약학박사 학위논문

Cancer big data analysis on acute myeloid leukemia for new drug target discovery

암 빅데이터 분석을 통한 급성 골수성 백혈병에서의 새로운 약물 타겟 발굴

2023년 8월

서울대학교 대학원

약학과 천연물과학전공

김 한 선

Cancer big data analysis on acute myeloid leukemia for new drug target discovery

암 빅데이터 분석을 통한 급성 골수성

백혈병에서의 새로운 약물 타겟 발굴

지도교수 박 성 혁

이 논문을 약학박사 학위논문으로 제출함
2023년 8월

서울대학교 대학원
약학과 천연물과학전공
김 한 선

김한선의 약학박사 학위논문을 인준함
2023년 8월

위 원 장 노 민 수 (인)

부위원장 이 상 국 (인)

위 원 이 주 용 (인)

위 원 조 병 식 (인)

위 원 박 성 혁 (인)

Abstract

Cancer big data analysis on acute myeloid leukemia for new drug target discovery

Han Sun Kim

Natural Products Science Major

College of Pharmacy

The Graduate School

Seoul National University

Acute myeloid leukemia (AML) generally has an unsatisfactory prognosis despite the recent introduction of new regimens including targeted agents and antibodies. Moreover, even though the European LeukemiaNet (ELN) 2017 criteria have been widely accepted as the risk classification of AML patients, their application in studying biological pathways related to risk categories has been limited, and they have not helped improve drug treatment options for high-risk patients.

To address those issues, analysis on cancer big data was used to identify new target candidates of AML. Initially, to find a new druggable pathway, integrated bioinformatic pathway screening was performed on large Oregon Health

& Science University (OHSU) and Microarray Innovations in Leukemia (MILE) AML databases. This analysis revealed the SUMOylation pathway, which was independently validated with an external dataset (totaling 2959 AML and 642 normal sample data in all databases). The clinical relevance of SUMOylation in AML was supported by its core gene expression, which correlated with patient survival, ELN2017 risk classification, and AML-relevant mutations. TAK-981, a first-in-class SUMOylation inhibitor currently under clinical trials for solid tumors, showed anti-leukemic effects with apoptosis induction, cell-cycle arrest, and induction of differentiation marker expression in leukemic cells. It exhibited potent nanomolar activity, often stronger than that of cytarabine, which is part of the standard-of-care. TAK-981's utility was further demonstrated in *in vivo* mouse and human leukemia models as well as patient-derived primary AML cells. The results also indicated TAK-981 exert direct anti-AML effects inherent to cancer cells, different from the IFN1 and immune-dependent mechanism in a previous solid tumor study. Overall, these findings provide a proof-of-concept for targeting SUMOylation as a new targetable pathway in AML, with TAK-981 showing a promising direct anti-AML agent. The data should prompt studies on optimal combination strategies and transitions to clinical trials in AML.

In addition, biological pathways whose upregulations are correlated with increased ELN2017 risks were investigated using a recent AML database. Filtering and validating with patient survival and other independent transcriptomics and proteomics AML database gave 'synthesis of unsaturated fatty acids' and 'metabolism of folate' pathways as candidates. Further refinement at the gene level, along with a literature search, identified *SCD* and *MTHFD2* as key targets relevant

to high-risk groups. Both the SCD inhibitor A939572 and MTHFD2 inhibitor DS18561882 showed cancer selectivity and synergy with cytarabine - a standard drug for induction therapy - in cell lines with relatively high IC_{50}s for cytarabine. It was also found that *SCD* gene expression correlated with the amount of unsaturated fatty acids. Overall, the suggested targets may be further exploited to find better therapeutic options and mechanistic insights in high-risk AML. Furthermore, the workflow could be readily applied to find other target genes/pathways or even to solid tumors.

Keywords: Acute myeloid leukemia, SUMOylation, TAK-981, immune-independent, ELN2017, high-risk, SCD, MTHFD2

Student Number: 2017-20187

Table of Contents

Part II : SCD and MTHFD2 inhibitors for high-risk acute myeloid leukemia patients, as suggested by ELN2017-pathway association

List of Figures

List of Tables

General Introduction

Acute myeloid leukemia (AML) is a heterogeneous disease characterized by an accumulation of immature progenitor cells with arrested differentiation leading to suppression of hematopoiesis [1]. In the United States, it had the highest percentage (62%) of leukemic deaths in 2019 [2]. Also, among all cancer types, AML had the 5th worst five-year overall survival in the United States, 2000-2016 (Surveillance, Epidemiology, and End Results (SEER) data) [2]. In addition, AML is typically a disease of older people, with the median age at diagnosis showing 68 years (2011-2016, SEER data) [2].

The treatment of AML typically divides into two phases, that is, induction therapy and consolidation therapy (Table 1). Initial assessment for deciding whether a patient is eligible for intensive induction therapy is needed [1]. If complete remission is achieved after induction therapy, appropriate consolidation therapy is required [1]. Current standard-of-care treatments for AML include combination chemotherapy with cytotoxic drugs, usage of hypomethylating agents, and/or hematopoietic stem cell transplantation (HSCT) [3]. The combination chemotherapy with cytotoxic drugs has not changed for nearly a half century; that is, the “7 + 3” induction therapy regimen comprised of cytarabine (days 1 to 7) plus anthracycline-based drugs (days 1 to 3). For patients ineligible for this therapy, i.e. in some of older patients, regimens based on low-dose cytarabine is used. Recent improvement in the understanding of AML pathogenesis has led to the introduction of several novel targeted agents since 2017 (Table 2) [4-12]. Nevertheless, long-

term survival is still suboptimal without allogeneic HSCT [13], and thus, more efforts should be done to unravel novel prognostic, predictive, and targetable molecular abnormalities. However, lack of prevailing driver genomic mutations and available unique markers for AML has made it quite difficult. In this context, investigations into post-genomic pathways relevant to AML pathogenesis and approaches to their targeting have been desired.

ELN2017, originally starting from ELN2010, is a recommendation for the diagnosis and management of AML patients suggested by an international expert panel on behalf of the European LeukemiaNet (ELN) [3]. ELN2017 classified three risk groups (‘Favorable,’ ‘Intermediate,’ and ‘Adverse’) based on molecular and genetic aberrations, and it has been widely used in many clinical trials and regulatory offices [3, 14]. ELN2017 has proven effective for risk management, including hematopoietic stem cell transplantation [15-19]. Some trials refined it to classify better the subpopulation of the risk categories [17, 20], particularly discovering patient groups with distinctly favorable and poor prognoses, suggesting the addition of two new categories (‘Very Favorable’ and ‘Very Adverse’) [17]. Even though ELN2017 or its revised versions have shown power in the risk classification of AML patients, there are only a few studies investigating which biological pathways are related to the categories (or subcategories) of ELN2017 [21-25]. Considering the clinical importance of the ELN2017, investigating which biological pathways are correlated is urgently needed.

In this study, using bioinformatics, I tried to find pathways or genes to target in AML, followed by experimental validation. By comparing gene expression from normal samples, I found SUMOylation as a targetable pathway,

and suggested TAK-981 as a direct anti-AML agent. In addition, by utilizing multi-omics database, I found ‘synthesis of unsaturated fatty acids’ and ‘folate metabolism’ as targetable pathways in high-risk AML patients in particular, found *SCD* and *MTHFD2* genes as target genes, and suggested their inhibitors as drug candidates.

Table 1. Conventional treatment regimens for AML patients

Patients criteria				Treatment
Eligible for intensive chemotherapy	Induction therapy	All ages		"7+3"
	Consolidation therapy	Younger patients (18-60/65 years)	Favorable-risk genetics	IDAC
			Intermediate-risk genetics	IDAC or allogeneic HSCT
			Adverse-risk genetics	Allogeneic HSCT
		Older patients (> 60/65 years)	Favorable-risk genetics	IDAC
			Intermediate/adverse-risk genetics	Consider allogeneic HSCT
Not eligible for intensive chemotherapy				Azacitidine Decitabine Low-dose cytarabine

The information was retrieved and summarized from [3]. “7+3”, cytarabine (days 1 to 7) plus anthracycline-based drugs (days 1 to 3); IDAC, intermediate-dose cytarabine; HSCT, hematopoietic stem cell transplant.

Table 2. Recent drugs (since 2017) in AML approved by Food and Drug Administration (FDA)

Treatment	Approval date	Description
Midostaurin	April 2017	Multikinase FLT3 inhibitor
Gemtuzumab ozogamycin	September 2017	Anti-CD33 antibody-drug conjugate
CPX-351	August 2017	Liposomal cytarabine and daunorubicin (5:1 molar ratio)
Glasdegib	November 2018	Hedgehog pathway inhibitor
Venetoclax	November 2018	BCL-2 inhibitor
Enasidenib	August 2017	IDH2 inhibitor
Ivosidenib	July 2018 May 2019	IDH1 inhibitor
Gilteritinib	November 2018	FLT3 inhibitor
CC-486	September 2020	Oral azacitidine hypomethylating agent
Oral Decitabine-cedazuridine	July 2020	Oral hypomethylating agent

The information was retrieved and summarized from [4] and [12].

Part I

TAK-981, a SUMOylation inhibitor, suppresses AML growth immune-independently

I. Introduction

SUMOylation is a post-translational modification (PTM) involved in the conjugation of small ubiquitin-like modifiers (SUMOs) to substrate proteins [26]. SUMO-activating enzyme E1 (SAE1 and SAE2 encoded by *SAE1* and *UBA2*, respectively), an E2 (ubiquitin-conjugating enzyme 9, UBC9 encoded by *UBE2I*), and a limited set of E3 ligases participate in this process [26, 27]. SUMOylation seems to be important in nuclear functions of proliferating or developing cells by regulating the mitotic cell cycle and DNA damage response [28-30]. Specific pathways affected by SUMOylation in cancer may include p53 [31, 32] and cMYC [33, 34], but more studies are needed to resolve some of the controversies [35, 36]. Additionally, innate immunity is mostly suppressed by SUMOylation, inhibition of which, therefore, might have implications for cancer therapy [26, 37]. There were some studies on SUMOylation in lymphoma [33] and solid tumors [28, 32, 34, 38, 39], including cervix, prostate, breast, pancreas, and colon. As for AML, only a few studies on the roles of SUMOylation have been published [40-42]. Therefore, concrete evidence of the therapeutic utility of SUMOylation or of specific inhibitors of SUMOylation in AML has been lacking. TAK-981 is an inhibitor of the SUMO-activating enzyme (SAE) that forms a SUMO-TAK-981 adduct [43].

As the first-in-class SAE inhibitor targeting cancers, it is currently in clinical trials for solid tumors or lymphomas (NCT03648372, NCT04074330, NCT04381650). In blood cancer, it has been known to shift the T cell balance toward healthy immune cell subsets in chronic lymphocytic leukemia [44]. To my knowledge, TAK-981 has not been studied for AML or evaluated in AML clinical trials.

For solid tumors, large-scale bioinformatic analysis has been successfully performed comparing normal and cancer samples thanks to The Cancer Genome Atlas (TCGA) data. TCGA also contains data on AML (TCGA-LAML [45] dataset), but it lacks the data for non-cancer controls, limiting its application in AML field. As of now, three large-scale gene expression databases contain both AML and normal data: 1) MILE study stage I data [46], 2) OHSU data from the BeatAML 1.0 program [47], and 3) the Gene Expression Omnibus (GEO) compilation [48]. Therefore, analysis of these large databases in all (totaling 2959 AML and 642 normal samples) might yield new and useful information on targets for broader AML patients.

Here, accessing large gene expression databases for AML, I evaluated the clinical relevance of the SUMOylation pathway and investigated the anti-leukemic effects of its inhibition by TAK-981.

II. Materials and methods

1. Bioinformatics analysis

1) Data download and preprocessing, GSEA, GSEAPreranked and Pathway Clustering

For MILE study stage I, the gene expression table and sample information were downloaded from from the National Center for Biotechnology Information's Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/, accession number GSE13159). The probe IDs were converted to gene symbols, and for those with duplicate matches, the probe with a maximum coefficient of variation (that is, standard deviation divided by mean) was selected. Then, for each gene, the difference of mean from the AML bone marrow samples (501 samples) to healthy bone marrow samples (73 samples) was calculated, and these numbers were used as input for GSEAPreranked.

For OHSU BeatAML 1.0 program, the raw counts and sample information were downloaded from GDC (https://portal.gdc.cancer.gov/). The raw counts were DESeq2-normalized and rlog-transformed by DESeq2 [49] package in R. The ensemble IDs were left as-is. For each gene, the difference of mean from the AML bone marrow samples (245 samples) to healthy bone marrow samples (21 samples) was calculated, and these numbers were used as input for GSEAPreranked. The survival information was downloaded from http://vizome.org/aml/.

For GEO collection database, the gene expression data and sample information were downloaded from https://doi.org/10.5281/zenodo.3257786.

Specifically, for the gene expression data, "All_2761_Corrected_for_All_Factors_SampleSource_DiseaseState_Batch_Datasetwise_2213_AML_1st_548_Healthy_2nd_and_removed_613_dummy_with_44754_probsets_RMA_Normalized_Log2Trans_Zscore_Standardized_Transposed_Data.csv" file was downloaded. The probe IDs were converted to gene symbols, and for those with duplicate matches, the same procedure was applied as in MILE study stage I above.

For TCGA-LAML database, the gene expression data and survival information were downloaded from https://gdc.cancer.gov/about-data/publications/pancanatlas. Sample IDs starting with "TCGA-AB" were considered as AML samples. The categorization of the TCGA samples by ELN2017 risk groups was kindly provided by the authors of a recent paper publication [50].

The GSE173116 dataset used in Fig. 16 was downloaded from GEO. GSEA and GSEAPreranked were run using GSEA software from the Broad Institute. For gene sets in GSEAPreranked, the gene set database (Human_GOBP_AllPathways_no_GO_iea_June_24_2019_symbol.gmt) was downloaded from http://baderlab.org/GeneSets. In order to avoid errors in later steps, some modifications including removing special characters were performed; the R packages GSA, stringr, and rowr were used in this process. The minimum and maximum gene set size filters were set to 10 and 500, respectively. As a result, 7036 gene sets were used in the analysis.

The AML-upregulated pathway result tables from both GSEAPreranked analyses were imported into R, and only the gene sets satisfying the following

criteria were selected; (i) nominal $p < 0.05$ in both databases, (ii) false discovery rate (FDR) < 0.25 in at least one database. This resulted in 154 gene sets. For pathway clustering analysis, GSCluster [51] package was used with the 154 gene sets and the leading-edge genes in both databases. For the q values, the FDR values from OHSU results were used. For the clustering method, 'Distance' parameter was set to pMM, 'Network weight' to 1, and 'Maximum gene-set distance' to 0.25.

For gene set database of Fig. 16, the Hallmark gene set in MSigDb (version 7.4) was used. For Fig. 18, Biocarta gene set in MSigDb (version 7.5.1) was used.

The following 17 genes were considered as related to SUMOylation in Figs. 2B and 5; *SAE1*, *UBA2*, *UBE2I*, *PIAS1*, *PIAS2*, *PIAS4*, *BMI1*, *PHC1*, *PHC2*, *PHC3*, *CBX2*, *CBX4*, *CBX8*, *RING1*, *RNF2*, *SUMO1*, and *SUMO2*.

2) Survival analysis

For the association between the SUMOylation pathway and overall survival, patient groups were stratified into high or low groups according to the expression levels of several important genes in the pathway using the best risk separation approach [52], and the survival difference between the two groups was evaluated with Cox regression.

For univariate and multivariate analysis in Table 9, survival R package was used. The patient information was retrieved from http://vizome.org/aml and original paper of OHSU BeatAML 1.0 program [47]. The following 12 parameters were included for the univariate analysis; ELN2017, isRelapse, consensus_sex,

cumulativeChemo, ageAtSpecimenAcquisition, CEBPA_Biallelic, FLT3-ITD, NPM1, RUNX1, ASXL1, TP53, and the converted Gene Set Variance Analysis (GSVA) pathway scores of BIOCARTA_SUMO_PATHWAY (termed "SUMOgene"). Of these, 6 parameters (ELN2017, cumulativeChemo, ageAtSpecimenAcquisition, NPM1, TP53, and SUMOgene) were included for the multivariate analysis.

3) Conversion of gene expression data to pathway scores data

To convert gene expression to pathway scores, GSVA R package was used. Hallmark gene sets ("h.all.v7.2.symbols.gmt"), canonical pathways which contain BioCarta, Kyoto Encyclopedia of Genes and Genomes (KEGG), Pathway Interaction Database (PID), Reactome and WikiPathways pathway database ("c2.cp.v7.2.symbols.gmt"), and Gene Ontology Biological Process gene sets ("c5.go.bp.v7.2.symbols.gmt") were downloaded from http://baderlab.org/GeneSets, read in R with GSA package and combined. With this combined gene set, DESeq2- and rlog-transformed OHSU gene expression data from above were used as input for the function gsva, with parameters "min.sz" set to 5, "max.sz" to 700, "method" to "gsva".

2. Cells – Reagents

MOLM-14 (DSMZ, Braunschweig, Germany), U937, THP-1, KG-1, and C1498 (ATCC, Manassas, VA, USA) were used in this study. Cells were cultured in RPMI 1640 media supplemented with 10% fetal bovine serum (FBS), 100 U/ml penicillin, and 100 μg/ml streptomycin, at 37 °C in a 5% CO_2 incubator. TAK-981 was purchased from MedChemExpress (Monmouth Junction, NJ, USA). Cytarabine was purchased from Sigma-Aldrich (St.Louis, MO, USA). The concentrations used in Figs. 14-24 (except Figs. 16 and 18) were based on the the results from the initial estimation of IC_{50} of TAK-981 for each cell line. Ficoll-Hypaque is from Sigma-Aldrich (St.Louis, MO, USA). All other chemical reagents were from Sigma-Aldrich unless otherwise noted.

3. Antibodies for flow cytometry

Harvested cells were stained with antibodies against human CD33-PE and CD34-PE-Cy7 purchased from BD Biosciences (BD Biosciences, San Jose, CA, USA). DAPI (Sigma Aldrich, St. Louis, MO, USA), propidium iodide (PI), and annexin-V-APC (BD Biosciences, San Jose, CA, USA) were used to stain dead and apoptotic cells, respectively. Mouse cells were stained with antibodies against mouse CD90.1 purchased from eBioscience (eBioscience, Waltham, MA, USA).

4. Cell viability with CCK-8 assay

Cells were seeded at 1×10^4 cells/well on 96-well plates and exposed to drugs (TAK-981 alone or in combination with cytarabine, azacitidine, quizartinib or venetoclax) at various concentrations for 48 h. Cell viability was determined with the D-plus CCK Cell Viability Assay Kit (Dongin Biotech, Seoul, South Korea). The IC_{50} value was identified using the GraphPad Prism 9.1.1 software.

5. Primary AML cells from patients

Bone marrow samples from patients with AML were collected during routine diagnostic procedures after informed consent was obtained in accordance with Institutional Review Board regulations of The Catholic University of Korea (KC20SISI0957) and the Declaration of Helsinki. Mononuclear cells were freshly isolated from 25 patients (BM, n = 13, PB, n = 12) with AML and 5 healthy controls (Tables 3-5) by Ficoll-Hypaque density gradient centrifugation. The cells were cultured with different doses of TAK-981, cytarabine, or both for 48 h. To compare cytotoxicity between groups, leukemic cells were gated with CD33 and/or CD34 by flow cytometry and viable cells were compared between groups according to DAPI negative/Annexin V negative status.

Table 3. Patient information for primary AML cells acquired from bone marrow

Sample number	Sex/Age	Type	Sample Status	ELN Risk	Karyotype	Mutations by NGS	FLT3-ITD	WBC (/μL)	Bone marrow blasts	Frontline treatment	Response	HSCT	Relapse after HSCT	Live or death
001	Male/78	MRC	Newly diagnosed	Adverse	46,XY[20]	*ASXL1/RUNX1/CEBPA*	Not mutated	37040	20%	Not treated	NA	NA	NA	Died
002	Female/61	De novo	Newly diagnosed	Favorable	46,XX[20]	*NPM1/TET2/TET2/ETV6/ETV6/WT1*	Not mutated	63090	98%	IDA/ARA	No response	NA	NA	Died
003	Female/21	De novo	Newly diagnosed	Favorable	46,XX,inv(16)(p13.1q22)[20]	*NRAS*	Not mutated	47010	84%	IDA/ARA	CR	MSD	No	Alive
004	Female/21	De novo	Newly diagnosed	Intermediate	46,XX[20]	*NRAS/NRAS/ETV6*	Not mutated	54480	60%	IDA/ARA	CR	MUD	No	Alive
005	Male/63	De novo	Relapsed	Adverse	46~47,XY,del(5)(q22q31),+6,-7,+8,der(11)add(11)(p13)t(4;11)(q12;q14),-12,add(12)(p13),add(15)(p11.2),+21,add(22)(p11.2)[cp10]/46~47,idem,der(1)ins(1;?)(q31;?),del(4)(q21),[cp6]/48,idem,der(11)r(11;?),+del(12)(q21q24.1)[4]	*NRAS/JAK1*	Not mutated	13910	72%	Not treated	NA	NA	NA	Died
006	Male/37	MRC	Newly diagnosed	Intermediate	45,X,-Y[17]/46,XY[3]	*NRAS/SF3B1/WT1*	Not mutated	12900	70%	IDA/ARA	CR	Haplo	NA	Alive
007	Male/62	De novo	Relapsed	Intermediate	46,XY,t(5;9)(q33;q34)[20]	*NPM1/DNMT3A*	Allelic ratio: 8.530	93990	76%	Not treated	NA	NA	NA	Died
008	Female/41	De novo	Newly diagnosed	Intermediate	46,XX,inv(9)(p12q13)[20]	*NPM1/DNMT3A*	Allelic ratio: 2.429	24910	71%	IDA/ARA+ Gilteretinib	CR	MUD	Yes	Alive
009	Male/28	De novo	Newly diagnosed	Adverse	46,XY[20]	*CEBPAsm*	Allelic ratio: 0.765	259920	99%	DNR/ARA+ Midostaurin	NR	Haplo	No	Alive
010	Female/46	MRC	Newly diagnosed	Adverse	46,XX[20]	*RUNX1*	Allelic ratio: 0.616	3120	73%	DNR/ARA+ Midostaurin	CR	MSD	No	Alive
011	Female/57	De novo	Newly diagnosed	Intermediate	46,XX,del(19)(q13.2)[15]/46,XX[5]	*STAG2/CEBPA*	Allelic ratio: 0.494	176340	93%	DNR/ARA+ Midostaurin	CR	Haplo	No	Died
012	Male/67	De novo	Newly diagnosed	Intermediate	47,XY,+15[20]	*BCOR*	Allelic ratio: 0.119	27250	95%	Decitabine+ Venetoclax	No response	NA	NA	Died
013	Female/40	De novo	Newly diagnosed	Favorable	47,XX,+4[4]/48,idem,+22[15]/46,XX[1]	*NPM1/FLT3-TKD*	Allelic ratio: 0.112	47390	98%	IDA/ARA+ Gilteretinib	CR	Haplo	No	Alive

Abbreviations: NA, non-available; NGS, next-generation sequencing; ELN, EuropeanLeukemia net; MRC, myelodysplasia-related change; IDA/ARA, idarubicin/cytarabine; DNR/AR, daunorubicin/cytarabine; CR, complete remission; MSD, matched sibling donor; MUD, matched unrelated donor; Haplo, haploidentical donor; HSCT hematopoietic stem cell transplantation

Table 4. Patient information for primary AML cells acquired from peripheral blood

Sample number	Sex/Age	Type	Sample Status	ELN Risk	Karyotype	Mutations by NGS	FLT3-ITD	WBC (/μL)	Peripheral blasts	Frontline treatment	Response	SAE1/ GAPDH	SAE2/ GAPDH	UBC9/ GAPDH
001	Male/53	De novo	Newly diagnosed	Favorable	46,XY[20]	*CEBPA/CEBPA /DNMT3A*	Not mutated	10180	57%	IDA/ARA	Pending	1.33	0.75	0.68
002	Female/65	De novo	Newly diagnosed	Favorable	46,XX[20]	*NPM1(TypeA) /TET2/CEBPA*	Not mutated	4030	37%	IDA/ARA	Pending	1.25	0.83	0.54
003	Male/57	De novo	Newly diagnosed	Intermediate	46,XY,t(11;19)(q23;p13.1)[20]	*NRAS/NRAS/ STAG2*	Allelic ratio: 0.031	6980	62%	DEC+VEN	Pending	1.27	0.70	0.65
004	Female/57	De novo	Newly diagnosed	Adverse	46,XX[20]	*Pending*	Allelic ratio: 0.796	15660	79%	DNR/ARA+ Midostaurin	Pending	1.25	0.60	0.56
005	Female/67	De novo	Newly diagnosed	Adverse	46,XX,t(9;22)(q34;q11.2)[20]	*Pending*	Not mutated	24710	77%	DEC+VEN	Pending	1.23	0.43	0.59
006	Female/62	De novo	Newly diagnosed	Pending	46,XX[20]	*Pending*	Allelic ratio: 0.067	6320	28%	DNR/ARA+ Midostaurin	Pending	0.71	0.17	0.24
007	Male/68	De novo	Newly diagnosed	Intermediate	46,XY[20]	*IDH2/DNMT3A*	Allelic ratio: 0.069	5630	1%	DEC+VEN	Pending	0.29	0.31	1.08
008	Male/61	De novo	Remission state	Favorable	45,X,-Y,t(8;21)(fq22:q22)[20]	*Pending*	Not mutated	2950	0%	IDA/ARA	Pending	0.14	0.03	0.14
009	Female/70	De novo	Remission state	NA	46,XX[20]	*Not done*	Not mutated	3420	0%	IDA/ARA	CR	0.11	0.02	0.14
010	Male/19	De novo	Remission state	Intermediate	46,XY[20]	*GATA2/NRAS/ TET2*	Not mutated	5340	0%	IDA/ARA	CR	0.09	0.06	0.16
011	Female/51	De novo	Remission state	Favorable	46,XX,t(8;21)(fq22:q22)[11]/36,idem,del(9)[q13q22)[8]/46,XX[1]	*Not done*	Not mutated	4410	0%	IDA/ARA	CR	0.17	0.00	0.14
012	Male/70	De novo	Remission state	Intermediate	47,XY,+8[8]/46,XY[12]	*IDH2/DDX41/ DDX41/TP53*	Not mutated	9440	0%	DEC+VEN	CR	0.06	0.03	0.20

Abbreviations: NA, non-available; NGS, next-generation sequencing; ELN, EuropeanLeukemia net; IDA/ARA, idarubicin/cytarabine; DNR/AR, daunorubicin/cytarabine; DEC/VEN, decitabine/venetoclax; CR, complete remission

Table 5. Information of healthy controls

Sample number	Sex/Age	WBC (/μL)	Sample Status
001	Male/29	6740	Healthy control
002	Female/30	5600	Healthy control
003	Male/26	4800	Healthy control
004	Male/35	7310	Healthy control
005	Female/25	6200	Healthy control

6. Flow cytometry

The expression of various target proteins was analyzed using an LSRFortessaTM flow cytometer (BD Biosciences, San Jose, CA, USA). The harvested cells were stained with antibodies against human and mouse cells targets and an appropriate isotype-matched antibody was used as a negative control. Flow cytometric data were analyzed with FlowJo vX.10 software.

7. Apoptosis analysis

Cell lines or primary AML cells were seeded at 0.2×10^6 cells/mL or 1.0×10^6 cells/mL, respectively, in 24-well plate. After 48 h of incubation with drugs, cells were harvested and stained with DAPI or PI and annexin-V, according to the manufacture's direction. Viable or apoptotic cells were quantified by flow cytometry. Data were analyzed with FlowJo vX.10 software.

8. Cell-cycle analysis

Cells were seeded on 60 mm dishes (1×10^6 cells/dish), then exposed to TAK-981 at indicated concentrations for each cell line. After 48 h incubation, cells were washed twice with DPBS and fixed with cold 70% ethanol, then stored at -20ºC for a minimum of 24 h. Before the analysis, ethanol was discarded completely through centrifugation and by washing the pellets with cold DPBS. Cellular RNA was removed by incubating the pellets with RNAase (200 µg/mL) at 37ºC for 30 min. PI (50 µg/mL) was used to stain the cellular DNA for another 30 min at room temperature. The analysis was conducted on a FACSCalibur flow cytometer (BD

Biosciences, San Jose, CA, USA). Signals were detected on the FL2 channel (ext. 488 nm, emi. 564–606 nm) and data were analyzed by FlowJo vX.10 software.

9. RT-qPCR validation

Total RNA was purified by using a Trizol reagent (Invitrogen, Carlsbad, CA, USA) followed by the cDNAs synthesis using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions. qPCR analysis was conducted on an Applied Biosystems Prism 7300 instrument, using the iTaq Universal SYBR Green Supermix kit (Bio-Rad, Hercules, CA, USA) and primers listed in Table 12. Data were normalized with *ACTB* mRNA level as an internal reference.

For primer efficiency in Table 12, cDNA samples from TAK-981-treated-cell lines were serially diluted, Ct values were obtained for each samples, standard curves were drawn, and the slope of the regression line was calculated. The primer efficiency was calculated using the following formula: (10^(-1/Slope)-1)*100.

10. Western blotting

Cells were seeded on 60 mm dishes (1×10^6 cells/dish), then exposed to TAK-981 at indicated concentrations for each cell line for 24 h or 48 h. The samples were homogenized in RIPA buffer with protease and phosphatase inhibitors (1 mM PMSF (phenylmethylsulfonyl fluoride) 2 μg/mL aprotinin, 1 μg/mL pepstatin A). Protein extracts were separated by SDS (sodium dodecyl sulfate) electrophoresis with 10% gel, then transferred to the PVDF

(polyvinylidene difluoride) membranes. Membranes were blocked with 5% skim milk in TBST (Tris-buffered saline with 0.1% tween) and incubated at 4ºC overnight with the following primary antibodies: β-actin (sc-47778, Santa Cruz), cleaved-caspase 3 (ab32042, Abcam), cytochrome C (1896-1, Epitomics), p21 (ab109520, Abcam), SUMO-2/3/4 (sc-393144, Santa Cruz), SUMO1 (21C7, Thermo Fisher Scientific), SUMO2/3 (8A2, Developmental Studies Hybridoma Bank), p53 (ab131442, Abcam), MDM2 (ab38618, Abcam), SAE1 (ab185949, Abcam), SAE2 (ab185955, Abcam), UBC9 (4930S, Cell Signaling Technology), GAPDH (2118S, Cell Signaling Technology). Anti-rabbit IgG-HRP (sc-2004, Santa Cruz) and anti-mouse IgG-HRP (sc-2005, Santa Cruz) were used as secondary antibodies. The protein bands were visualized by using an EZ-Western Detection kit (DoGen, Seoul, South Korea) and imaged on a LAS-4000 imaging system (GE Healthcare, Chicago, IL, USA).

For SUMOylation levels from TAK-981 treated animals, individual samples were pooled to one, since the amount of the live cancer cells from TAK-981-treated mouse was really small, due to the very high activity of TAK-981, and it was very difficult to get live AML cells for analysis. For all the other western blots, experiments were in at least three biological replicates.

11. Animal experiments

All animal experiments were done in accordance with a protocol approved by the Institutional Animal Care and Use Committee of The Catholic University of Korea (CUMC-2020-0318-01).

1) Syngeneic mouse AML models

C1498 cells labeled with Luc/CD90.1, (C1498/Luc/CD90.1) [53] were intravenously injected into C57BL/6 (female, 6–8-weeks-old, Orient-Bio, Korea) mice through tail vein at a concentration of 2×10^6 cells/mouse to investigate the *in vivo* effects of treatment with TAK-981. Bioluminescence imaging was used to monitor tumor burden. Briefly, mice were anesthetized and imaged noninvasively with an *in vivo* imaging system (Optical in vivo Imaging System-IVIS Lumina XRMS; PerkinElmer, Waltham, MA, USA) after injection with luciferase substrate coelenterazine (Biotium, Heyward, CA, USA). After confirming leukemia engraftment by bioluminescence imaging, mice were randomized into each group. TAK-981 (7.5 mg/kg) formulated in 20% 2-hydroxypropyl-β-cyclodextrin was administered intravenously three times a week for 3 weeks. The three mice for each group were euthanized to examine the extent of leukemic infiltration of different organs and femurs with flow cytometry. Remained mice for each group were monitored for survival.

2) AML xenograft mouse models

MOLM-14 cells labeled with Luc/GFP (MOLM-14/Luc/GFP) [54] were intravenously injected into NOD/SCID/IL-2rγnull (NSG) mice (NSG, female, 6–8-weeks-old, The Jackson Laboratory, Bar Harbor, ME, USA) through tail vein at a concentration of 0.5×10^6 cells/mouse to investigate the *in vivo* effects of treatment with TAK-981. Bioluminescence imaging was used to monitor tumor burden. Briefly, mice were anesthetized and imaged noninvasively with an *in vivo* imaging

system (Optical in vivo Imaging System-IVIS Lumina XRMS; PerkinElmer, Waltham, MA, USA) after injection with luciferase substrate coelenterazine (Biotium, Heyward, CA, USA). After confirming leukemia engraftment by bioluminescence imaging, mice were randomized into each group. TAK-981 (7.5 mg/kg) formulated in 20% 2-hydroxypropyl-β-cyclodextrin was administered intravenously three times a week for 3 weeks. The three mice for each group were euthanized to examine the extent of leukemic infiltration of different organs and femurs with flow cytometry. Western blot analysis for SUMOylation was performed with leukemic cells sorted and separated from bone marrow and spleen by flow cytometer in each group. Remained mice for each group were monitored for survival.

12. Statistical analysis

The Wilcoxon rank-sum test was used for most of the comparison of two groups. For Figs. 17, 20, 21, 23, and 24, one-way Analysis of Variance (ANOVA) and Student's *t*-test were used. For Figs. 28B, 29B, 30B and 31B, Student's *t*-test was used. For the ELN2017 analysis, the Jonckheere-Terpstra test from the DescTools package in R was used. All post hoc analyses were performed with the two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, as implemented in GraphPad Prism. Cox regression was used for survival analysis, with the minimum *p*-value determined by the surv_cutpoint function in the survminer package in R (minprop parameter set to 0.15) or with the median cutoff

approach. All of the statistical analyses were performed with GraphPad Prism 9.1.1 (GraphPad Software, San Diego, CA, USA) or R (version 4.1.1).

III. Results

A. Bioinformatic screening identifies SUMOylation pathway as AML-specific target

First, I performed an integrated analysis on large-scale databases (MILE study stage I and OHSU BeatAML 1.0 program) (Fig. 1A). Selection of significant pathways in the two GSEA results (AML vs. Normal) followed by their clustering based on common leading-edge genes and protein-protein interactions yielded 4-distinct pathway clusters: (i) Translation/rRNA/Mitochondria, (ii) Histone-related, (iii) SUMOylation, and (iv) Regulation of mRNA (Fig. 1B, and Appendix A and B). Interestingly, inhibitors targeting the first cluster, such as ribosome biogenesis inhibitors or tetracyclines, had shown both *in vitro* and *in vivo* anti-leukemic activities and were entered into clinical development [55-57]. These facts show that my bioinformatic results may have real relevance for AML targeting. Of the three remaining clusters, I focused on the (iii) SUMOylation cluster, because it had not been much explored for AML, and the other two were either difficult to establish the causality (‘Histone-related’) or too non-specific (‘Regulation of mRNA’). Most of the individual genes comprising the SUMOylation pathway were found to be upregulated in AML samples from both the MILE and OHSU databases (*SUMO1* and *UBA2* in Fig. 2A; all the others in Figs. 3 and 4). I further validated the results using another large independent dataset from the GEO collection of 2213 AML and

548 normal samples [48]. Consistently, I found that 11 of 17 genes related to SUMOylation were found to be significantly upregulated in AML samples (*SUMO1* and *UBA2* in Fig. 2B; all the others in Fig. 5). In particular, I observed higher protein levels of E1 (SAE1 and SAE2), targets for TAK-981, and E2 (UBC9) in AML patient cells than those in healthy control or patients with remission after therapy (Fig. 2C). I believe these provide further support for the involvement of SUMOylation at the protein level. The results also suggest that the upregulated SUMOylation pathway in AML may be a target for therapeutic intervention.

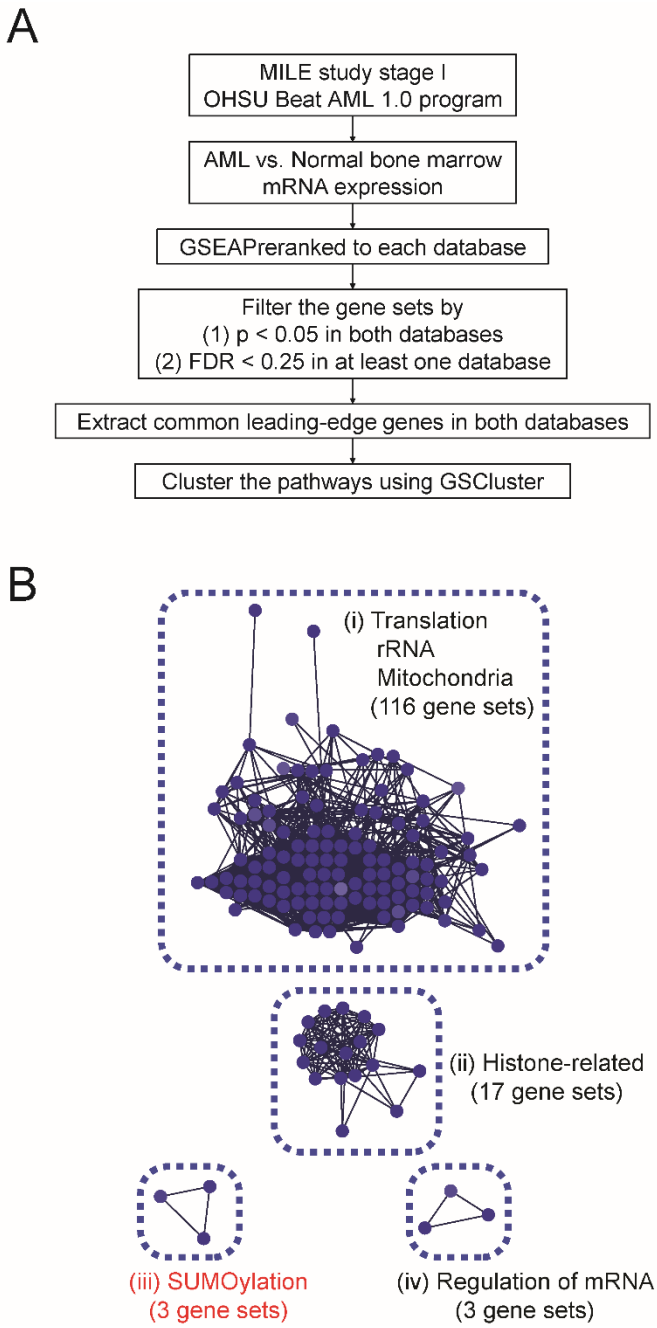


Figure 1. Bioinformatic screening to find AML-specific pathways

(A) Overall strategy for database screening. (B) Graphical illustration of 4 pathway clusters upregulated in AML bone marrow samples from (A), using GSCluster [51] R package. The number of connected gene sets in each cluster is indicated.

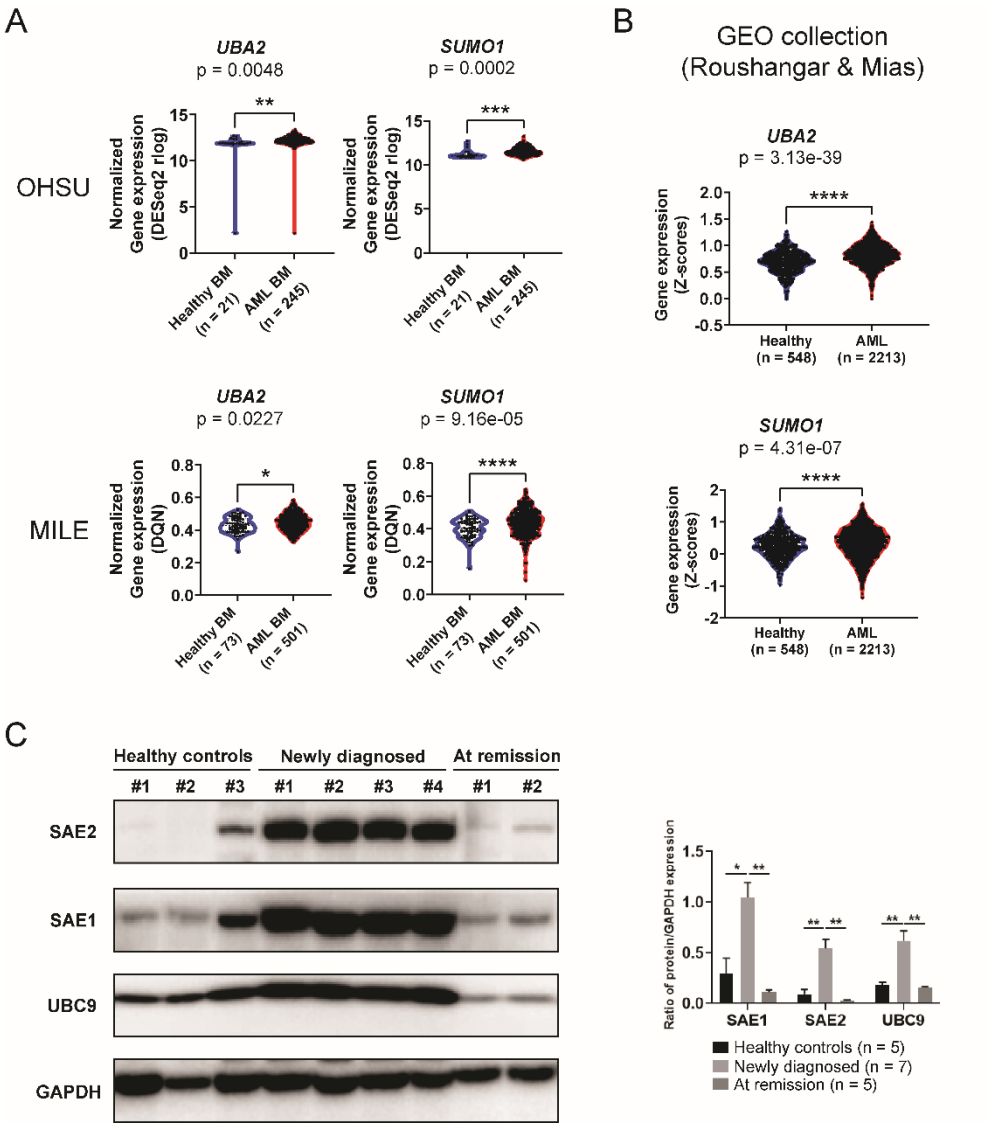


Figure 2. The expression of core genes/pathways of SUMOylation pathway in AML

(A) Comparison of *UBA2* and *SUMO1* gene expression between healthy and AML bone marrow samples in OHSU and MILE databases. (B) Comparison of *UBA2* and *SUMO1* gene expression between healthy and AML bone marrow/peripheral blood samples in GEO datasets by Roushangar and Mias [48]. (C) Left: Representative western blot for SAE2, SAE1, UBC9, and GAPDH in peripheral blood from healthy controls and AML patients at diagnosis or remission state after treatment. Right: The intensities of the bands from the all samples were quantified

by densitometry and displayed as the ratio of each protein to GAPDH (loading control). Newly diagnosed AML patients (n = 7), those at remission state (n = 5), and healthy controls (n = 5). Results are expressed as the mean ± SEM. *P*-values are from Wilcoxon rank-sum test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

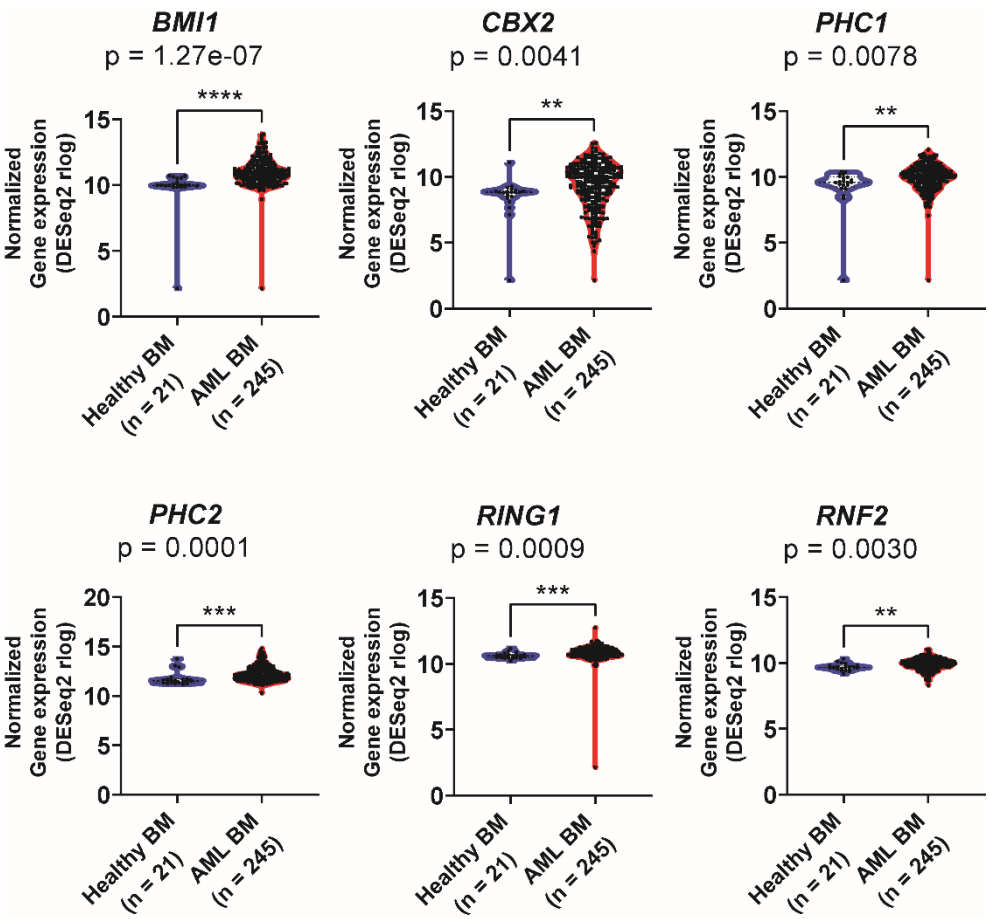


Figure 3. Comparison of the gene expression levels related to SUMOylation between healthy and AML samples in OHSU database

mRNA expression levels of genes related to SUMOylation (except *UBA2* and *SUMO1* that are shown in Fig. 2A) in healthy and AML samples from OHSU database. AML samples are from bone marrow (BM). *P*-values are from Wilcoxon rank-sum test. ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

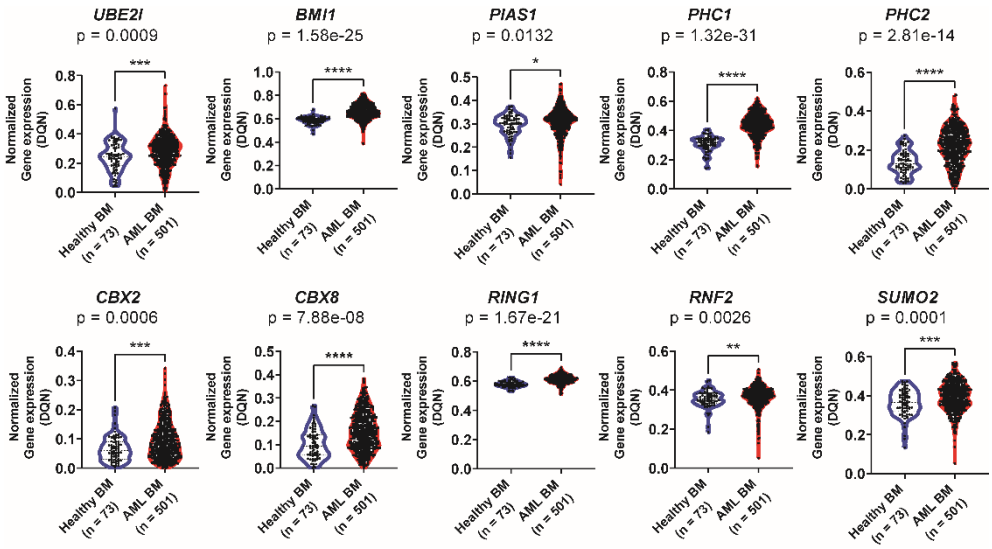


Figure 4. Comparison of the gene expression levels related to SUMOylation between healthy and AML samples in MILE database

mRNA expression levels of genes related to SUMOylation (except *UBA2* and *SUMO1* that are shown in Fig. 2A) in healthy and AML samples from MILE database. AML samples are from bone marrow (BM). *P*-values are from Wilcoxon rank-sum test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

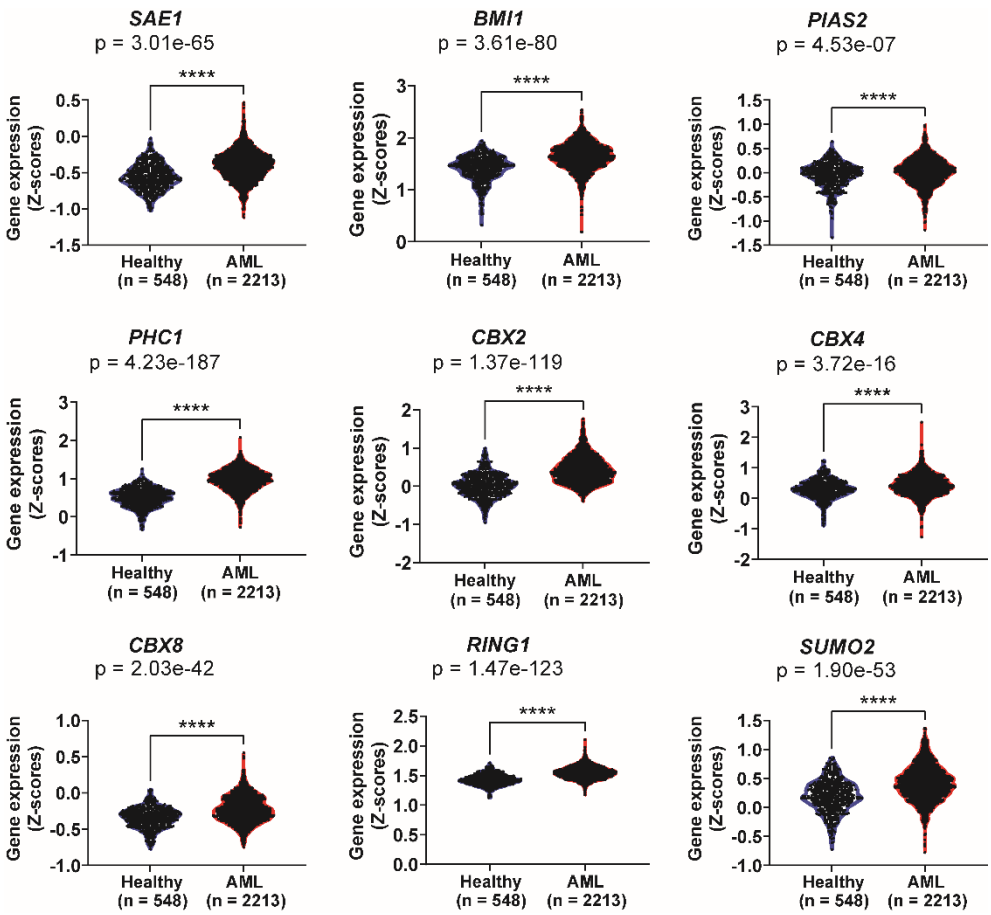


Figure 5. Comparison of the gene expression levels related to SUMOylation between healthy and AML samples in GEO compilation

mRNA expression levels of genes related to SUMOylation (except *UBA2* and *SUMO1* that are shown in Fig. 2B) in healthy and AML samples from GEO collection [48]. AML samples are from peripheral blood. *P*-values are from Wilcoxon rank-sum test. **** $p < 0.0001$.

