 + others	0 (0.0)	3 (4.6)	
	7 + 3 + others	0 (0.0)	8 (12.3)	

n, number of patients; BM_BLAST, bone marrow blast cell; WBC, peripheral blood WBC; PB_BLAST, peripheral blood blast cell; MUD, HSCT of matched unrelated donor; Haplo, haploidentical HSCT; WT, wild type. Unpaired t test (two sides), was used in two group measurement data. ANOVA test was used in in multi groups' measurement data. Fisher's exact test was used in enumeration data.

Characteristics	Level	RASGEF1A-low	RASGEF1A-high	P-value
	Cytarabine	0 (0.0)	1 (1.5)	
	Decitabine	1 (10.0)	4 (6.2)	
	Decitabine + others	0 (0.0)	1 (1.5)	
	Hydrea + others	0 (0.0)	1 (1.5)	
	no treatment	0 (0.0)	1 (1.5)	
	Others	0 (0.0)	1 (1.5)	
	Revlimid	0 (0.0)	4 (6.2)	
Transplant (%)	Auto	1 (10.0)	3 (4.6)	0.65
	Chemotherapy	5 (50.0)	32 (49.2)	
	Haplo	0 (0.0)	1 (1.5)	
	MUD	1 (10.0)	15 (23.1)	
	no treatment	0 (0.0)	1 (1.5)	
	sib Allo	3 (30.0)	13 (20.0)	
Before_transplant (%)	CR 1	3 (30.0)	15 (23.1)	1
	CR 2	1 (10.0)	6 (9.2)	
	CR 3	0 (0.0)	1 (1.5)	
	No transplant	5 (50.0)	32 (49.2)	
	no treatment	0 (0.0)	1 (1.5)	
	Others	1 (10.0)	7 (10.8)	
	Rel 1	0 (0.0)	2 (3.1)	
	Rel 2	0 (0.0)	1 (1.5)	
Relapse (%)	No	7 (70.0)	21 (32.3)	0.034
	Yes	3 (30.0)	44 (67.7)	
DNMT3A (%)	Mutation	4 (40.0)	25 (38.5)	1
	Unknown	0 (0.0)	1 (1.5)	

n, number of patients; BM_BLAST, bone marrow blast cell; WBC, peripheral blood WBC; PB_BLAST, peripheral blood blast cell; MUD, HSCT of matched unrelated donor; Haplo, haploidentical HSCT; WT, wild type. Unpaired t test (two sides), was used in two group measurement data. ANOVA test was used in in multi groups' measurement data. Fisher's exact test was used in enumeration data.

Characteristics	Level	RASGEF1A-low	RASGEF1A-high	P-value
	WT	6 (60.0)	39 (60.0)	
NPM1 (%)	Mutation	7 (70.0)	36 (55.4)	0.571
	Unknown	0 (0.0)	1 (1.5)	
	WT	3 (30.0)	28 (43.1)	
TET2 (%)	Mutation	2 (20.0)	7 (10.8)	0.653
	Unknown	0 (0.0)	1 (1.5)	
	WT	8 (80.0)	57 (87.7)	
FLT3 (%)	Mutation	2 (20.0)	28 (43.1)	0.295
	Unknown	0 (0.0)	1 (1.5)	
	WT	8 (80.0)	36 (55.4)	
IDH2 (%)	Mutation	3 (30.0)	7 (10.8)	0.245
	Unknown	0 (0.0)	1 (1.5)	
	WT	7 (70.0)	57 (87.7)	
IDH1 (%)	Mutation	3 (30.0)	6 (9.2)	0.218
	Unknown	0 (0.0)	1 (1.5)	
	WT	7 (70.0)	58 (89.2)	
RUNX1 (%)	Mutation	0 (0.0)	7 (10.8)	0.639
	Unknown	0 (0.0)	1 (1.5)	
	WT	10 (100.0)	57 (87.7)	
NRAS (%)	Mutation	2 (20.0)	4 (6.2)	0.293
	Unknown	0 (0.0)	1 (1.5)	
	WT	8 (80.0)	60 (92.3)	
WT1 (%)	Mutation	1 (10.0)	4 (6.2)	0.59
	Unknown	0 (0.0)	1 (1.5)	
	WT	9 (90.0)	60 (92.3)	

n, number of patients; BM_BLAST, bone marrow blast cell; WBC, peripheral blood WBC; PB_BLAST, peripheral blood blast cell; MUD, HSCT of matched unrelated donor; Haplo, haploidentical HSCT; WT, wild type. Unpaired t test (two sides), was used in two group measurement data. ANOVA test was used in in multi groups' measurement data. Fisher's exact test was used in enumeration data.

Characteristics	Level	RASGEF1A-low	RASGEF1A-high	P-value
CEBPA (%)	Mutation	0 (0.0)	8 (12.3)	0.643
	Unknown	0 (0.0)	1 (1.5)	
	WT	10 (100.0)	56 (86.2)	
PTPN11 (%)	Mutation	0 (0.0)	5 (7.7)	1
	Unknown	0 (0.0)	1 (1.5)	
	WT	10 (100.0)	59 (90.8)	
KRAS (%)	Mutation	1 (10.0)	2 (3.1)	0.443
	Unknown	0 (0.0)	1 (1.5)	
	WT	9 (90.0)	62 (95.4)	

n, number of patients; BM_BLAST, bone marrow blast cell; WBC, peripheral blood WBC; PB_BLAST, peripheral blood blast cell; MUD, HSCT of matched unrelated donor; Haplo, haploidentical HSCT; WT, wild type. Unpaired t test (two sides), was used in two group measurement data. ANOVA test was used in in multi groups' measurement data. Fisher's exact test was used in enumeration data.