B. SUMOylation pathway is associated with adverse risk features and poor survival in AML

I then explored the clinical relevance of SUMOylation. First, higher expression of most of the important genes in the SUMOylation pathway from the OHSU database (the survival analysis results of genes analyzed by the median cutoff approach in Fig. 6A; genes analyzed by the best risk separation approach in Fig. 6B) was significantly associated with shorter survival. Some of those negative correlations (for *SAE1*, *BMI1*, and *PHC2*) were validated with the TCGA database (Fig. 7), and all the results, along with those without correlations, are shown in Table 6. Second, the ELN2017 risk analysis on the four groups (healthy, favorable, intermediate, adverse in OHSU database) demonstrated that most of the core genes in the SUMOylation pathway expressed at higher levels in the high-risk groups ($p < 0.05$) (*SUMO1*, *UBA2*, *SAE1* in Fig. 8A, and all the others in Fig. 8B). Post hoc analysis showed that the difference concerning SUMOylation pathway between the healthy and adverse risk group was significant (except for *UBE2I* gene). This trend also was confirmed from the three patient risk groups (favorable, intermediate, adverse) in the TCGA database for several genes including *BMI1*, *CBX2*, and core genes such as *SAE1* and *UBA2*, and the results are shown in Table 7 along with the results for all the other genes without such confirmation [50]. As the above results are for individual gene levels, I further explored the pathway-specific relationship between SUMOylation and overall survival/ELN2017, by performing similar analyses with GSVA pathway scores [58]. Consistent with the results from

individual genes, higher scores of SUMOylation pathways were found to be significantly related with poorer prognosis in both survival analysis and ELN2017 risk analysis (Table 8). These relationships remained valid after adjusting for high-risk AML patient characteristics that might have confounding effects, as evidenced by multivariate analysis (Table 9).

Third, I tested if particular gene mutations are related to core SUMOylation gene expression. Among the four gene mutations (*FLT3*-ITD, *NPM1*, *TP53*, and *RUNX1*) that had enough patients (n > 5) for both mutated and wild-type groups, three mutations (*NPM1*, *TP53*, and *RUNX1*) exhibited consistent patterns between prognosis and core SUMOylation gene expression (*SUMO1* and *UBE2I* in Fig. 9A; all the others in Fig. 9B). Specifically, patients with the *NPM1* mutation associated with better prognosis had lower SUMOylation gene expression, whereas those with the *TP53* and *RUNX1* mutations associated with poor prognosis had higher SUMOylation gene expression. These results suggest that activation of the SUMOylation pathway is associated with adverse risk features and poorer survival.

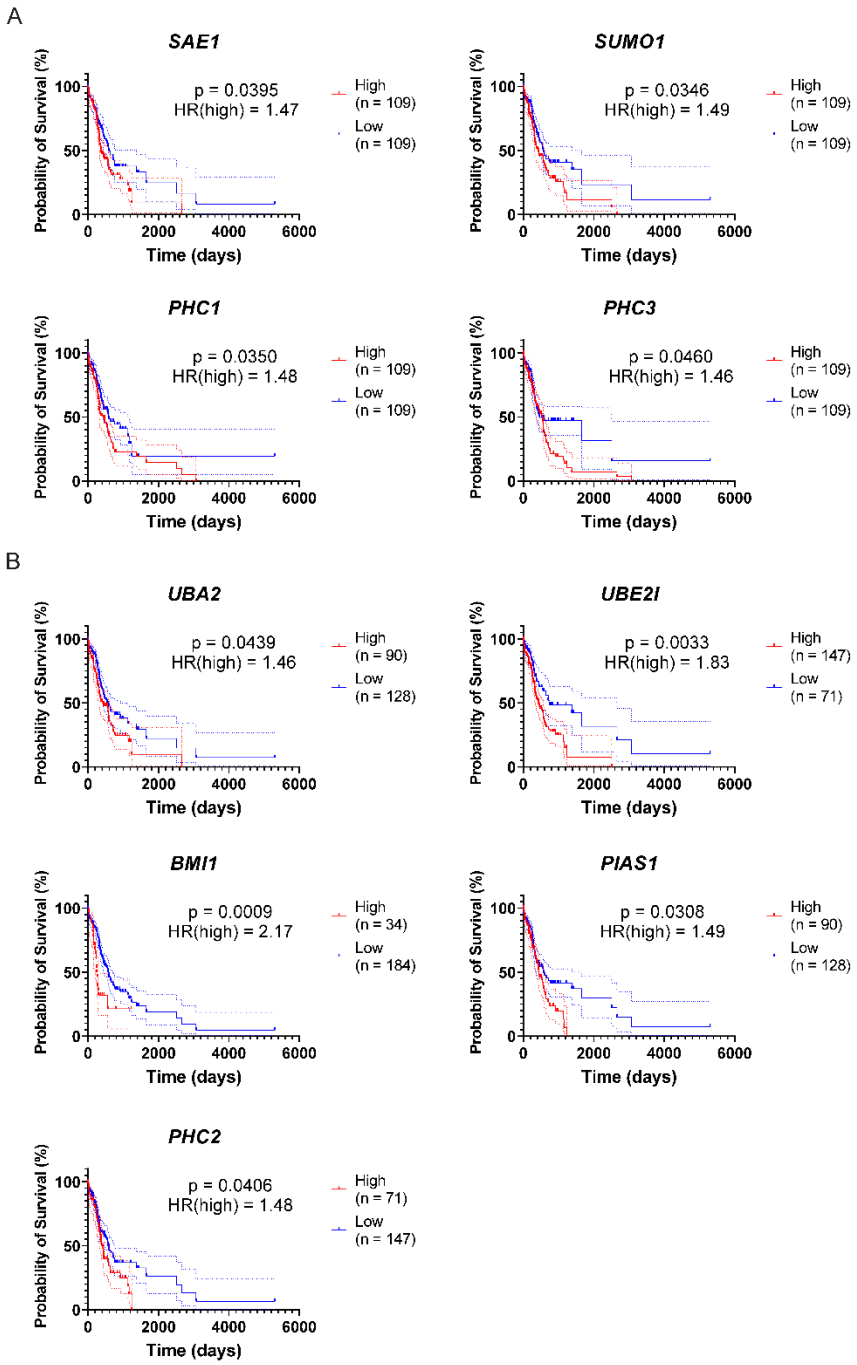


Figure 6. Survival analysis for genes in SUMOylation pathway in OHSU database

Kaplan-Meier curves with 95% confidence intervals (dotted lines) for overall survival of AML patients in OHSU, according to the expression levels of each

indicated gene. The division of the high- and low-expression groups was determined by the median cutoff or the best risk separation approach, and the genes in (A) are those which showed significant result for both approaches. Only the result of the median cutoff approach is shown in (A). The genes in (B) are those which showed significant result not for the median cutoff approach but for the best risk separation approach. HR(high), hazard ratio of high expression group.

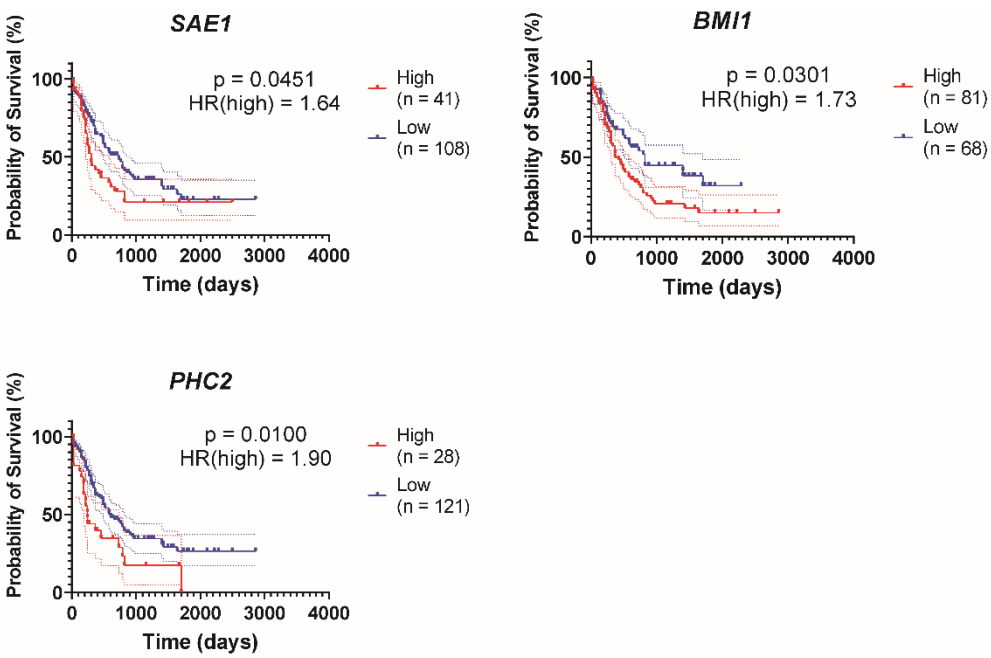


Figure 7. Survival analysis for genes in SUMOylation pathway in TCGA-LAML database

Kaplan-Meier curves with 95% confidence intervals (dotted lines) for genes related to SUMOylation that are significantly associated with poor overall survival in TCGA-LAML. The division of the high- and low-expression groups was determined by the best risk separation approach. HR(high), hazard ratio of high expression group.

Table 6. Survival analysis results in TCGA-LAML database for genes whose high expression showed significance with poor prognosis in OHSU database (Fig. 6)

Genes	*p*-value	HR(high)
SAE1	0.045	1.56
UBA2	0.332	1.28
UBE2I	0.370	1.25
PIAS1	0.334	0.81
BMI1	0.010	1.73
PHC1	0.040	0.65
PHC2	0.010	1.90
PHC3	0.003	0.53
SUMO1	0.149	1.43

The division of the high- and low-expression groups was determined by the best risk separation approach. HR(high), hazard ratio of high expression group.

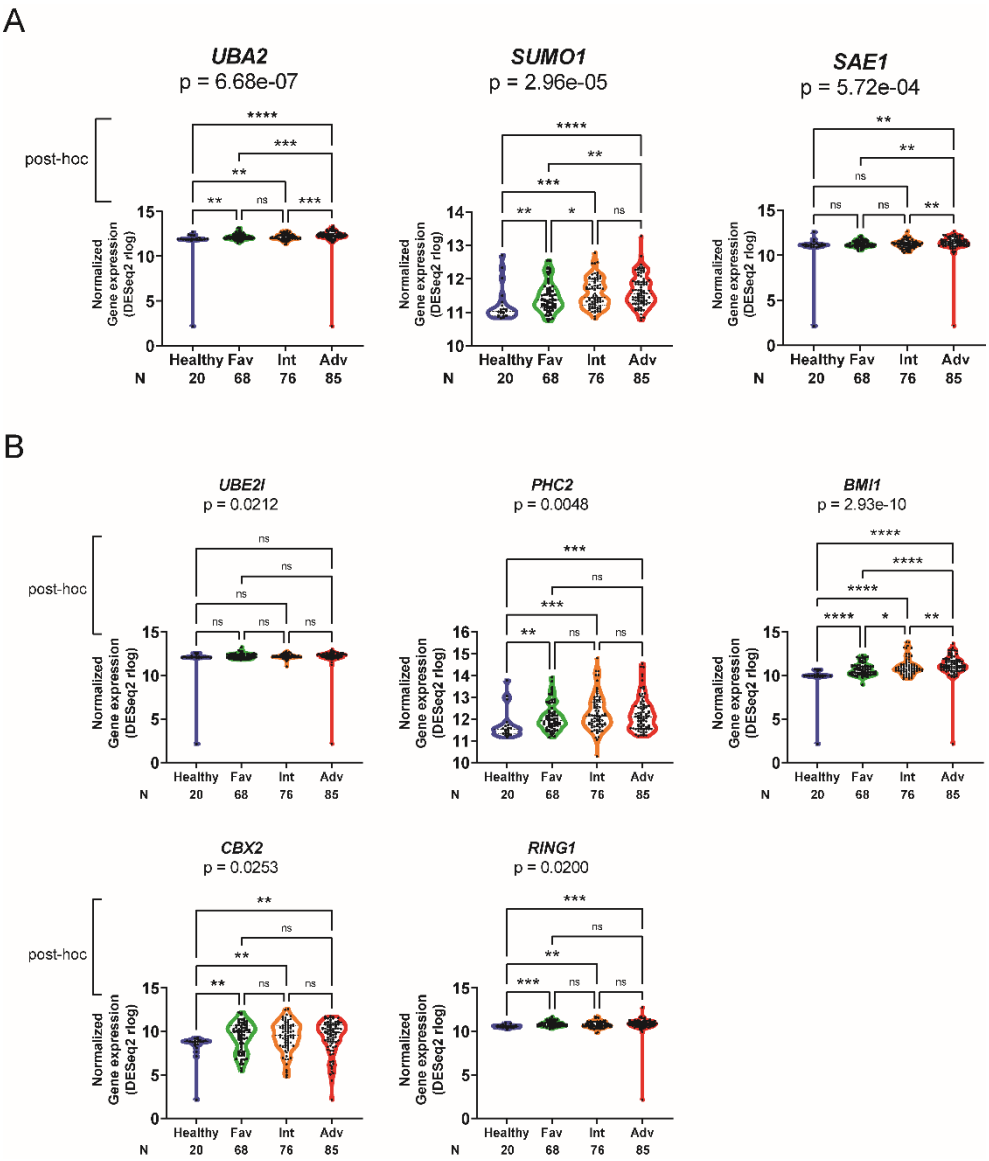


Figure 8. ELN2017 analysis for genes in SUMOylation pathway in OHSU database

Comparison of expression of (A) *UBA2*, *SUMO1*, and *SAE1* and (B) five other indicated genes in SUMOylation pathway across healthy and ELN2017 risk groups. *P*-values are from Jonckheere-Terpstra test. Subsequent post hoc analyses were performed with the two-stage linear step-up procedure, and the significance is indicated for each comparison. The number of subjects is indicated for each group. ns, not significant; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Table 7. ELN2017 analysis results in TCGA-LAML database for genes which showed significant increasing trend with ELN2017 risk categories in OHSU database (Fig. 8)

Genes	*p*-value
SAE1	3.32e-04
UBA2	4.22e-02
UBE2I	8.35e-01
BMI1	1.21e-02
PHC2	9.14e-01
CBX2	6.21e-04
RING1	1.17e-01
SUMO1	5.05e-01

Table 8. SUMOylation pathways whose GSVA pathway scores show significance in both overall survival and ELN2017 analysis in OHSU database

Pathways	Survival *p*-value	Survival HR(high)	ELN2017 *p*-value
BIOCARTA_SUMO_PATHWAY	0.0483	1.44	7.03e-05
REACTOME_SUMO_IS_PROTEOLYTICALLY_PROCESSED	0.0299	1.51	6.33e-06
REACTOME_SUMOYLATION_OF_DNA_METHYLATION_PROTEINS	0.0129	1.59	3.45e-05
REACTOME_SUMOYLATION_OF_IMMUNE_RESPONSE_PROTEINS	0.0249	1.52	7.33e-03

The division of the high- and low-expression groups was determined by the median cutoff. HR(high), hazard ratio of high expression group.

Table 9. Univariate and Multivariate analysis of overall survival utilizing the gene expression and clinical information in OHSU database

Clinical characteristic	Univariate analysis		Multivariate analysis	
	HR (95% CI)	***P*-value**	**HR (95% CI)**	***P*-value**
Sex (Male vs. Female)	1.40 (0.82-2.37)	0.333		
Relapse (TRUE vs. FALSE)	0.55 (0.19-1.56)	0.260		
Cumulative chemo (y vs. n)	0.27 (0.12-0.59)	0.001	0.15 (0.03-0.81)	0.027
ELN2017 (Favorable vs. Adverse)	0.32 (0.18-0.59)	< 0.001	0.17 (0.04-0.68)	0.012
Age at specimen acquisition	1.03 (1.01-1.04)	0.002	1.00 (0.98-1.02)	0.849
Mutation (positive vs. negative)				
FLT3-ITD	1.12 (0.58-2.16)	0.746		
RUNX1	2.44 (0.91-6.58)	0.077		
ASXL1	1.06 (0.12-9.14)	0.958		
NPM1	0.49 (0.25-0.96)	0.038	0.66 (0.11-4.10)	0.653
TP53	4.23 (2.00-8.92)	< 0.001	1.01 (0.39-2.64)	0.981
CEBPA_Biallelic (y vs. n)	3.80e-08 (0-Inf)	0.996		
SUMOgene	2.18 (1.35-3.52)	0.002	2.60 (1.09-6.22)	0.032

HR, Hazard Ratio. CI, Confidence Interval. 'SUMOgene' refers to GSVA pathway score of "BIOCARTA_SUMO_PATHWAY".

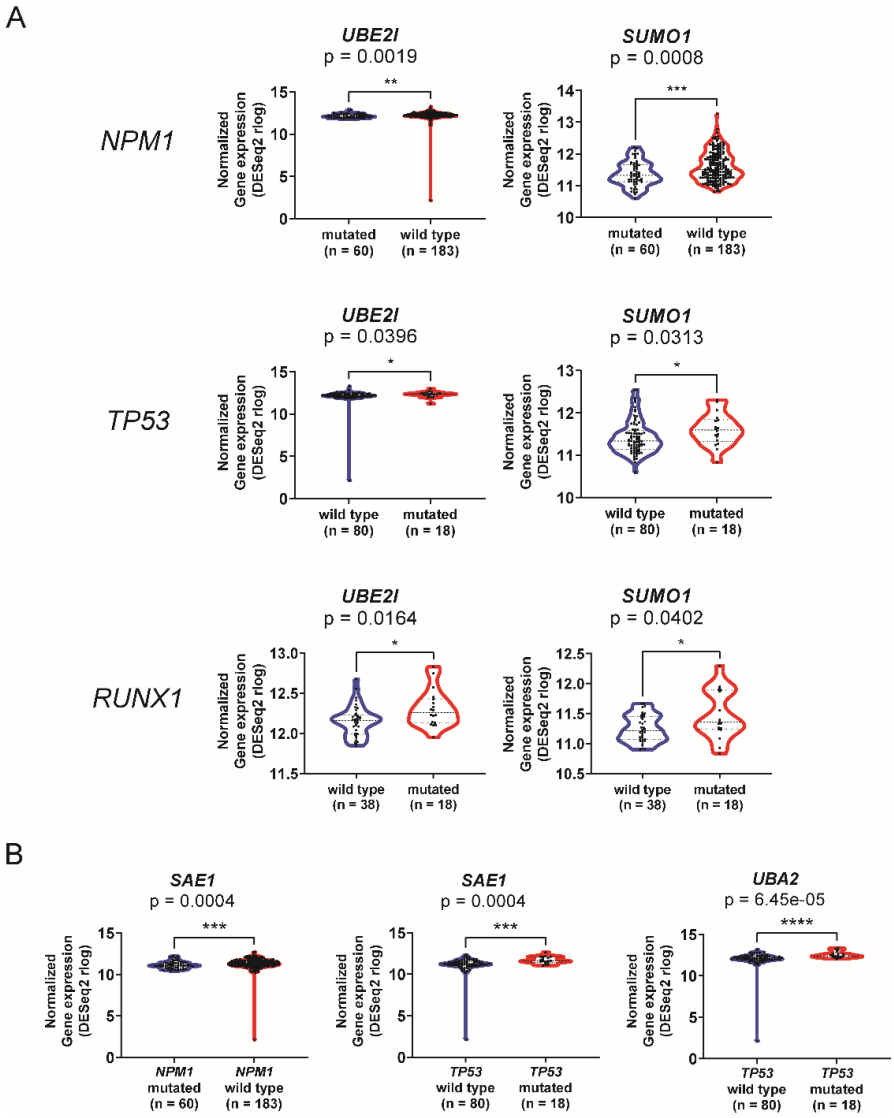


Figure 9. Comparison of gene expression between groups with or without AML-relevant mutations for genes in SUMOylation pathway in OHSU database

(A) Comparison of *UBE2I* and *SUMO1* gene expression between mutated and wild-type of *NPM1*, *TP53* and *RUNX1* genes in OHSU database. (B) Comparison of *SAE1* gene expression in patients with mutated and wild-type *NPM1* (left), and comparison of *SAE1* (middle) and *UBA2* (right) gene expression in patients with mutated and wild-type *TP53* in OHSU database. *P*-values are from Wilcoxon rank-sum test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

C. TAK-981, a new SUMOylation inhibitor, exhibits potent anti-leukemic effects *in vitro*

In my quest for an inhibitor of SUMOylation, I found TAK-981, which was developed very recently as a first-in-class inhibitor of SAE step [43] and is currently under clinical trials for various solid tumors. As its effects against AML are still unknown, I evaluated them *in vitro*.

Surprisingly, TAK-981 showed larger or similar potency compared with cytarabine (Ara-C), a standard drug used in clinics, against four AML cell lines (Fig. 10A). Notably, the IC_{50} values for TAK-981, all within a two-digit nanomolar range, were somewhat uniform across the cell lines. By contrast, those for cytarabine differed markedly (> 1 micromolar for KG-1 and THP-1; two-digit nanomolar range for U937). In comparison, tetracycline, targeting the "Translation/rRNA/Mitochondria" identified above, exhibited only several-hundred-micromolar potency (Fig. 10B).

Next, I tested TAK-981 for any synergistic or dose-reduction effect when used with cytarabine for the four cell lines (Fig. 11). In addition, TAK-981's synergy with two new targeted-therapy drugs, venetoclax and quizartinib, along with a demethylating drug, azacitidine, was tested for the MOLM-14 cell line having the *FLT3*-ITD mutation, which is associated with poor prognosis (Fig. 11). Synergy, as judged by the CompuSyn scores [59], varied substantially across cell lines, with U937 and MOLM-14 exhibiting significant synergy, while KG-1 and THP-1 showing little synergy in the combination with cytarabine. For MOLM-14,

TAK-981 exhibited significant synergy with azacitidine, some synergy at higher drug concentrations with venetoclax, but no synergy with quizartinib. In addition, TAK-981 showed similar and lower potency in comparison with venetoclax, a BCL2 inhibitor, and quizartinib, an FLT3 inhibitor, respectively (Figs. 12 and 13). Although I used only concentration values around IC_{50} for each drug, significant synergy might be observed with different concentration combinations. I also assessed the dosage reduction effects of TAK-981 (Table 10). Notably, even when there was no apparent synergy, the dose reduction indices (DRI) of the drugs combined with TAK-981 were above 1 for all of the drug-cell-line settings, indicating significant dosage reduction effects. This could be exploited to lower the toxicity of such drugs when combined with TAK-981. Overall, TAK-981's combination with conventional or targeted drugs holds promise for improved therapeutics.

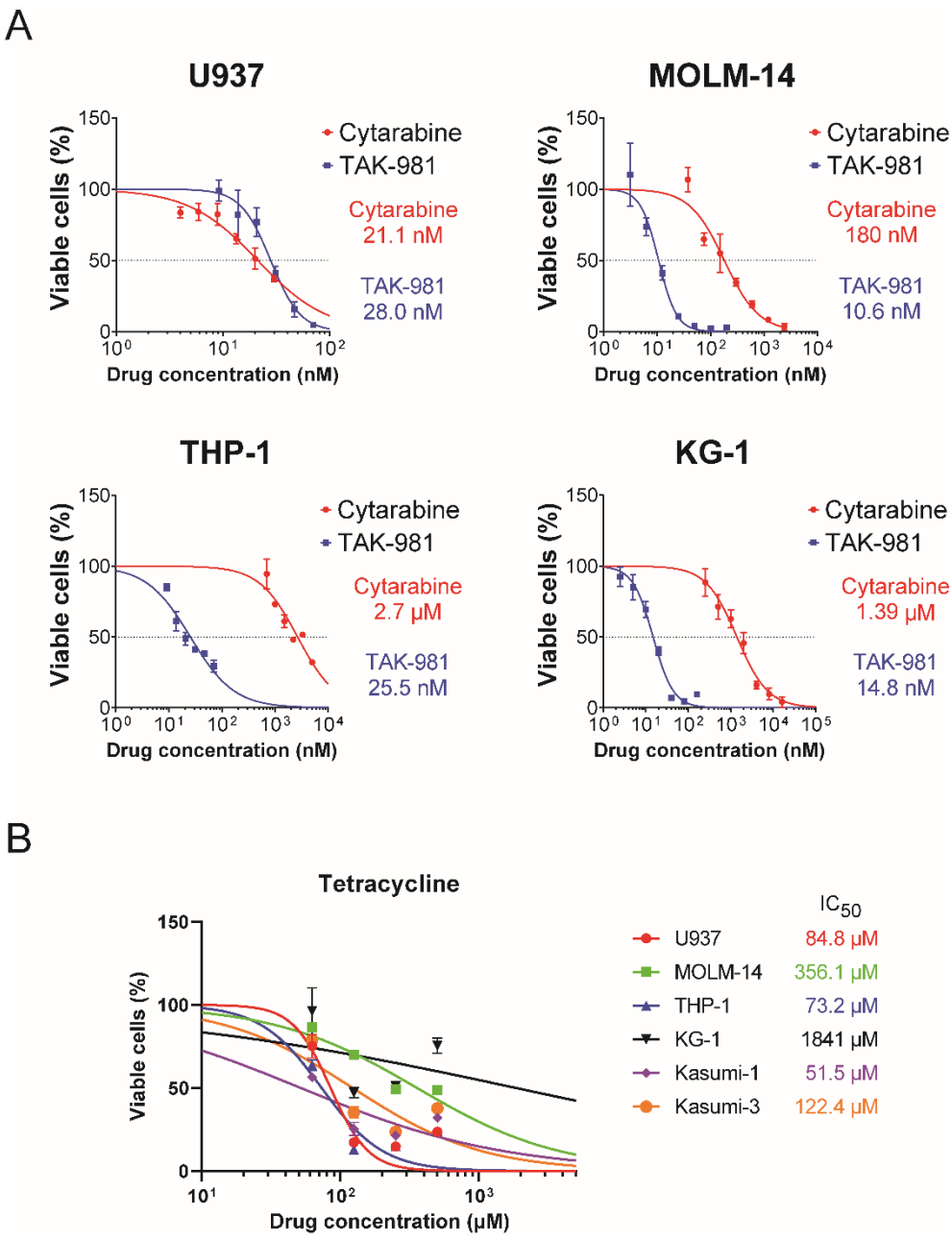


Figure 10. TAK-981 and tetracycline's potency for AML cells

(A) Dose-response curves of TAK-981 and cytarabine for four AML cell lines. The concentration values right beside each curve represent IC_{50} values. (B) Dose-response curves of tetracycline for six AML cell lines. Cell viability was measured by CCK-8 assay.

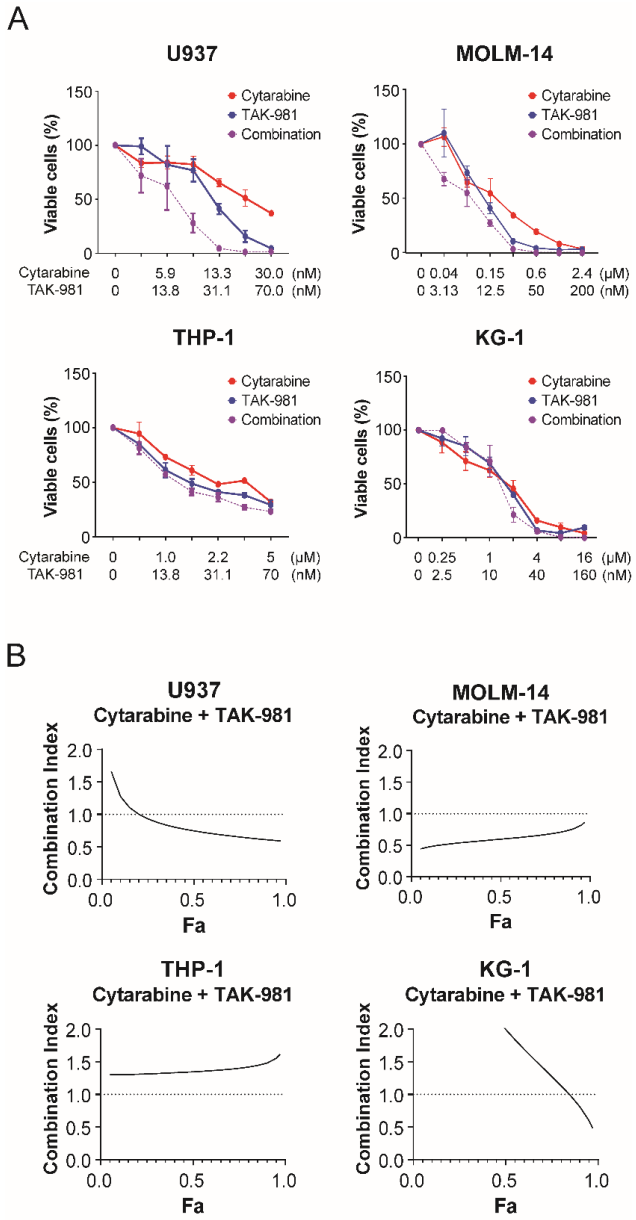


Figure 11. TAK-981's synergy with cytarabine or other drugs for AML cells

(A) Synergy between TAK-981 and cytarabine for four AML cell lines. (B) Combination index plots computed from the data in (A) by CompuSyn software. For (A), different concentration ranges were used for each drug, and the error bars indicate standard deviation. Also, cell viability was measured by CCK-8 assay. For (B), values below the dotted line at 1.0 indicate synergy, and Fa refers to 'Fractions affected'.

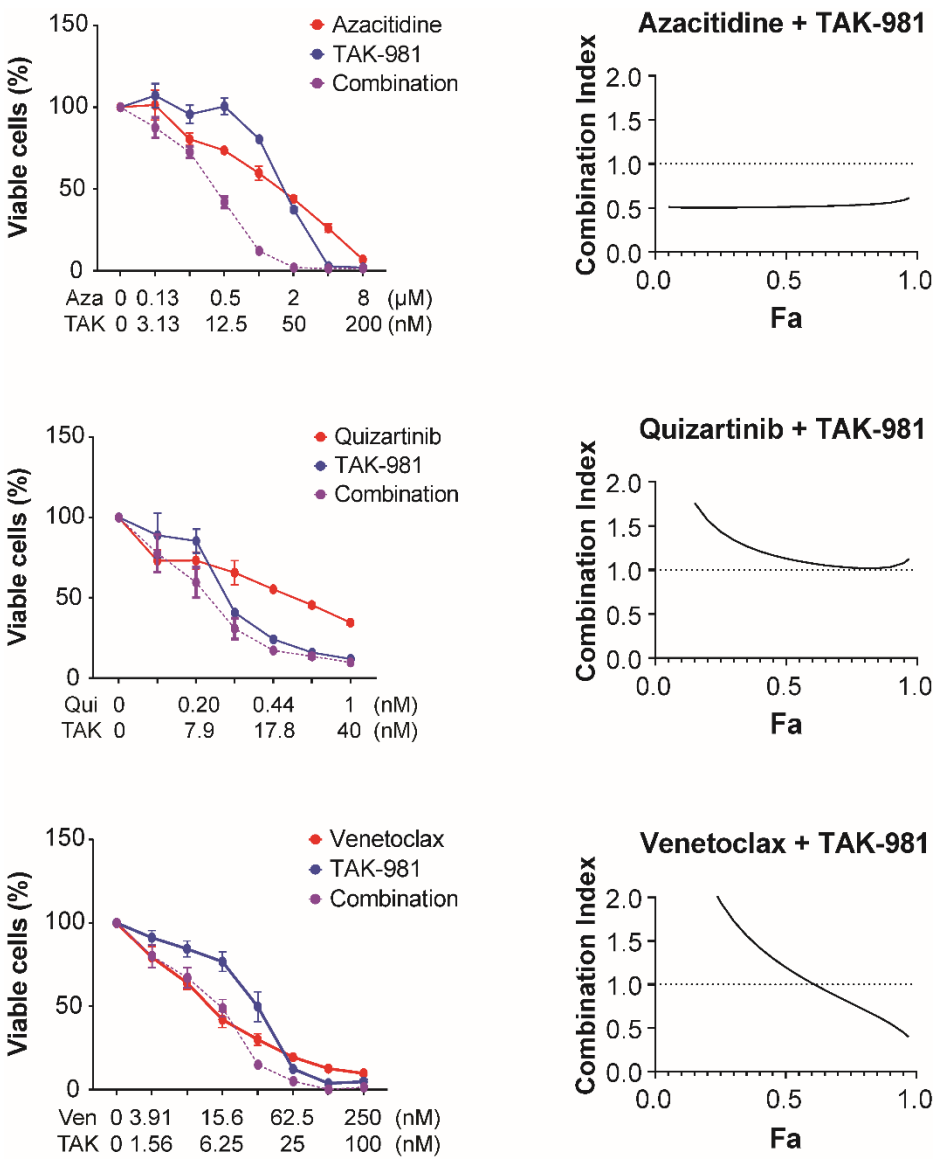


Figure 12. TAK-981's synergy with azacitidine, quizartinib or venetoclax for AML cells

Synergy between TAK-981 and several drugs for MOLM-14 cell line and combination index plots computed by CompuSyn software. For combination index plots, values below the dotted line at 1.0 indicate synergy, and Fa refers to 'Fractions affected'. Aza, Qui, Ven, and TAK refer to azacytidine, quizartinib, venetoclax, and TAK-981, respectively. Cell viability was measured by CCK-8 assay.

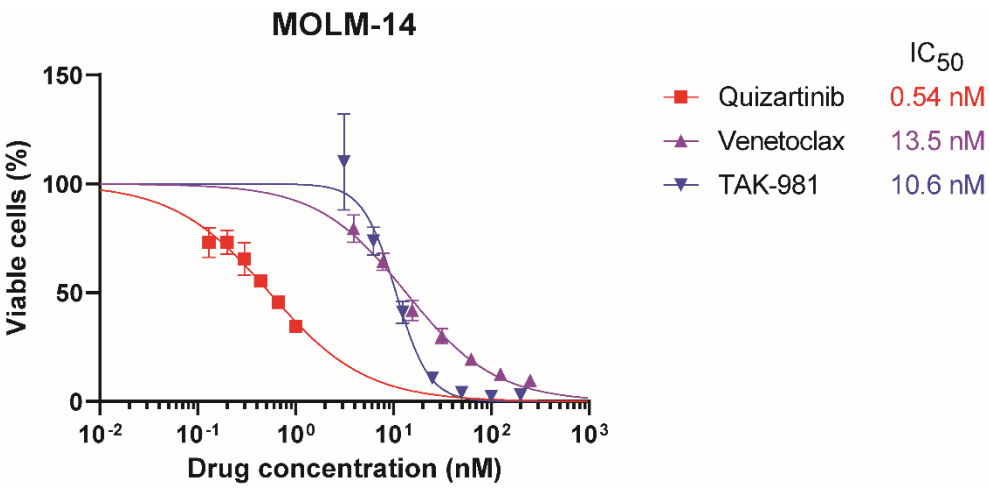


Figure 13. Comparison of potency of quizartinib, venetoclax and TAK-981 for MOLM-14 cell line

Dose-response curves of quizartinib, venetoclax, and TAK-981 for MOLM-14 cell line. Cell viability was measured with CCK-8 assay.

Table 10. Dose reduction index of cytarabine or other drugs when combined with TAK-981 in AML cell lines

Cell line	Combination Drug	Dose Reduction Index (DRI) at Fa = 0.9
U937	Cytarabine	12.03
THP-1	Cytarabine	1.25
KG-1	Cytarabine	2.61
MOLM-14	Cytarabine	2.94
MOLM-14	Azacitidine	4.87
MOLM-14	Quizartinib	8.64
MOLM-14	Venetoclax	4.98

Fa, Fractions affected. Fa = 0.9 refers to the point where the inhibition effect is 90%, i.e., when 90% of the cells are dead. The number 0.9 was chosen, since for cancer therapies, high effect levels are thought to be more therapeutically relevant than low effect levels [59].

D. TAK-981 induces apoptosis, cell-cycle arrest, and/or differentiation marker expression in AML cell lines

To study how TAK-981 exhibits anti-leukemic effects, I investigated cellular events upon drug treatment. As expected, TAK-981 reduced SUMOylation for some of the proteins, if not all, from the cell extracts (24 h or 48 h treatment, Figs. 14 and 15). Because SUMOylation plays a critical role in transcription regulation, I next analyzed gene expression profile changes by TAK-981 treatment (16 h) using GSEA (GSE173116 [39]: THP-1 cells; Fig. 16). The upregulated pathways included those for cell death and cell-cycle arrest, such as the p53 pathway and apoptosis. Experimentally, the mRNA expression of genes for apoptosis (*DDIT3*) and cell-cycle arrest (*P21* and *TP53*), known to be downregulated by SUMOylation in AML cells [32, 42, 60], were significantly higher in TAK-981-treated THP-1 cells (48 h) than in those from the control or cytarabine-treated group (Fig. 17). I also found that there was a trend that SUMO core pathway is downregulated in TAK-981-treated THP-1 cells (Fig. 18), although TAK-981's post-translational effect on SUMO may not necessarily involve the expression of SUMO core genes. Further analysis in several other AML cell lines with Western blot (p21, caspase 3, and cytochrome C; Fig. 19), flow cytometry for apoptosis (Fig. 20), and DNA content analysis (Fig. 21) showed that apoptosis and cell-cycle arrest were generally observed for the TAK-981-treated AML cells (48 h), with only minor variations. For example, G2/M phase arrest was observed for U937, THP-1, and KG-1 cells, whereas G0/G1 arrest was observed in MOLM-14

cells. Meanwhile, there is heterogeneity in terms of p53 mutations among the cell lines used in this study (Table 11). As p21 can be regulated either by p53 dependently or independently, I tested if the induction of p21 by TAK-981 is also reflected in the p53. TAK-981 treatment did not change the levels of either p53 or MDM2 (Fig. 22), suggesting that TAK-981-induced p21 change may not be related to p53. Possible mechanistic disconnection between p53 and p21 upon TAK-981 treatment could be an interesting topic for future research.

TAK-981 treatment (48 h) also affected the differentiation of leukemic cells dose-dependently, as shown by the increase in the differentiation markers for U937 (*CD15*) [60-62], THP-1 (*CD14*), and MOLM-14 (*CD11B*) cells (Fig. 23). Moreover, TAK-981 suppressed the expression of *CD39* (48 h, Fig. 24), which is known to be involved in AML chemoresistance [63], in both chemo-sensitive (U937) and chemo-resistant cells (KG-1, THP-1, MOLM-14). These data suggest that TAK-981 exhibits anti-leukemic effects by inducing apoptosis, cell-cycle arrest, differentiation, or lower chemoresistance.

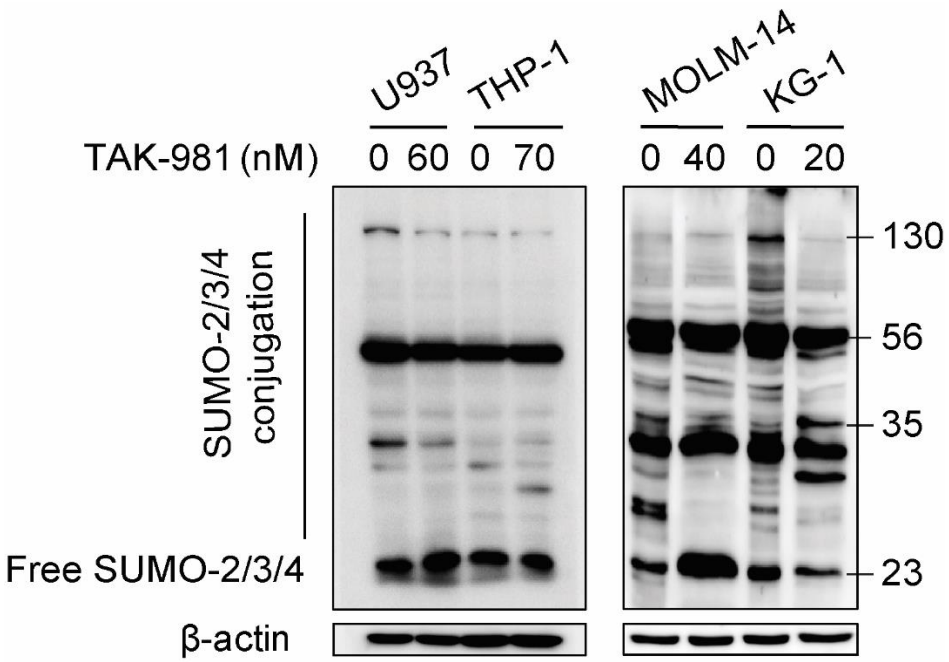


Figure 14. Effect of TAK-981 on protein SUMOylation

The effect of TAK-981 on protein SUMOylation in AML cells after 24 h (U937, THP-1) or 48 h (MOLM-14, KG-1) treatment. Western blot analysis was performed with the antibody for SUMO-2/3/4.

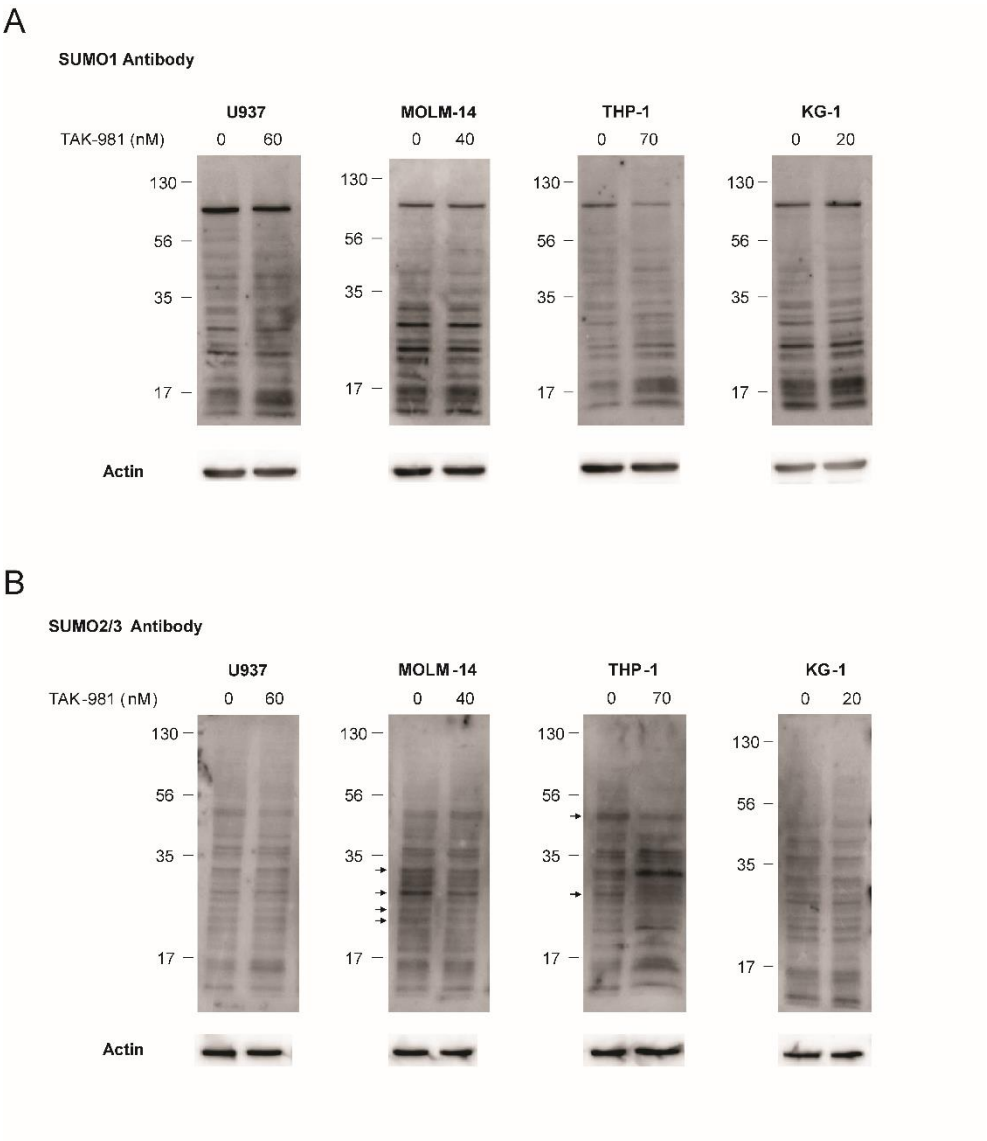


Figure 15. Western blot showing the effect of TAK-981 on protein SUMOylation with SUMO1- and SUMO2/3- specific antibodies

Effect of TAK-981 on protein SUMOylation in AML cells after 48 h treatment. Western blot analysis was performed with the antibodies specific for (A) SUMO1 and (B) SUMO2/3. For (B), the arrows indicate decreased protein SUMOylation in TAK-981-treated cells.

NAME OF PATHWAY	NES	*P*-value	FDR
TNF alpha signaling via NF-κB	1.981	0.000	0.000
Interferon alpha response	1.843	0.000	0.000
Interferon gamma response	1.721	0.000	0.000
Inflammatory response	1.698	0.000	0.000
IL6/ JAK/ STAT3 signaling	1.692	0.000	0.000
TGF beta signaling	1.589	0.000	0.001
P53 pathway	**1.567**	**0.000**	**0.002**
KRAS signaling up	1.450	0.000	0.016
IL2/ STAT5 signaling	1.437	0.000	0.017
Notch signaling	1.399	0.040	0.028
Apoptosis	**1.349**	**0.014**	**0.052**
Hypoxia	1.346	0.011	0.049

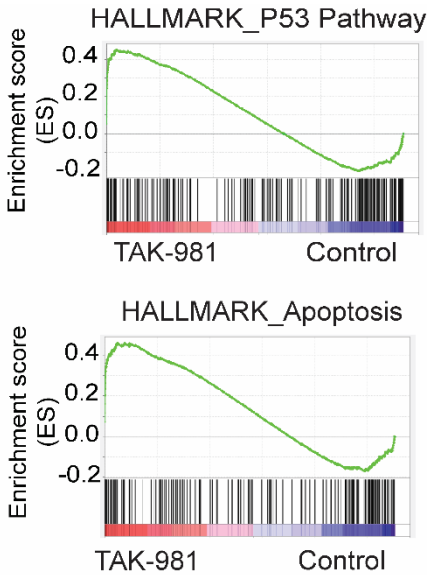


Figure 16. GSEA result from GSE173116 dataset

Left: Top 12 pathways with $p < 0.05$ from GSEA analysis of TAK-981-treated THP-1 cells from GSE173116 data set with the Hallmark gene set. The pathways are in the order of the normalization of the enrichment score (NES). Right: Enrichment score plots for genes belonging to p53 and apoptosis pathways from the GSEA analysis.

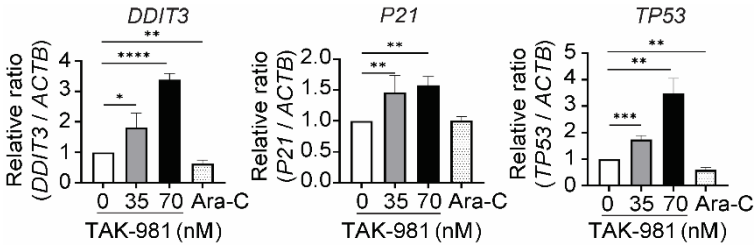


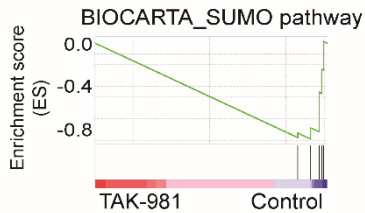
Figure 17. Comparison of mRNA expression of genes related with apoptosis or cell-cycle arrest after TAK-981 treatment

Relative mRNA expression of *DDIT3*, *P21*, and *TP53* in TAK-981 (indicated concentrations) and cytarabine (1 μM) in THP-1 cells after 48 h treatment, as measured by qRT-PCR. Two-tailed Student's *t*-test was used. Data are expressed as mean ± SD (n = 3), * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. The efficiencies of the primers used are listed in Table 12.

A

NAME OF PATHWAY	NES	*P*-value	FDR
SM pathway	-1.73	0.000	0.245
Proteasome pathway	-1.72	0.007	0.128
Eosinophils pathway	-1.69	0.000	0.135
NO1 pathway	-1.59	0.016	0.341
CTBP1 pathway	-1.53	0.047	0.548
PEPI pathway	-1.50	0.034	0.563
SUMO pathway	**-1.42**	**0.074**	**0.926**

B



C

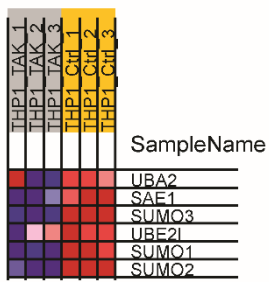


Figure 18. Result of GSEA analysis from GSE173116 data with Biocarta gene set

(A) Top 7 pathways which are downregulated in TAK-981-treated THP-1 cells. The pathways are in the order of the normalization of the enrichment score (NES). (B) Enrichment score plots for "BIOCARTA_SUMO pathway" from the GSEA analysis from (A). (C) Heatmap for the genes comprising the "BIOCARTA_SUMO pathway". The left 3 columns represent TAK-981-treated THP-1 cells, and the right 3 columns represent control THP-1 cells.

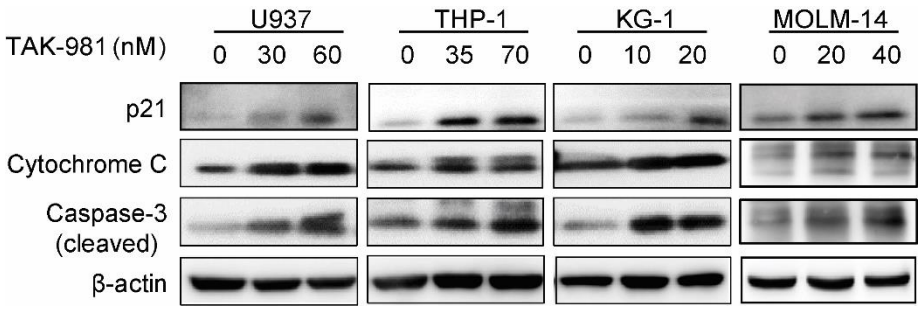


Figure 19. Comparison of protein expression of proteins related with apoptosis or cell-cycle arrest after TAK-981 treatment

Western blot for p21, cleaved caspase-3, and cytochrome C expression in AML cells after 48 h treatment with TAK-981. All experiments were done with n = 3.