RASGEF1A expression can forecast the prognosis of CN-AML

To address the correlation between RASGEF1A expression level and survival rate of AML samples, we conducted a study. We examine RASGEF1A expression level and prognosis in 75 TCGA CN-AML patients and 240 GSE12417 CN-AML patients. We found 10 patients with low RASGEF1A expression from TCGA CN-AML has a better EFS and OS ($P < 0.0001$, Log-rank test, Figure. 1A). In CN-AML patients from GSE12417, 21 patients with low RASGEF1A expression from GSE12417 U133B ($P < 0.0001$) and 12 patients with low RASGEF1A expression from GSE12417 U133 Plus dataset ($P < 0.0001$, Figure. 1B) have a improved OS.

The 75 TCGA CN-AML samples' multivariate analysis

Table 2 clearly illustrates that the expression level of RASGEF1A is a robust independent risk factor for event-free survival (HR = 5.5534, 95% CI: 1.2982–23.756, $P = 0.0208$), and overall survival (HR = 5.3615, 95% CI: 1.1014–26.099, $P = 0.0376$, Cox regression analysis, Table 2) of patients with CN-AML. Other relatively high risk factors are age (OS: HR = 2.4264, 95% CI: 1.1372–5.177, $P = 0.0219$), DNMT3A mutation (EFS: HR = 3.1185, 95% CI: 1.4403–6.752, $P = 0.0039$; OS: HR = 2.35, 95% CI: 1.1763–4.695, $P = 0.0155$), FLT3 mutation (EFS: HR = 3.5715, 95% CI: 1.3287–9.6, $P = 0.0116$), KRAS mutation (EFS: HR = 5.8517, 95% CI: 1.1686–29.302, $P = 0.0316$). Hazard ratios are not significantly high in clinical features like BM_BLAST, PB_BLAST and other gene mutations such as NPM1, IDH2, IDH1.

Table 2
Multivariate analysis for EFS and OS in the TCGA CN-AML.

Variables	EFS				OS			
	HR	Lower 95%	Upper 95%	P-value	HR	Lower 95%	Upper 95%	P-value
Age (≥ 60 vs. < 60 years)	1.9254	0.855	4.336	0.11372	2.4264	1.13724	5.177	0.0219
BM_BLAST ($\geq 70\%$ vs. $< 70\%$)	0.5537	0.2369	1.294	0.1724	0.9302	0.45074	1.92	0.8447
WBC (≥ 30 vs. $< 30 \times 10^9/L$)	1.019	0.4282	2.425	0.96612	1.1347	0.54283	2.372	0.7369
PB_BLAST ($\geq 50\%$ vs. $< 50\%$)	1.8751	0.7424	4.736	0.18356	1.2148	0.50991	2.894	0.6604
DNMT3A (Mutation vs. WT)	3.1185	1.4403	6.752	0.00391	2.35	1.17632	4.695	0.0155
NPM1 (Mutation vs. WT)	0.6069	0.2248	1.639	0.32449	1.2705	0.50758	3.18	0.609
TET2 (Mutation vs. WT)	0.8438	0.3097	2.298	0.73967	0.4217	0.13443	1.323	0.1389
FLT3 (Mutation vs. WT)	3.5715	1.3287	9.6	0.01162	0.8792	0.37612	2.055	0.7664
IDH2 (Mutation vs. WT)	1.0288	0.2805	3.774	0.96584	0.6492	0.20394	2.067	0.4646
IDH1 (Mutation vs. WT)	1.8848	0.403	8.816	0.42067	0.6076	0.14666	2.518	0.4922
RUNX1 (Mutation vs. WT)	1.5279	0.4063	5.746	0.53054	1.9235	0.5451	6.787	0.3093
NRAS (Mutation vs. WT)	0.5721	0.133	2.46	0.45303	0.3369	0.06415	1.77	0.1986

BM_BLAST, bone marrow blast cell; WBC, peripheral blood WBC; PB_BLAST, peripheral blood blast cell; RASGEF1A, RASGEF1A RNA expression; EFS, event-free survival time; OS, overall survival time;

Variables	EFS				OS			
	HR	Lower 95%	Upper 95%	P-value	HR	Lower 95%	Upper 95%	P-value
WT1 (Mutation vs. WT)	2.2595	0.5181	9.854	0.27798	1.3055	0.35175	4.845	0.6903
CEBPA (Mutation vs. WT)	1.3225	0.3567	4.904	0.67591	1.1164	0.31616	3.942	0.8642
PTPN11 (Mutation vs. WT)	0.9007	0.2056	3.947	0.88971	1.2873	0.36355	4.558	0.6954
KRAS (Mutation vs. WT)	5.8517	1.1686	29.302	0.03159	1.5696	0.34934	7.053	0.5565
RASGEF1A (High vs. Low)	5.5534	1.2982	23.756	0.02078	5.3615	1.10137	26.099	0.0376

BM_BLAST, bone marrow blast cell; WBC, peripheral blood WBC; PB_BLAST, peripheral blood blast cell; RASGEF1A, RASGEF1A RNA expression; EFS, event-free survival time; OS, overall survival time;

RASGEF1A expression level can forecast the prognosis of AML samples received chemotherapy or Allo-HSCT

To further understand the correlation between RASGEF1A expression and the prognosis of samples with AML received chemotherapy or Allo-HSCT, we compared the survival rates of patients with low and high RASGEF1A expression received treatment. In 67 AML patients after Allo-HSCT, RASGEF1A-low expression group had more favorable prognosis than RASGEF1A-high expression group (EFS: $P < 0.0001$, OS: $P < 0.0001$, Log-rank test, Figure. 2A). The same result was revealed in the study of 92 AML patients from TCGA database after chemotherapy, patients in RASGEF1A-low expression group had a higher survival rate than RASGEF1A-high expression group (EFS: $P < 0.0001$, OS: $P < 0.0001$, Log-rank test, Figure. 2B).

Methylation of RASGEF1A can forecast the prognosis of samples with CN-AML

We detected the methylation levels of RASGEF1A at the same time. In 85 TCGA CN-AML patients, the survival rate was elevated in RASGEF1A-high methylation group (EFS: $P < 0.0001$, OS: $P < 0.0001$, Log-rank test, Figure. 3A). Similarly in the study of 77 AML patients received Allo-HSCT, RASGEF1A-high methylation group had significant favorable prognosis (EFS: $P < 0.0001$, OS: $P < 0.0001$, Log-rank test, Figure. 3B). Additionally in the study of 101 patients received chemotherapy, RASGEF1A-high methylation

group had a higher survival rate than RASGEF1A-low methylation group (EPS: $P < 0.0001$, OS: $P < 0.0001$, Log-rank test, Figure. 3C).

Combination of expression and methylation levels of RASGEF1A provide a accurate prognostic classification for patients with CN-AML

We comprehensively examine *expression and methylation levels* of *RASGEF1A* of 75 CN-AML patients and obtained a more accurate method for prognosis classification. As shown in Fig. 4, 75 patients were divided into 4 groups, which are RASGEF1A-high expression RASGEF1A-high methylation group (G1), RASGEF1A-high expression RASGEF1A-low methylation group (G2), RASGEF1A-low expression RASGEF1A-high methylation group (G3) and RASGEF1A-low expression RASGEF1A-low methylation group (G4). In EFS, G3 had the best prognosis, G1 and G4 had moderate prognosis and G2 had the worst prognosis among the 4 groups (EFS: $P < 0.0001$, Log-rank test, Figure. 4). Due to the small sample size, G4 in right side of Fig. 4 had only 1 patient, which may have individual differences. Removing this point, G3 had the best prognosis, G1 had moderate prognosis and G2 had the worst prognosis among the 3 groups (OS: $P < 0.0001$, Log-rank test, Figure. 4).

The BeatAML CN-AML samples with RASGEF1A high expression level with better sensitivity to Crenolanib

To investigate the relationship between RASGEF1A expression level and sensitivity to Crenolanib in CN-AML patients, we integrated 191 CN-AML patients from BeatAML including 50 RASGEF1A low expression and 141 RASGEF1A high expression. We found the IC50 of Crenolanib of CN-AML patients with RASGEF1A high expression level is lower, which suggest RASGEF1A-high has better sensitive to Crenolanib (Figure S1).

Discussion

With the gradual development of research on AML, the heterogeneity in the disease is becoming more common, and cytogenetics is increasingly recognized as an independent predictor of prognosis in AML patients. However, though many prognostic factors have been identified, there is more than 50% of patients do not have cytogenetic marker, which are cytogenetically normal AML patients.[9, 10] The good news is many gene abnormalities that escape cytogenetic detection have been discovered, like CEBPA, NPM1, KIT, having great significance in predicting prognosis.[14, 15] In addition, the role of gene methylation in prognosis has also been discovered, like high GADD45A methylation predicting low survival of AML patients.[21] Therefore, new gene abnormalities are still required for further discovery, in order to provide clues for predicting prognosis, studying the pathogenesis and exploring new therapeutic targets. In this study, we integrated survival of 75 TCGA samples with CN-AML and 240 GEO CN-AML samples, examined the effect of RASGEF1A expression on CN-AML patients' prognosis, including

patients received chemotherapy or Allo-HSCT, and the influence of RASGEF1A methylation levels on CN-AML samples' prognosis.

In this study, we found that RASGEF1A can be used as a CN-AML patients' prognostic factor. We compared the OS and EFS of 75 TCGA samples with CN-AML and 240 CN-AML patients from GEO database with or without high RASGEF1A expression, including CN-AML patients' overall prognosis and respective prognosis of AML patients received chemotherapy or Allo-HSCT. The study showed that the prognosis of patients in RASGEF1A-high expression group of RASGEF1A was poorer than the RASGEF1A-low expression group ($P < 0.0001$, Fig. 1 and Fig. 2), indicating that the RASGEF1A gene expression level is a strong adverse factor for CN-AML samples' prognosis. We can examine the 75 CN-AML samples' baseline characteristics in Table 1. Two groups had the same level in gender, age, race, FAB type of AML, induction program, status before transplant, and other gene mutations, showing that the results of this study group are not affected by other factors and have high credibility. Furthermore, The BeatAML CN-AML patients with RASGEF1A-high have better sensitivity to Crenolanib and in multivariate analysis, we can see that the effect of RASGEF1A expression on prognosis is strong and meaningful.

In addition, we studied the association between the methylation of RASGEF1A gene and patients survival rates. Just as the high methylation level of GADD45A can indicate adverse prognosis, we discovered the high methylation of RASGEF1A also suggests unfavorable prognosis in CN-AML patients.[21] As can be seen from Fig. 3A, the EFS and OS of patients are much better in RASGEF1A-high methylation group than RASGEF1A-low methylation group ($P < 0.0001$, Figure. 3A). Not only that, but we also found the same results in patients receiving chemotherapy or Allo-HSCT, high level of methylation suggests favorable prognosis of AML patients receiving chemotherapy ($P < 0.0001$, Figure. 3B) or Allo-HSCT ($P < 0.0001$, Figure. 3C). Therefore, we can say that the level of RASGEF1A methylation is a new prognostic factor of AML patients.