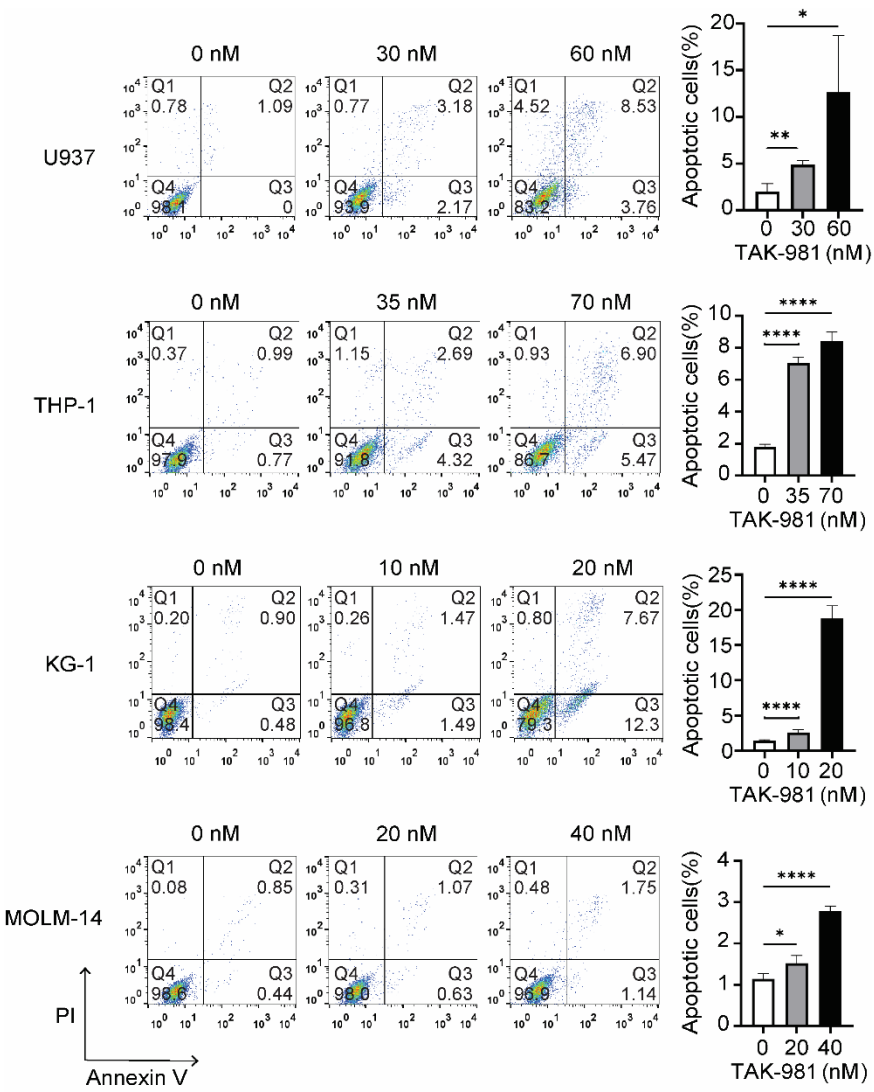


Figure 20. Apoptosis analysis for TAK-981-treated AML cells

Apoptosis analysis for TAK-981-treated AML cells after 48 h by flow cytometry with Annexin V/PI kit. Apoptotic cells (%) (right) is the sum of the early (Q3) and late (Q2) apoptosis percentages. Two-tailed Student's *t*-test was used. Data are expressed as mean ± SD (n = 3), * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

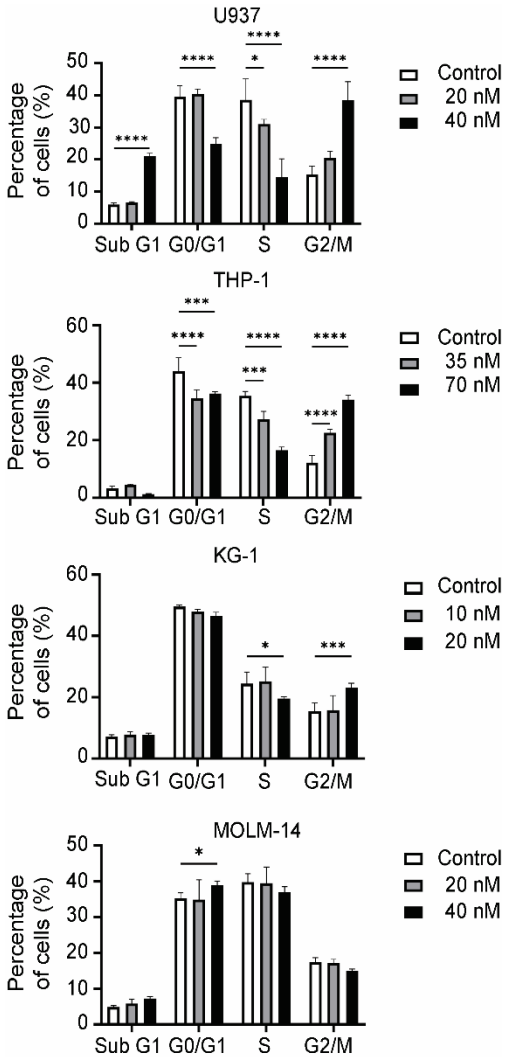


Figure 21. Cell-cycle analysis for TAK-981-treated AML cells

Cell-cycle analysis for TAK-981-treated AML cells after 48 h by flow cytometry. Each phase of cell cycle was analyzed with Cell-cycle platform in FlowJo software. Data are expressed as mean ± SD (n = 3). Two-tailed Student's *t*-test was used. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Table 11. The known *TP53* mutation status for AML cell lines used in this study

AML cell lines	*TP53* mutation status
U937	Positive
MOLM-14	Negative
THP-1	Positive
KG-1	Positive

The information was retrieved from Cell Model Passports (https://cellmodelpassports.sanger.ac.uk) and references [64, 65].

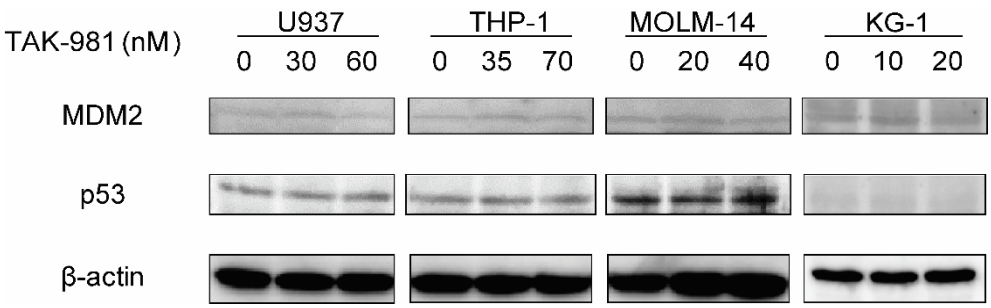


Figure 22. Western blot for MDM2 and p53 expression in AML cell lines

The four AML cell lines (U937, THP-1, MOLM-14, KG-1) were treated for 24 hours with TAK-981. All experiments were done with n = 3.

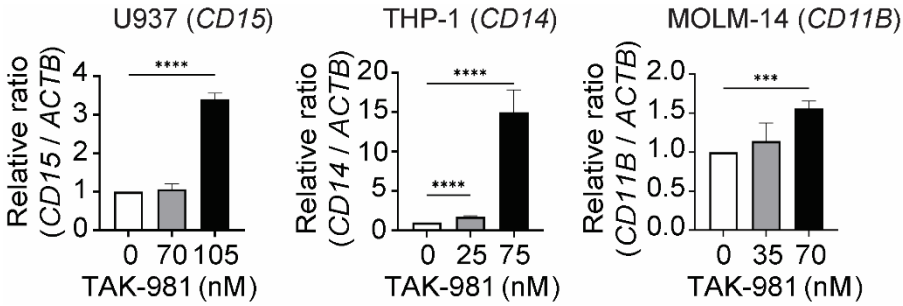


Figure 23. Comparison of gene expression of differentiation marker genes after TAK-981 treatment

qRT-PCR analysis of differentiation markers. mRNA expression in 48 h-TAK-981-treated AML cells for *CD15* in U937, *CD14* in THP-1, and *CD11B* in MOLM-14. Two-tailed Student's *t*-test was used. Data are expressed as mean ± SD (n = 3), * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. The efficiencies of the primers used are listed in Table 12.

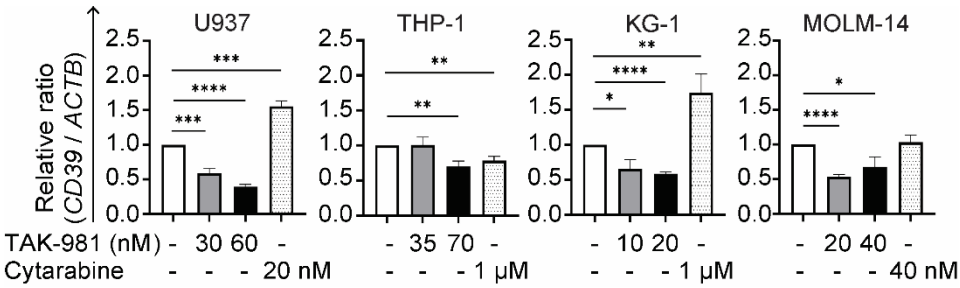


Figure 24. Comparison of *CD39* gene expression related with chemoresistance after TAK-981 treatment

qRT-PCR analysis of *CD39* gene. mRNA expression in 48 h-TAK-981-treated AML cells *CD39* in all cells. Data are expressed as mean ± SD (n = 3), * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. The efficiencies of the primers used are listed in Table 12.

Table 12. RT-qPCR primers (Bioneer, Daejeon, South Korea)

Genes	Primer (5'→3') Forward/Reverse	NCBI reference	Efficiency
ACTB (β-actin)	CATGTACGTTGCTATCCAGGC CTCCTTAATGTCACGCACGAT	NM_001101.5	90.12%
ITGAM (*CD11B*)	P201135 V	NM_000632.3	97.18%
CD14	P151972 V	NM_000591.3	96.81%
FUT4 (*CD15*)	P145260 V	NM_002033.3	109.16%
ENTPD1 (*CD39*)	GGAGACGGACCACAGCAAG TTGTTCTGGGTCAACCCCAC	NM_001776.6	96.78%
DDIT3	TTGCCTTTCTCCTTCGGGAC CAGTCAGCCAAGCCAGAGAA	NM_001195053.1	95.63%
CDKN1A (*P21*)	GGCATAGAAGAGGCTGGTGG CATTAGCGCATCACAGTCGC	NM_001220777.2	106.61%
TP53	TGACACGCTTCCCTGGATTG TCCGGGGACAGCATCAAATC	NM_001276695.3	103.77%

NCBI, National Center for Biotechnology Information

E. TAK-981 potency in primary AML cells *ex vivo*

The effects of TAK-981 also were evaluated *ex vivo* in primary AML cells from patient bone marrow (n = 13). TAK-981 exhibited higher inhibition of primary cell proliferation at equimolar concentrations than did cytarabine which did not appreciably inhibit the cells at up to ~50 micromolar concentrations (Fig. 25A). Interestingly, the inhibitory potencies of both compounds for the primary cells were much lower than those for the AML cell lines. Also, the SUMOylation status of primary AML cells from patients was lower than that in the cell lines (Fig. 26). The possible reasons for these differences between cell lines and primary cells are addressed in the discussion section.

Still, there was significant synergy between the two drugs against the primary cells (Fig. 25B), indicating the possible clinical utility of TAK-981. Consistently with the AML cell-line results, TAK-981 induced apoptosis in the primary AML cells, and this result suggests its direct effect on cancer cells independent of anti-tumor immunity (Fig. 27).

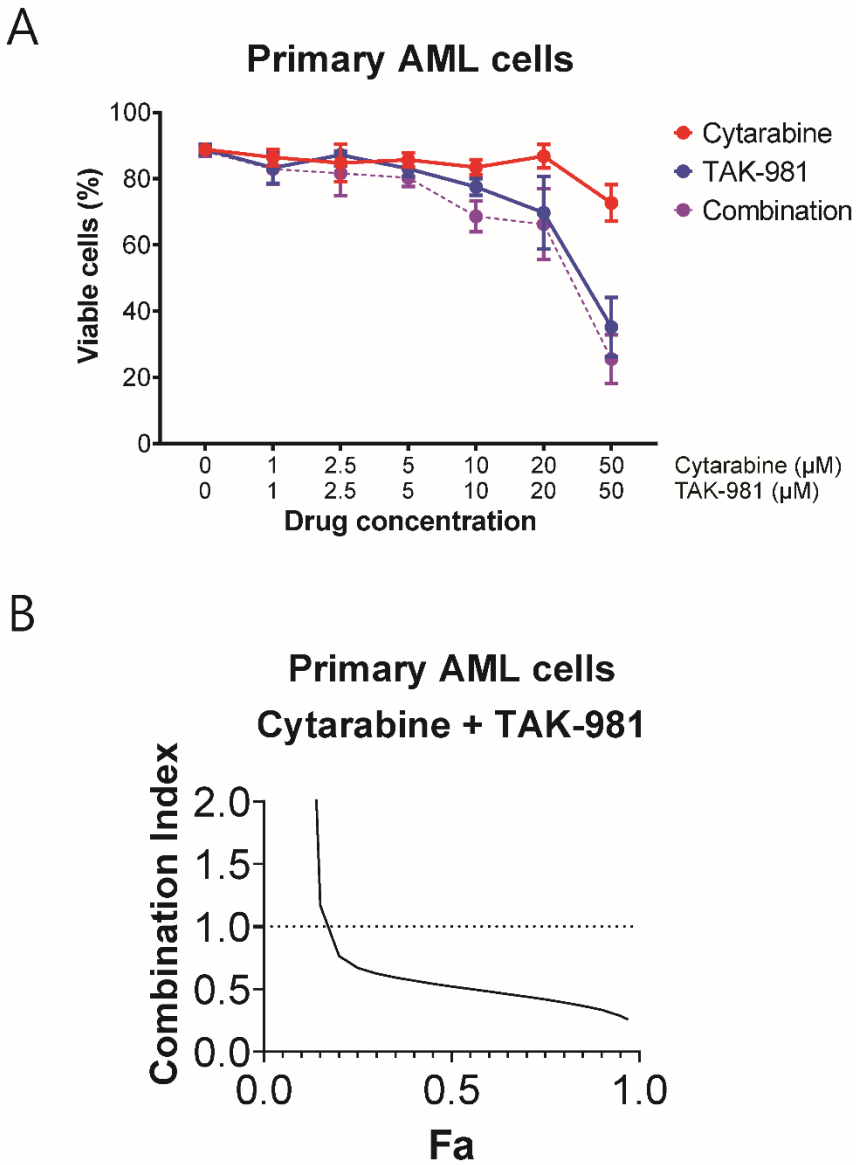


Figure 25. TAK-981's activity against primary AML cells *ex vivo*

Freshly isolated mononuclear cells from bone marrow of 13 patients with AML were cultured with different doses of TAK-981, cytarabine (Ara-C), or both for 48 h. (A) Potency and combination effects of TAK-981 and cytarabine. Viable cells were estimated by flow cytometric analysis of primary AML cells treated with TAK-981, cytarabine or both. Error bars are standard errors. (B) Synergistic combination index between TAK-981 and cytarabine from data in (A).

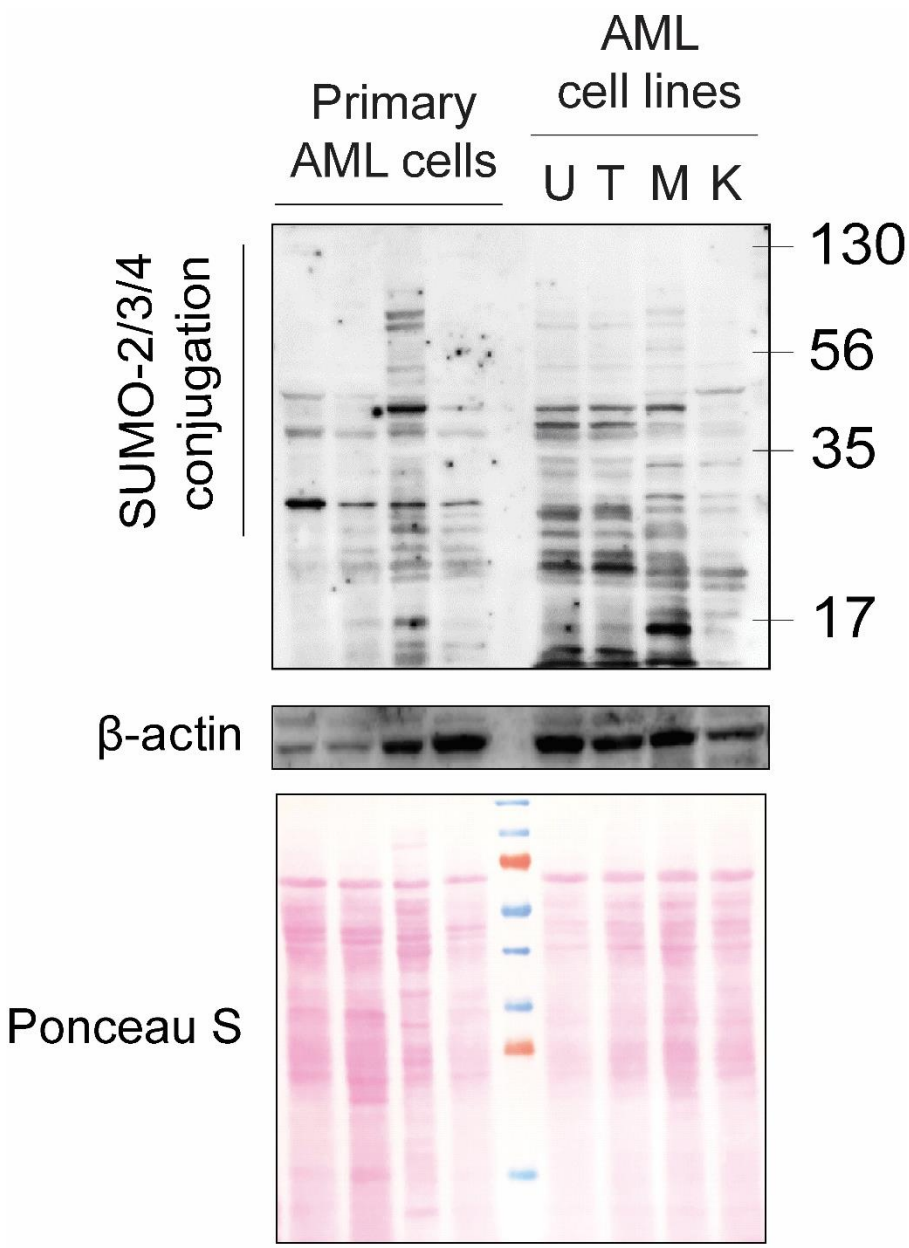


Figure 26. Western blot showing basal SUMOylation levels, and ponceau S staining in primary AML cells and AML cell lines

Western blot analysis was performed with the antibody for SUMO-2/3/4. All primary AML cells were from different patients, who had none p53-mutation status. U, T, M, and K in AML cell lines refer to U937, THP-1, MOLM-14, and KG-1, respectively.

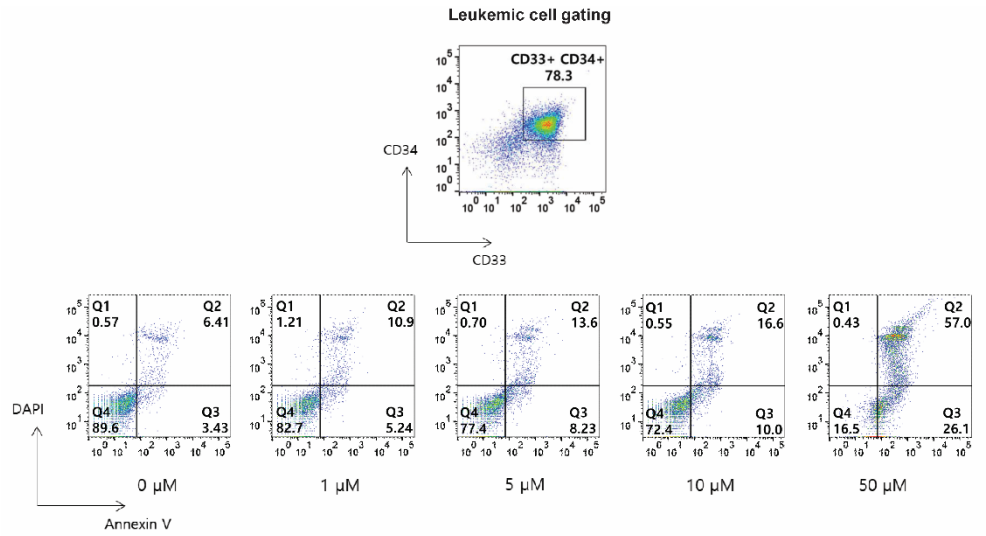


Figure 27. Apoptosis analysis for TAK-981-treated primary AML cells *ex vivo*

Leukemic cells were gated with CD33 and/or CD34 by flow cytometry and viable cells (DAPI negative/Annexin V negative) were compared between groups. Leukemic cell gating (upper part) and representative data of flow cytometry for apoptosis of primary AML cells at different concentrations of TAK-981 with DAPI and Annexin V.

F. TAK-981's anti-leukemic effects in both syngeneic AML mouse and human xenograft models

To assess TAK-981's anti-AML activity in an immune-competent environment, the mouse syngeneic AML model was employed using the C1498 cell line. For the mice injected with C1498/Luc/CD90.1 cells through tail veins, TAK-981 significantly reduced the leukemic burden on day 19 relative to the control group, as judged by the bioluminescence (Fig. 28). Flow cytometric analysis of leukemic cells from bone marrow and blood (from 3 euthanized animals from each group on day 19) showed much less leukemic cells in the TAK-981 group (Figs. 29A and B), consistent with the above imaging data on day 19. Significantly prolonged survival was also observed in the TAK-981 group relative to the controls (Fig. 29C). These data in the syngeneic immune-competent cancer model confirm TAK-981's *in vivo* anti-AML activity.

To confirm the human relevance of the anti-leukemic activity of TAK-981 and to evaluate the influence of anti-tumor immunity on its anti-AML effect, human AML cell MOLM-14/Luc/GFP (0.5×10^6) was injected into non-irradiated immune-deficient NSG mice (no T-cells and defective dendritic cells). Both the bioimaging data (Fig. 30) and the flow cytometric results on the blood and bone marrow cells (Figs. 31A and B) confirmed the lower leukemic burden in the TAK-981 group. Western blot with sorted leukemic cells showed a decreased level of SUMOylated proteins in the TAK-981 group, thereby confirming its *in vivo* deSUMOylation activity (Fig. 32). Significantly prolonged survival was also

observed in the TAK-981 group relative to the control (Fig. 31C). Therefore, the data confirm TAK-981's anti-human AML activity *in vivo*. Importantly, these data show that TAK-981's *in vivo* activity is independent of anti-tumor immunity, as it is lacking in the NSG mouse model.

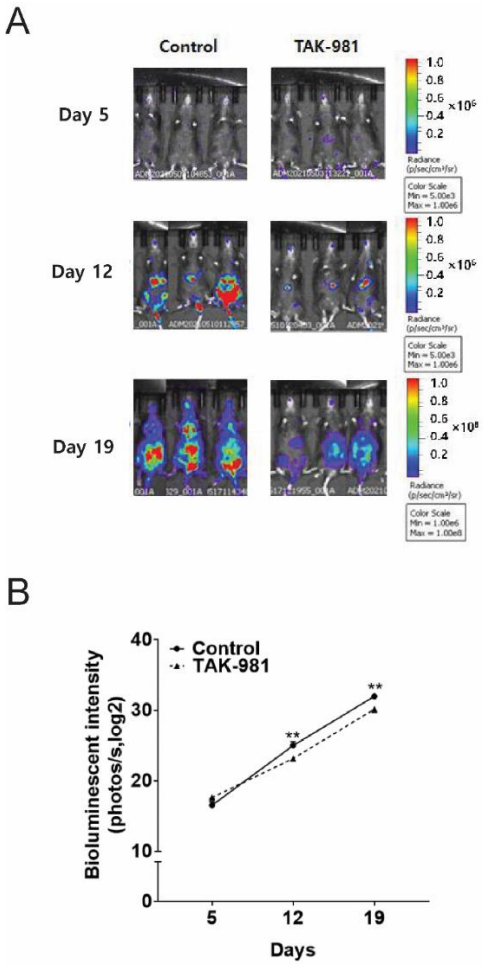


Figure 28. TAK-981's anti-leukemic effects confirmed by bioluminescence imaging in syngeneic AML mouse models (immune-competent mice)

Syngeneic mouse model was established by injecting C1498 cells labeled with Luc/CD90.1 (C1498/Luc/CD90.1) into C57BL/6 mice through tail vein. After confirming leukemia engraftment by bioluminescence imaging, the mice were divided into two groups (10 mice per group) and treatment began on day 5 until day 26: Control (no treatment) or TAK-981 (7.5 mg/kg formulated in 20% 2-hydroxypropyl-β-cyclodextrin, intravenously three times a week). Representative mice from each group were subjected to (A) serial bioluminescence images and (B) intensity quantitation on days 5, 12, and 19 after leukemic cell injection. For (B), the results are expressed as the mean ± SEM and Student's *t*-test was used. ** $p < 0.01$.

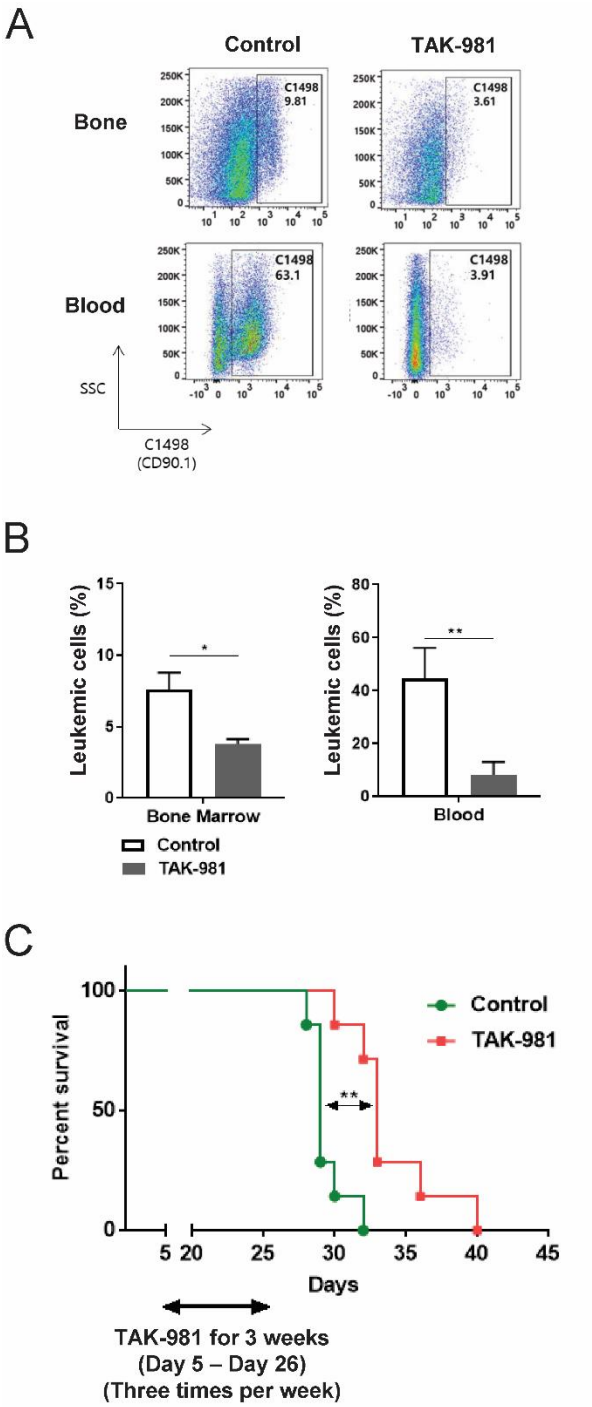


Figure 29. TAK-981's anti-leukemic effects confirmed by flow cytometry and survival analysis in syngeneic AML mouse models (immune-competent mice)

Syngeneic mouse model was established by injecting C1498 cells labeled with Luc/CD90.1 (C1498/Luc/CD90.1) into C57BL/6 mice through tail vein. After confirming leukemia engraftment by bioluminescence imaging, the mice were divided into two groups (10 mice per group) and treatment began on day 5 until day 26: Control (no treatment) or TAK-981 (7.5 mg/kg formulated in 20% 2-hydroxypropyl-β-cyclodextrin, intravenously three times a week). (A and B) Three representative mice per group were euthanized on day 19 to compare leukemic burdens in each group. Cells from the bone marrow and blood were analyzed by flow cytometry. The proportion of CD90.1-positive cells by flow cytometry to identify leukemic cells were compared between the groups. (C) Overall survival rate in each group (7 mice per group) was estimated by the Kaplan–Meier method. For (B), the results are expressed as the mean ± SEM and Student's *t*-test was used. * $p < 0.05$, ** $p < 0.01$.

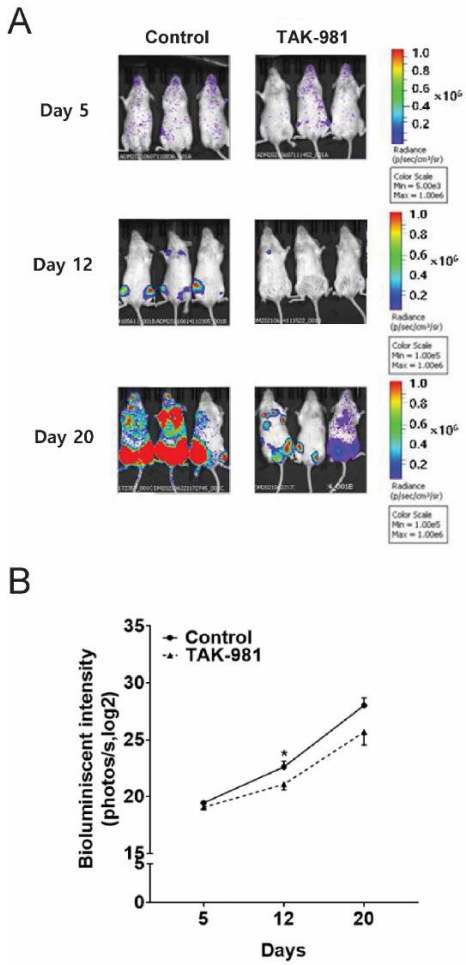


Figure 30. TAK-981's anti-leukemic effects confirmed by bioluminescence in human xenograft AML mouse models (immune-compromised mice)

Human AML mouse model was established by injecting MOLM-14 cells labeled with Luc/GFP (MOLM-14/Luc/GFP) into NOD/SCID/IL-2rγnull (NSG) mice through tail vein. After confirming leukemia engraftment by bioluminescence imaging, the mice were divided into two groups (10 mice per group) and treatment began on day 5 until day 26: Control (no treatment) or TAK-981 (7.5 mg/kg formulated in 20% 2-hydroxypropyl-β-cyclodextrin, intravenously three times a week). Representative mice from each group were subjected to (A) serial bioluminescence images and (B) intensity quantitation on days 5, 12, and 20 after leukemic cell injection. For (B), the results are expressed as the mean ± SEM and Student's *t*-test was used. * $p < 0.05$.

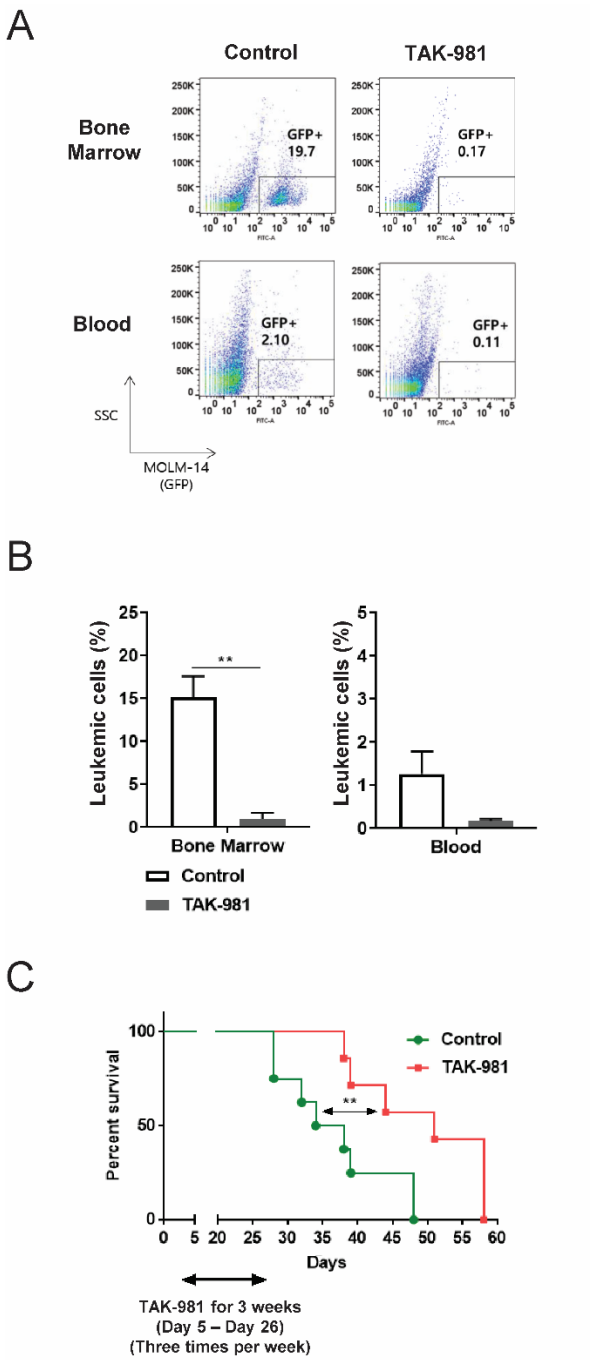


Figure 31. TAK-981's anti-leukemic effects confirmed by flow cytometry and survival analysis in human xenograft AML mouse models (immune-compromised mice)

Human AML mouse model was established by injecting MOLM-14 cells labeled with Luc/GFP (MOLM-14/Luc/GFP) into NOD/SCID/IL-2rγnull (NSG) mice through tail vein. After confirming leukemia engraftment by bioluminescence imaging, the mice were divided into two groups (10 mice per group) and treatment began on day 5 until day 26: Control (no treatment) or TAK-981 (7.5 mg/kg formulated in 20% 2-hydroxypropyl-β-cyclodextrin, intravenously three times a week). (A and B) Three representative mice per group were euthanized on day 20 to compare the leukemic burdens between the groups. Cells from the bone marrow and blood were analyzed by flow cytometry. The proportions of GFP-positive cells by flow cytometry to identify leukemic cells were compared between the groups. (C) The overall survival rate in each group (7 mice per group) was estimated by the Kaplan–Meier method. For (B), the results are expressed as the mean ± SEM and Student's *t*-test was used. ** $p < 0.01$.

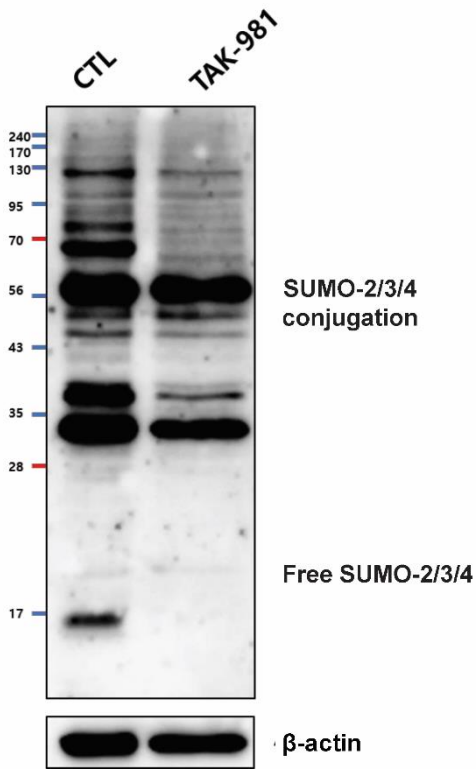


Figure 32. Western blot analysis for sorted leukemic cells from TAK-981-treated human xenograft AML mouse

Western blot was performed with sorted leukemic cells to evaluate SUMOylated proteins in each group. The sample was pooled from individual animals, representing the average levels (see the Materials and method section).

IV. Discussion

SUMOylation has not been much recognized in AML other than for cases of acute promyelocytic leukemia (APL), a minor (~10%) subset of AML with the characteristic chromosomal translocation generating the PML-RARα fusion protein [66]. The established therapy for APL, with all-trans-retinoic acid (ATRA) and As_2O_3, triggers SUMOylation and subsequent proteasomal degradation of PML-RARα, thus inducing APL differentiation [67]. Activities of ATRA-induced differentiation on some non-APL AML cell lines *in vitro* [68] led to clinical trials, but yielded overall disappointing outcomes [69]. In my results, TAK-981 could enhance *in vitro* differentiation of all AML cells tested. It will be interesting to revisit the issue of the differentiation of AML cells upon inhibition of SUMOylation *in vivo*. It is therefore worth noting that the addition of ATRA to decitabine improved clinical outcomes for treatment-difficult elderly patients in a phase II clinical trial [70]. There have also been a few reports on the SUMOylation of individual proteins involved in AML, such as iGF1R, sPRDM, and ERG [40, 41, 71]. In addition, a protein-array based screening on AML cell lines with acquired drug resistance vs. parental cell lines identified possible SUMOylation biomarkers related to drug resistance, which is yet to be validated *in vivo* [72]. However, considering the inhibition of the initial step of SUMOylation by TAK-981, it seems unlikely that one particular protein is responsible for TAK-981's anti-leukemic activity. Rather, TAK-981's activity should be contributed to by several SUMOylation-dependent processes [73]. The differential profiles of SUMOylation

dependency might explain why I observed a large variability in synergy between TAK-981 and cytarabine across the different AML cells. Inhibition of SUMOylation in general with different inhibitors also has been tested. Anacardic acid and/or 2-D08 induced apoptosis of leukemic cells through ROS-mediated deSUMOylation of *NOX* or *DDIT3* regulators [42, 74]. In addition, anacardic acid and 2-D08 sensitized non-APL AML cells to ATRA-based differentiation [60]. However, there is a conflicting report according to which, anacardic acid and ginkgolic acid alleviated ATRA-mediated inhibition of leukemic cell proliferation [75]. This shows that SUMOylation inhibition for AML therapy has not yet been well-established and that the existing literature may need to be considered with some caution. Particularly, most of these studies have employed cell lines *in vitro* or subcutaneous flank xenografts of AML cells and inhibitors with rather moderate micromolar activities without high specificity for SUMOylation [60, 75]. In comparison, I started from the clinical relevance of the SUMOylation pathway and investigated the association of core genes in the SUMOylation pathways and AML characteristics, rather than focusing on a single protein. Furthermore, I evaluated a highly specific SUMOylation inhibitor in multiple AML cell lines, patient-derived primary cells, and orthotopic leukemia models. Overall, after starting the study with bioinformatics using gene expression, I showed that the treatment of TAK-981 decreased SUMOylation in protein level with potent antileukemic effects resulting in prolonged survival in orthotopic models. My results should represent sufficient rationale for testing TAK-981 in AML treatment, as it is already being done in clinical trials for solid tumors.

TAK-981 is a highly specific inhibitor of SUMOylation having little effects on ubiquitination or neddylation [43]. Still, the mechanism of anti-cancer activity of TAK-981 may be multifaceted, due to the broad-reaching roles of SUMOylation in cancer [26, 73]. Interestingly, recent data suggested that TAK-981's activity against solid tumors is dependent on anti-tumor immunity, especially through IFN1 signaling regulated by SUMOylation [37, 39]. For an immune-competent syngeneic flank model, TAK-981's activity was abolished when the IFN1 receptor was knocked out [39]. In addition, in two different syngeneic flank models, a survival benefit was observed for the TAK-981-immune checkpoint inhibitor (ICI) combination groups but not for the TAK-981 monotherapy groups, suggesting a cancer-cell-extrinsic mechanism of TAK-981 [39]. In my orthotopic models for AML, a hematologic cancer, I observed significant inhibition of leukemia growth and survival benefits in both immune-competent syngeneic mouse transplant and human xenograft models with immune-deficient mice. It should be noted that the NSG immune-deficient mice used here lacked T lymphocytes and had defective dendritic cells that had proved critical to anti-tumor immunity by TAK-981 in the above solid-tumor settings. Additionally, I observed potent *in vitro* inhibitory effects of TAK-981 as well as induction of differentiation markers for various AML cell lines. Direct apoptotic effects of TAK-981 were also observed *ex vivo* for primary AML cells from patients. These results strongly suggest that TAK-981 exhibits cancer-cell-inherent anti-AML activity. The apparent discrepancy with the above study may be due to the fundamental differences between solid vs. AML cancer or the experimental settings (i.e., flank transplant vs. orthotopic (blood) xenograft).

Still, I do not exclude the possibility of anti-AML immunity by TAK-981 or synergy with ICIs in immune-competent human AML settings that I did not study. For acute leukemia, immunotherapy has been advanced and regularly used in clinics for acute lymphoblastic leukemia (ALL), and it has been also rapidly developing for AML [39, 76], as evidenced by the approval of Gemtuzumab-ozogamicin in 2017. At this point, ICI monotherapy for AML has been proved not to be very satisfactory [77, 78], and its combinations with hypomethylating agents, that have their own immune-modulatory effects [79, 80], have yielded mixed results [81, 82]. As for the positive ones, those from a phase 1b study on the combination of azacitidine and magrolimab on patients ineligible for intensive chemotherapy were quite encouraging [81]. Notably, this combination was effective even for therapy-refractory *TP53*-mutated AML patients, though the overall number of patients was small. Larger human clinical trials with TAK-981-ICI combinations are warranted to evaluate real effects in human AML [83].

I showed that TAK-981 exhibited stronger or similar potency than cytarabine in all of the AML cell lines tested as well as in patient-derived primary AML cells. Moreover, TAK-981 exhibited inhibition for cytarabine-resistant AML cell lines *in vitro* (KG-1, THP-1 cells; my results and other studies by Bossis [42], and Ma [84]) as well as in a therapy-resistant *in vivo* model (MOLM-14 orthotopic xenograft). TAK-981 also has decreased the expression of *CD39*, whose expression is mediated by SUMOylation [85]. *CD39* has been known to be overexpressed in both cytarabine-resistant AML cells and residual AML cells in patients after chemotherapy [63]. Enhancing *CD39* expression provoked resistance against cytarabine while inhibiting it improved the response to cytarabine in AML cells

[63]. These results might explain TAK-981's strong activity against cells with high IC_{50} values for cytarabine (> 100 nM) such as KG-1, THP-1, and MOLM-14. Considering the different of modes of action between TAK-981 and cytarabine and the differences in cell lines and primary cells, it will be interesting to see if their potency difference is maintained in real patient cases. Still, the different mode of action might explain the strong synergy of TAK-981 with current drugs in several settings shown in my study.

It is worthwhile to note that the IC_{50} values of both TAK-981 and cytarabine for the primary cells were much higher than those for the AML cell lines. With the lower SUMOylation status of primary AML cells than that in the cell lines (Fig. 26) being one explanation, an important consideration is that primary AML cells grow much slower than the established AML cell lines. It is possible that the high IC_{50} value of TAK-981 in primary AML cells may be due to the lower frequency of cell division. This is clearly the case with cytarabine that it almost completely lost its activity for the primary AML cells, even though it is a standard-of-care drug. Therefore, the absolute value of the IC_{50} may not be directly translated into the actually high *in vivo* toxicity. I believe the much slower proliferation of the primary AML cells should be considered seriously, and, therefore, a correlation analysis between SUMOylation extent and cytotoxicity across primary AML cells and cell lines might not be conclusive.

Further research should also be focused on (i) comparing the clustering result used in this study with other methods to check the consistency, although GSCluster was chosen since it could also reflect protein-protein interactions among

genes, and (ii) checking if TAK-981 also has immune-dependent anti-cancer activity, as in solid tumors.

Part II

SCD and MTHFD2 inhibitors for high-risk acute myeloid leukemia patients, as suggested by ELN2017-pathway association

I. Introduction

The European LeukemiaNet (ELN) 2017 criteria [3] is widely accepted as the risk classification of acute myeloid leukemia (AML) patients. However, their application to studying risk-related biological pathways is limited, failing to enhance treatment options for high-risk patients.

Folate metabolism, also known as one-carbon metabolism, aids in the *de novo* synthesis of nucleic acids or methylation reaction that supports the methionine cycle. Most studies connecting folate metabolism and AML have focused on drug treatment in AML [86-88]. In contrast, bioinformatic studies on folate metabolism and AML utilizing transcriptomics or proteomics are not common. Particularly, studies about the relationship between folate metabolism and prognosis in AML patients are very few [89]. In clinical practice, methotrexate, a well-known drug targeting folate metabolism, is not generally used for AML patients, except for very specific AML patient cases [90], notably because methotrexate has failed in the early clinical trials for AML [91]. Therefore, it may be worth investigating whether folate metabolism correlates with the prognosis of AML patients and whether drug candidates targeting this pathway might augment

the current therapy.

Unsaturated fatty acids (UFAs) are fatty acids that have at least one carbon-carbon double bond in their lipid chain, and they have not been studied much in AML [92-97]. However, they are essential lipid components with numerous biological functions, as exemplified by omega-3 polyunsaturated fatty acids (PUFAs) and omega-6 PUFAs. With a few studies for PUFA in AML [98, 99], one study reported PUFA plasma levels were positively correlated with bone marrow blasts at diagnosis and with ELN2017 risk categories [98], and another study reported that PUFA was elevated in AML serum compared to the controls [99]. However, the relationship between genes involving UFAs synthesis and the prognosis of AML patients has not been fully explored. From a pharmacological point of view, few drugs are available that target the synthesis of UFAs in AML.

Even though there were recent introductions of several targeted agents for AML [4], considering the seriously poor prognosis of the high-risk groups, the discovery of new drug candidates has been highly desired. It is also desired to find drugs synergizing with cytarabine, the current standard-of-care drug for AML. Here, I investigated which pathways are associated with higher risks according to the ELN2017 categories. My results suggest targetable pathways and genes in AML supported by multi-omics databases and experiments and suggest drugs with therapeutic potential.

II. Materials and methods

1. Databases used in the study

For OHSU BeatAML 1.0 database [47], gene expression data were obtained as follows; the raw counts were downloaded from the National Cancer Institute Genomic Data Commons (GDC, https://portal.gdc.cancer.gov/, downloaded on Nov. 2019), and summarized in a table, DESeq2-normalized and rlog-transformed using 'DESeq2' R package. The clinical information of patients was obtained from the original paper [47] and Vizome site (http://vizome.org/). For the information on the disease stage of the specimen, I referred to the "SpecimenGroups" parameter.

For TCGA-LAML database [45], pancancer gene expression data and clinical information of patients were downloaded from https://gdc.cancer.gov/about-data/publications/pancanatlas (downloaded on Feb. 2020). Gene expression data with only tumor samples was retrieved, and Entrez gene IDs were matched with gene symbols using the information downloaded from the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/, downloaded on Sep. 2020). TCGA-LAML had no normal samples. The values in the expression table were added by 2, followed by log2-transformation. Then, only the gene expression with sample IDs starting with 'TCGA-AB,' which means they are the samples from the TCGA-LAML cohort, were retrieved. TP53 mutation status was retrieved from cBioPortal (https://www.cbioportal.org/) and karyotype information from GDC.

For AML Proteomics database [100], tandem-mass-tag (TMT) protein abundance data and patients' clinical information were downloaded from the site https://proteomics.leylab.org/ (downloaded on Sep. 2022).

For the cell lines database, dependency scores in Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) screening and gene expression data were downloaded from Depmap (21Q2, https://depmap.org/). The designation of cell lines which are AML was based on the metadata from the same database.

2. Designation of canonical and revised ELN2017 risk criteria in the individual patients

For the designation of samples to canonical and revised ELN2017 risk groups, different approaches were used depending on the databases. For OHSU BeatAML 1.0 database, two files contained clinical information; one is from the Vizome site (http://vizome.org/), and the other is from the original paper [47]. Since the two files were complementary, I used the two files as needed. Samples with the 'ELN2017' parameter as 'Healthy, Individual BM MNC' were designated as 'Normal' samples. Samples with 'Favorable,' 'Intermediate,' and 'Adverse' were designated as is. I only used samples from bone marrow aspirate, except for Fig. 34. Also, by manually reviewing the 'specificDxAtAcquisition' parameter, samples from non-AML patients were excluded. Among the 'Favorable' samples, those who have 'inv(16)' in the 'specificDxAtAcquisition' parameter or 'y' in the 'CEBPA_Biallelic' parameter were designated as 'Very Favorable.' Among the 'Adverse' samples, firstly those with *TP53* mutation were found by excluding those

who have ‘negative’ or NA as ‘TP53’ parameter; among them, by manually reviewing the parameter ‘Karyotype,’ samples only with complex karyotype were left to be designated as ‘Very Adverse.’ In total, the number of samples per each group was as follows; for survival analysis in Fig. 34, ‘Very Favorable’ (n = 31), ‘Favorable’ (n = 113), ‘Intermediate’ (n = 170), ‘Adverse’ (n = 184), and ‘Very Adverse’ (n = 23); for testing trends in OHSU database, ‘Normal’ (n = 20), ‘Favorable’ (n = 56), ‘Intermediate’ (n = 68), ‘Adverse_nV’ (n = 68), and ‘Very Adverse’ (n = 14).

For TCGA-LAML database, the designation of samples to canonical ELN2017 risk groups was kindly provided by the authors of the previous publication [50]. For the designation of the ‘Very Adverse’ category, I utilized *TP53* mutation status information from cBioPortal, and karyotype information from GDC. The samples having ‘Complex’ in the ‘cytogenetic_abnormality_type’ parameter were regarded as having complex karyotypes, and the samples which both have *TP53* mutation and complex karyotype were designated as ‘Very Adverse.’ In total, the number of samples per each group was as follows; ‘Favorable’ (n = 59), ‘Intermediate’ (n = 38), ‘Adverse_nV’ (n = 45), and ‘Very Adverse’ (n = 11).

For AML Proteomics database, canonical ELN2017 risk groups were designated according to the ‘RISK (ELN2017)’ parameter. The samples which do not have NA as the ‘TP53’ parameter were regarded as having *TP53* mutation, and among them I determined complex karyotype by reviewing the ‘Cytogenetics’ parameter; those having both *TP53* mutation and complex karyotype were designated as ‘Very Adverse.’ In total, the number of samples per each group was

as follows; except for the *SCD* gene, ‘Favorable’ (n = 14), ‘Intermediate’ (n = 10), ‘Adverse_nV’ (n = 16), and ‘Very Adverse’ (n = 4); for *SCD* gene, ‘Favorable’ (n = 3), ‘Intermediate (n = 3)’, ‘Adverse_nV’ (n = 10), and ‘Very Adverse’ (n = 2).

In all databases, ‘Adverse_nV’ samples were designated by excluding ‘Very Adverse’ samples from ‘Adverse’ samples.