Furthermore, we found that the integrative analysis of gene expression and methylation can provide a more accurate method of prognostic classification. In Fig. 4, we can see that CN-AML samples can be separated into 4 groups by RNA level and methylation level, and the difference in EFS and OS among the 4 groups were statistically significant ($P < 0.0001$, Figure. 4). In the left side of Figure.4, it is easy to find that RASGEF1A-low expression RASGEF1A-high methylation group (G3) had the best EFS and RASGEF1A-high expression RASGEF1A-low methylation group (G2) had the worst EFS among 4 groups, consistent with previous conclusions. Due to the small overall sample size, there is only one patient in G1 group in the OS (Figure. 4, right side), we have to say individual factors may have a great impact on the outcome. Despite this, we can find that the G3 group had the best OS and G2 had the worst OS among other three groups similarly. It has been found that methylation can affect survival not only by regulating gene expression, but also through some other mechanisms.[34] This shows that the two may reflect different processes in the pathogenesis, while the integrative analysis may be more comprehensive. This method has been applied in other diseases and many new discoveries have been harvested. The new strategies to overcome tamoxifen resistance may be discovered by analyzing the DNA expression and methylation features of cancer stem cells.[35] By investigating the DNA imbalance and methylation

profiles of myeloma cells, it was found that the genomic heterogeneity always present from diagnosis to relapse.[36] The integrative analysis in soft tissue sarcomas, helped to find new biomarkers associated with pathogenesis.[37] The integrative analysis of gene expression level and methylation level of RASGEF1A allows CN-AML patients to be more accurately graded, making the prognosis prediction more accurate and even providing a more accurate direction for treatment. The CN-AML patients with RASGEF1A-high expression level have better sensitivity to Crenolanib.

RASGEF1A is an important guanine nucleotide exchange factors of Ras. Previous studies have exposed that RASGEF1A high expression is related to the survival and migration of intrahepatic cholangiocarcinoma cells, and our study has shown a correlation with the prognosis of patients with AML.[29] This may suggests that this gene further contribute to the pathogenesis of AML and is associated with some malignant diseases. At the same time, this gene was detected in the bone marrow and other 26 detected tissues, especially in brain, lymph nodes, spleen and testis.[25–30] We can guess that there are also potential correlations with RASGEF1A gene and diseases in those tissues, which requires further studies.

What is regrettable is that this experiment does not involve molecular mechanisms. To date, no prospective studies have proved the clinical significance of RASGEF1A in the treatment of CN-AML.

In conclusion, the high RASGEF1A expression level and low methylation level suggest poor survival and adverse prognosis for CN-AML patients. The integrative analysis of RNA and methylation level can provide a more accurate classification for prognosis. The CN-AML patients with RASGEF1A high expression level have better sensitivity to Crenolanib. Low RASGEF1A expression is a favorable prognostic factor for AML patients receiving chemotherapy or Allo-HSCT.

Declarations

Acknowledgements

The authors thank TCGA database, GSE12417 dataset and BeatAML dataset which are all available online..

Funding

This work was funded by National Natural Science Foundation of China (81800195 and 81460315), Key Clinical Projects of Peking University Third Hospital (BYSYZD2019026), interdisciplinary medicine Seed Fund of Peking University (BMU2018MB004), Beijing Natural Science Foundation (7132183 and 7182178), China Health Promotion Foundation (CHPF-zlkysx-001), Scientific Research Foundation (20141114) from Health Commission of Jiangxi Province, and Science and Technology Research Foundation (GJJ14676) from Educational Commission of Jiangxi Province, China.

Contributions

HMJ, YLZ and CJY conceived the project. XH, WLZ and WF analyzed the data. XH, WLZ, WF, XNL, PY, JW, SXL, XRZ, GHD, CJY, YLZ and HMJ contributed towards the interpretation of the data. All authors wrote and approved the final manuscript.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

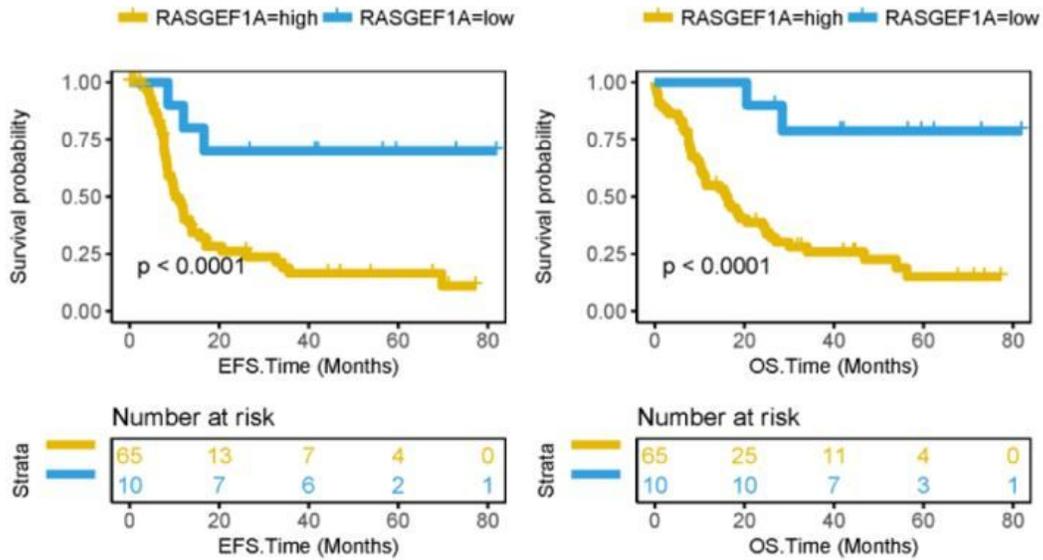
1. O'Donnell M, et al. Acute Myeloid Leukemia, Version 3.2017, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw*. 2017;15(7):926–57.
2. Burnett A, Wetzler M, Lowenberg B. Therapeutic advances in acute myeloid leukemia. *J Clin Oncol*. 2011;29(5):487–94.
3. Kadia T, et al. Progress in acute myeloid leukemia. *Clin Lymphoma Myeloma Leuk*. 2015;15(3):139–51.
4. Bell J, et al. Effectiveness and Safety of Therapeutic Regimens for Elderly Patients With Acute Myeloid Leukemia: A Systematic Literature Review. *Clinical Lymphoma Myeloma Leukemia*. 2018;18(7):e303–14.
5. Estey E. Acute myeloid leukemia: 2014 update on risk-stratification and management. *Am J Hematol*. 2014;89(11):1063–81.
6. H., A., et al., *Attempts to improve treatment outcomes in acute myeloid leukemia (AML) in older patients_ the results of the United Kingdom Medical Research Council AML11 trial*. *Blood*, 2001. 98: p. 1302–1311.
7. Wheatley K, et al. Prognostic factor analysis of the survival of elderly patients with AML in the MRC AML11 and LRF AML14 trials. *Br J Haematol*. 2009;145(5):598–605.
8. Byrd J, et al. Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with de novo acute myeloid leukemia: results from Cancer and Leukemia Group B (CALGB 8461). *Blood*. 2002;100(13):4325–36.
9. D G, et al. The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. *Blood*. 1998;92:2322–33.
10. Shen Y, et al. Gene mutation patterns and their prognostic impact in a cohort of 1185 patients with acute myeloid leukemia. *Blood*. 2011;118(20):5593–603.
11. Vardiman J, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*. 2009;114(5):937–51.
12. Mrozek K, et al. Clinical relevance of mutations and gene-expression changes in adult acute myeloid leukemia with normal cytogenetics: are we ready for a prognostically prioritized molecular

- classification? *Blood*. 2007;109(2):431–48.
13. Mannelli F, et al. CEBPA-double-mutated acute myeloid leukemia displays a unique phenotypic profile: a reliable screening method and insight into biological features. *Haematologica*. 2017;102(3):529–40.
 14. Weinberg O, et al. Diagnostic work-up of acute myeloid leukemia. *Am J Hematol*. 2017;92(3):317–21.
 15. Metzeler K, et al. Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. *Blood*. 2016;128(5):686–98.
 16. Loghavi S, et al. Clinical features of De Novo acute myeloid leukemia with concurrent DNMT3A, FLT3 and NPM1 mutations. *J Hematol Oncol*. 2014;7:74.
 17. Moore L, Le T, Fan G. *DNA methylation its basic function* *Neuropsychopharmacology*. 2013;38(1):23–38.
 18. Wan J, et al. Integrative analysis of tissue-specific methylation and alternative splicing identifies conserved transcription factor binding motifs. *Nucleic Acids Res*. 2013;41(18):8503–14.
 19. Lopez-Serra P, Esteller M. DNA methylation-associated silencing of tumor-suppressor microRNAs in cancer. *Oncogene*. 2012;31(13):1609–22.
 20. Ehrlich M. DNA hypomethylation in cancer cells. *Epigenomics*. 2009;1:239–59.
 21. Perugini M, et al. GADD45A methylation predicts poor overall survival in acute myeloid leukemia and is associated with IDH1/2 and DNMT3A mutations. *Leukemia*. 2013;27(7):1588–92.
 22. Kim S, et al. Level of HOXA5 hypermethylation in acute myeloid leukemia is associated with short-term outcome. *Korean J Lab Med*. 2010;30(5):469–73.
 23. ZH Z, et al. Decreased SCIN expression, associated with promoter methylation, is a valuable predictor for prognosis in acute myeloid leukemia. *Mol Carcinog*. 2018;57:735–44.
 24. Yaman E, et al. RasGEF1A and RasGEF1B are guanine nucleotide exchange factors that discriminate between Rap GTP-binding proteins and mediate Rap2-specific nucleotide exchange. *FEBS J*. 2009;276(16):4607–16.
 25. Wu J, et al. Rap2a serves as a potential prognostic indicator of renal cell carcinoma and promotes its migration and invasion through up-regulating p-Akt. *Sci Rep*. 2017;7(1):6623.
 26. JX W, et al. Rap2a is a novel target gene of p53 and regulates cancer cell migration and invasion. *Cell Signal*. 2015;27:1198–207.
 27. YE L, et al. The prognostic impact of RAP2A expression in patients with early and locoregionally advanced nasopharyngeal carcinoma in an endemic area. *Am J Transl Res*. 2015;7:912–21.
 28. L W, et al. Over-expression of Rap2a inhibits glioma migration and invasion by down-regulating p-AKT. *Cell Biol Int*. 2014;38:326–34.
 29. K U, et al. Enhanced RASGEF1A expression is involved in the growth and migration of intrahepatic cholangiocarcinoma. *Clin Cancer Res*. 2006;12:6611–6.
 30. Fagerberg L, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics*. 2014;13(2):397–406.