3. Transforming gene or protein expression data to pathway scores data

‘GSVA’ R package [58] was used to convert gene or protein expression data to pathway scores data. For input for pathways, the combination of curated canonical pathways (which includes pathways from BioCarta, KEGG, PID, Reactome, and WikiPathways) and hallmark gene sets from MSigDB (version 7.2) was used. The R package ‘GSA’ was used when reading the gene sets to R. For parameters when running GSVA, ‘min.sz’ was set to 5, ‘max.sz’ to 700, and ‘method’ to ‘gsva.’

4. Survival analysis and pathway clustering

For survival analysis, ‘survival’ and ‘survminer’ R packages were used. When conducting pairwise comparisons among more than two survival curves, pairwise_survdiff function was used, with p.adjust.method parameter set to ‘fdr.’ The stratification of two groups by gene expression or pathway scores was based on the best risk separation approach [52], using surv_cutpoint function with minprop parameter set to 0.1. Log-rank test and Cox regression were used for

comparison of the curves and obtaining hazard ratios, using GraphPad Prism 9.1.1 (GraphPad Software, San Diego, CA, USA) or R (version 4.1.1).

For disease-specific survival (DSS) analysis in OHSU database, only patients with the 'causeOfDeath' parameter as 'Dead-Disease' and 'Alive' were included. For overall survival (OS) analysis in the OHSU database, patients who did not have the parameter as NA were included.

For the clustering of pathways, 'GSCluster' R package [51] was used. For the q value input for GSCluster, FDR values from Jonckheere-Terpstra test screening result were used. When plotting the clusters, the Maximum gene-set distance parameter was set to 0.85.

5. Cell lines and reagents

MOLM-14 (DSMZ, Braunschweig, Germany), U937, THP-1, KG-1, HCC1954-BL (ATCC, Manassas, VA, USA), and HL-60 (Korean Cell Line Bank, Seoul, Korea) were used in this study. Cryopreserved peripheral blood mononuclear cells (PBMCs) were purchased from Zen-Bio (Research Triangle, NC, USA). Cells were cultured or incubated in RPMI 1640 media supplemented with 10% fetal bovine serum (FBS) (for KG-1 only, 20%), 100 U/mL penicillin, and 100 μg/mL streptomycin, at 37 °C in a 5% CO_2 incubator. A939572 and DS18561882 were purchased from MedChemExpress (Monmouth Junction, NJ, USA). Cytarabine was purchased from Sigma-Aldrich (St.Louis, MO, USA).

6. UFA measurement by NMR

U937, MOLM-14, THP-1, KG-1, and HL-60 cell lines were seeded 1×10^7 cells in 10 mL of media and cultured for 24 hours; to minimize potential factors affecting UFA amount in the cells, the media was unified to RPMI 1640 media supplemented with 20% fetal bovine serum (FBS), 100 U/mL penicillin, and 100 μg/mL streptomycin. For Fig. 48C, U937 cells were seeded 2×10^6 cells in 6 mL of media and cultured for 48 hours. Here, RPMI 1640 media supplemented with 10% fetal bovine serum (FBS), 100 U/mL penicillin, and 100 μg/mL streptomycin was used to make the media consistent with Fig. 48B. The cells were harvested and underwent standard two-phase extraction, and the lipid phase was dried by speedvac. The dried samples were dissolved in chloroform-d_6 (Cat. 151823, Sigma-Aldrich) and subjected to ^{1}H NMR. 800 MHz Bruker Avance III HD spectrometer equipped with a 5 mm CPTCI CryoProbe (Bruker BioSpin, Germany) was used. The spectra were processed with MestReNova software (version 12.0.1-20560). The UFA and PUFA amounts were measured as the area under 5.29-5.40 ppm and 2.70-2.90 ppm of the spectrum, respectively. Then, for each cell line, the numbers were normalized by the cell numbers counted at the time the cells were harvested, respectively.

7. Cell viability test and synergy test

For CCK-8 assay, cells were seeded at 1×10^4 cells/well on 96-well plates and treated with drugs (A939572, DS18561882, Cytarabine) at various concentrations for 48 hours, and measured with the D-plus CCK Cell Viability

Assay Kit (Dongin Biotech, Seoul, South Korea). For Trypan blue assay, 1-2 × 10^4 cells/well (for cell lines) or 1 × 10^5 cells/well (for PBMCs) were seeded on 96-well plates and treated with drugs (A939572, DS18561882) at various concentrations for 48 hours. The cells were counted with Countess II FL Automated Cell Counter (Thermo Fisher Scientific, Waltham, USA) or hemocytometer (for PBMCs). The IC_{50} value was obtained using the GraphPad Prism 9.1.1 software. The test for synergy and calculation of dose reduction indices were done using CompuSyn software [59].

8. Western blotting

The samples were homogenized in T-PER™ Tissue Protein Extraction Reagent (Thermo Scientific, USA) buffer with protease and phosphatase inhibitors (1 mM PMSF (phenylmethylsulfonyl fluoride) 2 μg/mL aprotinin, 1 μg/mL pepstatin A). 20 or 30 μg of protein extracts were loaded, and they were separated by SDS electrophoresis with 10% gel, then transferred to the NC (nitrocellulose) membranes. Membranes were blocked with 5% skim milk in TBST (Tris-buffered saline with 0.1% tween) and incubated at 4℃ overnight with the following primary antibodies: β-actin (sc-47778, Santa Cruz), SCD (A16429, Abclonal), MTHFD2 (A22653, Abclonal). Anti-rabbit IgG-HRP (31460, Invitrogen) and anti-mouse IgG-HRP (31430, Invitrogen) were used as secondary antibodies. The protein bands were visualized by using a Westsave star kit (Abfrontier, Seoul, South Korea) and imaged on a Fusion Solo Chemi-DOC (Vilber Lourmat, France). The

quantification of the bands was done with EvolutionCapt software (Vilber Lourmat, France).

9. Statistical analysis

Jonckheere-Terpstra test was used when testing for increasing trend across groups, using the 'DescTools' R package. Since one of the assumptions for the Jonckheere-Terpstra test is that the observations should be independent, 'Adverse_nV' instead of 'Adverse' was used when performing the test. Post hoc analyses were done with the two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli in GraphPad Prism. For Fig. 48C, Student's *t*-test was used. For survival analysis, details are in the 'Survival analysis and pathway clustering' section. For correlation analysis, Pearson's correlation was used. All statistical analyses were performed with GraphPad Prism 9.1.1 or R (version 4.1.1).

III. Results

A. Applicability of canonical and revised ELN2017 to the OHSU database

The overall scheme of my study is presented in Fig. 33. I first investigated if ELN2017 and its revision [17] are applicable to OHSU patient data for prognostic risk categorization. Kaplan-Meier analysis showed significantly different survival prognoses among risk groups (Fig. 34). In particular, the 'Very Adverse' group, specified by *TP53* mutation and complex karyotype within the 'Adverse' group, showed a much poorer prognosis than the other groups, consistent with previous report [17]. On the other hand, the 'Very Favorable' group, specified by biallelic CEBPA mutations or inv(16), could not be significantly discriminated from the 'Favorable' group (FDR-adjusted p = 0.152, Table 13). This analysis shows that four distinguished survival risk groups exist in OHSU data, the characterization of which may help improve current therapeutic options.

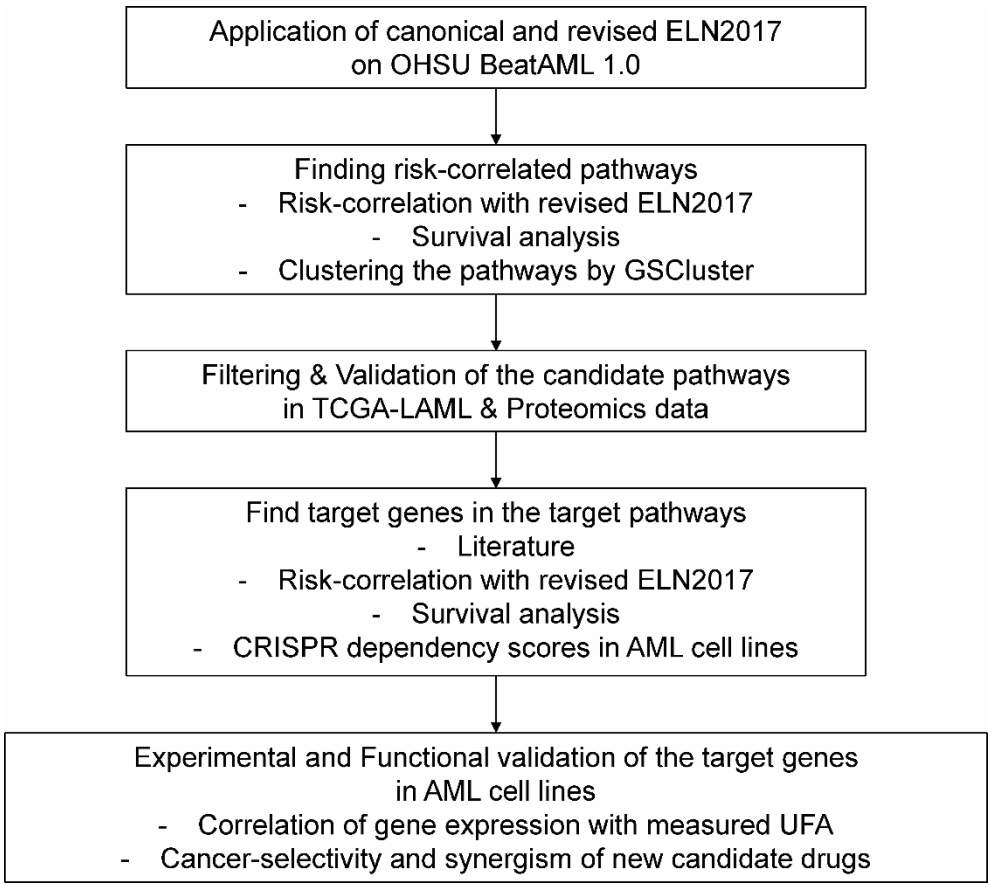


Figure 33. Overview of the study

Overall scheme of the study.

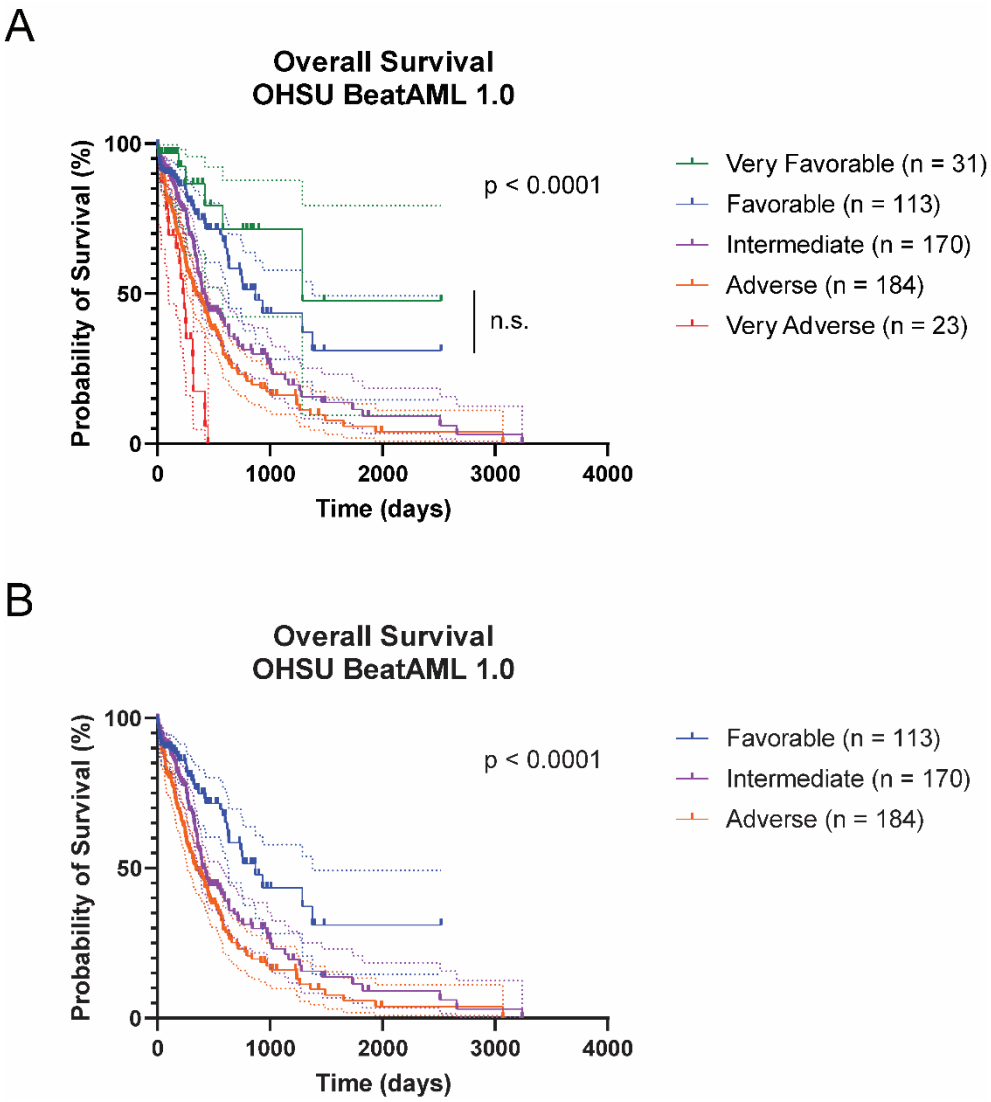


Figure 34. Application of revised and canonical ELN2017 to the OHSU database

Kaplan-Meier curves with 95% confidence intervals (dotted lines) for overall survival of AML patients in the OHSU BeatAML 1.0 database, according to the (A) revised and (B) canonical ELN2017 criteria. *P*-values are from the log-rank test. n.s. refers to not significant. For (A), complete comparisons among the groups are in Table 13.

Table 13. Results of pairwise comparisons among all groups in Fig. 34A

	Adverse	Favorable	Intermediate	Very Adverse
Favorable	< 0.001	-	-	-
Intermediate	0.030	0.002	-	-
Very Adverse	0.002	< 0.001	< 0.001	-
Very Favorable	< 0.001	0.152	0.003	< 0.001

B. Screening biological pathways that are risk-correlated with revised ELN2017 criteria

I then hypothesized that there should be biological pathways whose activities are upregulated as survival risk increases. Correlational screening between overall survival risks and pathway scores (Jonckheere-Terpstra test FDR < 0.05) resulted in 690 pathways. Further filtering the results based on disease-specific survival (DSS) with more stringent criteria (log-rank test FDR < 0.05, hazard ratio (HR) > 3, and FDR of HR < 0.05) resulted in 34 pathways that are correlated with patient survival. Analyzing these pathways based on gene constituents and protein-protein interactions using the GSCluster approach [51] generated three distinct clusters and three unclustered pathways related to survival (Fig. 35, and Appendix C and D). The largest cluster (26 pathways) was 'cell-cycle related'; considering 71.2% (146/205) of the samples were from the initial diagnosis stage, a connection between the proliferative state at diagnosis and patient prognosis is suggested. As relapsed AML cells after chemotherapy exhibit higher dormancy [101] and leukemia stem cells often remain in a quiescent state [102], it will be an interesting future topic to specifically compare the relationship between the prognosis and the cell cycle progression at initial vs. late-stage AML. Since the 'cell-cycle related' cluster already encompasses a standard regimen drug, cytarabine, I focused on other pathways. Of the remaining eight pathways, three pathways were related to the synthesis of unsaturated fatty acids, two pathways with cholesterol biosynthesis, and the unclustered three pathways with the

metabolism of folate, signaling by MST1, and NO metabolism in cystic fibrosis (Fig. 35). Actual correlations between the pathway scores and the survival risks are given in Figs. 36-39, and Table 14.

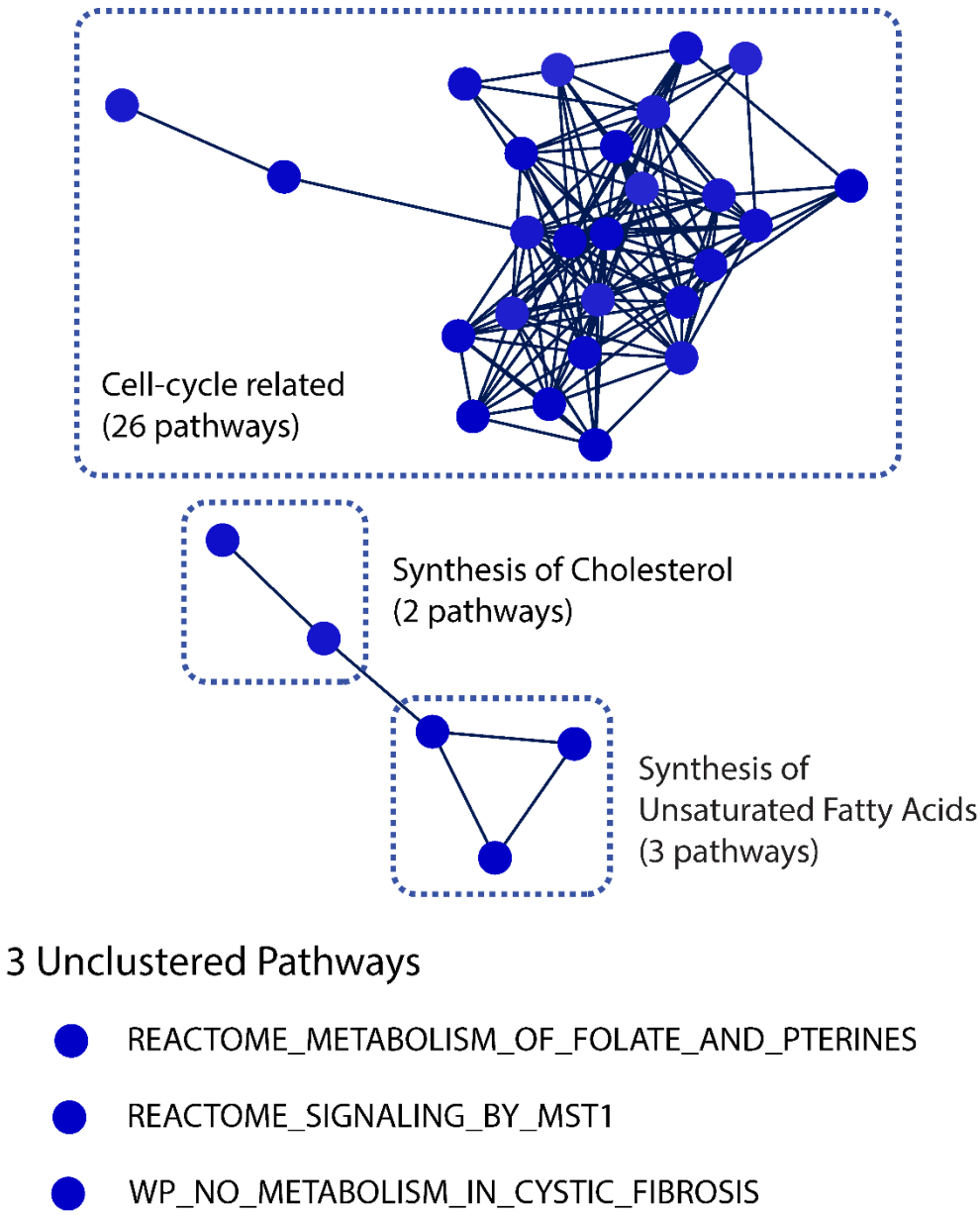


Figure 35. Pathways that are risk-correlated with revised ELN2017 criteria in OHSU database

Clustering result of 34 pathways filtered by risk-correlation with revised ELN2017 and survival analysis.

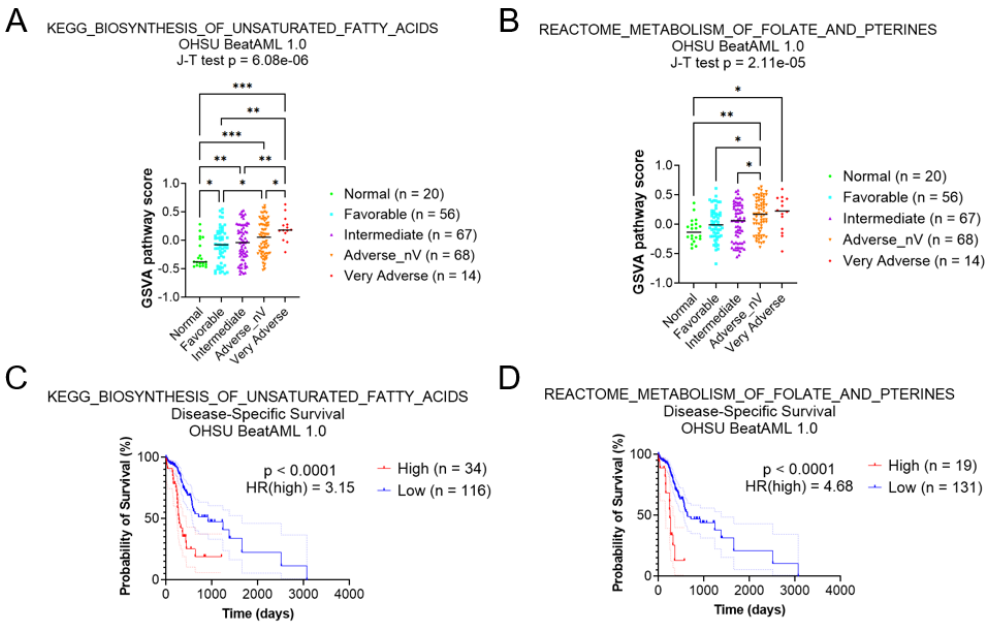


Figure 36. Identification of risk-correlated biological pathways

(A, B) The distributions of GSVA pathway scores of (A) KEGG_BIOSYNTHESIS_OF_UNSATURATED FATTY ACIDS pathway or (B) REACTOME_METABOLISM_OF_FOLATE_AND_PTERINES pathway, in each risk category of revised ELN2017 in OHSU BeatAML 1.0 database. (C, D) Kaplan-Meier curves with 95% confidence intervals (dotted lines) for disease-specific survival of AML patients in OHSU BeatAML 1.0 database, for (C) KEGG_BIOSYNTHESIS_OF_UNSATURATED FATTY ACIDS pathway or (D) REACTOME_METABOLISM_OF_FOLATE_AND_PTERINES pathway. For (A), and (B), 'Adverse_nV' refers to the patients in the 'Adverse' category but not in the 'Very Adverse' category. The black lines indicate medians for each group. *P*-values are from the Jonckheere-Terpstra test. Post hoc analyses were performed with a two-stage linear step-up procedure. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. For (C), and (D), *p*-values are from the log-rank test. The stratification of two groups in each graph was based on the best risk separation approach. HR(high) refers to the hazard ratio of the group with high pathway scores.

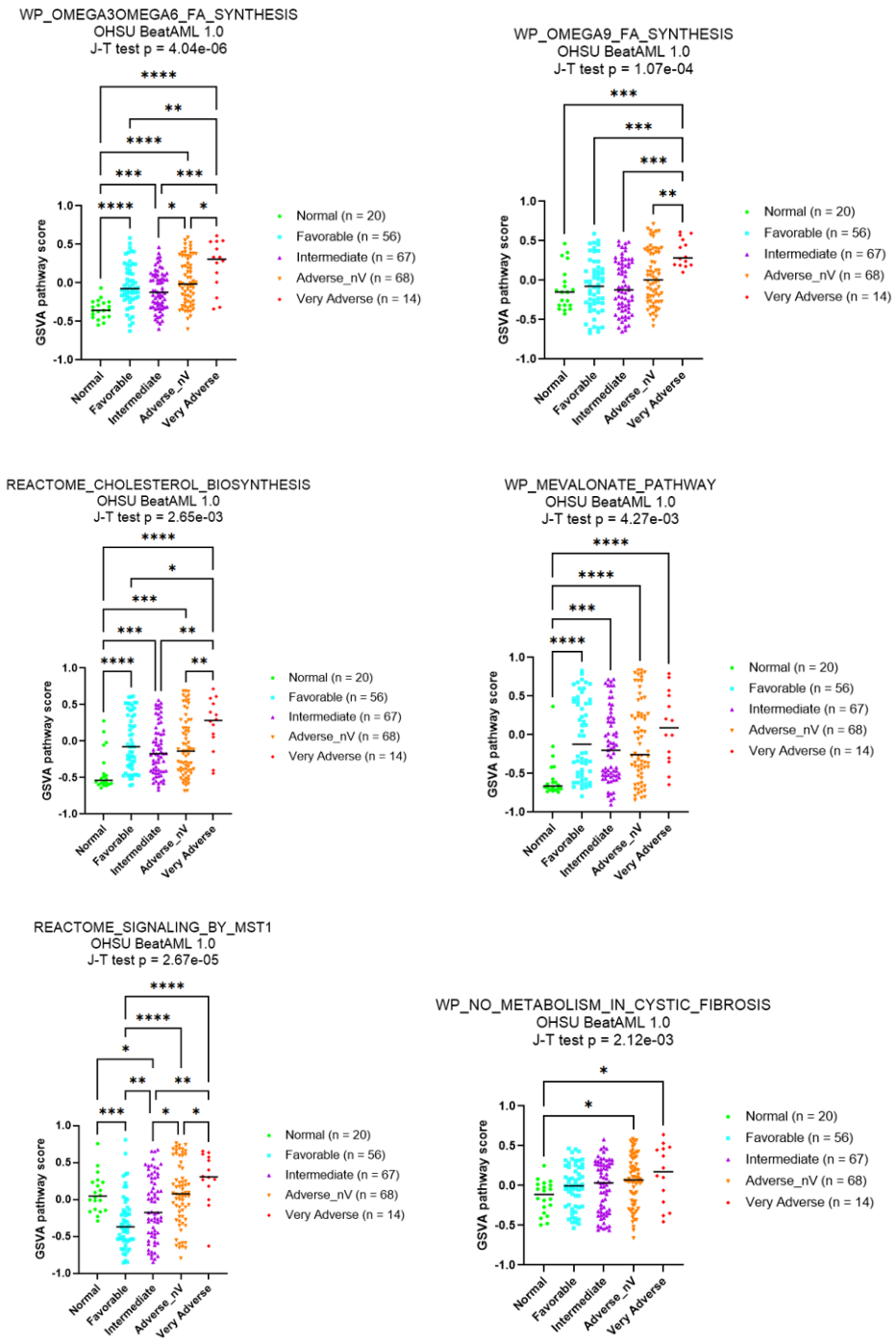


Figure 37. The distributions of GSVA pathway scores of high-risk pathways in the OHSU database

The GSVA pathway scores for each risk category of the revised ELN2017 are plotted. The pathways are from Fig. 35, except for the pathways in the 'cell-cycle related' cluster, KEGG_BIOSYNTHESIS_OF_UNSATURATED FATTY ACIDS pathway, and REACTOME_METABOLISM_OF_FOLATE_AND_PTERINES pathway. 'Adverse_nV' refers to patients in the 'Adverse' category but not the 'Very Adverse' category. The black lines indicate medians for each group. *P*-values are from Jonckheere-Terpstra test. Post hoc analyses were performed with a two-stage linear step-up procedure. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

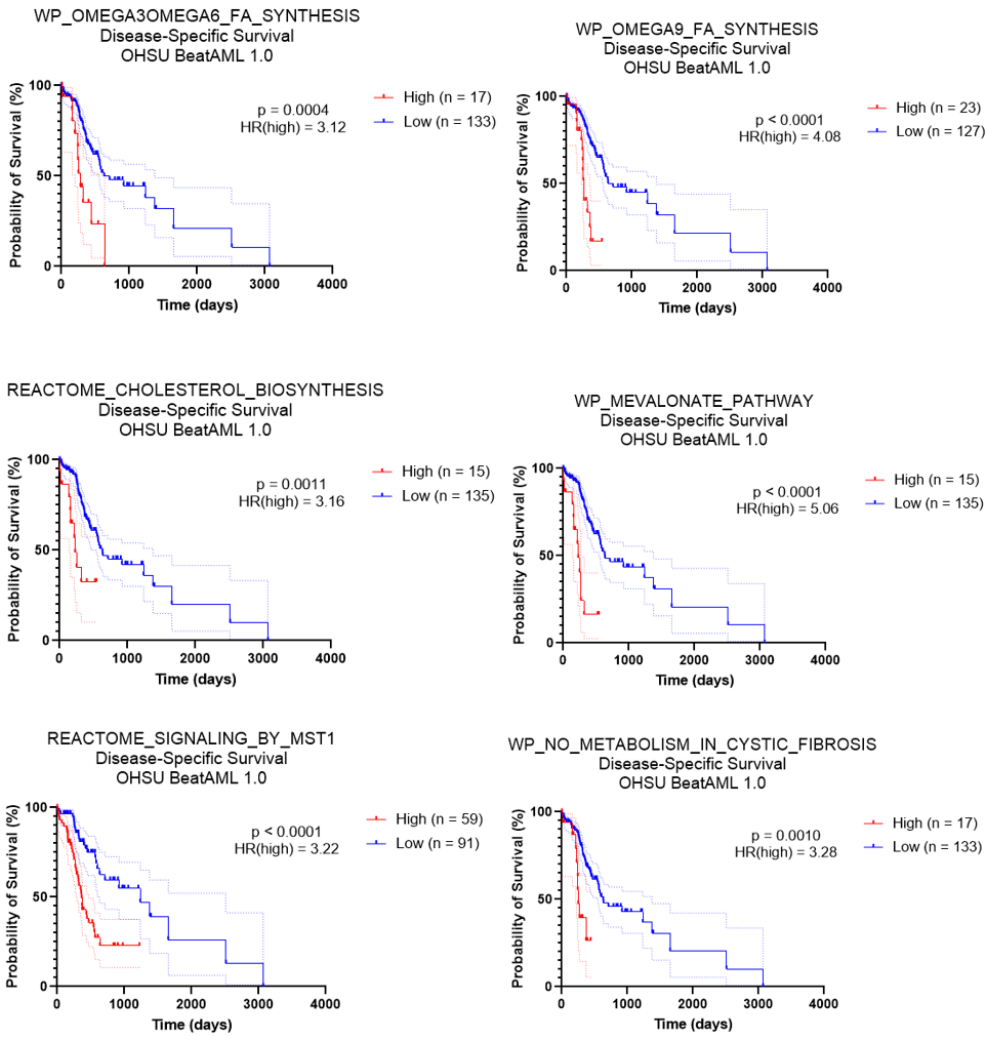


Figure 38. Disease-specific survival analysis of high-risk pathways in the OHSU database

Kaplan-Meier curves with 95% confidence intervals (dotted lines) for disease-specific survival of AML patients in OHSU BeatAML 1.0 database for pathways in Fig. 35, except for pathways in the 'cell-cycle related' cluster, KEGG_BIOSYNTHESIS_OF_UNSATURATED FATTY ACIDS pathway, and REACTOME_METABOLISM_OF_FOLATE_AND_PTERINES pathway. *P*-values are from the log-rank test. The stratification of two groups in each graph was based on the best risk separation approach. HR(high) refers to the hazard ratio of the group with high pathway scores.

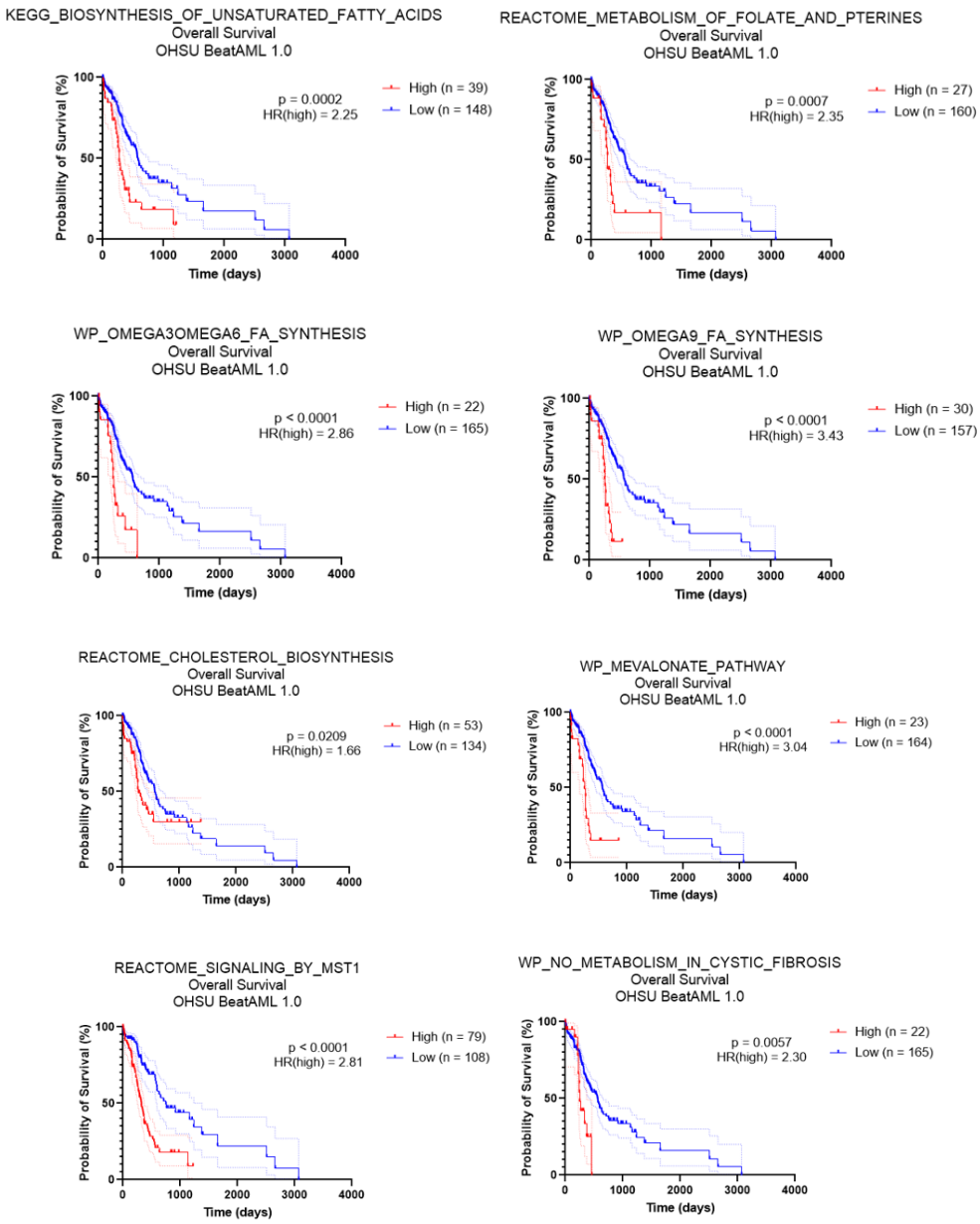


Figure 39. Overall survival analysis of high-risk pathways in OHSU database

Kaplan-Meier curves with 95% confidence intervals (dotted lines) for overall survival of AML patients in the OHSU BeatAML 1.0 database for pathways in Fig. 35, except for pathways in the 'cell-cycle related' cluster. *P*-values are from the log-rank test. The stratification of two groups in each graph was based on the best risk separation approach. HR(high) refers to the hazard ratio of the group with high pathway scores.

Table 14. Risk-correlation analysis of revised ELN2017 for pathways using multi-omics databases

Pathway	OHSU					TCGA-LAML	Proteomics
	Jonckheere-Terpstra test	Survival (DSS)	HR	Survival (OS)	HR	Jonckheere-Terpstra test	Jonckheere-Terpstra test
KEGG_BIOSYNTHESIS_OF_UNSATURATED_FATTY_ACIDS (‘UFA_Synthesis’)	< 0.001	< 0.001	3.15	< 0.001	2.25	0.007	0.013
WP_OMEGA3OMEGA6_FA_SYNTHESIS	< 0.001	< 0.001	3.12	< 0.001	2.86	0.196	0.036
WP_OMEGA9_FA_SYNTHESIS	< 0.001	< 0.001	4.08	< 0.001	3.43	0.068	0.067
REACTOME_CHOLESTEROL_BIOSYNTHESIS	0.003	0.001	3.16	0.021	1.66	0.466	0.195
WP_MEVALONATE_PATHWAY	0.004	< 0.001	5.06	< 0.001	3.04	0.622	0.471
REACTOME_METABOLISM_OF_FOLATE_AND_PTERINES (‘Folate_Metabolism’)	< 0.001	< 0.001	4.68	< 0.001	2.35	0.006	0.001
REACTOME_SIGNALING_BY_MST1	< 0.001	< 0.001	3.22	< 0.001	2.81	< 0.001	NA
WP_NO_METABOLISM_IN_CYSTIC_FIBROSIS	0.002	< 0.001	3.28	0.006	2.30	< 0.001	0.220

DSS, disease-specific survival; OS, overall survival; HR, the hazard ratio of the high pathway score group; NA, not available.

C. Validation of two targetable pathways with independent TCGA and proteomics database

For the eight pathways outside the 'cell-cycle related' cluster, I tried to validate the correlations between pathway scores and prognosis in another extensive AML database, TCGA-LAML. I divided the TCGA-LAML patients into four categories of revised ELN2017 criterion as above in reference to the karyotype information and *TP53* mutation status, and especially with the information provided by Straube et al [50]. Similar correlation analysis showed that all pathways, except those related to cholesterol and those with specific omega-3, 6, or 9 fatty acid synthesis, were still significant in TCGA-LAML (Table 14).

Further validation was performed at the protein level using recent proteomics AML database [100]. I leveraged Gene Set Variation Analysis (GSVA), originally used for transcriptomic data, to generate pathway scores from proteomic data. The trend analysis between risk groups and pathways showed that 'KEGG_BIOSYNTHESIS_OF_UNSATURATED_FATTY_ACIDS' (hereby termed 'UFA_Synthesis') and 'REACTOME_METABOLISM_OF_FOLATE_AND_PTERINES' (hereby termed 'Folate_Metabolism') were significant (Table 14). Combining all the above analysis suggested that these two pathways are significantly related to the revised ELN2017 risk groups.

D. Finding candidate genes in the risk-related pathways

Having found the risk-correlated pathways, I tried to find targetable genes in those pathways. For the Folate_Metabolism pathway, recent literatures directly suggest the *MTHFD2* gene as a target for AML [87, 88]; therefore, I tried to investigate this gene. Upon similar analysis used for pathways, *MTHFD2* gene expression exhibited a significant increasing trend according to the ELN2017 risk groups in OHSU (Fig. 40A). In survival analysis, patients with higher expression of *MTHFD2* gene exhibited significantly shorter overall survival than those with lower expression (Fig. 40B). This also shows that my bioinformatic screening approach is valid. Clinical trials of methotrexate, an antifolate drug, were unsuccessful in AML due to reduced polyglutamylation activity, which is essential for its effectiveness [103]. Therefore, alternative drugs targeting this pathway seem necessary.

Up to this date, little has been known about the UFA_Synthesis pathway in AML. Hence, I analyzed the expression of all 22 genes in the pathway according to the ELN2017 risk groups in OHSU, TCGA-LAML, and proteomics database (Table 15). Then, I looked for genes whose high expression is significantly associated with poor survival in OHSU (Table 15). The results showed that *ACOT7* and *SCD* genes were the only genes passing all the criteria (Table 15; Fig. 41 for the *SCD* gene, and Fig. 42 for the *ACOT7* gene). I also looked for dependency scores in CRISPR screening in AML cell lines from the Depmap database. *SCD* had the lowest median dependency scores compared to the other genes (Fig. 43),

indicating its essentiality in AML cell lines. On the other hand, the median dependency score of *ACOT7* was close to 0, indicating that this gene is not essential. Therefore, I selected *SCD* as the target. Notably, while *SCD* level was higher in the 'Normal' group than in the 'Favorable' group, there was a clear and significant upward trend of *SCD* expression correlating with worse ELN2017 criteria (Fig. 41A). This provides rationale for *SCD* as a target, aligning with my goal of finding targets for high-risk AML patients.

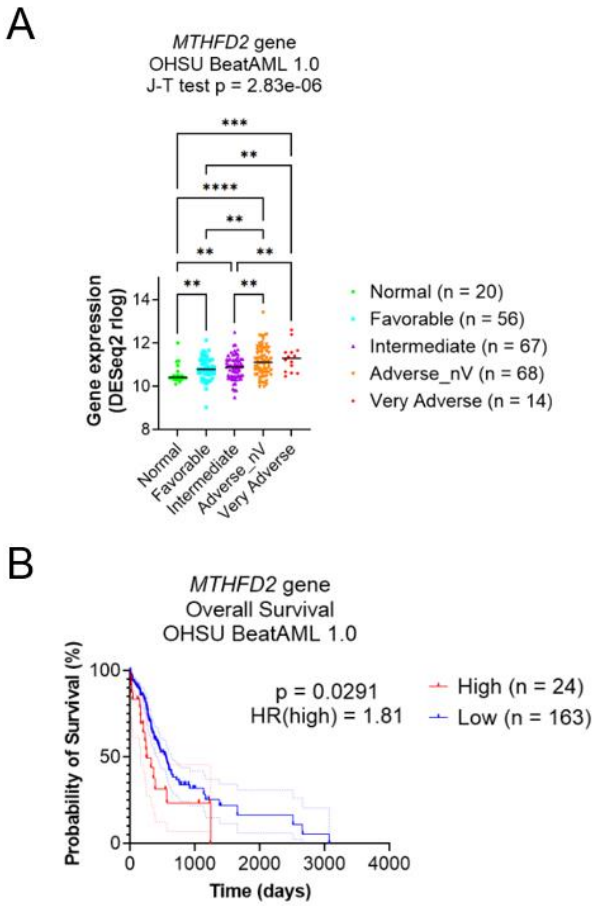


Figure 40. Risk-correlation and survival analysis for *MTHFD2* gene in the OHSU database

(A) The distributions of gene expression of *MTHFD2* gene in each risk category of revised ELN2017 in OHSU BeatAML 1.0 database. (B) Kaplan-Meier curves with 95% confidence intervals (dotted lines) for overall survival of AML patients in the OHSU BeatAML 1.0 database for *MTHFD2* gene. For (A), 'Adverse_nV' refers to the patients in the 'Adverse' category but not in the 'Very Adverse' category. The black lines indicate medians for each group. *P*-value is from the Jonckheere-Terpstra test. Post hoc analyses were performed with a two-stage linear step-up procedure. ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. For (B), *p*-value is from the log-rank test. The stratification of two groups in the graph was based on the best risk separation approach. HR(high) refers to the hazard ratio of the group with high gene expression.

Table 15. Risk-correlation analysis of revised ELN2017 for genes in the UFA_Synthesis pathway using multi-omics databases

Gene	OHSU					TCGA-LAML	Proteomics
	Jonckheere-Terpstra test	Survival (DSS)	HR for DSS	Survival (OS)	HR for OS	Jonckheere-Terpstra test	Jonckheere-Terpstra test
ACAA1	0.940	0.076	0.582	0.119	0.683	1.000	0.929
ACOT1	< 0.001	0.199	0.702	0.267	0.785	0.102	0.325
ACOT2	< 0.001	0.210	1.410	0.197	0.624	0.005	0.834
ACOT4	0.420	0.089	0.641	0.033	0.627	0.999	0.500
ACOT7	< 0.001	< 0.001	3.130	< 0.001	2.370	0.004	0.007
ACOX1	0.317	0.023	0.551	0.079	0.693	0.652	0.364
ACOX3	0.329	0.086	0.580	0.032	0.600	0.006	0.067
BAAT	0.014	< 0.001	3.090	< 0.001	2.510	0.001	NA
ELOVL2	0.067	0.004	2.090	0.005	1.780	< 0.001	NA
ELOVL5	0.878	0.005	0.413	0.016	0.530	0.023	0.013
ELOVL6	0.301	0.004	2.710	0.106	1.800	0.001	NA
FADS1	< 0.001	0.002	2.640	0.004	1.990	0.020	0.114
FADS2	0.007	0.048	1.800	0.048	1.610	0.084	0.013
HACD1	0.053	0.005	2.070	0.029	1.610	0.576	0.977
HACD2	0.724	0.019	1.870	0.009	1.720	0.001	0.021
HADHA	0.254	0.056	0.617	0.027	0.628	0.966	0.542
HSD17B12	0.089	0.049	1.770	0.068	1.500	0.895	0.774
PECR	0.389	< 0.001	2.930	< 0.001	2.110	< 0.001	0.001
SCD	< 0.001	< 0.001	3.430	< 0.001	2.320	< 0.001	0.010
SCD5	0.656	0.006	2.030	0.012	1.810	0.796	0.563
TECR	0.703	0.083	1.870	0.073	1.480	0.238	0.252
YOD1	0.019	0.029	2.000	0.046	1.650	< 0.001	0.075

DSS, disease-specific survival; OS, overall survival; HR, the hazard ratio of the high expression group; NA, not available.

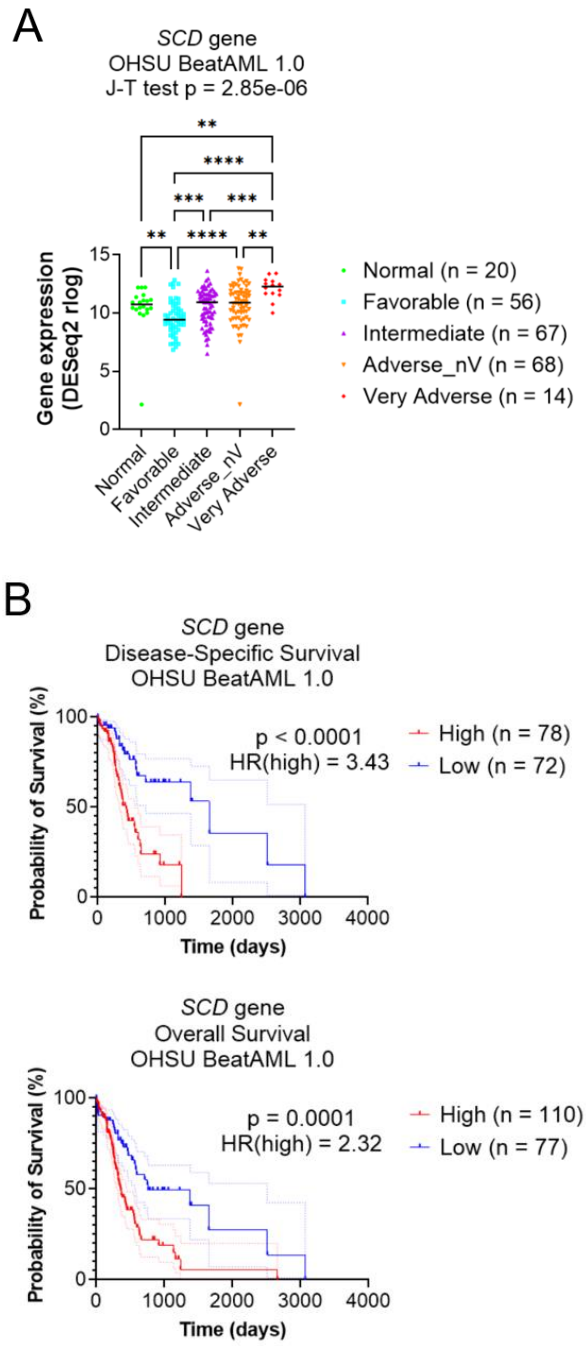


Figure 41. Risk-correlation and survival analysis for *SCD* gene in the OHSU database

(A) The distributions of gene expression of *SCD* gene in each risk category of revised ELN2017 in OHSU BeatAML 1.0 database. (B) Kaplan-Meier curves with

95% confidence intervals (dotted lines) for disease-specific survival and overall survival of AML patients in OHSU BeatAML 1.0 database for *SCD* gene. For (A), 'Adverse_nV' refers to the patients in the 'Adverse' category but not in the 'Very Adverse' category. The black lines indicate medians for each group. *P*-value is from the Jonckheere-Terpstra test. Post hoc analyses were performed with a two-stage linear step-up procedure. ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. For (B), *p*-values are from the log-rank test. The stratification of two groups in each graph was based on the best risk separation approach. HR(high) refers to the hazard ratio of the group with high gene expression.

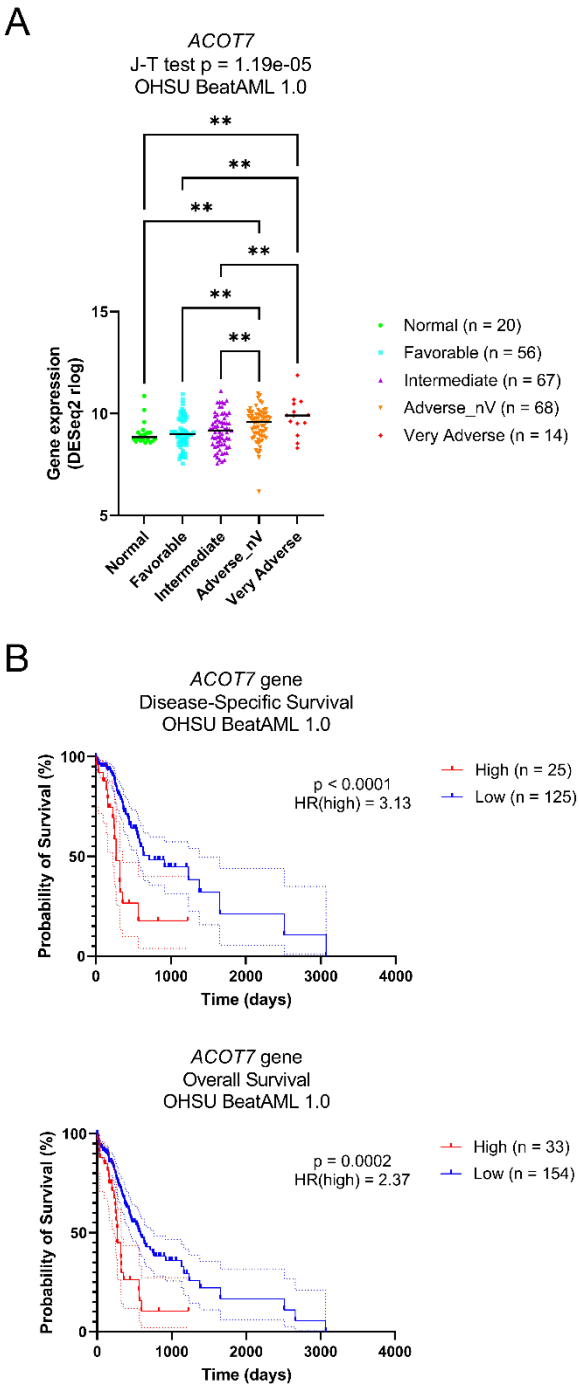


Figure 42. Identification of risk-correlated target gene *ACOT7*

(A) The distributions of gene expression of the *ACOT7* gene in each risk category of revised ELN2017 in the OHSU BeatAML 1.0 database. 'Adverse_nV' refers to

patients in the ‘Adverse’ category but not the ‘Very Adverse’ category. The black lines indicate medians for each group. *P*-value is from the Jonckheere-Terpstra test. Post hoc analyses were performed with a two-stage linear step-up procedure. ** $p < 0.01$. (B) Kaplan-Meier curves with 95% confidence intervals (dotted lines) for disease-specific survival and overall survival of AML patients in the OHSU BeatAML 1.0 database for the *ACOT7* gene. *P*-values are from the log-rank test. The stratification of two groups in each graph was based on the best risk separation approach. HR(high) refers to the hazard ratio of the group with high expression.