31. Cancer Genome Atlas Research. N., et al., Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med.* 2013;368(22):2059–74.
32. Metzeler K, et al. An 86-probe-set gene-expression signature predicts survival in cytogenetically normal acute myeloid leukemia. *Blood.* 2008;112(10):4193–201.
33. Tyner JW, et al. Functional genomic landscape of acute myeloid leukaemia. *Nature.* 2018;562(7728):526–31.
34. Smith A, et al. A novel approach to the discovery of survival biomarkers in glioblastoma using a joint analysis of DNA methylation and gene expression. *Epigenetics.* 2014;9(6):873–83.
35. Lin X, et al. Integrative analyses of gene expression and DNA methylation profiles in breast cancer cell line models of tamoxifen-resistance indicate a potential role of cells with stem-like properties. *Breast Cancer Res.* 2013;15:R119.
36. Krzeminski P, et al. Integrative analysis of DNA copy number, DNA methylation and gene expression in multiple myeloma reveals alterations related to relapse. *Oncotarget.* 2016;7:80664–79.
37. Renner M, et al. Integrative DNA methylation and gene expression analysis in high-grade soft tissue sarcomas. *Genome Biol.* 2013;14:r137.

Figures

A



B

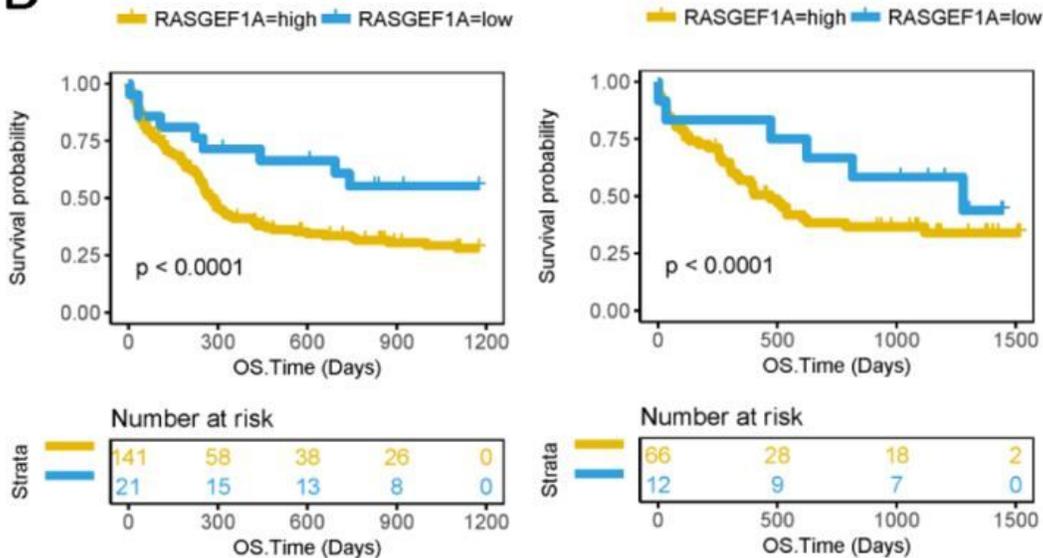
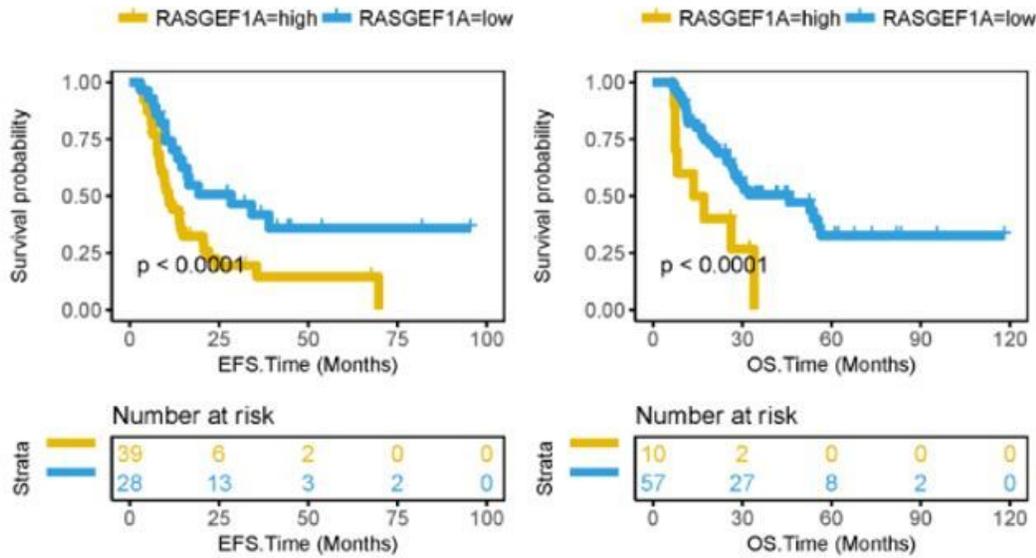