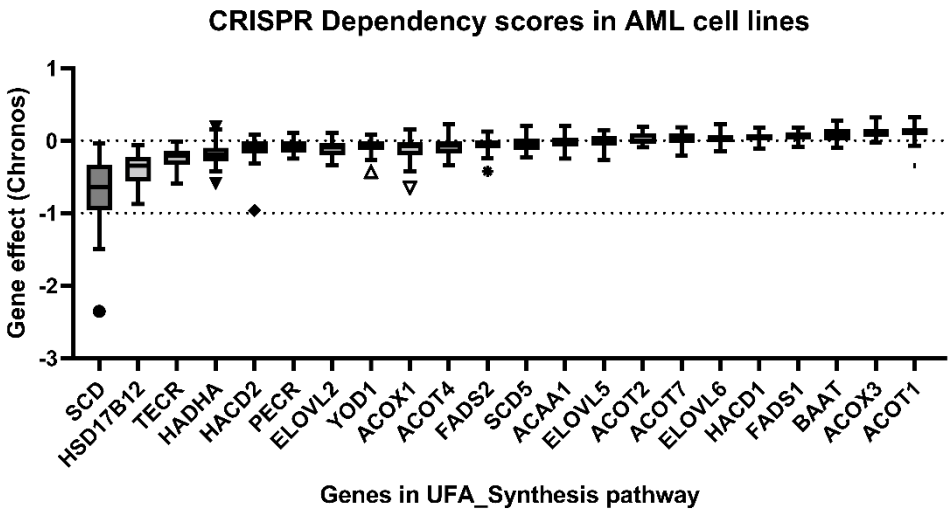


Figure 43. Identifying essentiality for genes in the 'UFA_Synthesis' pathway in AML

The distributions of CRISPR dependency scores in AML cell lines for each gene comprising the 'UFA_Synthesis' pathway. The scores were retrieved from Depmap. A score of 0 represents no viability effect, and a score of -1 corresponds to the median effect of known common-essential genes. A lower score indicates a higher likelihood that the gene of interest is essential in the given cell line. The genes are ordered by medians of the scores.

E. Experimental and functional validation of the target genes with inhibitors and AML cell lines

With *MTHFD2* and *SCD* bioinformatically suggested as risk-associated genes, I experimentally validated them using five AML (U937, MOLM-14, THP-1, KG-1, and HL-60) and one normal (HCC1954-BL) cell line. CCK-8 assay confirmed MTHFD2 inhibitor DS18561882 exhibit higher potency against all five AML cell lines compared to normal cell line (Fig. 44A). SCD inhibitor A939572 also showed ~8 times higher IC_{50}s for normal cell line (Fig. 44B). Trypan blue assay, along with normal peripheral blood mononuclear cells (PBMCs), also demonstrated the two inhibitors' cancer cell selectivity (Fig. 45). Additionally, SCD and MTHFD2 proteins' expression was higher in AML cell lines than normal PBMCs (Fig. 46). However, no significant correlations were observed among AML cell lines (Fig. 47). DS18561882 treatment reduced MTHFD2 protein levels (Fig. 48A), and A939572 treatment reduced unsaturated fatty acid levels without affecting SCD protein levels (Figs. 48B and 48C), confirming the targeted effects of both drugs on the target proteins. Due to limited references regarding SCD in AML, I further validated its functional relevance. Unsaturated fatty acids (PUFA and UFA) were measured in AML cell lines using nuclear magnetic resonance (NMR). Among the 22 genes, *SCD* highly correlated with unsaturated fatty acids (Fig. 49 and Table 16). My findings suggest the proposed inhibitors' selectivity toward AML cells, indicating their potential use for high-risk AML groups.

My bioinformatic screening showed the ‘cell-cycle related’ cluster significantly correlated with AML-risk groups. Cytarabine, a currently standard-of-care drug for AML, is a DNA replication inhibitor and inhibits the cell cycle [104]. Interestingly, it exhibited a remarkable difference in sensitivity across five AML cell lines (Fig. 50). For example, THP-1 and KG-1 were about 50 times less sensitive to cytarabine than U937. As the function of my target genes, *MTHFD2* and *SCD*, are orthogonal to the cytarabine’s mechanism, I hypothesized that their inhibitors might exhibit synergy when used with cytarabine. I tested the combination of A939572 or DS18561882 with cytarabine for THP-1 and KG-1 with high IC_{50} values for cytarabine (Fig. 51). The results showed that cytarabine combined with the SCD inhibitor or the MTHFD2 inhibitor exhibit synergy in AML cell survival inhibition (Fig. 52). In addition, dose-reduction of cytarabine was observed in all four combinations (Table 17), suggesting that A939572 or DS18561882 might be tested to alleviate cytarabine toxicity.

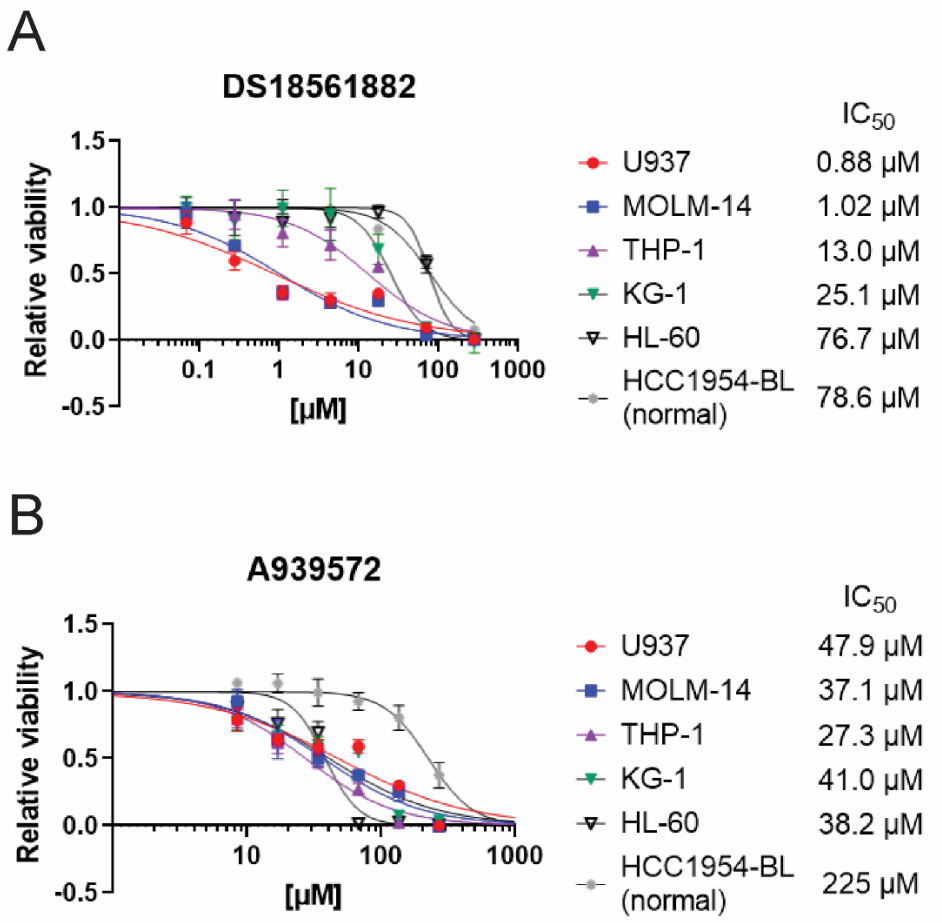


Figure 44. Cell viability assay for DS18561882 and A939572 drugs by CCK-8 assay

Dose-response curves and IC_{50}s for (A) DS18561882 drug and (B) A939572 drug to five AML cell lines (U937, MOLM-14, THP-1, KG-1, and HL-60) and one normal cell line (HCC1954-BL) by CCK-8 assay. The drugs were treated for 48 hours.

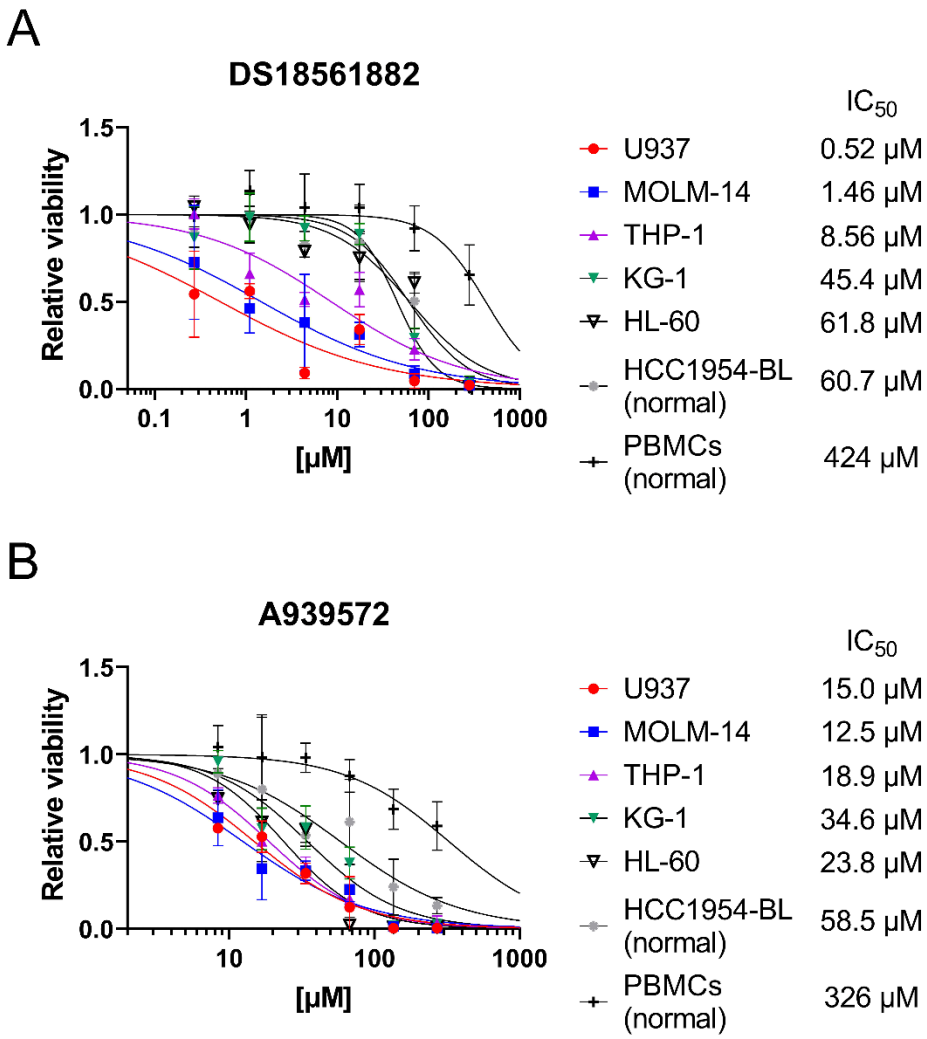


Figure 45. Cell viability assay for DS18561882 and A939572 drugs including normal PBMCs by Trypan blue assay

Dose-response curves and IC_{50}s for (A) DS18561882 and (B) A939572 to five AML cell lines (U937, MOLM-14, THP-1, KG-1, and HL-60), one normal cell line (HCC1954-BL) and normal PBMCs by Trypan blue assay. The drugs were treated for 48 hours.

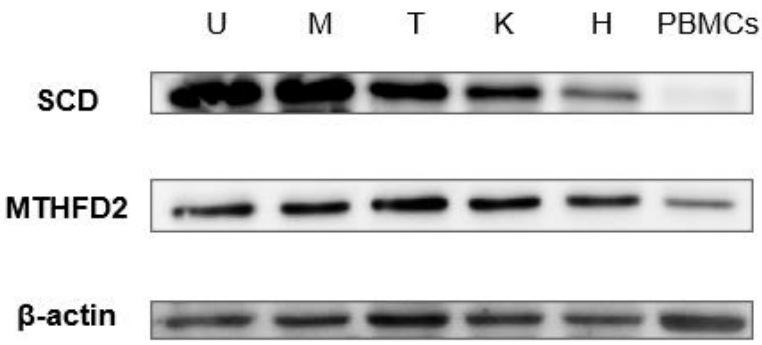


Figure 46. Basal SCD and MTHFD2 protein expression in AML cell lines and PBMCs

Western blot for SCD and MTHFD2 proteins in five AML cell lines (U937, MOLM-14, THP-1, KG-1, and HL-60) and normal PBMCs. U, M, T, K, and H refer to U937, MOLM-14, THP-1, KG-1, and HL-60, respectively.

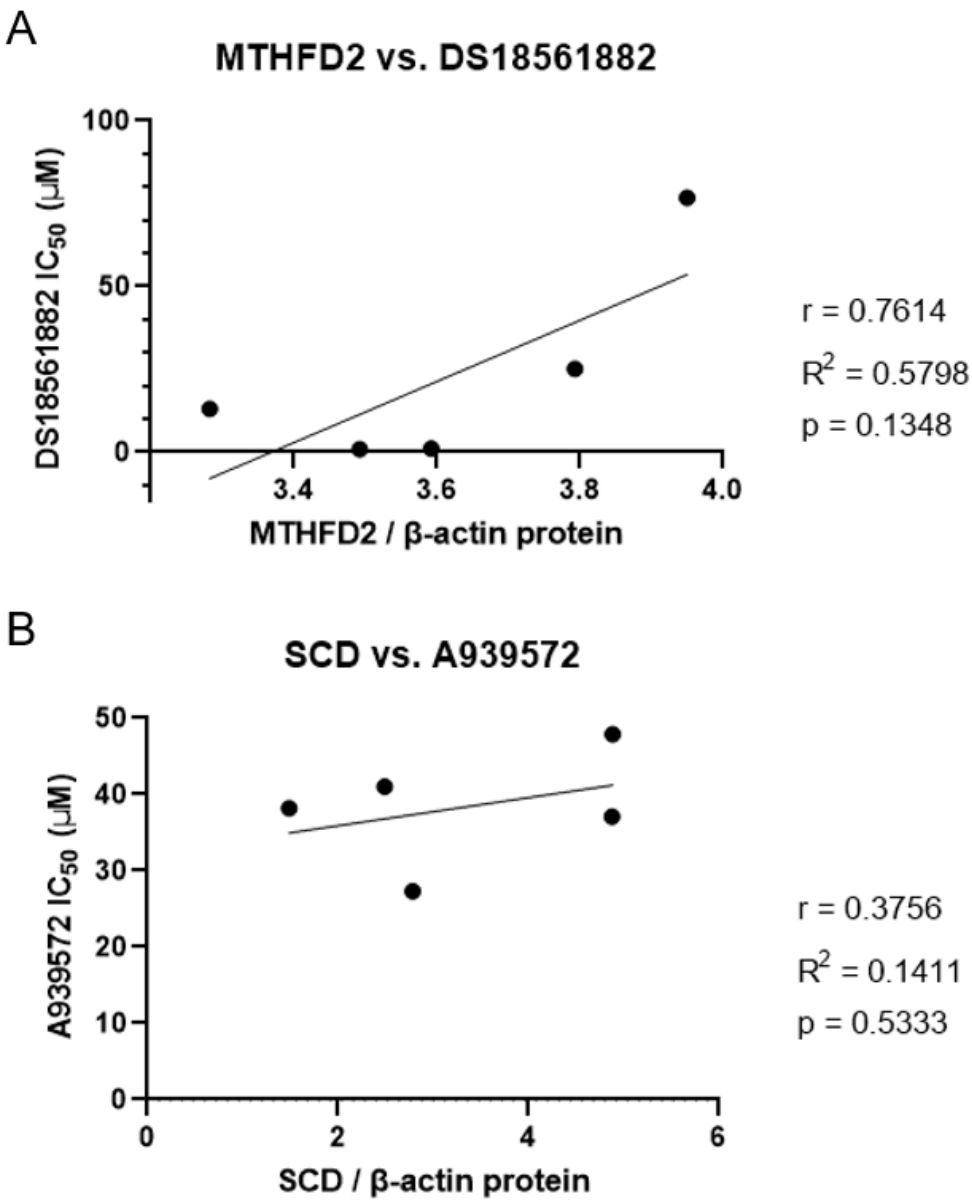


Figure 47. Correlation of MTHFD2 and SCD protein levels with their respective inhibitors

The relationship of (A) MTHFD2 protein level with DS18561882 IC_{50}s and (B) SCD protein level with A939572 IC_{50}s in the 5 AML cell lines (U937, MOLM-14, THP-1, KG-1, and HL-60). The IC_{50}s are from Fig. 44. Regression lines, r, R^2, and *p*-values from Pearson correlation analysis are shown.

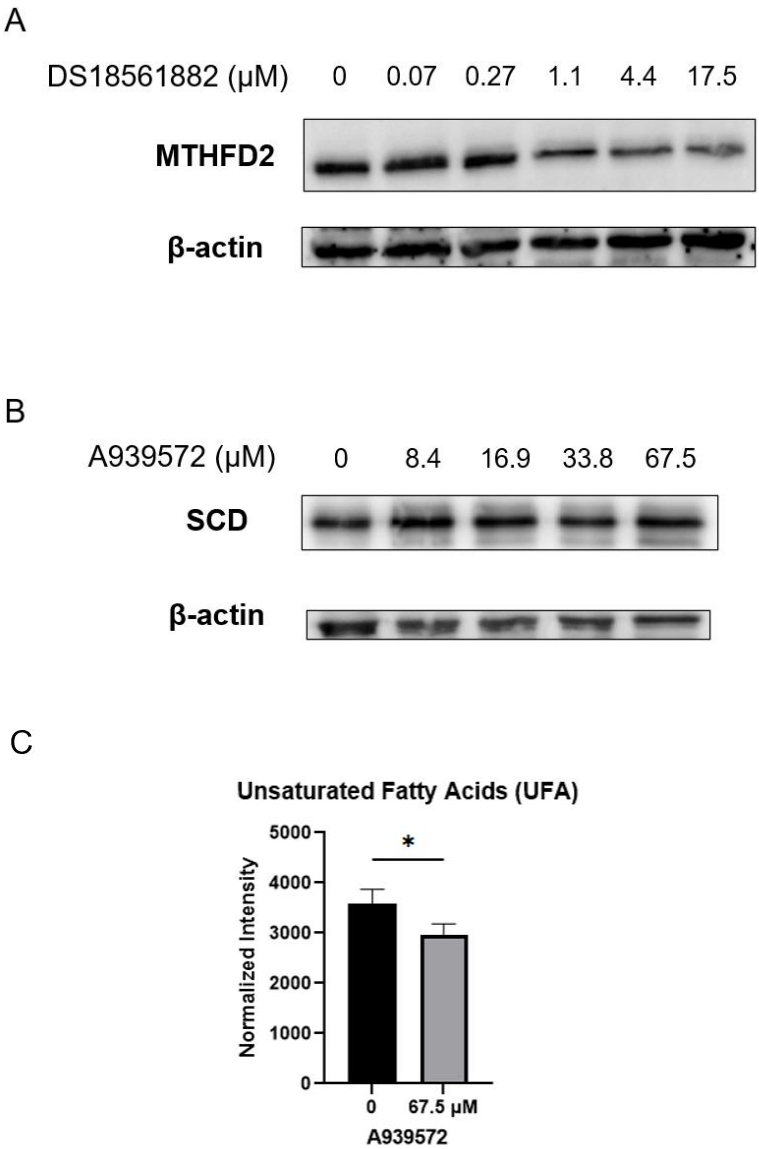


Figure 48. Effect of DS18561882 on MTHFD2 protein and A939572 on SCD protein

Western blot for (A) MTHFD2 protein in the MOLM-14 cell line treated with DS18561882 and (B) SCD protein in the U937 cell line in the indicated concentrations. The drugs were treated for 48 hours. (C) The amount of unsaturated fatty acids normalized by the cell numbers between A939572 non-treated and treated U937 cell lines. The drug was treated for 48 hours. *P*-value is from Student's *t*-test. * $p < 0.05$.

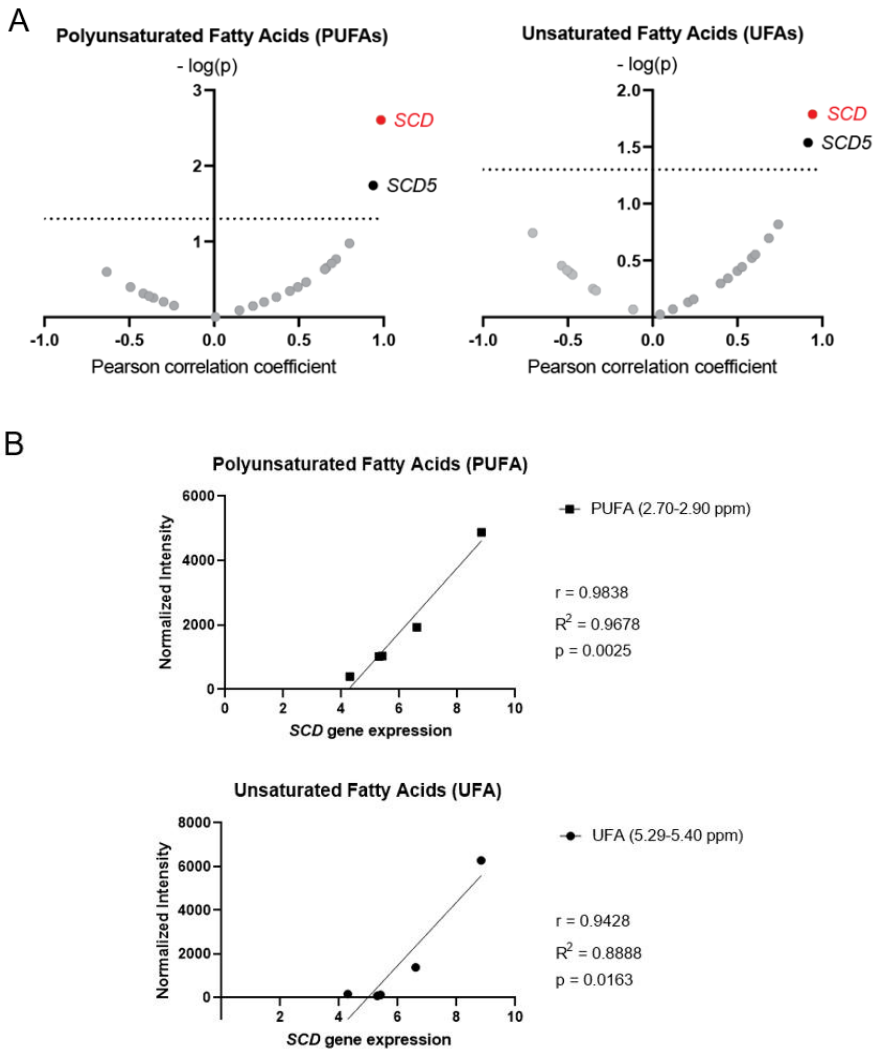


Figure 49. Functional validation of *SCD* gene

(A) Volcano plots for PUFAs and UFAs in the 5 AML cell lines (U937, MOLM-14, THP-1, KG-1, and HL-60). The x-axis refers to the Pearson correlation coefficient calculated by correlating gene expression of five AML cell lines in Depmap with measured PUFAs or UFAs by NMR for the genes only in the 'UFA_Synthesis' pathway. The y-axis refers to the minus log10-transformed *p*-value of the Pearson correlation coefficient. Non-significant (*p*-value 0.05 or higher) genes are indicated in grey. (B) The relationship of *SCD* gene expression and measured PUFAs or UFAs in the 5 AML cell lines. Regression lines, r, R^2, and *p*-values from Pearson correlation analysis are shown.

Table 16. UFA and PUFA correlation analysis for genes in the UFA_Synthesis pathway

Gene	UFA (5.29 – 5.40 ppm)		PUFA (2.70 – 2.90 ppm)	
	Pearson's r	*p*-value	Pearson's r	*p*-value
ACAA1	0.585	0.300	0.662	0.224
ACOT1	-0.536	0.352	-0.417	0.485
ACOT2	-0.487	0.406	-0.358	0.554
ACOT4	-0.471	0.423	-0.383	0.525
ACOT7	0.500	0.391	0.493	0.398
ACOX1	-0.116	0.853	0.008	0.989
ACOX3	0.401	0.504	0.447	0.451
BAAT	-0.505	0.386	-0.493	0.399
ELOVL2	-0.351	0.562	-0.297	0.627
ELOVL5	0.741	0.152	0.798	0.106
ELOVL6	-0.334	0.583	-0.236	0.702
FADS1	0.527	0.362	0.653	0.232
FADS2	0.605	0.280	0.719	0.171
HACD1	0.209	0.736	0.367	0.543
HACD2	0.242	0.695	0.295	0.629
HADHA	0.685	0.202	0.693	0.195
HSD17B12	-0.708	0.181	-0.633	0.252
PECR	0.119	0.849	0.230	0.710
SCD	0.943	0.016	0.984	0.002
SCD5	0.916	0.029	0.938	0.018
TECR	0.044	0.944	0.149	0.811
YOD1	0.444	0.454	0.543	0.344

UFA, unsaturated fatty acid; PUFA, polyunsaturated fatty acid.

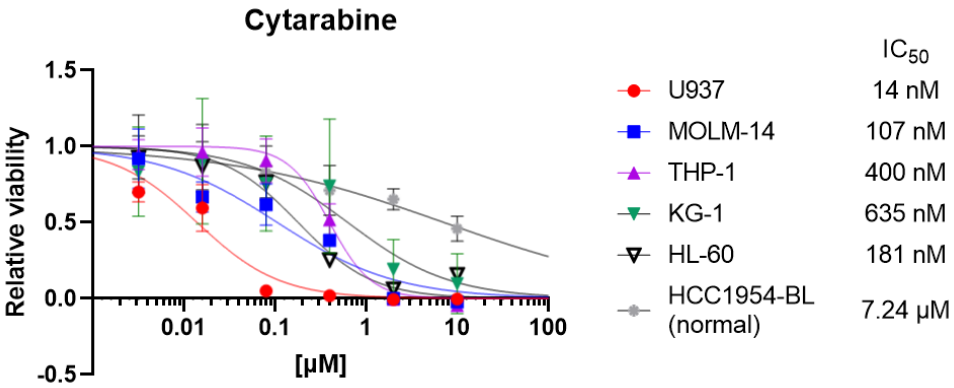


Figure 50. Dose-response curves for cytarabine

Dose-response curves and IC_{50}s for cytarabine to five AML cell lines (U937, MOLM-14, THP-1, KG-1, and HL-60) and one normal cell line (HCC1954-BL). The drugs were treated for 48 hours.

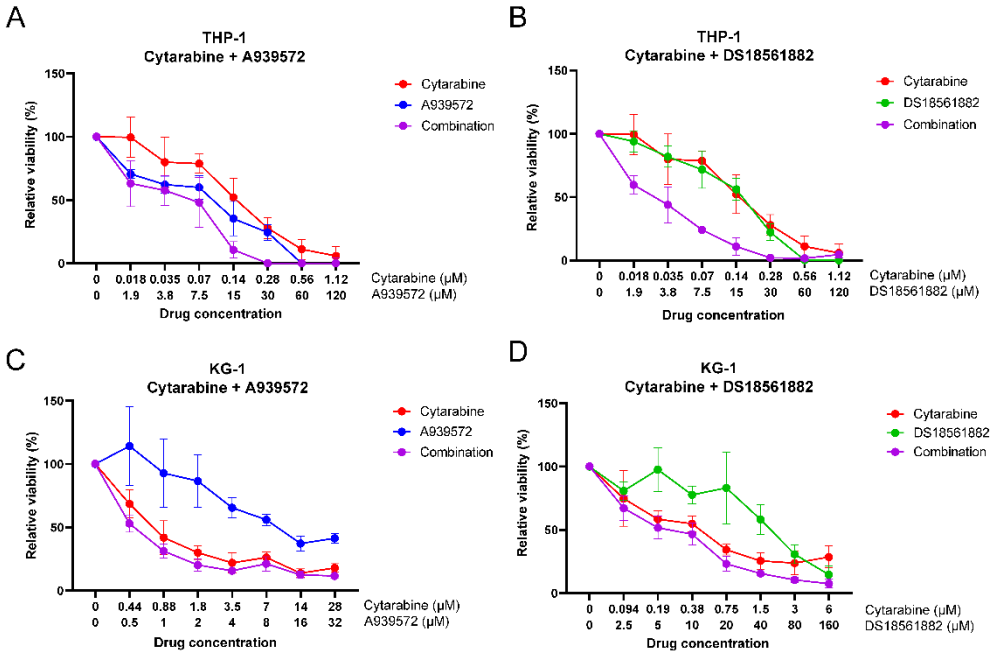


Figure 51. Combination of SCD and MTHFD2 inhibitor with cytarabine

(A, B) Dose-response curves for THP-1 cell line, testing synergy of cytarabine either with (A) A939572 or (B) DS18561882. (C, D) Dose-response curves for KG-1 cell line, the testing synergy of cytarabine either with (C) A939572 or (D) DS18561882.

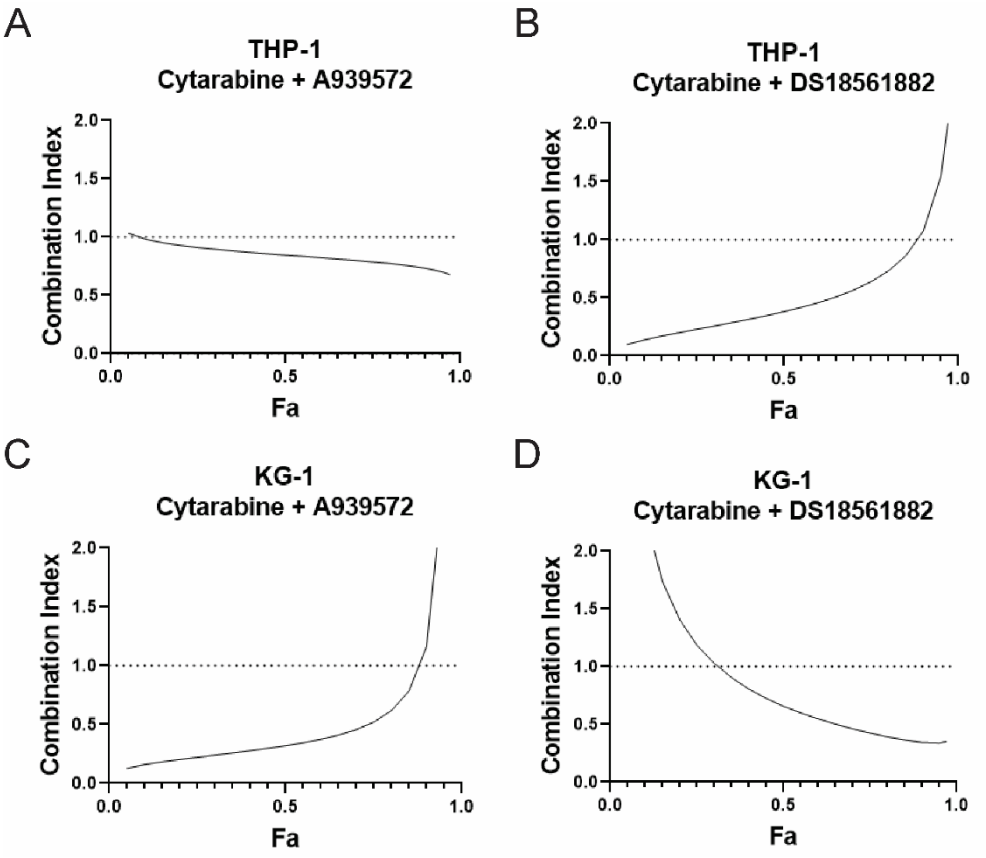


Figure 52. Synergy of SCD and MTHFD2 inhibitor with cytarabine

(A, B) Combination index plots for THP-1 cell line for the combination of cytarabine with (A) A939572 or (B) DS18561882. (C, D) Combination index plots for KG-1 cell line for the combination of cytarabine with (C) A939572 or (D) DS18561882. The combination index of less than 1 indicates synergy, and Fa refers to fractions affected by particular dose of a drug, herein the fraction of dead cells compared to non-treated samples.

Table 17. Dose reduction index of cytarabine at fractions affected (Fa) = 0.9 in cytarabine-resistant cell lines

	THP-1	KG-1
Cytarabine + A939572	6.65	1.76
Cytarabine + DS18561882	3.59	8.03

IV. Discussion

Since its inception in 2017, ELN2017 has been used for diagnosing and managing AML patients and has also been applied to research. Still, most studies have focused on validating ELN2017 risk criteria in terms of survival in individual hospitals. Only a few papers have addressed specific biological pathways dysregulated in high-risk groups or specific drugs targeting these pathways. Some focused on individual mutations comprising the 'Adverse' category of ELN2017, such as *RUNX1* mutation [21] or *TP53* mutation [22]. Another study found gene modules related to 14 markers, including ELN2017 itself and the contributing mutations, later focusing on modules correlated with *NPM1*/*FLT3*-ITD mutation [25]. Other researchers first sought to find particular genes related to survival and then constructed the prognostic risk scores correlated with ELN2017 [23, 24]. Of note, one of the studies found that the high-risk phenotype score is enriched with the biosynthesis of unsaturated fatty acid gene set [24], consistent with my result implicating high expression of the *SCD* gene in high-risk groups. Furthermore, a metabolomics study also showed that plasma levels of some species of PUFA were positively correlated with risk stratification [98] in AML patients.

Compared with the previously reported approaches, the critical difference in my current work is that I initiated the analysis by assessing pathways that directly correlate with ELN2017 using GSVA pathway scores. Other studies started with individual genes, rather than pathways, for survival relationships or focused on individual mutations comprising ELN2017. Then, I subsequently validated the

resulting pathways and genes with survival data and also with results from other multi-omics databases. I also added the 'Very Adverse' group and the 'Normal' group, which were not included in most of the prior studies, which, I believe, gave more reliable results. I also utilized the proteomics AML database, which has not been used in the above studies, adding confidence to my results. Probably most importantly, I carried out experimental and functional validation of the target genes after the bioinformatics screening, which is seldom the case for most related studies. My results showing the relationship between *SCD* expression and actual UFA and PUFA amounts suggest that *SCD* is a promising target for AML. I believe that these extra steps enabled my suggested drugs, A939572 and DS18561882, to exhibit selectivity toward cancer cells and efficacy against AML cells relatively resistant to cytarabine. These drugs' synergy and dose reduction for cytarabine against cell lines with high IC_{50} values, THP-1 and KG-1, are noteworthy for developing new combinational treatments. It is also important to note that these cell lines harbor *TP53* mutation (https://cellmodelpassports.sanger.ac.uk) and complex karyotype (https://www.dsmz.de/), which are the important characteristics of the 'Very Adverse' group of revised ELN2017 criteria usedf in my study. To my knowledge, only limited information about SCD inhibitors on AML is available. MTHFD2 inhibitors' effects on AML have been revealed only recently, particularly the study suggesting pyrimidine depletion and replication stress as the underlying mechanism [88]. Although the authors observed synergy between MTHFD2 inhibitors and ATR inhibitors or dUTPase inhibitors, they did not address the synergy with cytarabine, a standard-of-care drug. My results for synergy between

SCD/MTHFD2 inhibitors and cytarabine may have additional implications in this respect.

Furthermore, the significant advantage of my approach is that it can also be applied to find other target genes/pathways, or even for solid tumors. Even though I chose the *MTHFD2* gene in the Folate_Metabolism pathway due to its literature evidence, when I looked for another candidate gene using the same approach used in the UFA_Synthesis pathway, *ALDH1L2* gene was found to be the only gene significant in all three databases (Table 18). Unlike *MTHFD2*, in AML, no literature study and no inhibitors are available, but as seen in the case of the *SCD* gene, the *ALDH1L2* gene could be the possible target. In addition, in solid tumors, the ELN2017 category is analogous to stage or grade information in that it is closely related to patient prognosis. Other relevant parameters, such as recurrence status after initial treatment and lymph/distant metastasis status available in TCGA databases, could also be combined to build an ELN2017-like variable for prognosis categorization. Then, a simple Jonckheere-Terpstra test can be applied to correlate pathways to the variable, as it is a non-parametric test that can be used regardless of sample distributions. This correlation between the prognosis variable and pathways should give more valuable information related to cancer malignancy than conventional analysis of tumor vs. normal samples.

Very recently, a new version of ELN recommendation, ELN2022, was introduced [105]. While the stratification of patients into three categories was maintained, it now uses more information on patient genetics, such as bZIP in-frame mutated CEBPA, KAT6A::CREBBP fusion, or variant allele fraction of *TP53* mutation [105], some of which information is not registered in the present

databases. According to a study that attempted the validation of ELN2022 in clinic [106], 83%, 72%, and 90% of patients in each risk category kept their allocation in ELN2017, suggesting broadly similar categorization. Also, in this study, the authors concluded that the ELN2022 classification did not significantly perform better in outcome prognostication than ELN2017 classification, implying that much more validations are yet needed. Therefore, my analysis with ELN2017 should still be meaningful and valuable.

Table 18. Risk-correlation analysis of revised ELN2017 for genes in the Folate_metabolism pathway using multiomics databases

Gene	OHSU					TCGA-LAML	Proteomics
	Jonckheere-Terpstra test	Survival (DSS)	HR for DSS	Survival (OS)	HR for OS	Jonckheere-Terpstra test	Jonckheere-Terpstra test
ALDH1L1	0.965	0.012	2.031	0.016	1.749	0.241	0.093
ALDH1L2	0.027	0.004	2.303	< 0.001	2.119	< 0.001	0.001
DHFR	0.483	0.001	3.075	0.002	2.029	0.800	0.259
DHFR2	0.995	0.076	2.023	0.153	1.494	0.139	NA
FOLR2	0.353	0.035	1.889	0.121	1.432	< 0.001	0.002
FPGS	0.050	0.002	0.414	0.003	0.525	0.406	0.424
MTHFD1	0.008	0.035	1.865	0.009	2.149	0.004	0.275
MTHFD1L	< 0.001	0.043	1.805	0.056	1.561	0.003	0.816
MTHFD2	< 0.001	0.054	1.788	0.029	1.808	0.181	0.092
MTHFD2L	0.008	0.051	0.553	0.236	0.759	0.747	0.500
MTHFR	0.017	0.153	0.484	0.093	1.406	< 0.001	0.061
MTHFS	0.269	0.038	1.714	0.069	1.462	0.986	0.982
SHMT1	0.873	< 0.001	2.943	0.003	1.936	0.991	0.325
SHMT2	0.001	0.007	2.103	0.004	2.014	0.158	0.339
SLC19A1	0.062	0.066	1.639	0.060	1.490	0.001	0.002
SLC25A32	0.003	0.049	5.777	0.073	2.741	0.034	0.014
SLC46A1	0.841	0.072	1.780	0.235	1.363	0.111	NA

DSS, disease-specific survival; OS, overall survival; HR, the hazard ratio of the high expression group; NA, not available.

Conclusion

Overall, I used bioinformatics to find target genes or pathways in AML, looked for literature support, and performed experimental validations.

The study from Part I provides strong evidence for SUMOylation as a new targetable pathway for AML, based on integrated bioinformatic screening and validations with *in vitro*, *ex vivo*, and *in vivo* preclinical AML models. For toxicity, the longer survival of TAK-981-treated mice indicates a favorable therapeutic index. Consistent with this, a previous study with TAK-981 showed a good toxicity property up to 40 mg/kg in mice [39]. In addition, normal or patients with remission after therapy had lower SAE1/SAE2, the target of TAK-981, than patients with active AML (Fig. 2C), suggesting possible selectivity of the drug. These favorable efficacy and toxicity data should prompt further studies for its optimal combination and transitions to clinical trials with AML.

In the study from Part II, by utilizing transcriptomics and proteomics databases, I found pathways upregulated in correlation with increased risk, the specific genes to target in those pathways, and suggested drugs that might have a synergistic effect with standard-of-care drug cytarabine. Since not much is known about the roles of unsaturated fatty acids or folate metabolism in AML, my results could be further exploited to find a mechanistic relationship between those pathways and the malignancy of AML, and also to test the target drugs on high-risk AML animal models.

To summarize, in Part I, I found SUMOylation as AML-specific target pathway compared to normal bone marrow and suggested TAK-981 as the drug candidate. In Part II, I found *SCD* and *MTHFD2* gene as target for high-risk AML patients and suggested their respective inhibitors as the drug candidates.

References

1. Döhner H, Weisdorf DJ, Bloomfield CD. Acute Myeloid Leukemia. *N Engl J Med* 2015; 373: 1136-1152.
2. Shallis RM, Wang R, Davidoff A, Ma X, Zeidan AM. Epidemiology of acute myeloid leukemia: Recent progress and enduring challenges. *Blood Rev* 2019; 36: 70-87.
3. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T *et al*. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* 2017; 129: 424-447.
4. Kantarjian H, Kadia T, DiNardo C, Daver N, Borthakur G, Jabbour E *et al*. Acute myeloid leukemia: current progress and future directions. *Blood Cancer J* 2021; 11: 41.
5. Stone RM, Mandrekar SJ, Sanford BL, Laumann K, Geyer S, Bloomfield CD *et al*. Midostaurin plus Chemotherapy for Acute Myeloid Leukemia with a FLT3 Mutation. *N Engl J Med* 2017; 377: 454-464.
6. Lancet JE, Uy GL, Cortes JE, Newell LF, Lin TL, Ritchie EK *et al*. CPX-351 (cytarabine and daunorubicin) Liposome for Injection Versus Conventional Cytarabine Plus Daunorubicin in Older Patients With Newly Diagnosed Secondary Acute Myeloid Leukemia. *J Clin Oncol* 2018; 36: 2684-2692.
7. Castaigne S, Pautas C, Terré C, Raffoux E, Bordessoule D, Bastie JN *et al*.

Effect of gemtuzumab ozogamicin on survival of adult patients with de-novo acute myeloid leukaemia (ALFA-0701): a randomised, open-label, phase 3 study. *Lancet* 2012; 379: 1508-1516.

8. Perl AE, Martinelli G, Cortes JE, Neubauer A, Berman E, Paolini S *et al*. Gilteritinib or Chemotherapy for Relapsed or Refractory FLT3-Mutated AML. *N Engl J Med* 2019; 381: 1728-1740.

9. Lancet JE, Uy GL, Newell LF, Lin TL, Ritchie EK, Stuart RK *et al*. CPX-351 versus 7+3 cytarabine and daunorubicin chemotherapy in older adults with newly diagnosed high-risk or secondary acute myeloid leukaemia: 5-year results of a randomised, open-label, multicentre, phase 3 trial. *Lancet Haematol* 2021; 8: e481-e491.

10. Wei AH, Döhner H, Pocock C, Montesinos P, Afanasyev B, Dombret H *et al*. Oral Azacitidine Maintenance Therapy for Acute Myeloid Leukemia in First Remission. *N Engl J Med* 2020; 383: 2526-2537.

11. DiNardo CD, Jonas BA, Pullarkat V, Thirman MJ, Garcia JS, Wei AH *et al*. Azacitidine and Venetoclax in Previously Untreated Acute Myeloid Leukemia. *N Engl J Med* 2020; 383: 617-629.

12. Stubbins RJ, Francis A, Kuchenbauer F, Sanford D. Management of Acute Myeloid Leukemia: A Review for General Practitioners in Oncology. *Curr Oncol* 2022; 29: 6245-6259.

13. Kayser S, Levis MJ. Updates on targeted therapies for acute myeloid leukaemia. *Br J Haematol* 2022; 196: 316-328.

14. Döhner H, Estey EH, Amadori S, Appelbaum FR, Büchner T, Burnett AK *et al*. Diagnosis and management of acute myeloid leukemia in adults:

recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood* 2010; 115: 453-474.

15. Grimm J, Jentzsch M, Bill M, Goldmann K, Schulz J, Niederwieser D *et al*. Prognostic impact of the ELN2017 risk classification in patients with AML receiving allogeneic transplantation. *Blood Adv* 2020; 4: 3864-3874.

16. Jentzsch M, Grimm J, Bill M, Brauer D, Backhaus D, Goldmann K *et al*. ELN risk stratification and outcomes in secondary and therapy-related AML patients consolidated with allogeneic stem cell transplantation. *Bone Marrow Transplant* 2021; 56: 936-945.

17. Herold T, Rothenberg-Thurley M, Grunwald VV, Janke H, Goerlich D, Sauerland MC *et al*. Validation and refinement of the revised 2017 European LeukemiaNet genetic risk stratification of acute myeloid leukemia. *Leukemia* 2020; 34: 3161-3172.

18. Buccisano F, Palmieri R, Piciocchi A, Arena V, Candoni A, Melillo L *et al*. ELN2017 risk stratification improves outcome prediction when applied to the prospective GIMEMA AML1310 protocol. *Blood Adv* 2022; 6: 2510-2516.

19. Hansen DK, Kim J, Thompson Z, Hussaini M, Nishihori T, Ahmad A *et al*. ELN 2017 Genetic Risk Stratification Predicts Survival of Acute Myeloid Leukemia Patients Receiving Allogeneic Hematopoietic Stem Cell Transplantation. *Transplant Cell Ther* 2021; 27: 256.e251-256.e257.

20. Eisfeld AK, Kohlschmidt J, Mrózek K, Blachly JS, Walker CJ, Nicolet D *et al*. Mutation patterns identify adult patients with de novo acute myeloid leukemia aged 60 years or older who respond favorably to standard

chemotherapy: an analysis of Alliance studies. *Leukemia* 2018; 32: 1338-1348.

21. Zhu F, Huang R, Li J, Liao X, Huang Y, Lai Y. Identification of Key Genes and Pathways Associated with RUNX1 Mutations in Acute Myeloid Leukemia Using Bioinformatics Analysis. *Med Sci Monit* 2018; 24: 7100-7108.

22. Huang R, Liao X, Li Q. Identification of key pathways and genes in TP53 mutation acute myeloid leukemia: evidence from bioinformatics analysis. *Onco Targets Ther* 2018; 11: 163-173.

23. Tang Y, Xiao S, Wang Z, Liang Y, Xing Y, Wu J *et al*. A Prognostic Model for Acute Myeloid Leukemia Based on IL-2/STAT5 Pathway-Related Genes. *Front Oncol* 2022; 12: 785899.

24. Lai B, Lai Y, Zhang Y, Zhou M, OuYang G. Survival prediction in acute myeloid leukemia using gene expression profiling. *BMC Med Inform Decis Mak* 2022; 22: 57.

25. Guo C, Gao YY, Ju QQ, Zhang CX, Gong M, Li ZL. The landscape of gene co-expression modules correlating with prognostic genetic abnormalities in AML. *J Transl Med* 2021; 19: 228.

26. Kroonen JS, Vertegaal ACO. Targeting SUMO Signaling to Wrestle Cancer. *Trends Cancer* 2021; 7: 496-510.

27. Hay RT. SUMO: a history of modification. *Mol Cell* 2005; 18: 1-12.

28. Eifler K, Cuijpers SAG, Willemstein E, Raaijmakers JA, El Atmioui D, Ovaa H *et al*. SUMO targets the APC/C to regulate transition from metaphase to anaphase. *Nat Commun* 2018; 9: 1119.

29. Nacerddine K, Lehembre F, Bhaumik M, Artus J, Cohen-Tannoudji M, Babinet C *et al*. The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Dev Cell* 2005; 9: 769-779.

30. Flotho A, Melchior F. Sumoylation: a regulatory protein modification in health and disease. *Annu Rev Biochem* 2013; 82: 357-385.

31. Kim DH, Lee C, Kim B, Lee SH, Han KH. Rescuing p53 from mdm2 by a pre-structured motif in intrinsically unfolded SUMO specific protease 4. *BMB Rep* 2017; 50: 485-486.

32. Ding B, Sun Y, Huang J. Overexpression of SKI oncoprotein leads to p53 degradation through regulation of MDM2 protein sumoylation. *J Biol Chem* 2012; 287: 14621-14630.

33. Hoellein A, Fallahi M, Schoeffmann S, Steidle S, Schaub FX, Rudelius M *et al*. Myc-induced SUMOylation is a therapeutic vulnerability for B-cell lymphoma. *Blood* 2014; 124: 2081-2090.

34. Kessler JD, Kahle KT, Sun T, Meerbrey KL, Schlabach MR, Schmitt EM *et al*. A SUMOylation-dependent transcriptional subprogram is required for Myc-driven tumorigenesis. *Science* 2012; 335: 348-353.

35. Rabellino A, Melegari M, Tompkins VS, Chen W, Van Ness BG, Teruya-Feldstein J *et al*. PIAS1 Promotes Lymphomagenesis through MYC Upregulation. *Cell Rep* 2016; 15: 2266-2278.

36. González-Prieto R, Cuijpers SA, Kumar R, Hendriks IA, Vertegaal AC. c-Myc is targeted to the proteasome for degradation in a SUMOylation-dependent manner, regulated by PIAS1, SENP7 and RNF4. *Cell Cycle* 2015; 14: 1859-1872.

37. Hannoun Z, Maarifi G, Chelbi-Alix MK. The implication of SUMO in intrinsic and innate immunity. *Cytokine Growth Factor Rev* 2016; 29: 3-16.

38. Biederstädt A, Hassan Z, Schneeweis C, Schick M, Schneider L, Muckenhuber A *et al*. SUMO pathway inhibition targets an aggressive pancreatic cancer subtype. *Gut* 2020; 69: 1472-1482.

39. Lightcap ES, Yu P, Grossman S, Song K, Khattar M, Xega K *et al*. A small-molecule SUMOylation inhibitor activates antitumor immune responses and potentiates immune therapies in preclinical models. *Sci Transl Med* 2021; 13: eaba7791.

40. Dong S, Chen J. SUMOylation of sPRDM16 promotes the progression of acute myeloid leukemia. *BMC Cancer* 2015; 15: 893.

41. Zhang J, Huang FF, Wu DS, Li WJ, Zhan HE, Peng MY *et al*. SUMOylation of insulin-like growth factor 1 receptor, promotes proliferation in acute myeloid leukemia. *Cancer Lett* 2015; 357: 297-306.

42. Bossis G, Sarry JE, Kifagi C, Ristic M, Saland E, Vergez F *et al*. The ROS/SUMO axis contributes to the response of acute myeloid leukemia cells to chemotherapeutic drugs. *Cell Rep* 2014; 7: 1815-1823.

43. Langston SP, Grossman S, England D, Afroze R, Bence N, Bowman D *et al*. Discovery of TAK-981, a First-in-Class Inhibitor of SUMO-Activating Enzyme for the Treatment of Cancer. *J Med Chem* 2021; 64: 2501-2520.

44. Lam V, Best SR, Bruss N, Liu T, Rowland TH, Huszar D *et al*. Pharmacologic Inhibition of SUMO-Activating Enzyme (SAE) with TAK-981 Augments Interferon Signaling and Regulates T Cell Differentiation in Ex Vivo studies of Chronic Lymphocytic Leukemia (CLL). *Blood* 2019;

134: 1760.

45. Ley TJ, Miller C, Ding L, Raphael BJ, Mungall AJ, Robertson A *et al*. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* 2013; 368: 2059-2074.

46. Haferlach T, Kohlmann A, Wieczorek L, Basso G, Kronnie GT, Béné MC *et al*. Clinical utility of microarray-based gene expression profiling in the diagnosis and subclassification of leukemia: report from the International Microarray Innovations in Leukemia Study Group. *J Clin Oncol* 2010; 28: 2529-2537.

47. Tyner JW, Tognon CE, Bottomly D, Wilmot B, Kurtz SE, Savage SL *et al*. Functional genomic landscape of acute myeloid leukaemia. *Nature* 2018; 562: 526-531.

48. Roushangar R, Mias GI. Multi-study reanalysis of 2,213 acute myeloid leukemia patients reveals age- and sex-dependent gene expression signatures. *Sci Rep* 2019; 9: 12413.

49. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; 15: 550.

50. Straube J, Ling VY, Hill GR, Lane SW. The impact of age, NPM1(mut), and FLT3(ITD) allelic ratio in patients with acute myeloid leukemia. *Blood* 2018; 131: 1148-1153.

51. Yoon S, Kim J, Kim SK, Baik B, Chi SM, Kim SY *et al*. GScluster: network-weighted gene-set clustering analysis. *BMC Genomics* 2019; 20: 352.

52. Mizuno H, Kitada K, Nakai K, Sarai A. PrognoScan: a new database for

meta-analysis of the prognostic value of genes. *BMC Med Genomics* 2009; 2: 18.

53. Shin AR, Lee SE, Choi H, Sohn HJ, Cho HI, Kim TG. An effective peptide vaccine strategy circumventing clonal MHC heterogeneity of murine myeloid leukaemia. *Br J Cancer* 2020; 123: 919-931.

54. Chen Y, Jacamo R, Konopleva M, Garzon R, Croce C, Andreeff M. CXCR4 downregulation of let-7a drives chemoresistance in acute myeloid leukemia. *J Clin Invest* 2013; 123: 2395-2407.

55. Derenzini E, Rossi A, Treré D. Treating hematological malignancies with drugs inhibiting ribosome biogenesis: when and why. *J Hematol Oncol* 2018; 11: 75.

56. Skrtić M, Sriskanthadevan S, Jhas B, Gebbia M, Wang X, Wang Z *et al*. Inhibition of mitochondrial translation as a therapeutic strategy for human acute myeloid leukemia. *Cancer Cell* 2011; 20: 674-688.

57. Reed GA, Schiller GJ, Kambhampati S, Tallman MS, Douer D, Minden MD *et al*. A Phase 1 study of intravenous infusions of tigecycline in patients with acute myeloid leukemia. *Cancer Med* 2016; 5: 3031-3040.

58. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013; 14: 7.

59. Chou TC. Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacol Rev* 2006; 58: 621-681.

60. Baik H, Boulanger M, Hosseini M, Kowalczyk J, Zaghdoudi S, Salem T *et al*. Targeting the SUMO Pathway Primes All-trans Retinoic Acid-Induced

Differentiation of Nonpromyelocytic Acute Myeloid Leukemias. *Cancer Res* 2018; 78: 2601-2613.

61. Nakayama F, Nishihara S, Iwasaki H, Kudo T, Okubo R, Kaneko M *et al*. CD15 expression in mature granulocytes is determined by alpha 1,3-fucosyltransferase IX, but in promyelocytes and monocytes by alpha 1,3-fucosyltransferase IV. *J Biol Chem* 2001; 276: 16100-16106.

62. Geletu M, Balkhi MY, Peer Zada AA, Christopeit M, Pulikkan JA, Trivedi AK *et al*. Target proteins of C/EBPalphap30 in AML: C/EBPalphap30 enhances sumoylation of C/EBPalphap42 via up-regulation of Ubc9. *Blood* 2007; 110: 3301-3309.

63. Aroua N, Boet E, Ghisi M, Nicolau-Travers ML, Saland E, Gwilliam R *et al*. Extracellular ATP and CD39 Activate cAMP-Mediated Mitochondrial Stress Response to Promote Cytarabine Resistance in Acute Myeloid Leukemia. *Cancer Discov* 2020; 10: 1544-1565.

64. Kogan AA, Topper MJ, Dellomo AJ, Stojanovic L, McLaughlin LJ, Creed TM *et al*. Activating STING1-dependent immune signaling in TP53 mutant and wild-type acute myeloid leukemia. *Proc Natl Acad Sci U S A* 2022; 119: e2123227119.

65. Sugimoto K, Toyoshima H, Sakai R, Miyagawa K, Hagiwara K, Ishikawa F *et al*. Frequent mutations in the p53 gene in human myeloid leukemia cell lines. *Blood* 1992; 79: 2378-2383.

66. Daniel MT, Koken M, Romagné O, Barbey S, Bazarbachi A, Stadler M *et al*. PML protein expression in hematopoietic and acute promyelocytic leukemia cells. *Blood* 1993; 82: 1858-1867.

67. Lallemand-Breitenbach V, Jeanne M, Benhenda S, Nasr R, Lei M, Peres L *et al*. Arsenic degrades PML or PML-RARalpha through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nat Cell Biol* 2008; 10: 547-555.

68. Collins SJ. The role of retinoids and retinoic acid receptors in normal hematopoiesis. *Leukemia* 2002; 16: 1896-1905.

69. Johnson DE, Redner RL. An ATRActive future for differentiation therapy in AML. *Blood Rev* 2015; 29: 263-268.

70. Lübbert M, Grishina O, Schmoor C, Schlenk RF, Jost E, Crysandt M *et al*. Valproate and Retinoic Acid in Combination With Decitabine in Elderly Nonfit Patients With Acute Myeloid Leukemia: Results of a Multicenter, Randomized, 2 × 2, Phase II Trial. *J Clin Oncol* 2020; 38: 257-270.

71. Chen X, Qin Y, Zhang Z, Xing Z, Wang Q, Lu W *et al*. Hyper-SUMOylation of ERG Is Essential for the Progression of Acute Myeloid Leukemia. *Front Mol Biosci* 2021; 8: 652284.

72. Gâtel P, Brockly F, Reynes C, Pastore M, Hicheri Y, Cartron G *et al*. Ubiquitin and SUMO conjugation as biomarkers of acute myeloid leukemias response to chemotherapies. *Life Sci Alliance* 2020; 3.

73. Seeler JS, Dejean A. SUMO and the robustness of cancer. *Nat Rev Cancer* 2017; 17: 184-197.

74. Zhou P, Chen X, Li M, Tan J, Zhang Y, Yuan W *et al*. 2-D08 as a SUMOylation inhibitor induced ROS accumulation mediates apoptosis of acute myeloid leukemia cells possibly through the deSUMOylation of NOX2. *Biochem Biophys Res Commun* 2019; 513: 1063-1069.

75. Zhou Q, Zhang L, Chen Z, Zhao P, Ma Y, Yang B *et al*. Small ubiquitin-related modifier-1 modification regulates all-trans-retinoic acid-induced differentiation via stabilization of retinoic acid receptor α. *Febs j* 2014; 281: 3032-3047.

76. Daver N. Immune checkpoint inhibitors in acute myeloid leukemia. *Best Pract Res Clin Haematol* 2021; 34: 101247.

77. Reville PK, Kantarjian HM, Ravandi F, Jabbour E, DiNardo CD, Daver N *et al*. Nivolumab maintenance in high-risk acute myeloid leukemia patients: a single-arm, open-label, phase II study. *Blood Cancer J* 2021; 11: 60.

78. Berger R, Rotem-Yehudar R, Slama G, Landes S, Kneller A, Leiba M *et al*. Phase I safety and pharmacokinetic study of CT-011, a humanized antibody interacting with PD-1, in patients with advanced hematologic malignancies. *Clin Cancer Res* 2008; 14: 3044-3051.

79. Daver N, Boddu P, Garcia-Manero G, Yadav SS, Sharma P, Allison J *et al*. Hypomethylating agents in combination with immune checkpoint inhibitors in acute myeloid leukemia and myelodysplastic syndromes. *Leukemia* 2018; 32: 1094-1105.

80. Wang LX, Mei ZY, Zhou JH, Yao YS, Li YH, Xu YH *et al*. Low dose decitabine treatment induces CD80 expression in cancer cells and stimulates tumor specific cytotoxic T lymphocyte responses. *PLoS One* 2013; 8: e62924.

81. Sallman D, Asch A, Kambhampati S, Al Malki M, Zeidner J, Donnellan W *et al*. The first-in-class anti-CD47 antibody magrolimab combined with azacitidine is well-tolerated and effective in AML patients: phase 1b results.

Blood 2020; 136: 330.

82. Zeidan AM, Cavenagh J, Voso MT, Taussig D, Tormo M, Boss I *et al*. Efficacy and Safety of Azacitidine (AZA) in Combination with the Anti-PD-L1 Durvalumab (durva) for the Front-Line Treatment of Older Patients (pts) with Acute Myeloid Leukemia (AML) Who Are Unfit for Intensive Chemotherapy (IC) and Pts with Higher-Risk Myelodysplastic Syndromes (HR-MDS): Results from a Large, International, Randomized Phase 2 Study. *Blood* 2019; 134: 829.

83. Kapoor S, Champion G, Basu A, Mariampillai A, Olnes MJ. Immune Therapies for Myelodysplastic Syndromes and Acute Myeloid Leukemia. *Cancers (Basel)* 2021; 13.

84. Ma J, Li X, Su Y, Zhao J, Luedtke DA, Epshteyn V *et al*. Mechanisms responsible for the synergistic antileukemic interactions between ATR inhibition and cytarabine in acute myeloid leukemia cells. *Sci Rep* 2017; 7: 41950.

85. Ding X, Wang A, Ma X, Demarque M, Jin W, Xin H *et al*. Protein SUMOylation Is Required for Regulatory T Cell Expansion and Function. *Cell Rep* 2016; 16: 1055-1066.

86. Su A, Ling F, Vaganay C, Sodaro G, Benaksas C, Dal Bello R *et al*. The Folate Cycle Enzyme MTHFR Is a Critical Regulator of Cell Response to MYC-Targeting Therapies. *Cancer Discov* 2020; 10: 1894-1911.

87. Pikman Y, Puissant A, Alexe G, Furman A, Chen LM, Frumm SM *et al*. Targeting MTHFD2 in acute myeloid leukemia. *J Exp Med* 2016; 213: 1285-1306.

88. Bonagas N, Gustafsson NMS, Henriksson M, Marttila P, Gustafsson R, Wiita E *et al*. Pharmacological targeting of MTHFD2 suppresses acute myeloid leukemia by inducing thymidine depletion and replication stress. *Nat Cancer* 2022; 3: 156-172.

89. Organista-Nava J, Gómez-Gómez Y, Del Moral-Hernandez O, Illades-Aguiar B, Gómez-Santamaria J, Rivera-Ramírez AB *et al*. Deregulation of folate pathway gene expression correlates with poor prognosis in acute leukemia. *Oncol Lett* 2019; 18: 3115-3127.

90. National Comprehensive Cancer Network. Acute Myeloid Leukemia (Version 2.2022), 2022.

91. Rots MG, Pieters R, Jansen G, Kaspers GJ, Van Zantwijk CH, Noordhuis P *et al*. A possible role for methotrexate in the treatment of childhood acute myeloid leukaemia, in particular for acute monocytic leukaemia. *Eur J Cancer* 2001; 37: 492-498.

92. Fiancette R, Vincent C, Donnard M, Bordessoule D, Turlure P, Trimoreau F *et al*. Genes encoding multiple forms of phospholipase A(2) are expressed in immature forms of human leukemic blasts. *Leukemia* 2009; 23: 1196-1199.

93. Tcheng M, Roma A, Ahmed N, Smith RW, Jayanth P, Minden MD *et al*. Very long chain fatty acid metabolism is required in acute myeloid leukemia. *Blood* 2021; 137: 3518-3532.

94. Slagsvold JE, Pettersen CH, Follestad T, Krokan HE, Schønberg SA. The antiproliferative effect of EPA in HL60 cells is mediated by alterations in calcium homeostasis. *Lipids* 2009; 44: 103-113.

95. Yamagami T, Porada CD, Pardini RS, Zanjani ED, Almeida-Porada G. Docosahexaenoic acid induces dose dependent cell death in an early undifferentiated subtype of acute myeloid leukemia cell line. *Cancer Biol Ther* 2009; 8: 331-337.

96. Picou F, Debeissat C, Bourgeais J, Gallay N, Ferrié E, Foucault A *et al*. n-3 Polyunsaturated fatty acids induce acute myeloid leukemia cell death associated with mitochondrial glycolytic switch and Nrf2 pathway activation. *Pharmacol Res* 2018; 136: 45-55.

97. Southam AD, Khanim FL, Hayden RE, Constantinou JK, Koczula KM, Michell RH *et al*. Drug Redeployment to Kill Leukemia and Lymphoma Cells by Disrupting SCD1-Mediated Synthesis of Monounsaturated Fatty Acids. *Cancer Res* 2015; 75: 2530-2540.

98. Pabst T, Kortz L, Fiedler GM, Ceglarek U, Idle JR, Beyoğlu D. The plasma lipidome in acute myeloid leukemia at diagnosis in relation to clinical disease features. *BBA Clin* 2017; 7: 105-114.

99. Simonetti G, Mengucci C, Padella A, Fonzi E, Picone G, Delpino C *et al*. Integrated genomic-metabolic classification of acute myeloid leukemia defines a subgroup with NPM1 and cohesin/DNA damage mutations. *Leukemia* 2021; 35: 2813-2826.

100. Kramer MH, Zhang Q, Sprung R, Day RB, Erdmann-Gilmore P, Li Y *et al*. Proteomic and phosphoproteomic landscapes of acute myeloid leukemia. *Blood* 2022; 140: 1533-1548.

101. Niu J, Peng D, Liu L. Drug Resistance Mechanisms of Acute Myeloid Leukemia Stem Cells. *Front Oncol* 2022; 12: 896426.

102. Vetrie D, Helgason GV, Copland M. The leukaemia stem cell: similarities, differences and clinical prospects in CML and AML. *Nat Rev Cancer* 2020; 20: 158-173.

103. Lin JT, Tong WP, Trippett TM, Niedzwiecki D, Tao Y, Tan C *et al*. Basis for natural resistance to methotrexate in human acute non-lymphocytic leukemia. *Leuk Res* 1991; 15: 1191-1196.

104. Kufe DW, Munroe D, Herrick D, Egan E, Spriggs D. Effects of 1-beta-D-arabinofuranosylcytosine incorporation on eukaryotic DNA template function. *Mol Pharmacol* 1984; 26: 128-134.

105. Döhner H, Wei AH, Appelbaum FR, Craddock C, DiNardo CD, Dombret H *et al*. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. *Blood* 2022; 140: 1345-1377.

106. Jentzsch M, Bischof L, Ussmann J, Backhaus D, Brauer D, Metzeler KH *et al*. Prognostic impact of the AML ELN2022 risk classification in patients undergoing allogeneic stem cell transplantation. *Blood Cancer J* 2022; 12: 170.

국문초록

암 빅데이터 분석을 통한 급성 골수성 백혈병에서의 새로운 약물 타겟 발굴

김한선
천연물과학 전공
서울대학교 약학과

급성 골수성 백혈병 (AML)은 최근 새로운 치료법으로 표적 치료제와 항체 치료제 등을 도입했음에도 불구하고 일반적으로 예후가 만족스럽지 못하다. 또한, AML 환자들의 예후 위험도를 분류하기 위한 European LeukemiaNet (ELN) 2017 분류법이 널리 사용되고 있지만, 각 위험도 카테고리와 관련된 생물학적 경로에 관한 연구는 부족한 상태이며, 이로 인해 고위험 환자군에 대한 약물 치료 옵션에 대한 개선에도 도움이 되지 못하였다.

이러한 문제에 대처하기 위해, AML에서의 새로운 타겟 후보를 찾기 위해 암 빅데이터 분석을 이용하였다. 우선, 약물로 타겟 가능한 새로운 경로를 찾기 위해, 대규모 AML 데이터베이스인 Oregon Health & Science University (OHSU)와 Microarray Innovations in Leukemia

(MILE)를 활용하여 종합적인 생물정보학적 경로 스크리닝을 수행하였다. 이를 통해 SUMOylation 경로를 찾아내었으며, 이 결과를 외부 데이터베이스에서 독립적으로 검증하였다 (총 2959개의 AML 샘플과 642개의 정상 샘플 데이터 사용). AML에서의 SUMOylation의 임상적 연관성은 주요 유전자의 발현량과 환자의 생존율, ELN2017 위험도 분류, AML과 연관된 돌연변이들과의 관련성을 통해 확인되었다. 현재 고형암에서 임상 시험 중에 있는 새로운 SUMOylation 억제제인 TAK-981은 AML 세포에서 세포 사멸을 유도하고, 세포 주기를 정지시키며 분화 마커 유전자 발현을 유도함으로써 항백혈병 효과를 보였다. TAK-981은 강력한 nanomolar 활성을 나타냈고, 표준 치료제인 cytarabine보다도 강한 활성을 보이기도 하였다. TAK-981의 유용성은 환자 유래 primary AML 세포 및 *in vivo* 마우스 그리고 사람의 백혈병 모델에서도 확인되었다. 또한 이전 고형암 연구와는 다르게 IFN1 및 면역 의존적인 메커니즘 대신, TAK-981은 직접적이고 암세포에 고유한 항백혈병 효과를 보여주었다. 요약하면, AML에서 SUMOylation이 새로운 타겟 가능한 경로임을 확인하고, TAK-981을 유망하고 직접적인 항백혈병 약물로 제시하였다. 이러한 근거를 바탕으로, AML에서 약물 병용 투여 전략의 최적화에 대한 연구와 임상 시험을 기대하는 바이다.

또한, 최근의 AML 데이터베이스를 사용하여, ELN2017 위험도가 증가함에 따라 함께 증가하는 생물학적 경로를 조사하였다. 환자의 생존 분석과 다른 독립적인 전사체학, 단백질학의 AML 데이터베이스를 이용하여

필터링하고 검증한 결과, ‘불포화 지방산의 합성’과 ‘엽산 대사’ 경로가 후보로 도출되었다. 문헌 조사와 유전자 수준에서의 추가적인 조사를 통해, 고위험 환자군과 연관된 핵심 타겟으로 *SCD*와 *MTHFD2* 유전자가 확인되었다. SCD 억제제인 A939572와 MTHFD2 억제제인 DS18561882는 암 선택성을 보였고, 또한 관해 유도요법에 사용되는 표준 약인 cytarabine에 비교적 높은 IC_{50}를 가진 세포주들에서 cytarabine과 시너지를 보였다. *SCD* 유전자 발현은 불포화 지방산의 양과 상관성이 있는 것으로 나타났다. 요약하면, 본 연구에서 제시된 타겟들은 고위험 AML 환자군에서 보다 더 나은 치료 옵션을 찾거나 메커니즘의 이해를 위해 연구될 수 있으며, 이와 같은 연구 접근법은 다른 타겟 유전자나 경로를 찾거나 심지어 고형암에서도 용이하게 적용될 수 있다.

주요어: 급성 골수성 백혈병, SUMOylation, TAK-981, 면역 독립적, ELN2017, 고위험, SCD, MTHFD2

학번: 2017-20187

Acknowledgements

The work in Part I has been published in the following: Kim HS, Kim BR, Dao TTP, Kim JM, Kim YJ, Son H *et al*. TAK-981, a SUMOylation inhibitor, suppresses AML growth immune-independently. *Blood Adv* 2023; 7: 3155-3168. It was done in collaboration with Professor Byung-Sik Cho's lab from the Catholic University of Korea, and with the help of Thien T. P. Dao.

The work in Part II has been published in the following: Kim HS, Kim D, Kim J, Park S, de Guzman ACV. SCD and MTHFD2 inhibitors for high-risk acute myeloid leukaemia patients, as suggested by ELN2017-pathway association. *Clin Transl Med* 2023; 13: e1311.

I want to thank Professor Sunghyouk Park and the colleague lab members for helping me throughout the whole Ph.D. course. Also, I want to thank Professor Byung-Sik Cho and his lab members for the collaboration. Finally, I want to say thank you to my family.

Appendix

This appendix includes supporting information for the main text (A-D), followed by two of my published papers. I contributed one paper as the first author and the other as the co-first author. These works were done during my Ph.D. course under my supervisor, Professor Sunghyouk Park.

A. List of pathways and their sources in each cluster of Fig. 1B

Cluster	Pathway	Source
(i)	ACTIVATION.OF.THE.MRNA.UPON.BINDING.OF.THE.CAP_BINDING.COMPLEX.AND.EIFS_.AND.SUBSEQUENT.BINDING.TO.43S	REACTOME.DATABASE.ID.RELEASE.69
	AMIDE.BIOSYNTHETIC.PROCESS	GOBP
	AMINO.ACID.ACTIVATION	GOBP
	CAP_DEPENDENT.TRANSLATION.INITIATION	REACTOME.DATABASE.ID.RELEASE.69
	COTRANSLATIONAL.PROTEIN.TARGETING.TO.MEMBRANE	GOBP
	CYTOPLASMIC.RIBOSOMAL.PROTEINS	WIKIPATHWAYS_20190610
	CYTOPLASMIC.TRANSLATION	GOBP
	DEADENYLATION.OF.MRNA	REACTOME
	DEADENYLATION_DEPENDENT.MRNA.DECAY	REACTOME
	ESTABLISHMENT.OF.PROTEIN.LOCALIZATION.TO.ENDOPLASMIC.RETICULUM	GOBP
	ESTABLISHMENT.OF.PROTEIN.LOCALIZATION.TO.MEMBRANE	GOBP
	ESTABLISHMENT.OF.PROTEIN.LOCALIZATION.TO.ORGANELLE	GOBP
	ESTABLISHMENT.OF.RNA.LOCALIZATION	GOBP
	EUKARYOTIC.TRANSLATION.ELONGATION	REACTOME.DATABASE.ID.RELEASE.69
	EUKARYOTIC.TRANSLATION.INITIATION	REACTOME
	EUKARYOTIC.TRANSLATION.TERMINATION	REACTOME.DATABASE.ID.RELEASE.69
	FORMATION.OF.A.POOL.OF.FREE.40S.SUBUNITS	REACTOME.DATABASE.ID.RELEASE.69
	FORMATION.OF.THE.TERNARY.COMPLEX_.AND.SUBSEQUENTLY_.THE.43S.COMPLEX	REACTOME
	GTP.HYDROLYSIS.AND.JOINING.OF.THE.60S.RIBOSOMAL.SUBUNIT	REACTOME

	IMPORT.INTO.NUCLEUS	GOBP
	INFLUENZA.INFECTION	REACTOME.DATABASE.ID.RELEASE.69
	INFLUENZA.LIFE.CYCLE	REACTOME
	INFLUENZA.VIRAL.RNA.TRANSCRIPTION.AND.REPLICATION	REACTOME.DATABASE.ID.RELEASE.69
	L13A_MEDIATED.TRANSLATIONAL.SILENCING.OF.CERULOPLASMIN.EXPRESSION	REACTOME.DATABASE.ID.RELEASE.69
	MAJOR.PATHWAY.OF.RRNA.PROCESSING.IN.THE.NUCLEOLUS.AND.CYTOSOL	REACTOME
	MATURATION.OF.5_8S.RRNA	GOBP
	MATURATION.OF.5_8S.RRNA.FROM.TRICISTRONIC.RRNA.TRANSCRIPT._SSU_RRNA_.5_8S.RRNA_.LSU_RRNA_	GOBP
	MATURATION.OF.LSU_RRNA	GOBP
	MATURATION.OF.LSU_RRNA.FROM.TRICISTRONIC.RRNA.TRANSCRIPT._SSU_RRNA_.5_8S.RRNA_.LSU_RRNA_	GOBP
	MATURATION.OF.SSU_RRNA	GOBP
	MATURATION.OF.SSU_RRNA.FROM.TRICISTRONIC.RRNA.TRANSCRIPT._SSU_RRNA_.5_8S.RRNA_.LSU_RRNA_	GOBP
	MIRNA.METABOLIC.PROCESS	GOBP
	MITOCHONDRIAL.GENE.EXPRESSION	GOBP
	MITOCHONDRIAL.RNA.METABOLIC.PROCESS	GOBP
	MITOCHONDRIAL.TRANSLATION	GOBP
	MITOCHONDRIAL.TRANSLATION	REACTOME.DATABASE.ID.RELEASE.69
	MITOCHONDRIAL.TRANSLATION.ELONGATION	REACTOME.DATABASE.ID.RELEASE.69
	MITOCHONDRIAL.TRANSLATION.TERMINATION	REACTOME
	MITOCHONDRIAL.TRANSLATIONAL.TERMINATION	GOBP
	MITOCHONDRIAL.TRNA.AMINOACYLATION	REACTOME.DATABASE.ID.RELEASE.69
	MRNA.CATABOLIC.PROCESS	GOBP
	MRNA.EXPORT.FROM.NUCLEUS	GOBP
	MRNA.PROCESSING	WIKIPATHWAYS_20190610
	MRNA.PROCESSING	GOBP
	MRNA.SPLICING	REACTOME.DATABASE.ID.RELEASE.69
	MRNA.SPLICING_.VIA.SPLICEOSOME	GOBP
	MRNA.TRANSPORT	GOBP
	MRNA_CONTAINING.RIBONUCLEOPROTEIN.COMPLEX.EXPORT.FROM.NUCLEUS	GOBP

	NCRNA.3__END.PROCESSING	GOBP
	NCRNA.METABOLIC.PROCESS	GOBP
	NCRNA.PROCESSING	GOBP
	NONSENSE.MEDIATED.DECAY._NMD_.ENHANCED.BY.THE.EXON.JUNCTION.COMPLEX._EJC_	REACTOME
	NONSENSE.MEDIATED.DECAY._NMD_.INDEPENDENT.OF.THE.EXON.JUNCTION.COMPLEX._EJC_	REACTOME.DATABASE.ID.RELEASE.69
	NONSENSE_MEDIATED.DECAY._NMD_	REACTOME.DATABASE.ID.RELEASE.69
	NUCLEAR.TRANSPORT	GOBP
	NUCLEAR_TRANSCRIBED.MRNA.CATABOLIC.PROCESS	GOBP
	NUCLEAR_TRANSCRIBED.MRNA.CATABOLIC.PROCESS_.NONSENSE_MEDIATED.DECAY	GOBP
	NUCLEIC.ACID.TRANSPORT	GOBP
	NUCLEOBASE_CONTAINING.COMPOUND.CATABOLIC.PROCESS	GOBP
	NUCLEOBASE_CONTAINING.COMPOUND.TRANSPORT	GOBP
	NUCLEOCYTOPLASMIC.TRANSPORT	GOBP
	PEPTIDE.BIOSYNTHETIC.PROCESS	GOBP
	PEPTIDE.CHAIN.ELONGATION	REACTOME
	PEPTIDE.METABOLIC.PROCESS	GOBP
	PRE_MRNA.SPLICING	REACTOME
	PROCESSING.OF.CAPPED.INTRON_CONTAINING.PRE_MRNA	REACTOME.DATABASE.ID.RELEASE.69
	PRODUCTION.OF.MIRNAS.INVOLVED.IN.GENE.SILENCING.BY.MIRNA	GOBP
	PROTEIN.IMPORT	GOBP
	PROTEIN.LOCALIZATION.TO.ENDOPLASMIC.RETICULUM	GOBP
	PROTEIN.TARGETING	GOBP
	PROTEIN.TARGETING.TO.ER	GOBP
	PROTEIN.TARGETING.TO.MEMBRANE	GOBP
	REGULATION.OF.EXPRESSION.OF.SLITS.AND.ROBOS	REACTOME.DATABASE.ID.RELEASE.69
	RIBONUCLEOPROTEIN.COMPLEX.ASSEMBLY	GOBP
	RIBONUCLEOPROTEIN.COMPLEX.BIOGENESIS	GOBP
	RIBONUCLEOPROTEIN.COMPLEX.EXPORT.FROM.NUCLEUS	GOBP
	RIBONUCLEOPROTEIN.COMPLEX.LOCALIZATION	GOBP
	RIBONUCLEOPROTEIN.COMPLEX.SUBUNIT.ORGANIZATION	GOBP
	RIBOSOMAL.LARGE.SUBUNIT.ASSEMBLY	GOBP
	RIBOSOMAL.LARGE.SUBUNIT.BIOGENESIS	GOBP

	RIBOSOMAL.SCANNING.AND.START.CODON.RECOGNITION	REACTOME.DATABASE.ID.RELEASE.69
	RIBOSOMAL.SMALL.SUBUNIT.BIOGENESIS	GOBP
	RIBOSOME.ASSEMBLY	GOBP
	RIBOSOME.BIOGENESIS	GOBP
	RNA.3__END.PROCESSING	GOBP
	RNA.CATABOLIC.PROCESS	GOBP
	RNA.EXPORT.FROM.NUCLEUS	GOBP
	RNA.LOCALIZATION	GOBP
	RNA.METHYLATION	GOBP
	RNA.POLYMERASE.II.TRANSCRIPTION.TERMINATION	REACTOME.DATABASE.ID.RELEASE.69
	RNA.SPLICING	GOBP
	RNA.SPLICING_.VIA.TRANSESTERIFICATION.REACTIONS	GOBP
	RNA.SPLICING_.VIA.TRANSESTERIFICATION.REACTIONS.WITH.BULGED.ADENOSINE.AS.NUCLEOPHILE	GOBP
	RNA.TRANSPORT	GOBP
	RRNA.METABOLIC.PROCESS	GOBP
	RRNA.MODIFICATION.IN.THE.NUCLEUS.AND.CYTOSOL	REACTOME
	RRNA.PROCESSING	REACTOME
	RRNA.PROCESSING	GOBP
	RRNA.PROCESSING.IN.THE.NUCLEUS.AND.CYTOSOL	REACTOME.DATABASE.ID.RELEASE.69
	SELENOAMINO.ACID.METABOLISM	REACTOME.DATABASE.ID.RELEASE.69
	SELENOCYSTEINE.SYNTHESIS	REACTOME.DATABASE.ID.RELEASE.69
	SIGNALING.BY.ROBO.RECEPTORS	REACTOME
	SPLICEOSOMAL.SNRNP.ASSEMBLY	GOBP
	SRP_DEPENDENT.COTRANSLATIONAL.PROTEIN.TARGETING.TO.MEMBRANE	GOBP
	SRP_DEPENDENT.COTRANSLATIONAL.PROTEIN.TARGETING.TO.MEMBRANE	REACTOME.DATABASE.ID.RELEASE.69
	TRANSLATION	GOBP
	TRANSLATION	REACTOME
	TRANSLATION.FACTORS	WIKIPATHWAYS_20190610
	TRANSLATION.INITIATION.COMPLEX.FORMATION	REACTOME.DATABASE.ID.RELEASE.69
	TRANSLATIONAL.ELONGATION	GOBP

	TRANSLATIONAL.INITIATION	GOBP
	TRNA.AMINOACYLATION	GOBP
	TRNA.METABOLIC.PROCESS	GOBP
	VIRAL.GENE.EXPRESSION	GOBP
	VIRAL.MRNA.TRANSLATION	REACTOME
	VIRAL.TRANSCRIPTION	GOBP
(ii)	ACTIVATED.PKN1.STIMULATES.TRANSCRIPTION.OF.AR._ANDROGEN.RECEPTOR_.REGULATED.GENES.KLK2.AND.KLK3	REACTOME
	ACTIVATION.OF.ANTERIOR.HOX.GENES.IN.HINDBRAIN.DEVELOPMENT.DURING.EARLY.EMBRYOGENESIS	REACTOME
	ACTIVATION.OF.HOX.GENES.DURING.DIFFERENTIATION	REACTOME
	ACTIVATION.OF.RRNA.EXPRESSION.BY.ERCC6._CSB_.AND.EHMT2._G9A_	REACTOME.DATABASE.ID.RELEASE.69
	B_WICH.COMPLEX.POSITIVELY.REGULATES.RRNA.EXPRESSION	REACTOME.DATABASE.ID.RELEASE.69
	DNA.METHYLATION	REACTOME
	EPIGENETIC.REGULATION.OF.GENE.EXPRESSION	REACTOME
	HDACS.DEACETYLATE.HISTONES	REACTOME
	NEGATIVE.EPIGENETIC.REGULATION.OF.RRNA.EXPRESSION	REACTOME
	NORC.NEGATIVELY.REGULATES.RRNA.EXPRESSION	REACTOME
	OXIDATIVE.STRESS.INDUCED.SENESCENCE	REACTOME.DATABASE.ID.RELEASE.69
	PACKAGING.OF.TELOMERE.ENDS	REACTOME
	POSITIVE.EPIGENETIC.REGULATION.OF.RRNA.EXPRESSION	REACTOME.DATABASE.ID.RELEASE.69
	PRC2.METHYLATES.HISTONES.AND.DNA	REACTOME.DATABASE.ID.RELEASE.69
	RNA.POLYMERASE.I.PROMOTER.CLEARANCE	REACTOME
	RNA.POLYMERASE.I.PROMOTER.ESCAPE	REACTOME.DATABASE.ID.RELEASE.69
	RNA.POLYMERASE.I.TRANSCRIPTION	REACTOME
(iii)	SUMOYLATION.OF.CHROMATIN.ORGANIZATION.PROTEINS	REACTOME
	SUMOYLATION.OF.RNA.BINDING.PROTEINS	REACTOME
	SUMOYLATION.OF.TRANSCRIPTION.COFACTORS	REACTOME.DATABASE.ID.RELEASE.69
(iv)	REGULATION.OF.MRNA.METABOLIC.PROCESS	GOBP
	REGULATION.OF.MRNA.POLYADENYLATION	GOBP
	REGULATION.OF.MRNA.PROCESSING	GOBP

GOBP, Gene Ontology Biological Process.

B. List of genes used in the clustering in Fig. 1B and their expression differences between AML and normal samples in OHSU and MILE databases

Cluster	Gene	OHSU		MILE	
		Median_diff	***P*-value**	**Median_diff**	***P*-value**
(i)	*AAMP*	0.282	< 0.001	0.038	< 0.001
	AAR2	0.399	0.010	0.043	< 0.001
	AARS2	0.045	0.250	0.028	0.023
	ACE	0.245	0.324	0.036	< 0.001
	ACSF3	0.088	0.010	0.034	< 0.001
	ADA	1.215	< 0.001	0.133	< 0.001
	ADPRM	0.396	0.017	0.022	< 0.001
	AGXT	0.614	< 0.001	0.030	0.002
	AIMP1	0.587	< 0.001	0.035	< 0.001
	AMN	1.578	< 0.001	0.095	< 0.001
	ANKRD16	0.463	0.002	0.018	< 0.001
	AP3M1	-0.039	0.902	0.025	< 0.001
	AP4M1	0.578	< 0.001	0.047	< 0.001
	AP4S1	0.295	0.001	0.032	0.001
	B4GALT6	0.556	0.007	0.066	< 0.001
	BAD	0.560	< 0.001	0.033	0.026
	BDH2	0.357	0.002	0.044	< 0.001
	BICD1	0.122	0.247	0.042	< 0.001
	BMF	-0.074	0.989	0.031	< 0.001
	BMS1	-0.044	0.359	0.052	< 0.001
	BOP1	0.423	< 0.001	0.048	< 0.001
	BRF1	0.165	0.024	0.025	0.021
	CALR	0.328	0.022	0.044	< 0.001
	CCAR2	0.218	< 0.001	0.039	< 0.001
	CCNT1	0.054	0.938	0.076	< 0.001
	CDK9	0.451	< 0.001	0.038	< 0.001
	CELF1	-0.203	0.353	0.016	0.012
	CERS1	0.534	0.004	0.026	< 0.001
	CERS6	0.948	< 0.001	0.077	< 0.001
	CLASRP	-0.108	0.202	0.020	< 0.001
	CLN8	0.189	0.077	0.028	0.006

	CLNS1A	0.410	< 0.001	0.056	< 0.001
	CLP1	0.341	0.115	0.048	< 0.001
	CLUH	-0.063	0.665	0.038	< 0.001
	CNOT7	0.202	0.002	0.033	< 0.001
	COL4A5	1.445	0.350	0.030	0.015
	CPSF1	0.322	< 0.001	0.023	0.003
	CPSF6	0.199	0.047	0.032	< 0.001
	CPSF7	0.027	0.524	0.023	< 0.001
	CPXM1	1.946	< 0.001	0.184	< 0.001
	CSNK1E	0.560	0.003	0.045	< 0.001
	CT45A1	0.000	0.017	0.065	< 0.001
	DARS	0.526	< 0.001	0.035	< 0.001
	DCAF13	0.003	0.867	0.019	0.003
	DCTPP1	0.731	< 0.001	0.046	< 0.001
	DDO	1.891	< 0.001	0.042	< 0.001
	DDX17	-0.560	0.003	0.011	0.294
	DDX21	0.284	0.016	0.019	< 0.001
	DDX28	0.214	0.429	0.032	0.020
	DDX31	0.275	0.031	0.051	< 0.001
	DDX42	-0.045	0.511	0.025	0.013
	DDX49	0.462	< 0.001	0.024	0.009
	DDX5	0.148	0.014	0.037	< 0.001
	DDX52	0.288	< 0.001	0.052	< 0.001
	DEGS2	1.015	< 0.001	0.040	0.001
	DENR	0.641	< 0.001	0.067	< 0.001
	DERA	0.421	< 0.001	0.035	< 0.001
	DHPS	0.204	< 0.001	0.028	< 0.001
	DHRS4	0.569	< 0.001	0.045	< 0.001
	DHX37	0.326	< 0.001	0.031	< 0.001
	DIS3	0.096	0.501	0.025	< 0.001
	DIS3L	0.203	0.005	0.028	< 0.001
	DKC1	0.285	0.022	0.015	< 0.001
	DNPEP	0.508	< 0.001	0.025	< 0.001
	DTD2	-0.074	0.895	0.026	< 0.001
	DYNC1H1	0.186	0.022	0.027	0.007

	EDC3	0.138	0.223	0.012	0.028
	EDC4	0.147	0.001	0.026	< 0.001
	EEF1A1	0.888	< 0.001	0.040	< 0.001
	EEF2K	0.186	0.001	0.020	< 0.001
	EFTUD2	0.218	0.002	0.032	0.038
	EIF1AX	0.208	0.003	0.039	< 0.001
	EIF2A	0.246	0.010	0.031	< 0.001
	EIF2S1	0.231	0.189	0.011	0.005
	EIF2S2	0.068	0.251	0.027	0.002
	EIF2S3	0.443	< 0.001	0.042	< 0.001
	EIF3C	0.466	< 0.001	0.020	0.010
	EIF3D	0.401	< 0.001	0.032	< 0.001
	EIF3E	0.580	< 0.001	0.022	< 0.001
	EIF3G	0.066	0.433	0.038	< 0.001
	EIF3H	0.373	0.001	0.078	< 0.001
	EIF3I	0.506	< 0.001	0.025	< 0.001
	EIF3L	0.681	< 0.001	0.014	< 0.001
	EIF3M	0.445	< 0.001	0.051	< 0.001
	EIF4A1	0.670	< 0.001	0.028	< 0.001
	EIF4B	0.459	< 0.001	0.019	< 0.001
	EIF4EBP1	0.297	0.025	0.030	< 0.001
	ELAC2	0.129	0.027	0.024	0.024
	ELAVL1	0.082	0.149	0.028	0.001
	ELP3	0.135	0.188	0.021	< 0.001
	ENDOG	0.489	0.002	0.036	< 0.001
	ENO1	1.075	< 0.001	0.024	0.006
	ENY2	0.337	0.010	0.019	0.061
	ERCC2	0.187	0.055	0.022	< 0.001
	ERI3	0.360	< 0.001	0.045	< 0.001
	ETF1	0.583	0.001	0.018	0.001
	EXOSC2	0.072	0.210	0.039	0.001
	EXOSC5	0.495	0.004	0.046	< 0.001
	EXOSC6	0.222	0.020	0.062	< 0.001
	FAM98B	0.101	0.326	0.012	0.045
	FARS2	0.369	< 0.001	0.014	0.053

	FARSA	0.199	0.139	0.031	< 0.001
	FASTKD1	0.464	0.003	0.027	0.002
	FASTKD5	0.179	0.110	0.030	0.013
	FPGS	0.510	< 0.001	0.072	< 0.001
	FURIN	0.414	0.003	0.012	0.098
	FUS	0.129	0.109	0.036	0.003
	FYTTD1	0.466	< 0.001	0.048	< 0.001
	GATB	0.280	0.065	0.022	< 0.001
	GEMIN8	0.365	< 0.001	0.032	< 0.001
	GET4	0.234	0.079	0.038	0.002
	GGA1	0.076	0.429	0.031	< 0.001
	GNMT	0.438	0.006	0.022	0.006
	GNPTAB	0.574	< 0.001	0.052	< 0.001
	GOLPH3	0.533	< 0.001	0.031	< 0.001
	GRSF1	0.420	< 0.001	0.014	< 0.001
	GSS	0.404	< 0.001	0.042	< 0.001
	GSTK1	0.152	0.017	0.041	< 0.001
	GSTP1	0.736	< 0.001	0.020	< 0.001
	GSTZ1	0.397	< 0.001	0.018	< 0.001
	GTF3A	0.823	< 0.001	0.037	< 0.001
	GTF3C3	0.239	0.001	0.025	0.001
	GTF3C4	0.452	0.001	0.039	< 0.001
	GTF3C6	0.394	< 0.001	0.030	< 0.001
	HACD1	1.408	< 0.001	0.095	< 0.001
	HENMT1	0.569	0.001	0.026	0.001
	HHEX	1.048	< 0.001	0.088	< 0.001
	HINT1	0.822	< 0.001	0.071	< 0.001
	HM13	0.002	0.982	0.020	0.038
	HNRNPA0	0.690	< 0.001	0.046	< 0.001
	HNRNPA1	0.593	< 0.001	0.032	< 0.001
	HNRNPC	0.280	0.002	0.015	0.014
	HNRNPL	0.323	0.003	0.042	< 0.001
	HNRNPM	0.399	< 0.001	0.020	0.002
	HOMER3	1.366	< 0.001	0.136	< 0.001
	HOXA2	1.634	< 0.001	0.090	< 0.001

	HPGDS	-0.014	0.679	0.041	0.007
	HSP90AB1	0.908	< 0.001	0.053	< 0.001
	HSPA14	0.499	< 0.001	0.033	0.003
	HSPA5	0.680	0.003	0.020	0.007
	HSPD1	0.550	< 0.001	0.039	0.004
	IARS2	0.500	< 0.001	0.016	0.004
	IDH1	0.201	0.077	0.051	< 0.001
	IMMP2L	0.849	< 0.001	0.053	< 0.001
	INTS8	0.286	< 0.001	0.034	< 0.001
	IPO4	0.353	0.001	0.034	< 0.001
	IPO5	0.507	< 0.001	0.032	< 0.001
	IPO7	0.336	0.003	0.029	< 0.001
	ITPA	0.427	< 0.001	0.053	< 0.001
	KARS	0.338	< 0.001	0.017	< 0.001
	KDELR1	0.452	< 0.001	0.073	< 0.001
	KIN	-0.016	0.646	0.010	0.018
	KPNA3	0.371	0.001	0.016	< 0.001
	KPNA5	0.336	0.008	0.064	< 0.001
	KPNB1	0.340	< 0.001	0.050	< 0.001
	KRI1	-0.297	0.052	0.027	< 0.001
	KTI12	0.593	< 0.001	0.048	< 0.001
	LAMP1	0.761	< 0.001	0.034	< 0.001
	LAMP2	0.694	< 0.001	0.017	0.017
	LAS1L	-0.024	0.907	0.047	< 0.001
	LSM1	0.138	0.101	0.022	< 0.001
	LSM2	0.491	< 0.001	0.046	< 0.001
	LSM3	0.654	< 0.001	0.044	< 0.001
	LTV1	-0.159	0.060	0.037	< 0.001
	LUC7L	-0.216	0.094	0.035	< 0.001
	MAEL	0.541	0.010	0.031	< 0.001
	MCTS1	0.408	< 0.001	0.007	0.101
	MDN1	-0.212	0.297	0.016	< 0.001
	MGST1	0.570	0.012	0.032	< 0.001
	MICALL1	0.234	0.117	0.064	< 0.001
	MRPL15	0.416	0.003	0.030	< 0.001

	MRPL21	0.346	0.012	0.019	< 0.001
	MRPL23	0.558	0.005	0.031	0.008
	MRPL30	0.332	0.001	0.031	< 0.001
	MRPL33	0.805	< 0.001	0.037	< 0.001
	MRPL38	0.043	0.235	0.012	0.060
	MRPL4	0.191	0.146	0.021	< 0.001
	MRPL42	-0.051	0.926	0.023	< 0.001
	MRPL48	0.104	0.044	0.018	< 0.001
	MRPL54	0.353	0.001	0.021	< 0.001
	MRPS2	0.064	0.470	0.032	< 0.001
	MRPS27	0.516	< 0.001	0.017	0.029
	MRPS30	0.373	< 0.001	0.016	0.023
	MRPS33	0.571	< 0.001	0.009	0.181
	MRPS35	0.346	0.001	0.031	< 0.001
	MRPS6	-0.098	0.596	0.009	0.011
	MRRF	0.035	0.320	0.019	0.026
	MTPAP	0.293	0.002	0.040	< 0.001
	MTRF1L	0.088	0.695	0.085	< 0.001
	NARS	-0.118	0.766	0.013	0.002
	NARS2	0.797	< 0.001	0.033	< 0.001
	NAT10	-0.056	0.698	0.036	< 0.001
	NCBP2	0.211	0.008	0.056	< 0.001
	NDUFA7	0.212	0.029	0.023	0.039
	NEIL2	0.155	0.271	0.031	0.001
	NLE1	0.328	0.016	0.037	< 0.001
	NOA1	0.617	< 0.001	0.039	< 0.001
	NOB1	0.500	0.001	0.095	< 0.001
	NOC4L	0.057	0.693	0.025	< 0.001
	NOL8	-0.016	0.769	0.017	< 0.001
	NOLC1	-0.018	0.717	0.031	< 0.001
	NOP2	-0.062	0.481	0.025	< 0.001
	NPEPPS	0.151	0.005	0.022	0.048
	NSUN4	0.580	< 0.001	0.033	0.002
	NT5C	0.561	< 0.001	0.045	< 0.001
	NTHL1	0.691	< 0.001	0.078	< 0.001

	NTN1	1.500	< 0.001	0.036	0.001
	NUDT15	0.563	0.001	0.068	< 0.001
	NUDT5	0.523	< 0.001	0.082	< 0.001
	NUP54	0.389	0.002	0.040	0.005
	NUTF2	0.382	< 0.001	0.035	< 0.001
	NXT1	0.849	< 0.001	0.040	< 0.001
	PABPN1	0.325	< 0.001	0.030	< 0.001
	PAK6	0.503	0.002	0.078	< 0.001
	PAN3	0.433	< 0.001	0.052	< 0.001
	PARN	0.245	0.001	0.031	< 0.001
	PDCD11	0.081	0.210	0.028	< 0.001
	PDCD7	0.246	0.068	0.014	0.058
	PES1	0.053	1.000	0.045	< 0.001
	PEX26	0.082	0.010	0.030	0.021
	PEX3	0.110	0.329	0.044	< 0.001
	PEX6	0.469	0.001	0.059	< 0.001
	PFKP	1.148	< 0.001	0.152	< 0.001
	PHYH	0.950	< 0.001	0.051	< 0.001
	PIH1D2	0.430	< 0.001	0.027	< 0.001
	PITRM1	0.220	0.001	0.029	< 0.001
	PMPCA	0.380	< 0.001	0.029	< 0.001
	PNN	-0.278	0.213	0.031	< 0.001
	PPIL4	0.114	0.051	0.021	0.001
	PPP2R1A	0.483	< 0.001	0.021	< 0.001
	PRICKLE1	0.781	0.039	0.073	< 0.001
	PRKACB	0.260	0.015	0.069	< 0.001
	PRKCI	0.309	0.002	0.048	< 0.001
	PRKRA	0.538	< 0.001	0.061	< 0.001
	PRMT5	0.358	0.006	0.035	< 0.001
	PRMT7	0.030	0.477	0.028	< 0.001
	PRPF4	0.232	0.054	0.031	< 0.001
	PSMB1	0.591	< 0.001	0.054	< 0.001
	PSMB10	0.494	< 0.001	0.026	0.001
	PSMB5	0.533	< 0.001	0.019	0.002
	PTGES3	0.562	< 0.001	0.015	< 0.001

	PTTG1IP	0.728	< 0.001	0.037	< 0.001
	PUS1	0.123	0.120	0.023	< 0.001
	QARS	0.576	< 0.001	0.028	< 0.001
	RAB34	1.339	< 0.001	0.134	< 0.001
	RAN	0.118	0.234	0.021	< 0.001
	RANBP6	0.768	< 0.001	0.037	< 0.001
	RBM15B	0.430	< 0.001	0.032	< 0.001
	RBM28	-0.194	0.443	0.023	< 0.001
	RBM39	0.448	< 0.001	0.043	< 0.001
	RBM4	0.445	< 0.001	0.023	< 0.001
	RBMX	0.112	0.316	0.021	0.003
	RBMX2	0.134	0.212	0.022	< 0.001
	REXO4	0.515	< 0.001	0.051	< 0.001
	RHOA	0.497	< 0.001	0.024	0.032
	RNASEH2B	0.519	< 0.001	0.035	< 0.001
	ROBO2	0.000	0.199	0.013	0.090
	RPF2	-0.033	0.731	0.036	< 0.001
	RPL10A	0.995	< 0.001	0.010	0.018
	RPL13	0.846	< 0.001	0.009	0.130
	RPL14	0.476	0.008	0.051	< 0.001
	RPL15	0.842	< 0.001	0.052	< 0.001
	RPL17	0.068	0.436	0.034	< 0.001
	RPL18	0.475	0.001	0.013	0.001
	RPL18A	0.448	0.009	0.045	< 0.001
	RPL22	0.665	< 0.001	0.049	< 0.001
	RPL27	0.462	0.001	0.031	0.002
	RPL27A	0.401	0.006	0.041	< 0.001
	RPL28	-0.046	0.421	0.075	< 0.001
	RPL31	0.942	< 0.001	0.058	< 0.001
	RPL37A	0.679	< 0.001	0.033	< 0.001
	RPL7	0.949	< 0.001	0.027	< 0.001
	RPP40	0.770	0.001	0.066	< 0.001
	RPS11	0.294	0.145	0.082	< 0.001
	RPS15A	0.625	< 0.001	0.023	0.001
	RPS16	0.670	< 0.001	0.020	< 0.001

	RPS19	1.149	< 0.001	0.047	< 0.001
	RPS20	0.605	< 0.001	0.045	< 0.001
	RPS21	1.090	< 0.001	0.046	< 0.001
	RPS23	1.035	< 0.001	0.085	< 0.001
	RPS24	0.741	< 0.001	0.043	< 0.001
	RPS27	0.467	0.022	0.032	0.145
	RPS27L	0.884	< 0.001	0.032	< 0.001
	RPS28	0.429	0.004	0.025	0.002
	RPS29	0.498	< 0.001	0.073	< 0.001
	RPS6KA3	0.018	0.196	0.049	< 0.001
	RRP1B	-0.189	0.055	0.014	0.012
	RRP7A	0.030	0.682	0.031	0.027
	RSL1D1	0.597	< 0.001	0.048	< 0.001
	RSL24D1	1.083	< 0.001	0.025	< 0.001
	RTCB	0.324	< 0.001	0.018	< 0.001
	SAGE1	-1.208	0.194	0.028	< 0.001
	SARS2	0.244	0.009	0.021	< 0.001
	SART1	0.124	0.426	0.032	< 0.001
	SCAF8	0.089	0.427	0.018	< 0.001
	SCARB2	0.375	< 0.001	0.025	0.013
	SCRIB	0.310	0.003	0.042	< 0.001
	SEC61A2	0.134	0.132	0.067	< 0.001
	SEC61B	0.553	< 0.001	0.054	< 0.001
	SF3B3	-0.093	0.429	0.027	< 0.001
	SLC25A19	0.229	0.134	0.029	< 0.001
	SLC25A33	0.921	< 0.001	0.078	< 0.001
	SLC25A36	0.679	< 0.001	0.064	< 0.001
	SLC35B2	0.554	< 0.001	0.036	< 0.001
	SLC35B4	0.018	0.536	0.106	< 0.001
	SLC35C2	-0.121	0.336	0.047	< 0.001
	SMG1	-0.087	0.905	0.026	< 0.001
	SMG5	0.381	< 0.001	0.047	< 0.001
	SMG6	0.176	0.008	0.025	0.004
	SMNDC1	0.366	< 0.001	0.016	< 0.001
	SNAPC1	0.177	0.058	0.027	< 0.001