Figure 1

Compare the survival levels of the RASGEF1A-high expression group and the RASGEF1A-low group in CN-AML patients. A: EFS and OS in 75 CN-AML patients from the TCGA database ($P < 0.0001$). B, OS in 162 CN-AML patients from the GSE12417 dataset ($P < 0.0001$) and OS in 78 CN-AML patients from the GSE12417 dataset ($P < 0.0001$). A Log-rank test was used to compare the survival curves of high and low gene expression. CN-AML, cytogenetically normal acute myeloid leukemia. EFS, Event-free survival time (months); OS, Overall survival time (months). Left side of A: the x-axis represents the EFS time (months); the y-axis represents the survival probability; Right side of A: the x-axis represents the OS time (months);

the y-axis represents the survival probability. B: the x-axis represents the OS time (months); the y-axis represents the survival probability

A



B

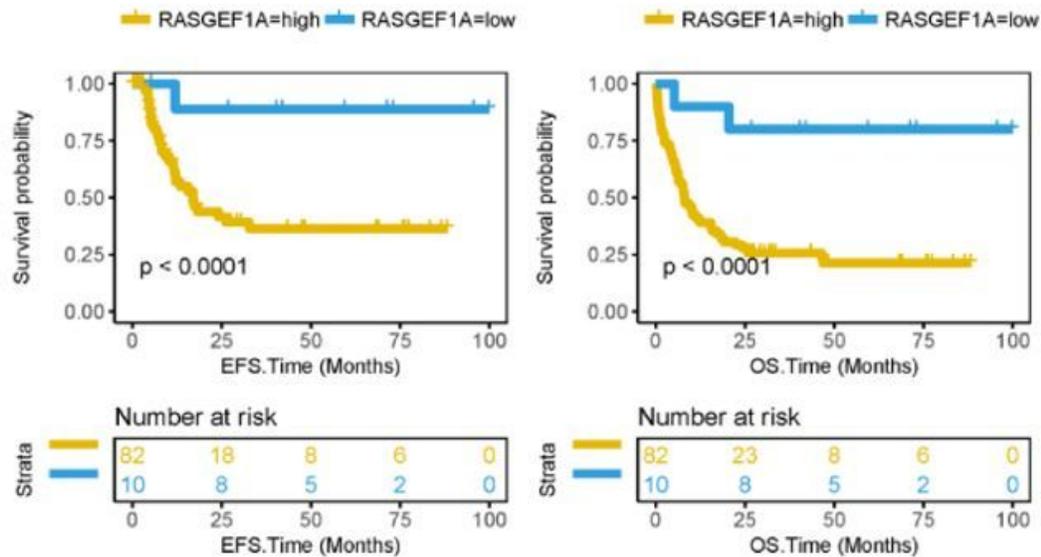
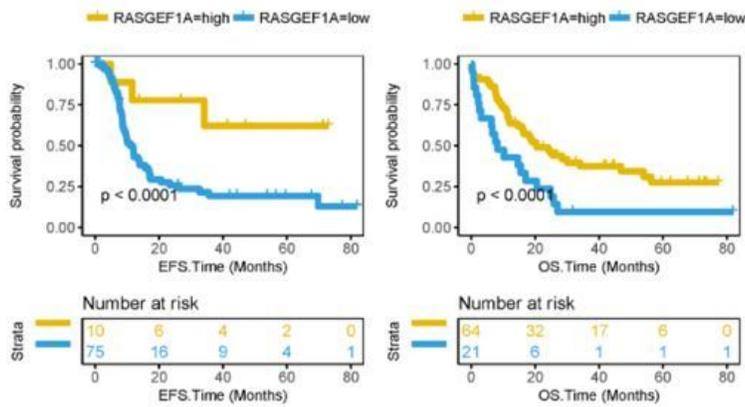


Figure 2

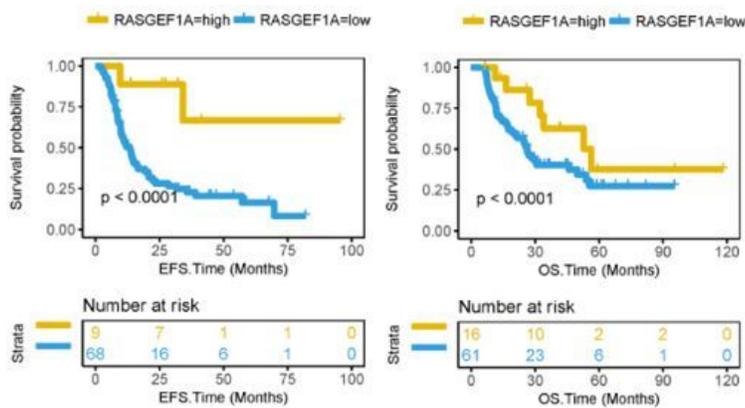
Compare the survival levels of the RASGEF1A-high expression group and the RASGEF1A-low expression group in AML patients receiving chemotherapy or Allo-HSCT. A, EFS and OS in 67 AML patients receiving Allo-HSCT from the TCGA database ($P < 0.0001$). B, EFS and OS in 92 AML patients receiving

chemotherapy from the TCGA database ($P < 0.0001$). A Log-rank test was used to compare the survival curves of high and low gene expression. Event-free survival time (months); OS, Overall survival time (months). Left side: the x-axis represents the EFS time (months); the y-axis represents the survival probability; Right side: the x-axis represents the OS time (months); the y-axis represents the survival probability.