	SNAPC2	0.464	< 0.001	0.034	< 0.001
	SNRNP70	-0.351	0.013	0.096	< 0.001
	SNRPA	0.112	0.518	0.042	< 0.001
	SNRPA1	0.014	0.971	0.019	0.010
	SNRPD2	0.133	0.304	0.015	< 0.001
	SNRPE	0.952	< 0.001	0.077	< 0.001
	SNX33	0.794	0.002	0.038	< 0.001
	SPIN1	0.591	< 0.001	0.060	< 0.001
	SPOP	0.140	0.107	0.025	0.097
	SPPL3	0.288	0.002	0.019	< 0.001
	SPTLC2	0.360	0.001	0.031	< 0.001
	SRBD1	0.332	0.001	0.038	< 0.001
	SRSF12	0.697	0.002	0.028	< 0.001
	STRAP	0.360	0.006	0.022	0.040
	SUGP2	0.033	0.046	0.018	< 0.001
	SYMPK	-0.130	0.163	0.067	< 0.001
	SYNCRIP	0.102	0.192	0.030	< 0.001
	TAF11	0.272	< 0.001	0.028	0.094
	TARBP1	0.501	0.001	0.082	< 0.001
	TARS	0.497	< 0.001	0.020	0.023
	TARSL2	-0.043	0.840	0.044	< 0.001
	TCP1	0.475	< 0.001	0.052	< 0.001
	TEFM	0.229	0.014	0.031	0.018
	TEX10	0.186	0.023	0.071	< 0.001
	TFAM	0.333	0.007	0.040	< 0.001
	TFB1M	0.137	0.018	0.033	< 0.001
	THOC6	0.416	< 0.001	0.020	< 0.001
	THUMPD1	-0.011	0.945	0.021	< 0.001
	THUMPD2	0.336	0.090	0.035	< 0.001
	TIMM44	0.262	0.001	0.029	< 0.001
	TIMM50	0.333	0.003	0.082	< 0.001
	TIMM9	0.629	< 0.001	0.043	< 0.001
	TMED10	0.033	0.956	-0.006	0.417
	TOE1	0.351	0.002	0.032	< 0.001
	TOMM34	0.331	0.009	0.030	< 0.001

	TPI1	0.654	< 0.001	0.067	< 0.001
	TRA2A	-0.009	0.780	0.028	0.195
	TRA2B	0.755	< 0.001	0.027	0.007
	TRIM71	1.839	< 0.001	0.160	< 0.001
	TRMT10B	-0.171	0.131	0.017	0.011
	TRMT44	0.145	0.004	0.036	< 0.001
	TRMT61A	0.308	0.018	0.025	< 0.001
	TRNT1	0.090	0.149	0.013	0.006
	TSC2	0.009	0.447	0.030	0.001
	TSR1	0.146	0.092	0.058	< 0.001
	TUFM	0.392	< 0.001	0.034	< 0.001
	TXNL4A	0.253	0.004	0.034	0.002
	TYW3	0.244	0.002	0.052	< 0.001
	U2AF1	-0.114	0.975	0.019	0.018
	U2AF1L4	-0.111	0.608	0.047	< 0.001
	UPP2	0.248	0.051	0.036	0.002
	UQCRC2	0.495	< 0.001	0.044	< 0.001
	URB1	0.261	0.063	0.028	< 0.001
	USPL1	0.250	0.093	0.026	0.015
	UTP14A	0.072	0.485	0.037	< 0.001
	UTP6	0.297	0.033	0.023	0.005
	VARS2	0.122	0.018	0.024	< 0.001
	WARS2	0.476	< 0.001	0.018	0.008
	WBP11	-0.043	0.288	0.016	0.003
	WDR12	0.206	0.065	0.028	< 0.001
	WDR46	0.378	< 0.001	0.027	< 0.001
	XAB2	0.259	0.001	0.014	< 0.001
	XPNPEP1	0.534	< 0.001	0.024	0.021
	XPOT	0.094	0.313	0.014	0.003
	YARS	0.365	< 0.001	0.010	0.013
	YARS2	0.410	< 0.001	0.031	< 0.001
	YTHDF2	0.744	< 0.001	0.037	< 0.001
	ZC3H15	0.359	< 0.001	0.022	0.055
	ZDHHC2	0.589	< 0.001	0.030	0.034
	ZDHHC21	0.121	0.056	0.038	< 0.001

	ZDHHC4	1.002	< 0.001	0.041	< 0.001
	ZFAND1	0.365	0.001	0.042	< 0.001
	ZNF593	0.491	< 0.001	0.068	< 0.001
	ZNF598	0.652	< 0.001	0.062	< 0.001
	ZPR1	0.220	0.043	0.050	< 0.001
(ii)	*AEBP2*	0.133	0.152	0.039	0.001
	AR	-0.432	0.645	0.043	< 0.001
	BMI1	0.868	< 0.001	0.067	< 0.001
	CBX2	0.912	0.004	0.024	0.001
	CBX3	0.560	< 0.001	0.021	0.002
	CDK6	1.467	< 0.001	0.129	< 0.001
	DDX21	0.284	0.016	0.019	< 0.001
	ERCC2	0.187	0.055	0.022	< 0.001
	H3F3B	1.439	< 0.001	0.049	< 0.001
	HIST1H2AD	2.497	< 0.001	0.045	0.002
	HIST1H2AE	2.572	< 0.001	0.033	0.073
	HIST1H2AG	0.930	< 0.001	0.035	0.001
	HIST1H2AM	1.241	< 0.001	0.051	< 0.001
	HIST1H2BB	0.675	< 0.001	0.006	0.137
	HIST1H2BC	2.798	< 0.001	0.093	< 0.001
	HIST1H2BG	2.200	< 0.001	0.055	0.032
	HIST1H2BK	1.833	< 0.001	0.094	< 0.001
	HIST1H2BL	1.043	< 0.001	0.032	< 0.001
	HIST1H2BN	1.636	< 0.001	0.022	0.024
	HIST1H2BO	1.089	< 0.001	0.038	0.033
	HIST1H3A	2.214	< 0.001	0.043	0.002
	HIST1H4A	1.306	< 0.001	0.024	0.085
	HIST1H4D	0.645	< 0.001	0.026	0.002
	HIST1H4E	2.407	< 0.001	0.026	0.008
	HOXA2	1.634	< 0.001	0.090	< 0.001
	HOXA3	2.346	0.002	0.164	< 0.001
	HOXA4	1.626	< 0.001	0.162	< 0.001
	HOXB3	0.844	0.073	0.135	< 0.001
	HOXB4	0.829	0.108	0.071	< 0.001
	IFNB1	0.697	< 0.001	0.013	0.031

	KAT2A	0.607	< 0.001	0.096	< 0.001
	MAP2K7	0.064	0.224	0.036	< 0.001
	MAPK11	1.335	< 0.001	0.032	< 0.001
	MAPK8	-0.048	0.858	0.017	0.001
	MBD3	0.470	< 0.001	0.038	< 0.001
	MEIS1	2.064	< 0.001	0.128	< 0.001
	PHC1	0.486	0.008	0.129	< 0.001
	PHC2	0.525	< 0.001	0.099	< 0.001
	PHF19	-0.018	0.412	0.017	0.031
	PKN1	0.662	< 0.001	0.049	< 0.001
	POLR1E	0.447	< 0.001	0.018	0.097
	RRN3	0.505	< 0.001	0.014	0.120
	SAP18	0.601	< 0.001	0.023	< 0.001
	SIN3A	-0.015	0.733	0.024	0.015
	TAF1D	0.410	< 0.001	0.038	< 0.001
	TERF2	0.221	0.049	0.006	0.141
	TWISTNB	0.713	< 0.001	0.092	< 0.001
	TXN	0.811	< 0.001	0.025	< 0.001
	YY1	0.036	0.769	0.036	< 0.001
(iii)	*BMI1*	0.868	< 0.001	0.067	< 0.001
	CBX2	0.912	0.004	0.024	0.001
	DDX5	0.148	0.014	0.037	< 0.001
	HIST1H4E	2.407	< 0.001	0.026	0.008
	NUP54	0.389	0.002	0.040	0.005
	PHC1	0.486	0.008	0.129	< 0.001
	PHC2	0.525	< 0.001	0.099	< 0.001
	RING1	0.219	0.001	0.036	< 0.001
	SUMO1	0.405	< 0.001	0.037	< 0.001
	SUMO2	0.190	0.077	0.034	< 0.001
	TRIM28	0.663	< 0.001	0.042	< 0.001
	UBE2I	0.091	0.127	0.044	0.001
	ZBED1	0.630	< 0.001	0.035	< 0.001
(iv)	*CCNT1*	0.054	0.938	0.076	< 0.001
	CDK9	0.451	< 0.001	0.038	< 0.001
	CNOT7	0.202	0.002	0.033	< 0.001

	CPSF6	0.199	0.047	0.032	< 0.001
	CPSF7	0.027	0.524	0.023	< 0.001
	CTR9	-0.063	0.633	0.018	0.001
	DDX17	-0.560	0.003	0.011	0.294
	DDX5	0.148	0.014	0.037	< 0.001
	DIS3	0.096	0.501	0.025	< 0.001
	ELAVL1	0.082	0.149	0.028	0.001
	EXOSC2	0.072	0.210	0.039	0.001
	EXOSC5	0.495	0.004	0.046	< 0.001
	EXOSC6	0.222	0.020	0.062	< 0.001
	FASTKD1	0.464	0.003	0.027	0.002
	FTO	-0.232	0.371	0.033	< 0.001
	FXR1	0.648	< 0.001	0.030	0.014
	HNRNPA0	0.690	< 0.001	0.046	< 0.001
	HNRNPA1	0.593	< 0.001	0.032	< 0.001
	HNRNPL	0.323	0.003	0.042	< 0.001
	HNRNPM	0.399	< 0.001	0.020	0.002
	HSF1	0.272	< 0.001	0.053	< 0.001
	IGF2BP2	0.848	< 0.001	0.092	< 0.001
	KHDRBS1	0.537	< 0.001	0.012	0.009
	MAPKAPK2	0.561	0.002	0.014	0.003
	NANOS3	1.061	< 0.001	0.019	0.009
	PABPN1	0.325	< 0.001	0.030	< 0.001
	PARN	0.245	0.001	0.031	< 0.001
	PSMB1	0.591	< 0.001	0.054	< 0.001
	PSMB10	0.494	< 0.001	0.026	0.001
	PSMB5	0.533	< 0.001	0.019	0.002
	RBFOX3	0.240	0.165	0.036	< 0.001
	RBM15B	0.430	< 0.001	0.032	< 0.001
	RBM3	0.406	0.055	0.028	< 0.001
	RBM4	0.445	< 0.001	0.023	< 0.001
	RBMX	0.112	0.316	0.021	0.003
	SAP18	0.601	< 0.001	0.023	< 0.001
	SET	0.754	< 0.001	0.040	< 0.001
	SRSF12	0.697	0.002	0.028	< 0.001

	SYNCRIP	0.102	0.192	0.030	< 0.001
	TRA2A	-0.009	0.780	0.028	0.195
	TRA2B	0.755	< 0.001	0.027	0.007
	TRIM71	1.839	< 0.001	0.160	< 0.001
	YTHDF2	0.744	< 0.001	0.037	< 0.001
	YTHDF3	0.369	0.033	0.023	< 0.001
	YWHAZ	0.342	0.007	0.027	< 0.001
	ZC3H10	0.121	0.088	0.031	0.001

The genes are common leading-edge genes from the result of GSEAPreranked from both OHSU and MILE databases. Median_diff refers to the median difference from AML samples to normal samples. *P*-values are from Wilcoxon rank-sum test.

C. List of pathways and their sources in each cluster of Fig. 35

Cluster	Pathway	Source
Cell-cycle related	HALLMARK_E2F_TARGETS	HALLMARK
	KEGG_OOCYTE_MEIOSIS	KEGG
	PID_RANBP2_PATHWAY	PID
	REACTOME_ABERRANT_REGULATION_OF_MITOTIC_EXIT_IN_CANCER_DUE_TO_RB1_DEFECTS	REACTOME
	REACTOME_AURKA_ACTIVATION_BY_TPX2	REACTOME
	REACTOME_CELL_CYCLE	REACTOME
	REACTOME_CELL_CYCLE_CHECKPOINTS	REACTOME
	REACTOME_CHROMOSOME_MAINTENANCE	REACTOME
	REACTOME_CONVERSION_FROM_APC_C_CDC20_TO_APC_C_CDH1_IN_LATE_ANAPHASE	REACTOME
	REACTOME_G2_M_CHECKPOINTS	REACTOME
	REACTOME_GABA_RECEPTOR_ACTIVATION	REACTOME
	REACTOME_GENOME_REPLICATION_AND_TRANSCRIPTION	REACTOME
	REACTOME_INHIBITION_OF_THE_PROTEOLYTIC_ACTIVITY_OF_APC_C_REQUIRED_FOR_THE_ONSET_OF_ANAPHASE_BY_MITOTIC_SPINDLE_CHECKPOINT_COMPONENTS	REACTOME
	REACTOME_MITOTIC_PROMETAPHASE	REACTOME
	REACTOME_MITOTIC_SPINDLE_CHECKPOINT	REACTOME
	REACTOME_PHASE_4_RESTING_MEMBRANE_POTENTIAL	REACTOME
	REACTOME_PHOSPHORYLATION_OF_THE_APC_C	REACTOME
	REACTOME_RECRUITMENT_OF_MITOTIC_CENTROSOME_PROTEINS_AND_COMPLEXES	REACTOME
	REACTOME_RECRUITMENT_OF_NUMA_TO_MITOTIC_CENTROSOMES	REACTOME
	REACTOME_REGULATION_OF_PLK1_ACTIVITY_AT_G2_M_TRANSITION	REACTOME
	REACTOME_RHO_GTPASES_ACTIVATE_FORMINS	REACTOME
	REACTOME_TELOMERE_EXTENSION_BY_TELOMERASE	REACTOME
	REACTOME_TRANSCRIPTION_COUPLED_NUCLEOTIDE_EXCISION_REPAIR_TC_NER_	REACTOME
	WP_HIJACK_OF_UBIQUITINATION_BY_SARSCOV2	WP
	WP_PARKINUBIQUITIN_PROTEASOMAL_SYSTEM_PATHWAY	WP
	WP_RETINOBLASTOMA_GENE_IN_CANCER	WP
Synthesis of cholesterol	REACTOME_CHOLESTEROL_BIOSYNTHESIS	REACTOME
	WP_MEVALONATE_PATHWAY	WP

Synthesis of unsaturated fatty acids	KEGG_BIOSYNTHESIS_OF_UNSATURATED_FATTY_ACIDS	KEGG
	WP_OMEGA3OMEGA6_FA_SYNTHESIS	WP
	WP_OMEGA9_FA_SYNTHESIS	WP

KEGG, Kyoto Encyclopedia of Genes and Genomes; PID, the Pathway Interaction Database; WP, WikiPathways.

D. List of genes used in the clustering in Fig. 35 including the unclustered pathways

Cluster or pathway	Genes
Cell-cycle related	*AAAS, ABL1, ABRAXAS1, AC023512.1, ACD, ACTB, ACTG1, ACTR1A, ADCY1, ADCY2, ADCY3, ADCY4, ADCY5, ADCY6, ADCY7, ADCY8, ADCY9, AHCTF1, AJUBA, AK2, AKAP9, AKT1, AKT2, AKT3, ALMS1, ANAPC1, ANAPC10, ANAPC11, ANAPC13, ANAPC15, ANAPC16, ANAPC2, ANAPC4, ANAPC5, ANAPC7, ANKLE2, ANKRD28, ANLN, ANP32E, AQR, AR, ARHGEF9, ARPP19, ASF1A, ASF1B, ATAD2, ATM, ATR, ATRIP, ATRX, AURKA, AURKB, B9D2, BABAM1, BABAM2, BANF1, BARD1, BIRC5, BLM, BLZF1, BORA, BRCA1, BRCA2, BRCC3, BRIP1, BRMS1L, BTRC, BUB1, BUB1B, BUB3, CABLES1, CALM1, CALM2, CALM3, CALML3, CALML5, CALML6, CAMK2A, CAMK2B, CAMK2D, CAMK2G, CASK, CASP1, CASP8, CBFB, CBX5, CC2D1B, CCDC6, CCNA1, CCNA2, CCNB1, CCNB2, CCND1, CCND2, CCND3, CCNE1, CCNE2, CCNH, CCP110, CDC14A, CDC16, CDC20, CDC23, CDC25A, CDC25B, CDC25C, CDC26, CDC27, CDC42, CDC45, CDC6, CDC7, CDCA3, CDCA5, CDCA8, CDK1, CDK11A, CDK11B, CDK2, CDK4, CDK5RAP2, CDK6, CDK7, CDKN1A, CDKN1B, CDKN1C, CDKN2A, CDKN2B, CDKN2C, CDKN2D, CDKN3, CDT1, CENPA, CENPC, CENPE, CENPF, CENPH, CENPI, CENPJ, CENPK, CENPL, CENPM, CENPN, CENPO, CENPP, CENPQ, CENPS, CENPT, CENPU, CENPW, CENPX, CEP131, CEP135, CEP152, CEP164, CEP192, CEP250, CEP290, CEP41, CEP43, CEP57, CEP63, CEP70, CEP72, CEP76, CEP78, CETN2, CHEK1, CHEK2, CHMP2A, CHMP2B, CHMP3, CHMP4A, CHMP4B, CHMP4C, CHMP6, CHMP7, CHP1, CHP2, CHTF18, CHTF8, CIT, CKAP5, CKS1B, CKS2, CLASP1, CLASP2, CLIP1, CLSPN, CNEP1R1, CNOT9, CNTRL, COP1, COPS2, COPS3, COPS4, COPS5, COPS6, COPS7A, COPS7B, COPS8, CPEB1, CSE1L, CSNK1D, CSNK1E, CSNK2A1, CSNK2A2, CSNK2B, CTC1, CTCF, CTDNEP1, CTPS1, CUL1, CUL2, CUL4A, CUL4B, DAAM1, DAXX, DBF4, DCK, DCLRE1B, DCTN1, DCTN2, DCTN3, DCTPP1, DDB1, DDX39A, DDX5, DEK, DEPDC1, DHFR, DIAPH1, DIAPH2, DIAPH3, DIDO1, DKC1, DLGAP5, DMC1, DNA2, DNMT1, DONSON, DSCC1, DSN1, DUT, DVL1, DVL2, DVL3, DYNC1H1, DYNC1I1, DYNC1I2, DYNC1LI1, DYNC1LI2, DYNLL1, DYNLL2, DYRK1A, E2F1, E2F2, E2F3, E2F4, E2F5, E2F6, E2F8, EED, EIF2S1, ELL, ELOB, ELOC, EMD, EML4, ENSA, EP300, ERCC1, ERCC2, ERCC3, ERCC4, ERCC5, ERCC6, ERCC6L, ERCC8, ESCO1, ESCO2, ESPL1, EVL, EXO1, EXOSC8, EZH2, FAF1, FANCG, FBXL18, FBXL7, FBXO43, FBXO5, FBXW11, FBXW7, FEN1, FKBP6, FKBPL, FMNL1, FMNL2, FMNL3, FOXM1, FZR1, GABBR1, GABBR2, GABRA1, GABRA2, GABRA3, GABRA4, GABRA5, GABRA6, GABRB1, GABRB2, GABRB3, GABRG2, GABRG3, GABRQ, GABRR1, GABRR2, GABRR3, GAR1, GINS1, GINS2, GINS3, GINS4, GMNN, GNAI1, GNAI2, GNAI3, GNAL, GNAT3, GNB1, GNB2, GNB3, GNB4, GNB5, GNG10, GNG11, GNG12, GNG13, GNG2, GNG3, GNG4, GNG5, GNG7, GNG8, GNGT1, GNGT2, GOLGA2, GORASP1, GORASP2, GPR37, GPS1, GSK3A, GSK3B, GSPT1, GTF2H1, GTF2H2, GTF2H3, GTF2H4, GTF2H5, GTSE1, H2AB1, H2AC14, H2AC18, H2AC19, H2AC20, H2AC4, H2AC6, H2AC7, H2AC8, H2AJ, H2AX, H2AZ1, H2AZ2, H2BC1, H2BC10, H2BC11, H2BC12, H2BC13, H2BC14, H2BC15, H2BC17, H2BC21, H2BC3, H2BC4, H2BC5, H2BC6, H2BC7, H2BC8, H2BC9, H2BS1, H2BU1, H3-3A, H3-3B, H3-4, H3C1,*

	H3C10, H3C11, H3C12, H3C13, H3C14, H3C15, H3C2, H3C3, H3C4, H3C6, H3C7, H3C8, H4-16, H4C1, H4C11, H4C12, H4C13, H4C14, H4C15, H4C2, H4C3, H4C4, H4C5, H4C6, H4C8, H4C9, HAUS1, HAUS2, HAUS3, HAUS4, HAUS5, HAUS6, HAUS7, HAUS8, HDAC1, HDAC4, HDAC8, HELLS, HERC2, HJURP, HLTF, HMGA1, HMGB1, HMGB2, HMGB3, HMGN1, HMMR, HNRNPD, HSP90AA1, HSP90AB1, HSPA14, HSPA1A, HSPA1B, HSPA1L, HSPA2, HSPA4, HSPA5, HSPA6, HSPA8, HSPA9, HUS1, IGF1, IGF1R, ILF3, INCENP, ING3, INS, IPO7, IST1, ISY1, ITGB1, ITGB3BP, ITPR1, ITPR2, ITPR3, JAK2, JPT1, KAT5, KCNJ10, KCNJ12, KCNJ14, KCNJ15, KCNJ16, KCNJ2, KCNJ3, KCNJ4, KCNJ5, KCNJ6, KCNJ9, KCNK1, KCNK10, KCNK12, KCNK13, KCNK15, KCNK16, KCNK17, KCNK18, KCNK2, KCNK3, KCNK4, KCNK5, KCNK6, KCNK7, KCNK9, KIF18A, KIF18B, KIF20A, KIF22, KIF23, KIF2A, KIF2B, KIF2C, KIF4A, KMT5A, KNL1, KNTC1, KPNA2, KPNB1, LBR, LCMT1, LEMD2, LEMD3, LIG1, LIG3, LIN37, LIN52, LIN54, LIN9, LMNA, LMNB1, LPIN1, LPIN2, LPIN3, LUC7L3, LYAR, LYN, MAD1L1, MAD2L1, MAD2L2, MAP2K1, MAPK1, MAPK12, MAPK13, MAPK3, MAPRE1, MASTL, MAU2, MAX, MCM10, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MCM8, MCPH1, MDC1, MDM2, MDM4, MELK, MIR3917, MIS12, MIS18A, MIS18BP1, MKI67, MLH1, MLH3, MMS22L, MNAT1, MND1, MOS, MRE11, MRTFA, MSH2, MSH4, MSH5, MSH6, MTHFD2, MXD3, MYBL2, MYC, MZT1, MZT2A, MZT2B, NAA38, NAE1, NAP1L1, NASP, NBN, NCAPD2, NCAPD3, NCAPG, NCAPG2, NCAPH, NCAPH2, NDC1, NDC80, NDE1, NDEL1, NEDD1, NEK2, NEK6, NEK7, NEK9, NHP2, NINL, NIPBL, NME1, NME7, NOLC1, NOP10, NOP56, NPAT, NPM1, NPTN, NSD2, NSL1, NUDC, NUDT21, NUF2, NUMA1, NUP107, NUP133, NUP153, NUP155, NUP160, NUP188, NUP205, NUP210, NUP214, NUP35, NUP37, NUP42, NUP43, NUP50, NUP54, NUP58, NUP62, NUP85, NUP88, NUP93, NUP98, ODF2, OFD1, OIP5, OPTN, ORC1, ORC2, ORC3, ORC4, ORC5, ORC6, PA2G4, PAFAH1B1, PAICS, PAN2, PCBP4, PCM1, PCNA, PCNT, PDS5A, PDS5B, PFN1, PFN2, PGR, PHF20, PHF5A, PHF8, PHLDA1, PIAS1, PIAS2, PIAS4, PIF1, PKMYT1, PLCZ1, PLK1, PLK4, PMF1, PMS2, PNN, POLA1, POLA2, POLD1, POLD2, POLD3, POLD4, POLE, POLE2, POLE3, POLE4, POLK, POLR2A, POLR2B, POLR2C, POLR2D, POLR2E, POLR2F, POLR2G, POLR2H, POLR2I, POLR2J, POLR2K, POLR2L, POM121, POM121C, POP7, POT1, PPIE, PPM1D, PPME1, PPP1CA, PPP1CB, PPP1CC, PPP1R12A, PPP1R12B, PPP1R8, PPP2CA, PPP2CB, PPP2R1A, PPP2R1B, PPP2R2A, PPP2R2D, PPP2R3B, PPP2R5A, PPP2R5B, PPP2R5C, PPP2R5D, PPP2R5E, PPP3CA, PPP3CB, PPP3CC, PPP3R1, PPP3R2, PPP6C, PPP6R3, PRDM9, PRDX4, PRIM1, PRIM2, PRKACA, PRKACB, PRKACG, PRKAR2B, PRKCA, PRKCB, PRKDC, PRKN, PRKX, PRMT2, PRPF19, PRPS1, PSIP1, PSMA1, PSMA2, PSMA3, PSMA4, PSMA5, PSMA6, PSMA7, PSMA8, PSMB1, PSMB10, PSMB11, PSMB2, PSMB3, PSMB4, PSMB5, PSMB6, PSMB7, PSMB8, PSMB9, PSMC1, PSMC2, PSMC3, PSMC3IP, PSMC4, PSMC5, PSMC6, PSMD1, PSMD10, PSMD11, PSMD12, PSMD13, PSMD14, PSMD2, PSMD3, PSMD4, PSMD5, PSMD6, PSMD7, PSMD8, PSMD9, PSME1, PSME2, PSME3, PSME4, PSMF1, PTK6, PTTG1, PTTG2, RAB1A, RAB1B, RAB2A, RAB8A, RABIF, RAC1, RACGAP1, RAD1, RAD17, RAD21, RAD50, RAD51, RAD51AP1, RAD51C, RAD9A, RAD9B, RAE1, RAF1, RAN, RANBP1, RANBP2, RANGAP1, RB1, RBBP4, RBBP7, RBBP8, RBL1, RBL2, RBP1, RBX1, RCC1, RCC2, REC8, RFC1, RFC2, RFC3, RFC4, RFC5, RHNO1, RHOA, RHOB, RHOC, RHOD, RMI1, RMI2, RNASEH2A, RNF168, RNF19A, RNF8, RPA1, RPA2, RPA3, RPA4, RPS27, RPS27A, RPS6KA1, RPS6KA2, RPS6KA3, RPS6KA6, RRM1,

	RRM2, RSF1, RTEL1, RUVBL1, RUVBL2, SAP30, SCAI, SDCCAG8, SEC13, SEH1L, SEM1, SEPTIN5, SET, SFI1, SFN, SGO1, SGO2, SHMT1, SHQ1, SIAH1, SIAH2, SIN3A, SIRT2, SKA1, SKA2, SKP1, SKP2, SLBP, SLK, SMARCA2, SMARCA5, SMC1A, SMC1B, SMC2, SMC3, SMC4, SMC6, SNCA, SNCAIP, SNRPB, SPAG5, SPAST, SPC24, SPC25, SPDL1, SPDYA, SPDYC, SPO11, SRC, SRF, SRGAP2, SRSF1, SRSF2, SSNA1, SSRP1, STAG1, STAG2, STAG3, STMN1, STN1, STUB1, SUMO1, SUN1, SUN2, SUV39H1, SYCE1, SYCE2, SYCE3, SYCP1, SYCP2, SYCP3, SYNCRIP, SYNE1, SYNE2, TACC3, TAOK1, TBRG4, TCEA1, TCF19, TEN1, TERF1, TERF2, TERF2IP, TERT, TEX12, TEX15, TFDP1, TFDP2, TFRC, TIMELESS, TINF2, TIPIN, TK1, TMPO, TNPO1, TOP2A, TOP3A, TOPBP1, TP53, TP53BP1, TPR, TPX2, TRA2B, TRIP13, TTK, TUBA1A, TUBA1B, TUBA1C, TUBA3C, TUBA3D, TUBA3E, TUBA4A, TUBA4B, TUBA8, TUBAL3, TUBB, TUBB1, TUBB2A, TUBB2B, TUBB3, TUBB4A, TUBB4B, TUBB6, TUBB8, TUBB8B, TUBG1, TUBG2, TUBGCP2, TUBGCP3, TUBGCP4, TUBGCP5, TUBGCP6, TYMS, UBA1, UBA3, UBA52, UBB, UBC, UBE2C, UBE2D1, UBE2E1, UBE2G1, UBE2G2, UBE2I, UBE2J1, UBE2J2, UBE2L3, UBE2L6, UBE2N, UBE2S, UBE2T, UBE2V2, UBR7, UIMC1, UNG, USO1, USP1, USP7, UVSSA, VHL, VPS4A, VRK1, VRK2, WAPL, WDR90, WEE1, WRAP53, WRN, XAB2, XPA, XPO1, XRCC1, XRCC6, YWHAB, YWHAE, YWHAG, YWHAH, YWHAQ, YWHAZ, ZCRB1, ZNF385A, ZNF655, ZNF830, ZW10, ZWILCH, ZWINT, ZYG11B
Synthesis of cholesterol	*ACAT2, ARV1, CYP51A1, DHCR24, DHCR7, EBP, FDFT1, FDPS, GGPS1, HMGCR, HMGCS1, HSD17B7, IDI1, IDI2, LBR, LSS, MSMO1, MVD, MVK, NSDHL, PLPP6, PMVK, SC5D, SQLE, TM7SF2*
Synthesis of unsaturated fatty acids	*ACAA1, ACOT1, ACOT2, ACOT4, ACOT7, ACOX1, ACOX3, ACSL1, ACSL3, ACSL4, BAAT, ELOVL1, ELOVL2, ELOVL3, ELOVL5, ELOVL6, FADS1, FADS2, FASN, HACD1, HACD2, HADHA, HSD17B12, MIR1908, MIR6734, PECR, PLA2G4A, PLA2G4B, PLA2G5, PLA2G6, SCD, SCD5, TECR, YOD1*
Metabolism of folate and pteridines [Pathway]	*ALDH1L1, ALDH1L2, DHFR, DHFR2, FOLR2, FPGS, MTHFD1, MTHFD1L, MTHFD2, MTHFD2L, MTHFR, MTHFS, SHMT1, SHMT2, SLC19A1, SLC25A32, SLC46A1*
Signaling by MST1 [Pathway]	*HPN, MST1, MST1R, SPINT1, SPINT2*
NO metabolism in cystic fibrosis [Pathway]	*CARM1, DDAH1, DDAH2, NOS1, NOS2, NOS3, PRMT1, PRMT2, PRMT3, PRMT5, PRMT6, PRMT7, PRMT8*

NO, nitric oxide.

REGULAR ARTICLE

blood advances

TAK-981, a SUMOylation inhibitor, suppresses AML growth immune-independently

Han Sun Kim,[1,*] Bo-Reum Kim,[2,*] Thien T. P. Dao,[1,*] Jin-Mo Kim,[1] Yoon-Ju Kim,[2] Hyunsong Son,[2] Sihyang Jo,[1] Doyeon Kim,[1] Jiwoo Kim,[1] Young Ju Suh,[3] Hee-Je Kim,[2,4] Byung-Sik Cho,[2,4] and Sunghyouk Park[1]

[1]Natural Products Research Institute, Department of Manufacturing Pharmacy, College of Pharmacy, Seoul National University, Seoul, Republic of Korea; [2]Leukemia Research Institute, Catholic Hematology Hospital, Seoul St. Mary's Hospital, The Catholic University of Korea, Seoul, Republic of Korea; [3]Department of Biomedical Sciences, College of Medicine, and Program in Biomedical Science and Engineering, Inha University, Incheon, Republic of Korea; and [4]Department of Hematology, Catholic Hematology Hospital, Seoul St. Mary's Hospital, The Catholic University of Korea, Seoul, South Korea

Key Points

- High SUMOylation is a clinically relevant vulnerability in AML, associated with higher European LeukemiaNet 2017 risk, poorer survival, and AML-specific mutations.
- TAK-981, a new SUMOylation inhibitor, shows nanomolar and immune-independent anti-AML activity, exhibiting synergy with other AML drugs.

Acute myeloid leukemia (AML) generally has an unsatisfactory prognosis despite the recent introduction of new regimens, including targeted agents and antibodies. To find a new druggable pathway, we performed integrated bioinformatic pathway screening on large OHSU and MILE AML databases, discovered the SUMOylation pathway, and validated it independently with an external data set (totaling 2959 AML and 642 normal sample data). The clinical relevance of SUMOylation in AML was supported by its core gene expression which is correlated with patient survival, European LeukemiaNet 2017 risk classification, and AML-relevant mutations. TAK-981, a first-in-class SUMOylation inhibitor currently under clinical trials for solid tumors, showed antileukemic effects with apoptosis induction, cell-cycle arrest, and induction of differentiation marker expression in leukemic cells. It exhibited potent nanomolar activity, often stronger than that of cytarabine, which is part of the standard of care. TAK-981's utility was further demonstrated in in vivo mouse and human leukemia models as well as patient-derived primary AML cells. Our results also indicate direct and cancer cell-inherent anti-AML effects by TAK-981, different from the type 1 interferon and immune-dependent mechanism in a previous solid tumor study. Overall, we provide a proof-of-concept for SUMOylation as a new targetable pathway in AML and propose TAK-981 as a promising direct anti-AML agent. Our data should prompt studies on optimal combination strategies and transitions to clinical trials in AML.

Introduction

Acute myeloid leukemia (AML) is a heterogeneous disease characterized by an accumulation of immature progenitor cells with arrested differentiation leading to suppression of hematopoiesis.[1] Current standard of care treatments include combination chemotherapy with cytotoxic drugs, usage of hypomethylating agents, and/or hematopoietic stem cell transplantation.[2] Recent improvement in our understanding of AML pathogenesis has led to the introduction of several novel targeted agents since 2017.[3] Nevertheless, long-term survival is still suboptimal without allogeneic hematopoietic stem cell transplantation,[4] and thus, more efforts should be done to unravel novel prognostic, predictive, and targetable molecular abnormalities. However,

Submitted 28 April 2022; accepted 10 February 2023; prepublished online on *Blood Advances* First Edition 21 February 2023; final version published online 3 July 2023. https://doi.org/10.1182/bloodadvances.2022007956.

*H.S.K., B.-R.K., and T.T.P.D. contributed equally to this study.

Data are available on request from the corresponding author, Sunghyouk Park (psh@snu.ac.kr).

The full-text version of this article contains a data supplement.

lack of prevailing driver genomic mutations and available unique markers for AML has made it quite difficult. In this context, investigations into postgenomic pathways relevant to AML pathogenesis and approaches to their targeting have been desired.

SUMOylation is a posttranslational modification involved in the conjugation of small ubiquitin-like modifiers (SUMOs) to substrate proteins.[5] SUMO-activating enzyme E1 (SAE1 and SAE2 encoded by *SAE1* and *UBA2*, respectively), an E2 (ubiquitin-conjugating enzyme 9 (UBC9) encoded by *UBE2I*), and a limited set of E3 ligases participate in this process.[5,6] SUMOylation seems to be important in the nuclear functions of proliferating or developing cells by regulating the mitotic cell cycle and DNA damage response.[7-9] Specific pathways affected by SUMOylation in cancer may include p53[10,11] and cMYC,[12,13] but more studies are needed to resolve some of the controversies.[14,15] In addition, innate immunity is mostly suppressed by SUMOylation, the inhibition of which, therefore, might have implications for cancer therapy.[5,16] As for AML, only a few studies on the roles of SUMOylation have been published.[17-19] Therefore, concrete evidence of the therapeutic utility of SUMOylation or of specific inhibitors of SUMOylation in AML has been lacking. TAK-981 is an inhibitor of the SUMO-activating enzyme (SAE) that forms a SUMO-TAK-981 adduct.[20] As the first-in-class SAE inhibitor targeting cancers, it is currently in clinical trials for solid tumors or lymphomas (#NCT03648372, #NCT04074330, and #NCT04381650). In blood cancer, it has been known to shift the T-cell balance toward healthy immune cell subsets in chronic lymphocytic leukemia.[21] To our knowledge, TAK-981 has not been studied for AML or evaluated in AML clinical trials.

For solid tumors, large-scale bioinformatic analysis has been successfully performed comparing normal and cancer samples, thanks to The Cancer Genome Atlas (TCGA) data. TCGA also contains data on AML (TCGA-LAML[22] data set), but it lacks the data for noncancer controls, limiting its application in the AML field. As of now, 3 large-scale gene expression databases contain both AML and normal data: (1) MILE study stage I data,[23] (2) OHSU data from the Beat AML 1.0 program,[24] and (3) the Gene Expression Omnibus (GEO) compilation.[25] Therefore, analysis of these large databases (totaling 2959 AML and 642 normal samples) might yield new and useful information on targets for patients with broader AML.

Here, accessing large gene expression databases for AML, we evaluated the clinical relevance of the SUMOylation pathway and investigated the antileukemic effects of its inhibition by TAK-981.

Methods

More detailed information is in supplemental Methods.

Bioinformatic analysis

Details are provided in supplemental Data.

Cells - reagents, antibodies for flow cytometry, and cell viability with cell counting kit-8 (CCK-8) assay

Details are provided in supplemental Data.

Primary AML cells from patients

The information of the samples, including mutation status, is listed in supplemental Table 7.

Flow cytometry, Apoptosis analysis, and Cell-cycle analysis

Details are provided in supplemental Data.

Quantitative reverse transcription polymerase chain reaction validation, Western blotting

The efficiencies of the primers used are listed in supplemental Table 8.

Animal experiments

Details are provided in supplemental Data.

Statistical analysis

The Wilcoxon rank-sum test, one-way analysis of variance, Student *t* test, and Jonckheere-Terpstra test were used when necessary. Details are provided in supplemental Data.

All animal experiments were performed in accordance with a protocol approved by the Institutional Animal Care and Use Committee of The Catholic University of Korea (CUMC-2020-0318-01). Bone marrow (BM) samples from patients with AML were collected during routine diagnostic procedures after informed consent was obtained in accordance with Institutional Review Board regulations of The Catholic University of Korea (KC20SISI0957) and the Declaration of Helsinki.

Results

Bioinformatic screening identifies SUMOylation pathway as AML-specific target

First, we performed an integrated analysis on large-scale databases (MILE study stage I and OHSU Beat AML 1.0 program) (Figure 1A). Selection of significant pathways in the 2 gene set enrichment analysis results (AML vs normal) (supplemental Table 1) followed by their clustering based on common leading-edge genes and protein-protein interactions yielded 4 distinct pathway clusters: (1) translation/ribosomal RNA/mitochondria, (2) histone-related, (3) SUMOylation, and (4) regulation of messenger RNA (mRNA) (Figure 1B). Interestingly, inhibitors targeting the first cluster, such as ribosome biogenesis inhibitors or tetracyclines, had shown both in vitro and in vivo antileukemic activities and were entered into clinical development.[27-29] These facts show that our bioinformatic results may have real relevance for targeting AML. Of the 3 remaining clusters, we focused on SUMOylation cluster because it had not been explored much for AML, and the other 2 were either difficult to establish the causality (Histone-related) or too nonspecific (Regulation of mRNA). Most of the individual genes comprising the SUMOylation pathway were found to be upregulated in AML samples from both the MILE and OHSU databases (*SUMO1* and *UBA2* in Figure 1C; all the others in supplemental Figure 1A). We further validated the results using another large independent data set from the GEO collection of 2213 AML and 548 normal samples.[25] Consistently, we found that 11 of 17 genes related to SUMOylation were found to be significantly upregulated in AML samples (*SUMO1* and *UBA2* in Figure 1D; all the others in supplemental Figure 1B). In particular, we observed higher protein levels of E1 (SAE1 and

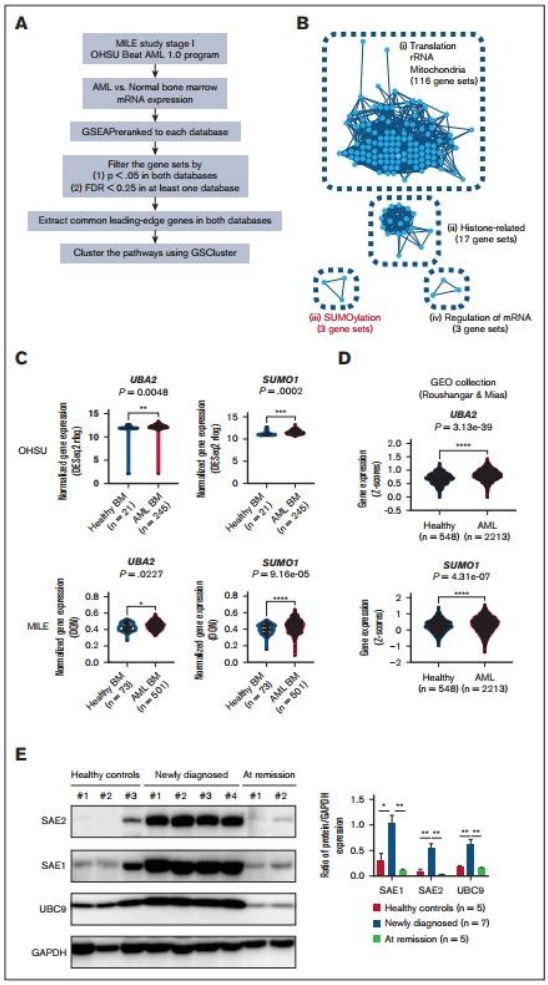


Figure 1. Bioinformatic screening to find AML-specific pathways. (A) Overall strategy for database screening. (B) Graphical illustration of 4 pathway clusters upregulated in AML BM samples from panel A, using GSCluster[26] R package. The number of connected gene sets in each cluster is indicated. (C) Comparison of *UBA2* and *SUMO1* gene expression between healthy and AML BM samples in OHSU and MILE databases. (D) Comparison of *UBA2* and *SUMO1* gene expression between healthy and AML BM/peripheral blood samples in Gene Expression Omnibus (GEO) data sets by Roushangar and Mias.[25] (E) Representative western blot for SAE2, SAE1, UBC9, and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) in peripheral blood from healthy controls and patients with AML at diagnosis or remission state after treatment (left). The intensities of the bands from all the samples were quantified by densitometry and displayed as the ratio of each protein to GAPDH (loading control) (right). Newly diagnosed patients with AML (n = 7), those at remission state (n = 5), and healthy controls (n = 5). Results are expressed as the mean ± standard error of the mean. For panels C-E, *P* values are from Wilcoxon rank-sum test; **P* < .05, ***P* < .01, ****P* < .001, *****P* < .0001. BM, bone marrow; GEO, Gene Expression Omnibus; FDR, false discovery rate; GSEA, gene set enrichment analysis; rRNA, ribosomal RNA.

SAE2), targets for TAK-981, and E2 (UBC9) in cells from patient with AML than those in healthy control or patients with remission after therapy (Figure 1E). We believe these provide further support for the involvement of SUMOylation at the protein level. The results also suggest that the upregulated SUMOylation pathway in AML may be a target for therapeutic intervention.

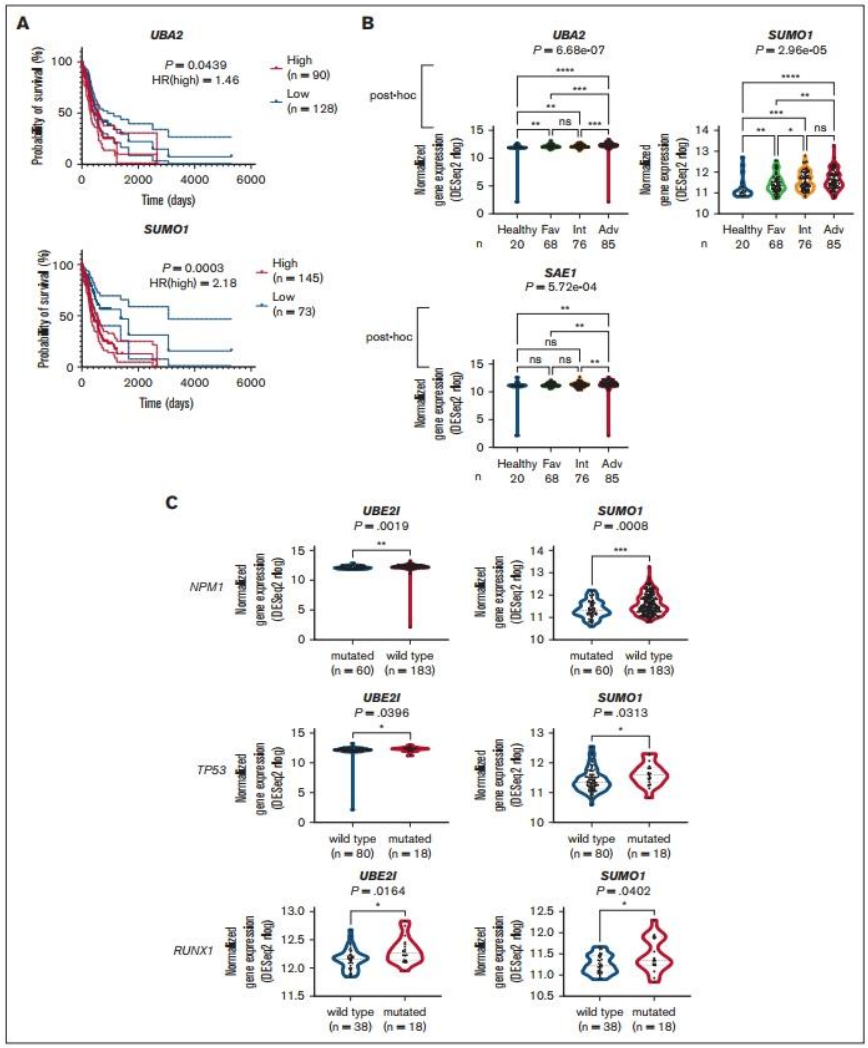


Figure 2.

SUMOylation pathway is associated with adverse risk features and poor survival in AML

We then explored the clinical relevance of SUMOylation. First, higher expression of most of the important genes in the SUMOylation pathway from the OHSU database (*SUMO1* and *UBA2* in Figure 2A; all the others in supplemental Figure 2A) was significantly associated with shorter survival. Some of those negative correlations (for *SAE1*, *BMI1*, and *PHC2*) were validated with the TCGA database (supplemental Figure 2A), and all the results, along with those without correlations, are shown in supplemental Table 2. Second, the ELN2017 risk analysis on the 4 groups (healthy, favorable, intermediate, and adverse in OHSU database) demonstrated that most of the core genes in the SUMOylation pathway expressed at higher levels in the high-risk groups ($P < .05$) (*SUMO1*, *UBA2*, and *SAE1* in Figure 2B; all the others in supplemental Figure 2B). Post hoc analysis showed that the difference concerning the SUMOylation pathway between the healthy and adverse risk groups was significant (except for *UBE2I* gene). This trend was also confirmed from the 3 patient risk groups (favorable, intermediate, and adverse) in the TCGA database for several genes, including *BMI1*, *CBX2*, and core genes such as *SAE1* and *UBA2*, and the results are shown in supplemental Table 3 along with the results for all the other genes without such confirmation.[30] As the above results are for individual gene levels, we further explored the pathway-specific relationship between SUMOylation and overall survival/ELN2017, by performing similar analyses with gene set variation analysis (GSVA) pathway scores.[31] Consistent with the results from individual genes, higher scores of SUMOylation pathways were found to be significantly related with a poorer prognosis in both the survival analysis and the ELN2017 risk analysis (supplemental Table 4). These relationships remained valid after adjusting for characteristics of patients with high-risk AML that might have confounding effects, as evidenced by multivariate analysis (supplemental Table 5).

Third, we tested if particular gene mutations are related to core SUMOylation gene expression. Among the 4 gene mutations (*FLT3*-ITD, *NPM1*, *TP53*, and *RUNX1*) that had enough patients (n > 5) for both mutated and wild-type groups, 3 mutations (*NPM1*, *TP53*, and *RUNX1*) exhibited consistent patterns between prognosis and core SUMOylation gene expression (*SUMO1* and *UBE2I* in Figure 2C; all the others in supplemental Figure 2C). Specifically, patients with the *NPM1* mutation associated with better prognosis had lower SUMOylation gene expression, whereas those with the *TP53* and *RUNX1* mutations associated with poor prognosis had higher SUMOylation gene expression. These results suggest that activation of the SUMOylation pathway is associated with adverse risk features and poorer survival.

TAK-981, a new SUMOylation inhibitor, exhibits potent antileukemic effects in vitro

In our quest for an inhibitor of SUMOylation, we found TAK-981, which was developed very recently as a first-in-class inhibitor of SAE step[20] and is currently under clinical trial for various solid tumors. As its effects against AML are still unknown, we evaluated them in vitro.

Surprisingly, TAK-981 showed greater or similar potency compared with cytarabine (Ara-C), a standard drug used in clinics, against 4 AML cell lines (Figure 3A). Notably, the 50% inhibitory concentration (IC_{50}) values for TAK-981, all within a 2-digit nanomolar range, were somewhat uniform across the cell lines. By contrast, those for cytarabine differed markedly (>1 μM for KG-1 and THP-1; 2-digit nM range for U937). In comparison, tetracycline, targeting the "translation/ribosomal RNA/mitochondria" identified above, exhibited only several-hundred-μM potency (supplemental Figure 3A).

Next, we tested TAK-981 for any synergistic or dose reduction effect when used with cytarabine in the 4 cell lines (Figure 3B-C). In addition, TAK-981's synergy with 2 new targeted-therapy drugs, venetoclax and quizartinib, along with a demethylating drug, azacitidine, was tested for the MOLM-14 cell line having the *FLT3*-ITD mutation, which is associated with poor prognosis (Figure 3D-E). Synergy, as judged by the CompuSyn scores,[32] varied substantially across cell lines, with U937 and MOLM-14 exhibiting significant synergy, whereas KG-1 and THP-1 showing little synergy in the combination with cytarabine. For MOLM-14, TAK-981 exhibited significant synergy with azacitidine, some synergy at higher drug concentrations with venetoclax, but no synergy with quizartinib. In addition, TAK-981 showed similar and lower potency in comparison with venetoclax, a BCL2 inhibitor, and quizartinib, an FLT3 inhibitor, respectively (Figure 3D; supplemental Figure 3B). Although we used only concentration values around IC_{50} for each drug, significant synergy might be observed with different concentration combinations. We also assessed the dosage reduction effects of TAK-981 (Table 1). Notably, even when there was no apparent synergy, the dose reduction indices of the drugs combined with TAK-981 were above 1 for all of the drug cell line settings, indicating significant dosage reduction effects. This could be exploited to lower the toxicity of such drugs when combined with TAK-981. Overall, TAK-981's combination with conventional or targeted drugs holds promise for improved therapeutics.