A



B



C

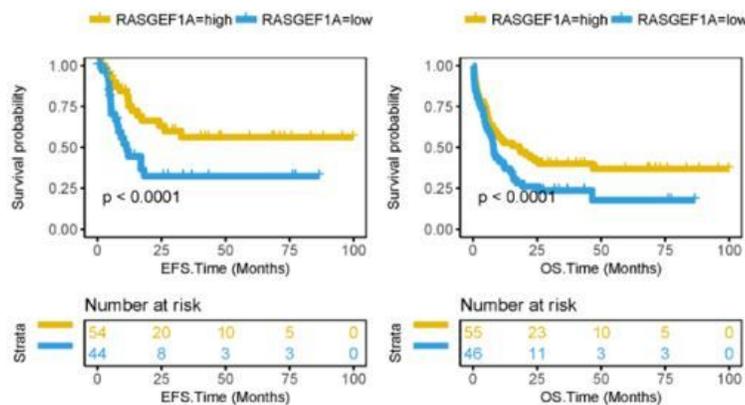


Figure 3

Compare the survival levels of the RASGEF1A-high methylation group and the RASGEF1A-low methylation group in CN-AML patients, and AML patients receiving chemotherapy or Allo-HSCT. A, EFS and OS in 85 CN-AML patients from the TCGA database ($P < 0.0001$). B, EFS and OS in 77 AML patients receiving Allo-HSCT from the TCGA database ($P < 0.0001$). C, EFS and OS in 101 AML patients receiving chemotherapy from the TCGA database ($P < 0.0001$). A Log-rank test was used to compare the survival curves of high and low gene expression. CN-AML, cytogenetically normal acute myeloid leukemia. Event-free survival time (months); OS, Overall survival time (months). Left side: the x-axis represents the EFS time (months); the y-axis represents the survival probability; Right side: the x-axis represents the OS time (months); the y-axis represents the survival probability.

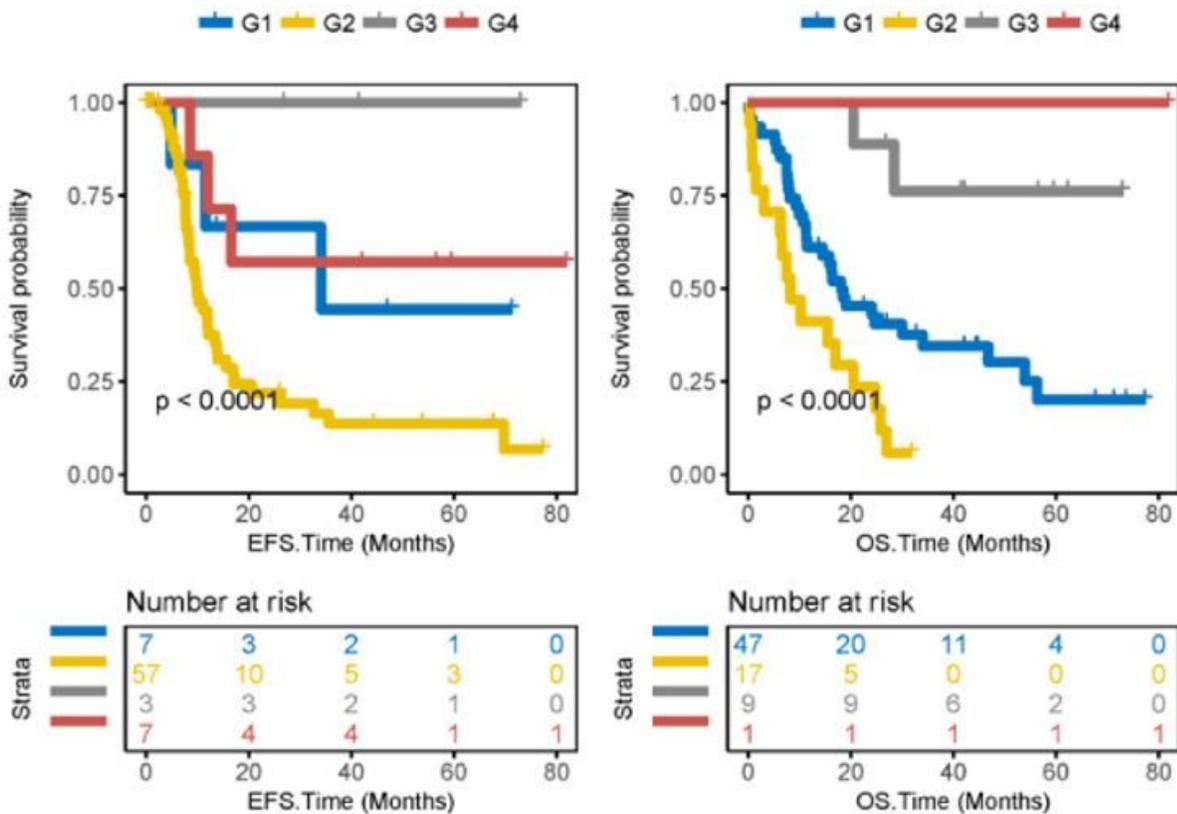


Figure 4

Compare the survival levels of the RASGEF1A-high expression RASGEF1A-high methylation group, the RASGEF1A-high expression RASGEF1A-low methylation group, the RASGEF1A-low expression RASGEF1A-high methylation group and the RASGEF1A-low expression RASGEF1A-low methylation group in CN-AML patients. EFS and OS in 75 CN-AML patients from the TCGA database ($P < 0.0001$). A Log-rank test was used to compare the survival curves of high and low gene expression. CN-AML,

cytogenetically normal acute myeloid leukemia. Event-free survival time (months); OS, Overall survival time (months). Left side: the x-axis represents the EFS time (months); the y-axis represents the survival probability; Right side: the x-axis represents the OS time (months); the y-axis represents the survival probability.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SuppFigureS1andTables.docx](#)