TAK-981 induces apoptosis, cell-cycle arrest, and/or differentiation marker expression in AML cell lines

To study how TAK-981 exhibits antileukemic effects, we investigated cellular events upon drug treatment. As expected, TAK-981 reduced SUMOylation for some of the proteins, if not all, from the cell extracts (24- or 48-hour treatment) (Figure 4A; supplemental Figure 4). Because SUMOylation plays a critical role in transcription regulation, we next analyzed gene expression profile changes by TAK-981 treatment (16 hours) using gene set enrichment analysis (GSE173116:[33] THP-1 cells) (Figure 4B). The upregulated pathways included those for cell death and cell-cycle arrest, such as the p53 pathway and apoptosis. Experimentally, the mRNA

Figure 2. Clinical relevance of SUMOylation pathway in AML. (A) Kaplan-Meier curves with 95% confidence intervals (dotted lines) for overall survival of patients with AML in OHSU, according to the gene expression levels of *UBA2* or *SUMO1*. The division of the high- and low-expression groups was determined by the best risk separation approach. (B) Comparison of *UBA2*, *SUMO1*, and *SAE1* gene expression across healthy and ELN2017 risk groups. *P* values are from Jonckheere-Terpstra test. Subsequent post hoc analyses were performed with the two-stage linear step-up procedure, and the significance is indicated for each comparison. The number of participants is indicated for each group. (C) Comparison of *UBE2I* and *SUMO1* gene expression between mutated and wild-type of *NPM1*, *TP53*, and *RUNX1* genes in OHSU database. *P* values are from Wilcoxon rank-sum test. For panels B-C, *$P < .05$, **$P < .01$, ***$P < .001$, ****$P < .0001$. Adv, adverse; Fav, favorable; HR(high), hazard ratio of high expression group; Int, intermediate; ns, not significant.

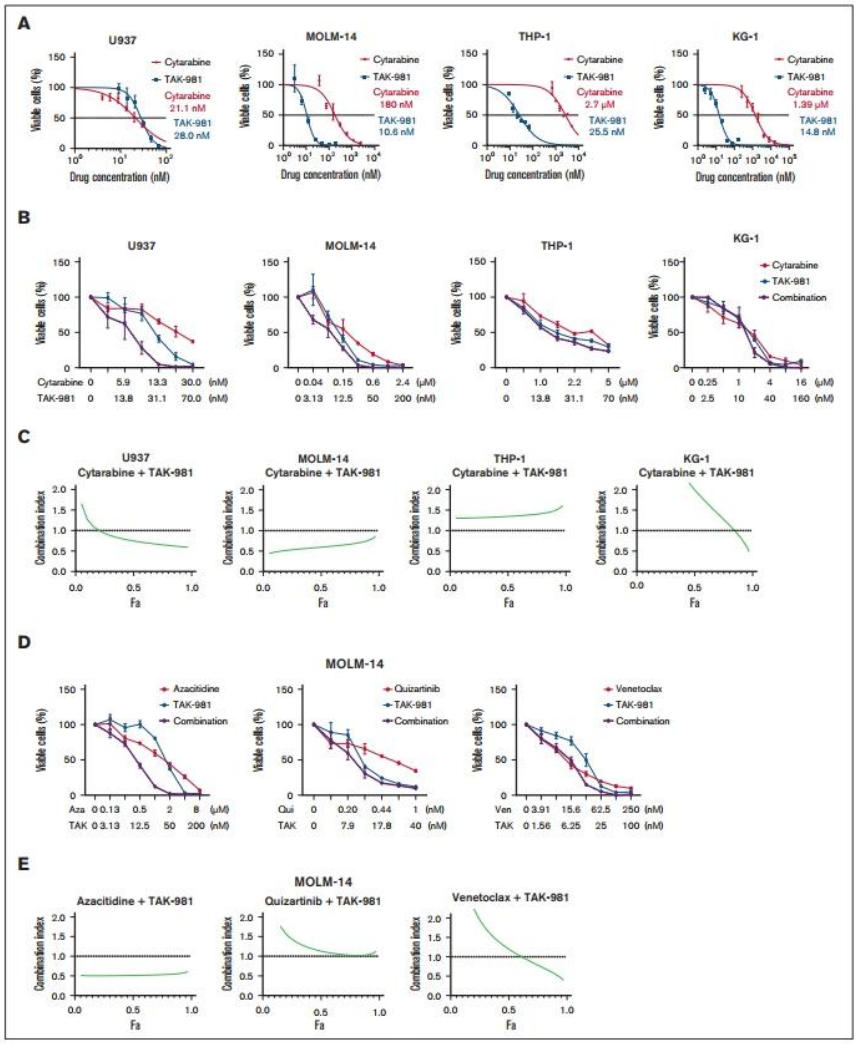


Figure 3.

Table 1. Dose reduction index of cytarabine or other drugs when combined with TAK-981 in AML cell lines

Cell line	Combination drug	Dose reduction index at fraction affected (Fa) = 0.9
U937	Cytarabine	12.03
THP-1	Cytarabine	1.25
KG-1	Cytarabine	2.61
MOLM-14	Cytarabine	2.94
MOLM-14	Azacitidine	4.87
MOLM-14	Quizartinib	8.64
MOLM-14	Venetoclax	4.98

Fa = 0.9 refers to the point where the inhibition effect is 90%, that is, when 90% of the cells are dead. The number 0.9 was chosen, because for cancer therapies, high effect levels are thought to be more therapeutically relevant than low effect levels.[32]

expression of genes for apoptosis (*DDIT3*) and cell-cycle arrest (*P21* and *TP53*), known to be downregulated by SUMOylation in AML cells,[11,19,34] were significantly higher in TAK-981-treated THP-1 cells (48 hours) than in those from the control or cytarabine–treated group (Figure 4C). We also found that there was a trend that SUMO core pathway is downregulated in TAK-981–treated THP-1 cells (supplemental Figure 5), although TAK-981's posttranslational effect on SUMO may not necessarily involve the expression of SUMO core genes. Further analysis in several other AML cell lines with western blot (p21, caspase 3, and cytochrome C) (Figure 4D), flow cytometry for apoptosis (Figure 4E), and DNA content analysis (Figure 4F) showed that apoptosis and cell-cycle arrest were generally observed for the TAK-981–treated AML cells (48 hours), with only minor variations. For example, G2/M phase arrest was observed for U937, THP-1, and KG-1 cells, whereas G0/G1 arrest was observed in MOLM-14 cells. Meanwhile, there is heterogeneity in terms of p53 mutations among the cell lines used in this study (supplemental Table 6). As p21 can be regulated either by p53 dependently or independently, we tested if the induction of p21 by TAK-981 is also reflected in the p53. TAK-981 treatment did not change the levels of either p53 or MDM2 (supplemental Figure 6), suggesting that the TAK-981–induced p21 change may not be related to p53. Possible mechanistic disconnection between p53 and p21 upon TAK-981 treatment could be an interesting topic for future research.

TAK-981 treatment (48 hours) also affected the differentiation of leukemic cells dose-dependently, as shown by the increase in the differentiation markers for U937 (*CD15*),[34-36] THP-1 (*CD14*), and MOLM-14 (*CD11B*) cells (Figure 4G). Moreover, TAK-981 suppressed the expression of *CD39* (48 hours) (Figure 4H), which is known to be involved in AML chemoresistance,[37] in both chemosensitive (U937) and chemoresistant cells (KG-1, THP-1, and MOLM-14). These data suggest that TAK-981 exhibits antileukemic effects by inducing apoptosis, cell-cycle arrest, differentiation, or lower chemoresistance.

TAK-981 potency in primary AML cells ex vivo

The effects of TAK-981 were also evaluated ex vivo in primary AML cells from patient B (n = 13). TAK-981 exhibited higher inhibition of primary cell proliferation at equimolar concentrations than did cytarabine, which did not appreciably inhibit the cells at up to ~50 μM concentrations (Figure 5A). Interestingly, the inhibitory potencies of both compounds for the primary cells were much lower than those for the AML cell lines. In addition, the SUMOylation status of primary AML cells from patients was lower than that in the cell lines (supplemental Figure 7). The possible reasons for these differences between cell lines and primary cells are addressed in the discussion section.

Still, there was significant synergy between the 2 drugs against the primary cells (Figure 5B), indicating the possible clinical utility of TAK-981. Consistently with the AML cell line results, TAK-981 induced apoptosis in the primary AML cells, and this result suggests its direct effect on cancer cells independent of antitumor immunity (Figure 5C).

TAK-981's antileukemic effects in both syngeneic AML mouse and human xenograft models

To assess TAK-981's anti-AML activity in an immune-competent environment, we used the mouse syngeneic AML model using the C1498 cell line. For the mice injected with C1498/luciferase/CD90.1 cells through tail veins, TAK-981 significantly reduced the leukemic burden on day 19 relative to the control group, as judged by the bioluminescence (supplemental Figure 8A-B). Flow cytometric analysis of leukemic cells from BM and blood (from 3 euthanized animals from each group on day 19) showed much fewer leukemic cells in the TAK-981 group (supplemental Figure 8C-D), consistent with the above imaging data on day 19. Significantly prolonged survival was also observed in the TAK-981 group relative to the controls (supplemental Figure 8E). These data in the syngeneic, immune-competent cancer model confirm TAK-981's in vivo anti-AML activity.

To confirm the human relevance of the antileukemic activity of TAK-981 and to evaluate the influence of antitumor immunity on its anti-AML effect, we injected human AML cell MOLM-14/luciferase/green fluorescent protein (0.5×10^6) into nonirradiated, immune-deficient NOD/SCID/IL-2rγ^{null} (NSG) mice (no T cells and defective dendritic cells). Both the bioimaging data (Figure 6A-B) and the flow cytometric results on the blood and BM cells (Figure 6C-D) confirmed the lower leukemic burden in the TAK-981 group. Western blot with sorted leukemic cells showed a decreased level of SUMOylated proteins in the TAK-981 group, thereby confirming its in vivo deSUMOylation activity (Figure 6E). Significantly prolonged survival was also observed in the TAK-981 group relative to the control (Figure 6F). Therefore, the data confirm TAK-981's anti-human AML activity in vivo. Importantly, these data show that TAK-981's in vivo activity is independent of antitumor immunity, as it is lacking in the NSG mouse model.

Figure 3. TAK-981's potency and its synergy with cytarabine for AML cells. (A) Dose-response curves of TAK-981 and cytarabine for 4 AML cell lines. The concentration values right beside each curve represent IC_{50} values. (B) Synergy between TAK-981 and cytarabine for 4 AML cell lines. (C) Combination index plots computed from the data in panel B by CompuSyn software. (D) Synergy between TAK-981 and several drugs for MOLM-14 cell line. (E) Combination index plots computed from the data in panel D by CompuSyn software. For panels B and D, different concentration ranges were used for each drug, and the error bars indicate standard deviation. For panels C and E, values below the dotted line at 1.0 indicate synergy. For panels A, B, and D, cell viability was measured by CCK-8 assay. Aza, azacitidine; Fa, fractions affected; Qui, quizartinib; TAK, TAK-981; Ven, venetoclax.

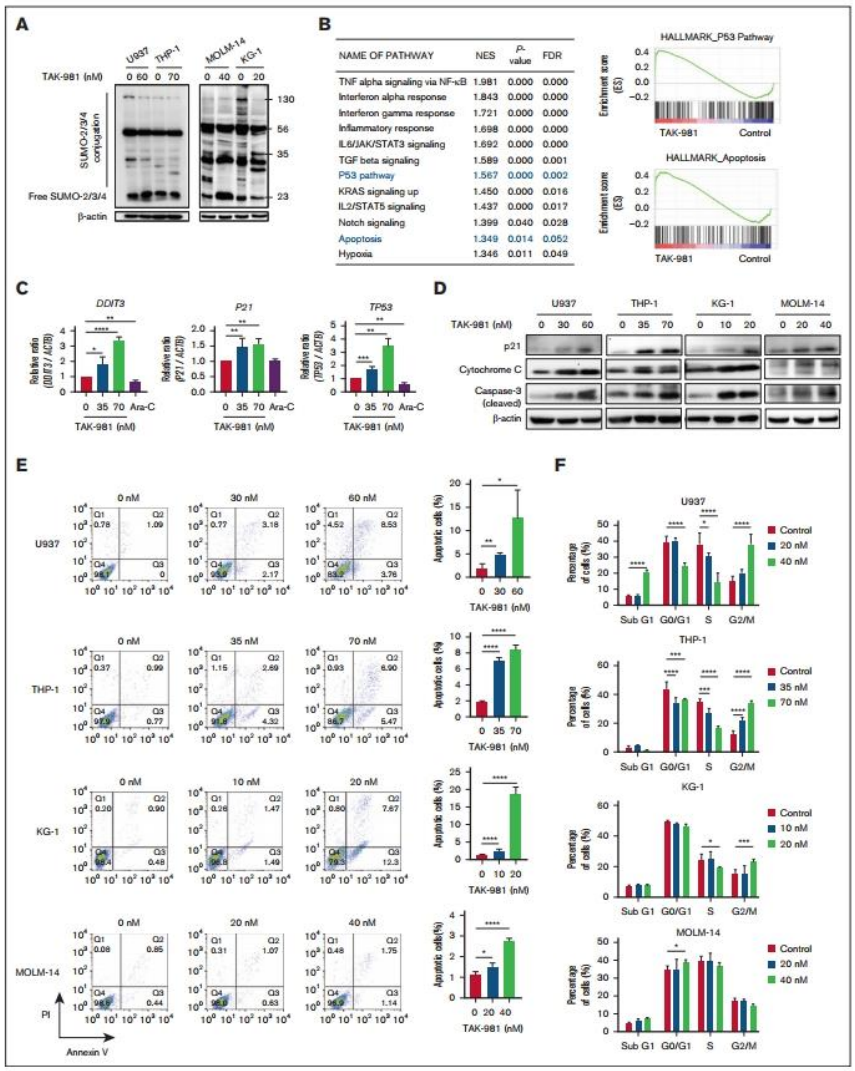


Figure 4. Apoptosis, cell-cycle arrest, and differentiation induced by TAK-981 in AML cells. (A) Effect of TAK-981 on protein SUMOylation in AML cells after 24- (U937, THP-1) or 48-hour (MOLM-14, KG-1) treatment. Western blot analysis was performed with the antibody for SUMO-2/3/4. (B) Top 12 pathways with $P < .05$ from GSEA analysis

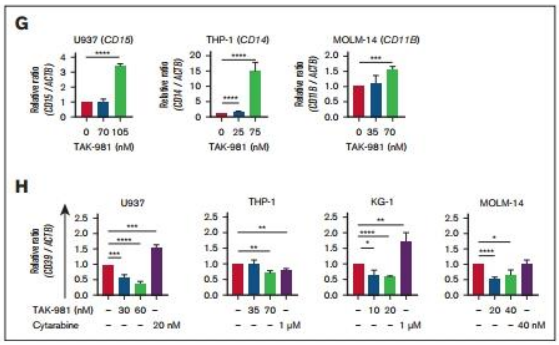


Figure 4 (continued) of TAK-981–treated THP-1 cells from GSE173116 data set with the Hallmark gene set (left). The pathways are in the order of the normalization of the enrichment score (NES). Enrichment score plots for genes belonging to p53 and apoptosis pathways from the GSEA analysis (right). (C) Relative mRNA expression of *DDIT3*, *P21*, and *TP53* in TAK-981 (indicated concentrations) and cytarabine (1 μM) in THP-1 cells after 48-hour treatment, as measured by quantitative reverse transcription polymerase chain reaction. (D) Western blot for p21, cleaved caspase-3, and cytochrome C expression in AML cells after 48-hour treatment with TAK-981. (E) Apoptosis analysis for TAK-981–treated AML cells (48 hours) by flow cytometry with Annexin V/propidium iodide (PI) kit. Apoptotic cells (%) (right) is the sum of the early (Q3) and late (Q2) apoptosis percentages. (F) Cell-cycle analysis for TAK-981–treated AML cells after 48 hours by flow cytometry. Each phase of cell cycle was analyzed with cell-cycle platform in FlowJo software. Quantitative reverse transcription polymerase chain reaction analysis of differentiation markers and *CD39* gene. mRNA expression in 48-hour TAK-981–treated AML cells for *CD15* in U937, *CD14* in THP-1, and *CD11B* in MOLM-14 (G) and *CD39* in all cells (H). Two-tailed Student *t* test was used for panels C, E, G, and H, and one-way analysis of variance was used for panel F. Data are expressed as mean ± standard deviation (n = 3); $^{*}P<.05$, $^{**}P<.01$, $^{***}P<.001$, $^{****}P<.0001$. For all except panel B, the drug concentrations were selected so that TAK-981 did not kill all the cells but had detectable effects on cells, based on the results from the initial estimation of IC_{50} of TAK-981 for each cell line. The time points are different because the emergence of the particular effects was expressed at different time points according to cell lines. In addition, all experiments were done with n = 3. For panels C, G, and H, the expression values were normalized against that of β-actin (*ACTB*). The efficiencies of the primers used are listed in supplemental Table 8. Ara-C, cytarabine; IL6, interleukin 6; JAK, Janus kinase; NF-κB, nuclear factor kappa B; STAT, signal transducer and activator of transcription; TGF, transforming growth factor; TNF, tumor necrosis factor.

Discussion

SUMOylation has not been much recognized in AML other than in cases of acute promyelocytic leukemia (APL), a minor (~10%) subset of AML with the characteristic chromosomal translocation generating the PML-RARα fusion protein.[38] The established therapy for APL, with all-trans retinoic acid (ATRA) and As_2O_3, triggers SUMOylation and subsequent proteasomal degradation of PML-RARα, thus inducing APL differentiation.[39] Activities of ATRA-induced differentiation on some non-APL AML cell lines in vitro[40] led to clinical trials, but yielded overall disappointing outcomes.[41] In our results, TAK-981 could enhance the in vitro differentiation of all AML cells tested. It will be interesting to revisit the issue of the differentiation of AML cells upon inhibition of SUMOylation in vivo. It is therefore worth noting that the addition of ATRA to decitabine improved clinical outcomes for older patients who are difficult to treat in a phase 2 clinical trial.[42] There have also been a few reports on the SUMOylation of individual proteins involved in AML, such as iGF1R, sPRDM, and ERG.[17,18,43] In addition, a protein array–based screening on AML cell lines with acquired drug resistance vs parental cell lines identified possible SUMOylation biomarkers related to drug resistance, which is yet to be validated in vivo.[44] However, considering the inhibition of the initial step of SUMOylation by TAK-981, it seems unlikely that one particular protein is responsible for TAK-981's antileukemic activity. Rather, TAK-981's activity should be contributed to by several SUMOylation-dependent processes.[45] The differential profiles of SUMOylation dependency might explain why we observed a large variability in synergy between TAK-981 and cytarabine across the different AML cells. Inhibition of SUMOylation in general with different inhibitors has also been tested. Anacardic acid and/or 2-D08 induce apoptosis of leukemic cells through reactive oxygen species–mediated deSUMOylation of *NOX* or *DDIT3* regulators.[19,46] In addition, anacardic acid and 2-D08 sensitized non-APL AML cells to ATRA-based differentiation.[34] However, there is a conflicting report according to which anacardic acid and ginkgolic acid alleviated ATRA-mediated inhibition of leukemic cell proliferation.[47] This shows that SUMOylation inhibition for AML therapy has not yet been well established and that the existing literature may need to be considered with some caution. Particularly, most of these studies have used cell lines in vitro or subcutaneous flank xenografts of AML cells and inhibitors with rather moderate micromolar activities without high specificity for SUMOylation.[34,47] In comparison, we started from the clinical relevance of the SUMOylation pathway and investigated the association of core genes in the SUMOylation pathways and AML characteristics, rather than focusing on a single protein. Furthermore, we evaluated a highly specific SUMOylation inhibitor in multiple AML cell lines, patient-derived primary cells, and orthotopic leukemia models. Overall,

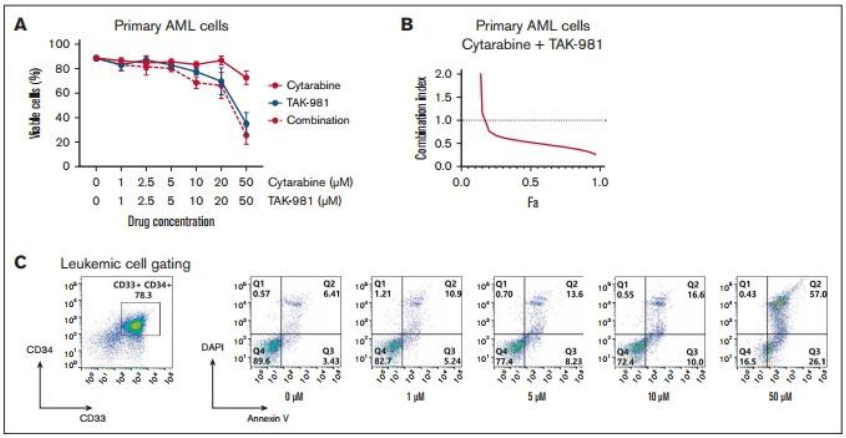


Figure 5. TAK-981's activity against primary AML cells ex vivo. Freshly isolated mononuclear cells from BM of 13 patients with AML were cultured with different doses of TAK-981, cytarabine, or both for 48 hours. Leukemic cells were gated with CD33 and/or CD34 by flow cytometry and viable cells (4′,6-diamidino-2-phenylindole [DAPI]–negative/Annexin V–negative) were compared between groups. (A) Potency and combination effects of TAK-981 and cytarabine. Viable cells were estimated by flow cytometric analysis of primary AML cells treated with TAK-981, cytarabine, or both. Error bars are standard errors. (B) Synergistic combination index between TAK-981 and cytarabine from data in panel A. (C) Leukemic cell gating (left) and representative data of flow cytometry for apoptosis of primary AML cells at different concentrations of TAK-981 with DAPI and Annexin V.

after starting the study with bioinformatics using gene expression, we showed that the treatment of TAK-981 decreased SUMOylation at the protein level with potent antileukemic effects, resulting in prolonged survival in orthotopic models. Our results should represent sufficient rationale for testing TAK-981 in AML treatment, as it is already being done in clinical trials for solid tumors.

TAK-981 is a highly specific inhibitor of SUMOylation having little effect on ubiquitination or neddylation.[20] Still, the mechanism of anticancer activity of TAK-981 may be multifaceted because of the broad-reaching roles of SUMOylation in cancer.[5,45] Interestingly, recent data suggest that TAK-981's activity against solid tumors is dependent on antitumor immunity, especially through type 1 interferon signaling regulated by SUMOylation.[16,33] For an immune-competent syngeneic flank model, TAK-981's activity was abolished when the type 1 interferon receptor was knocked out.[33] In addition, in 2 different syngeneic flank models, a survival benefit was observed for the TAK-981–immune checkpoint inhibitor (ICI) combination groups but not for the TAK-981 monotherapy groups, suggesting a cancer cell–extrinsic mechanism of TAK-981[33]. In our orthotopic models for AML, a hematologic cancer, we observed significant inhibition of leukemia growth and survival benefits in both immune-competent syngeneic mouse transplants and human xenograft models with immune-deficient mice. It should be noted that the NSG immune-deficient mice used here lacked T lymphocytes and had defective dendritic cells that had proved critical to antitumor immunity by TAK-981 in the above solid tumor settings. In addition, we observed potent in vitro inhibitory effects of TAK-981 as well as the induction of differentiation markers for various AML cell lines. Direct apoptotic effects of TAK-981 were also observed ex vivo for primary AML cells from patients. These results strongly suggest that TAK-981 exhibits cancer cell–inherent anti-AML activity. The apparent discrepancy with the above study may be because of the fundamental differences between solid tumors vs AML cancer or the experimental settings (ie, flank transplant vs orthotopic [blood] xenograft).

Still, we do not exclude the possibility of anti-AML immunity by TAK-981 or synergy with ICIs in immune-competent human AML settings that we did not study. For acute leukemia, immunotherapy has been advanced and regularly used in clinics for acute lymphoblastic leukemia, and it has also been rapidly developing for AML,[33,48] as evidenced by the approval of gemtuzumab ozogamicin in 2017. At this point, ICI monotherapy for AML has been proven not to be very satisfactory,[49,50] and its combinations with hypomethylating agents that have their own immune-modulatory effects[51,52] have yielded mixed results.[53,54] As for the positive ones, those from a phase 1b study on the combination of azacitidine and magrolimab in patients ineligible for intensive chemotherapy were quite encouraging.[53] Notably, this combination was effective even for therapy-refractory patients with *TP53*-mutated AML, though the overall number of patients was small. Larger human clinical trials with TAK-981-ICI combinations are warranted to evaluate their real effects in human AML.[55]

We showed that TAK-981 exhibited stronger or similar potency than cytarabine in all the AML cell lines tested as well as in patient-derived primary AML cells. Moreover, TAK-981 exhibited inhibition

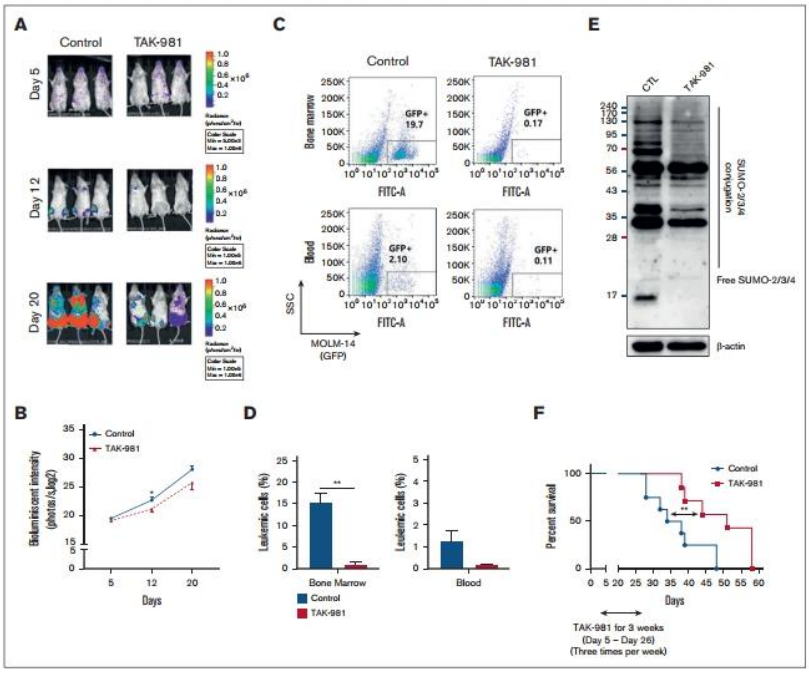


Figure 6. TAK-981's antileukemic effects in human xenograft AML mouse models (immune-compromised mice). (A-E) Human AML mouse model was established by injecting MOLM-14 cells labeled with luciferase/green fluorescent protein (GFP) (MOLM-14/luciferase/GFP) into NSG mice through tail vein. After confirming leukemia engraftment by bioluminescence imaging, the mice were divided into 2 groups (10 mice per group) and treatment began on day 5 until day 26: control (no treatment) or TAK-981 (7.5 mg/kg formulated in 20% 2-hydroxypropyl-β-cyclodextrin, IV 3 times a week). Representative mice from each group were subjected to serial bioluminescence images (A) and intensity quantitation on days 5, 12, and 20 after leukemic cell injection (B). (C-D) Three representative mice per group were euthanized on day 20 to compare the leukemic burdens between the groups. Cells from the BM and blood were analyzed using flow cytometry. The proportions of GFP+ cells by flow cytometry to identify leukemic cells were compared between the groups. (E) Western blot was performed with sorted leukemic cells to evaluate SUMOylated proteins in each group. The sample was pooled from individual animals, representing the average levels (supplemental Methods). (F) The overall survival rate in each group (7 mice per group) was estimated by the Kaplan-Meier method. The results are expressed as the mean ± standard error of the mean; $^{*}P < .05$, $^{**}P < .01$. CTL, control; SSC, side scatter.

for cytarabine-resistant AML cell lines in vitro (KG-1 and THP-1 cells; our results and other studies by Bossis et al[19] and Ma et al[56]) as well as in a therapy-resistant in vivo model (MOLM-14 orthotopic xenograft). TAK-981 has also decreased the expression of *CD39*, whose expression is mediated by SUMOylation.[57] *CD39* has been known to be overexpressed in both cytarabine-resistant AML cells and residual AML cells in patients after chemotherapy.[37] Enhancing *CD39* expression provoked resistance against cytarabine, whereas inhibiting it improved the response to cytarabine in AML cells.[37] These results might explain TAK-981's strong activity against cells with high IC_{50} values for cytarabine (>100 nM), such as KG-1, THP-1, and MOLM-14. Considering the different modes of action between TAK-981 and cytarabine and the differences in cell lines and primary cells, it will be interesting to see if their potency difference is maintained in real patient cases. Still, the different mode of action might explain the strong synergy of TAK-981 with current drugs in several settings shown in our study.

It is worthwhile to note that the IC_{50} values of both TAK-981 and cytarabine for the primary cells were much higher than those for the AML cell lines. With the lower SUMOylation status of primary AML cells than that of the cell lines (supplemental Figure 7) being one explanation, an important consideration is that primary AML cells

grow much slower than the established AML cell lines. It is possible that the high IC_{50} value of TAK-981 in primary AML cells may be because of the lower frequency of cell division. This is clearly the case with cytarabine that it almost completely lost its activity for the primary AML cells, even though it is a standard of care drug. Therefore, the absolute value of the IC_{50} may not be directly translated into high in vivo toxicity. We believe the much slower proliferation of the primary AML cells should be considered seriously, and, therefore, a correlation analysis between SUMOylation extent and cytotoxicity across primary AML cells and cell lines might not be conclusive.

Overall, the current study provides strong evidence for SUMOylation as a new targetable pathway for AML, based on integrated bioinformatic screening and validations with in vitro, ex vivo, and in vivo preclinical AML models. For toxicity, the longer survival of TAK-981–treated mice indicates a favorable therapeutic index. Consistent with this, a previous study with TAK-981 showed a good toxicity property up to 40 mg/kg in mice.[33] In addition, healthy patients or patients with remission after therapy had lower SAE1/SAE2, the target of TAK-981, than patients with active AML (Figure 1E), suggesting possible selectivity of the drug. These favorable efficacy and toxicity data should prompt further studies for its optimal combination and transitions to clinical trials with AML.

Acknowledgments

This work was supported by the Basic Science Research Program through National Research Foundation of Korea grants funded by the Korean Government Ministry of Science and ICT (NRF-2018R1A3B1052328) (S.P.), (MSIT-2022R1A2C2004517) (B.-S.C.), and (NRF-2020R1I1A1A01073124) (J.-M.K.).

Authorship

Contribution: H.S.K., B.-R.K., T.T.P.D., Y.-J.K., H.S., J.-M.K., S.J., D.K., J.K., and Y.J.S. established the methodology, obtained the experimental data, performed the formal analysis, and reviewed the manuscript; J.-M.K. obtained funding, supervised the study, and reviewed the manuscript; H.-J.K. obtained funding, recruited patients, supervised the study, and reviewed the manuscript; and B.-S.C. and S.P. conceptualized the study, obtained funding, supervised the study, and wrote the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: H.S.K., 0009-0007-8229-3933; T.T.P.D., 0000-0002-5408-7442; Y.J.S., 0000-0002-6816-8067; H.-J.K., 0000-0003-4098-3366; B.-S.C., 0000-0002-4524-6616.

Correspondence: Byung-Sik Cho; Leukemia Research Institute and Department of Hematology, Catholic Hematology Hospital, Seoul St. Mary's Hospital, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul 06591, South Korea; email: cbscho@catholic.ac.kr; and Sunghyouk Park; Natural Products Research Institute, College of Pharmacy, Seoul National University, 1 Gwanak-Ro, Gwanak-gu, Seoul 08826, South Korea; email: psh@snu.ac.kr.

References

1. Döhner H, Weisdorf DJ, Bloomfield CD. Acute myeloid leukemia. *N Engl J Med*. 2015;373(12):1136-1152.
2. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424-447.
3. Kantarjian H, Kadia T, DiNardo C, et al. Acute myeloid leukemia: current progress and future directions. *Blood Cancer J*. 2021;11(2):41.
4. Kayser S, Levis MJ. Updates on targeted therapies for acute myeloid leukaemia. *Br J Haematol*. 2022;196(2):316-328.
5. Kroonen JS, Vertegaal ACO. Targeting SUMO signaling to wrestle cancer. *Trends Cancer*. 2021;7(6):496-510.
6. Hay RT. SUMO: a history of modification. *Mol Cell*. 2005;18(1):1-12.
7. Eifler K, Cuijpers SAG, Willemstein E, et al. SUMO targets the APC/C to regulate transition from metaphase to anaphase. *Nat Commun*. 2018;9(1):1119.
8. Nacerddine K, Lehembre F, Bhaumik M, et al. The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Dev Cell*. 2005;9(6):769-779.
9. Flotho A, Melchior F. Sumoylation: a regulatory protein modification in health and disease. *Annu Rev Biochem*. 2013;82:357-385.
10. Kim D-H, Lee C, Kim B, Lee S-H, Han K-H. Rescuing p53 from mdm2 by a pre-structured motif in intrinsically unfolded SUMO specific protease 4. *BMB Rep*. 2017;50(10):485-486.
11. Ding B, Sun Y, Huang J. Overexpression of SKI oncoprotein leads to p53 degradation through regulation of MDM2 protein sumoylation. *J Biol Chem*. 2012;287(18):14621-14630.
12. Hoellein A, Fallahi M, Schoeffmann S, et al. Myc-induced SUMOylation is a therapeutic vulnerability for B-cell lymphoma. *Blood*. 2014;124(13):2081-2090.
13. Kessler JD, Kahle KT, Sun T, et al. A SUMOylation-dependent transcriptional subprogram is required for Myc-driven tumorigenesis. *Science*. 2012;335(6066):348-353.
14. Rabellino A, Melegari M, Tompkins VS, et al. PIAS1 promotes lymphomagenesis through MYC upregulation. *Cell Rep*. 2016;15(10):2266-2278.
15. González-Prieto R, Cuijpers SA, Kumar R, Hendriks IA, Vertegaal AC. c-Myc is targeted to the proteasome for degradation in a SUMOylation-dependent manner, regulated by PIAS1, SENP7 and RNF4. *Cell Cycle*. 2015;14(12):1859-1872.

16. Hannoun Z, Maarifi G, Chelbi-Alix MK. The implication of SUMO in intrinsic and innate immunity. *Cytokine Growth Factor Rev.* 2016;29:3-16.

17. Dong S, Chen J. SUMOylation of sPRDM16 promotes the progression of acute myeloid leukemia. *BMC Cancer.* 2015;15(1):893.

18. Zhang J, Huang F-F, Wu D-S, et al. SUMOylation of insulin-like growth factor 1 receptor, promotes proliferation in acute myeloid leukemia. *Cancer Lett.* 2015;357(1):297-306.

19. Bossis G, Sarry J-E, Kifagi C, et al. The ROS/SUMO axis contributes to the response of acute myeloid leukemia cells to chemotherapeutic drugs. *Cell Rep.* 2014;7(6):1815-1823.

20. Langston SP, Grossman S, England D, et al. Discovery of TAK-981, a first-in-class inhibitor of SUMO-activating enzyme for the treatment of cancer. *J Med Chem.* 2021;64(5):2501-2520.

21. Lam V, Best SR, Bruss N, et al. Pharmacologic inhibition of SUMO-activating enzyme (SAE) with TAK-981 augments interferon signaling and regulates T cell differentiation in ex vivo studies of chronic lymphocytic leukemia (CLL). *Blood.* 2019;134(suppl 1):1760.

22. Cancer Genome Atlas Research Network, Ley TJ, Miller C, Ding L, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med.* 2013;368(22):2059-2074.

23. Haferlach T, Kohlmann A, Wieczorek L, et al. Clinical utility of microarray-based gene expression profiling in the diagnosis and subclassification of leukemia: report from the International Microarray Innovations in Leukemia Study Group. *J Clin Oncol.* 2010;28(15):2529-2537.

24. Tyner JW, Tognon CE, Bottomly D, et al. Functional genomic landscape of acute myeloid leukaemia. *Nature.* 2018;562(7728):526-531.

25. Roushangar R, Mias GI. Multi-study reanalysis of 2,213 acute myeloid leukemia patients reveals age-and sex-dependent gene expression signatures. *Sci Rep.* 2019;9(1):12413.

26. Yoon S, Kim J, Kim S-K, et al. GScluster: network-weighted gene-set clustering analysis. *BMC Genomics.* 2019;20(1):352.

27. Derenzini E, Rossi A, Treré D. Treating hematological malignancies with drugs inhibiting ribosome biogenesis: when and why. *J Hematol Oncol.* 2018; 11(1):75.

28. Skrtić M, Sriskanthadevan S, Jhas B, et al. Inhibition of mitochondrial translation as a therapeutic strategy for human acute myeloid leukemia. *Cancer Cell.* 2011;20(5):674-688.

29. Reed GA, Schiller GJ, Kambhampati S, et al. A phase 1 study of intravenous infusions of tigecycline in patients with acute myeloid leukemia. *Cancer Med.* 2016;5(11):3031-3040.

30. Straube J, Ling VY, Hill GR, Lane SW. The impact of age, NPM1mut, and FLT3ITD allelic ratio in patients with acute myeloid leukemia. *Blood.* 2018; 131(10):1148-1153.

31. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics.* 2013;14(1):7-15.

32. Chou T-C. Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacol Rev.* 2006;58(3):621-681.

33. Lightcap ES, Yu P, Grossman S, et al. A small-molecule SUMOylation inhibitor activates antitumor immune responses and potentiates immune therapies in preclinical models. *Sci Transl Med.* 2021;13(611):eaba7791.

34. Baik H, Boulanger M, Hosseini M, et al. Targeting the SUMO pathway primes all-trans retinoic acid–induced differentiation of nonpromyelocytic acute myeloid Leukemias. *Cancer Res.* 2018;78(10):2601-2613.

35. Nakayama F, Nishihara S, Iwasaki H, et al. CD15 expression in mature granulocytes is determined by α1, 3-fucosyltransferase IX, but in promyelocytes and monocytes by α1, 3-fucosyltransferase IV. *J Biol Chem.* 2001;276(19):16100-16106.

36. Geletu M, Balkhi MY, Peer Zada AA, et al. Target proteins of C/EBPαp30 in AML: C/EBPαp30 enhances sumoylation of C/EBPαp42 via up-regulation of Ubc9. *Blood.* 2007;110(9):3301-3309.

37. Aroua N, Boet E, Ghisi M, et al. Extracellular ATP and CD39 activate cAMP-mediated mitochondrial stress response to promote cytarabine resistance in acute myeloid leukemia. *Cancer Discov.* 2020;10(10):1544-1565.

38. Daniel MT, Koken M, Romagne O, et al. PML protein expression in hematopoietic and acute promyelocytic leukemia cells. *Blood.* 1993;82(6): 1858-1867.

39. Lallemand-Breitenbach V, Jeanne M, Benhenda S, et al. Arsenic degrades PML or PML–RARα through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nat Cell Biol.* 2008;10(5):547-555.

40. Collins SJ. The role of retinoids and retinoic acid receptors in normal hematopoiesis. *Leukemia.* 2002;16(10):1896-1905.

41. Johnson DE, Redner RL. An ATRActive future for differentiation therapy in AML. *Blood Rev.* 2015;29(4):263-268.

42. Lübbert M, Grishina O, Schmoor C, et al. Valproate and retinoic acid in combination with decitabine in elderly nonfit patients with acute myeloid leukemia: results of a multicenter, randomized, 2× 2, phase II trial. *J Clin Oncol.* 2020;38(3):257-270.

43. Chen X, Qin Y, Zhang Z, et al. Hyper-SUMOylation of ERG is essential for the progression of acute myeloid leukemia. *Front Mol Biosci.* 2021;8:652284.

44. Gåtel P, Brockly F, Reynes C, et al. Ubiquitin and SUMO conjugation as biomarkers of acute myeloid leukemias response to chemotherapies. *Life Sci Alliance.* 2020;3(6):e201900577.

45. Seeler J-S, Dejean A. SUMO and the robustness of cancer. *Nat Rev Cancer.* 2017;17(3):184-197.

46. Zhou P, Chen X, Li M, et al. 2-D08 as a SUMOylation inhibitor induced ROS accumulation mediates apoptosis of acute myeloid leukemia cells possibly through the deSUMOylation of NOX2. *Biochem Biophys Res Commun.* 2019;513(4):1063-1069.

47. Zhou Q, Zhang L, Chen Z, et al. Small ubiquitin-related modifier-1 modification regulates all-trans-retinoic acid-induced differentiation via stabilization of retinoic acid receptor α. *FEBS J.* 2014;281(13):3032-3047.

48. Daver N. Immune checkpoint inhibitors in acute myeloid leukemia. *Best Pract Res Clin Haematol.* 2021;34(1):101247.

49. Reville PK, Kantarjian HM, Ravandi F, et al. Nivolumab maintenance in high-risk acute myeloid leukemia patients: a single-arm, open-label, phase II study. *Blood Cancer J.* 2021;11(3):60.

50. Berger R, Rotem-Yehudar R, Slama G, et al. Phase I safety and pharmacokinetic study of CT-011, a humanized antibody interacting with PD-1, in patients with advanced hematologic malignancies. *Clin Cancer Res.* 2008;14(10):3044-3051.

51. Daver N, Boddu P, Garcia-Manero G, et al. Hypomethylating agents in combination with immune checkpoint inhibitors in acute myeloid leukemia and myelodysplastic syndromes. *Leukemia.* 2018;32(5):1094-1105.

52. Wang L-X, Mei Z-Y, Zhou J-H, et al. Low dose decitabine treatment induces CD80 expression in cancer cells and stimulates tumor specific cytotoxic T lymphocyte responses. *PLoS One.* 2013;8(5):e62924.

53. Sallman D, Asch A, Kambhampati S, et al. The first-in-class anti-CD47 antibody magrolimab combined with azacitidine is well-tolerated and effective in AML patients: phase 1b results. *Blood.* 2020;136(suppl 1):330.

54. Zeidan AM, Cavenagh J, Voso MT, et al. Efficacy and safety of azacitidine (AZA) in combination with the anti-PD-L1 durvalumab (durva) for the front-line treatment of older patients (pts) with acute myeloid leukemia (AML) who are unfit for intensive chemotherapy (IC) and pts with higher-risk myelodysplastic syndromes (HR-MDS): results from a large, international, randomized phase 2 study. *Blood.* 2019;134(suppl 1):829.

55. Kapoor S, Champion G, Basu A, Mariampillai A, Olnes MJ. Immune therapies for myelodysplastic syndromes and acute myeloid leukemia. *Cancers (Basel).* 2021;13(19):5026.

56. Ma J, Li X, Su Y, et al. Mechanisms responsible for the synergistic antileukemic interactions between ATR inhibition and cytarabine in acute myeloid leukemia cells. *Sci Rep.* 2017;7(1):41950.

57. Ding X, Wang A, Ma X, et al. Protein SUMOylation is required for regulatory T cell expansion and function. *Cell Rep.* 2016;16(4):1055-1066.

Received: 30 March 2023 | Revised: 5 June 2023 | Accepted: 13 June 2023

DOI: 10.1002/ctm2.1311

CLINICAL AND TRANSLATIONAL MEDICINE WILEY

LETTER TO THE EDITOR

SCD and MTHFD2 inhibitors for high-risk acute myeloid leukaemia patients, as suggested by ELN2017-pathway association

Dear Editor,

The European LeukemiaNet (ELN) 2017 criteria[1] is widely accepted as the risk classification of acute myeloid leukaemia (AML) patients. However, their application to studying risk-related biological pathways is limited, failing to enhance treatment options for high-risk patients. Using multiomics databases, biological pathways and genes whose upregulations correlate with increased ELN2017 risks and poorer survival were investigated, followed by experimental and functional validation, exhibiting SCD and MTHFD2 inhibitors as potential therapeutics for treating high-risk AML patients.

As shown in Figure 1A, we initially evaluated ELN2017 and its revision[2] for prognostic risk categorisation in OHSU patient data.[3] Survival prognoses showed significant disparity among risk groups (Figures 1B and S1). Specifically, the 'Very Adverse' group exhibited a much poorer prognosis than other groups, consistent with a previous report.[2] However, the 'Very Favourable' and 'Favourable' groups did not reveal significant differences (Table S1). This analysis shows four distinguished survival risk groups existing in OHSU data.

We then performed correlational screening between revised ELN2017 risks and pathway scores (FDR < 0.05), resulting in 690 pathways. Further refinement based on disease-specific survival (DSS) criteria (FDR < 0.05, hazard ratio (HR) > 3, and FDR of HR < 0.05) resulted in 34 pathways. These pathways were categorised into three distinct clusters and three unclustered pathways (Figure 1C). The largest cluster (26 pathways) was 'cell-cycle related'; considering 71.2% (146/205) of the samples were from the initial diagnosis stage, a connection between the proliferative state at diagnosis and patient prognosis is suggested. As relapsed AML cells after chemotherapy exhibit higher dormancy[4] and leukaemia stem cells often remain in a quiescent state,[5] it will be an interesting future topic to specifically compare the relationship between the prognosis and the cell cycle progression at initial versus late-stage

TABLE 1 Risk-correlation analysis of revised ELN2017 for pathways using multiomics databases.

	OHSU						
Pathway	J-T test	Survival (DSS)	HR for DSS	Survival (OS)	HR for OS	TCGA-LAML J-T test	Proteomics J-T test
KEGG_BIOSYNTHESIS_OF_UNSATURATED_FATTY_ACIDS ('UFA_Synthesis')	< 0.001	< 0.001	3.15	< 0.001	2.25	0.007	0.013
WP_OMEGA3OMEGA6_FA_SYNTHESIS	< 0.001	< 0.001	3.12	< 0.001	2.86	0.196	0.036
WP_OMEGA9_FA_SYNTHESIS	< 0.001	< 0.001	4.08	< 0.001	3.43	0.068	0.067
REACTOME_CHOLESTEROL_BIOSYNTHESIS	0.003	0.001	3.16	0.021	1.66	0.466	0.195
WP_MEVALONATE_PATHWAY	0.004	< 0.001	5.06	< 0.001	3.04	0.622	0.471
REACTOME_METABOLISM_OF_FOLATE_AND_PTERINES ('Folate_Metabolism')	< 0.001	< 0.001	4.68	< 0.001	2.35	0.006	0.001
REACTOME_SIGNALING_BY_MST1	< 0.001	< 0.001	3.22	< 0.001	2.81	< 0.001	NA
WP_NO_METABOLISM_IN_CYSTIC_FIBROSIS	0.002	< 0.001	3.28	0.006	2.30	< 0.001	0.220

J-T test, Jonckheere-Terpstra test; DSS, disease-specific survival; OS, overall survival; HR, the hazard ratio of the high pathway score group; NA, not available.

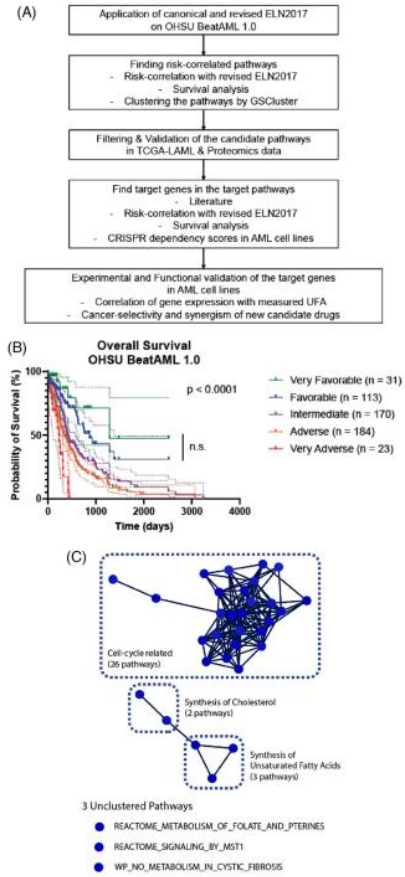


FIGURE 1 Overview of the study, application of revised ELN2017 to the OHSU database, and clustering for risk-correlated biological pathways. (A) Overall scheme of the study. (B) Kaplan–Meier curves with 95% confidence intervals (dotted lines) for overall survival of AML patients in the OHSU BeatAML 1.0 database, according to the revised ELN2017 criteria. *p* value is from the log-rank test. n.s. refers to not significant. Complete comparisons among the groups are in Table S1. (C) Clustering result of 34 pathways filtered by risk-correlation with revised ELN2017 and survival analysis (see main text for details).

AML. Since the 'cell-cyclerelated' cluster already encompasses a standard regimen drug, cytarabine, we focused on other pathways (Figures 2A–D and S2–S4; Table 1).

Our correlations were further studied in two other large AML databases (TCGA-LAML[6] and proteomics database[7]) for eight pathways outside the 'cell-cyclerelated' cluster. Patients were categorised into four groups using the revised ELN2017 as above. For the proteomics database, we leveraged Gene Set Variation Analysis (GSVA), originally used for transcriptomic data, to generate pathway scores. Combining analyses in all three databases revealed significant relationship between revised ELN2017 risk groups and 'KEGG_BIOSYNTHESIS_OF_UNSATURATED_FATTY_ACIDS' ('UFA_Synthesis') and 'REACTOME_METABOLISM_OF_FOLATE_AND_PTERINES' ('Folate_Metabolism') (Table 1).

After identifying risk-correlated pathways, we examined targetable genes within those pathways. MTHFD2, a previously reported AML target,[8,9] was investigated for the Folate_Metabolism pathway. *MTHFD2* gene expression exhibited a significant increasing trend across ELN2017 risk groups (Figure 2E), and patients with higher gene expression experienced significantly shorter overall survival (Figure 2F). Clinical trials of methotrexate, an antifolate drug, were unsuccessful in AML due to reduced polyglutamylation activity, which is essential for its effectiveness.[10] Therefore, alternative drugs targeting this pathway seem necessary.

For the UFA_Synthesis pathway, we analysed all 22 genes in the pathway in OHSU, TCGA-LAML, and proteomics databases since little has been known for this pathway (Table S2). *ACOT7* (Figure S5) and *SCD* (Figure 2G and H) emerged as candidates. Among these, *SCD* exhibited the highest vulnerability upon knockout in AML cells (lowest median dependency score in the Depmap database; Figure S6). Therefore, we selected *SCD* as our target. Notably, while *SCD* level was higher in the 'Normal' group than in the 'Favourable' group, there was a clear and significant upward trend of *SCD* expression correlating with worse ELN2017 criteria (Figure 2G). This provides rationale for *SCD* as a target, aligning with our goal of finding targets for high-risk AML patients.

With *MTHFD2* and *SCD* bioinformatically suggested as risk-associated genes, we experimentally validated them using five AML (U937, MOLM-14, THP-1, KG-1 and HL-60) and one normal (HCC1954-BL) cell line. CCK-8 assay confirmed MTHFD2 inhibitor DS18561882 exhibited higher potency against all five AML cell lines compared to normal cell line (Figure 3A). SCD inhibitor A939572 also showed ~8 times higher IC_{50}s for normal cell line (Figure 3B). Trypan blue assay, along with normal peripheral blood mononuclear cells (PBMCs), also demonstrated the two inhibitors' cancer cell selectivity (Figure S7).

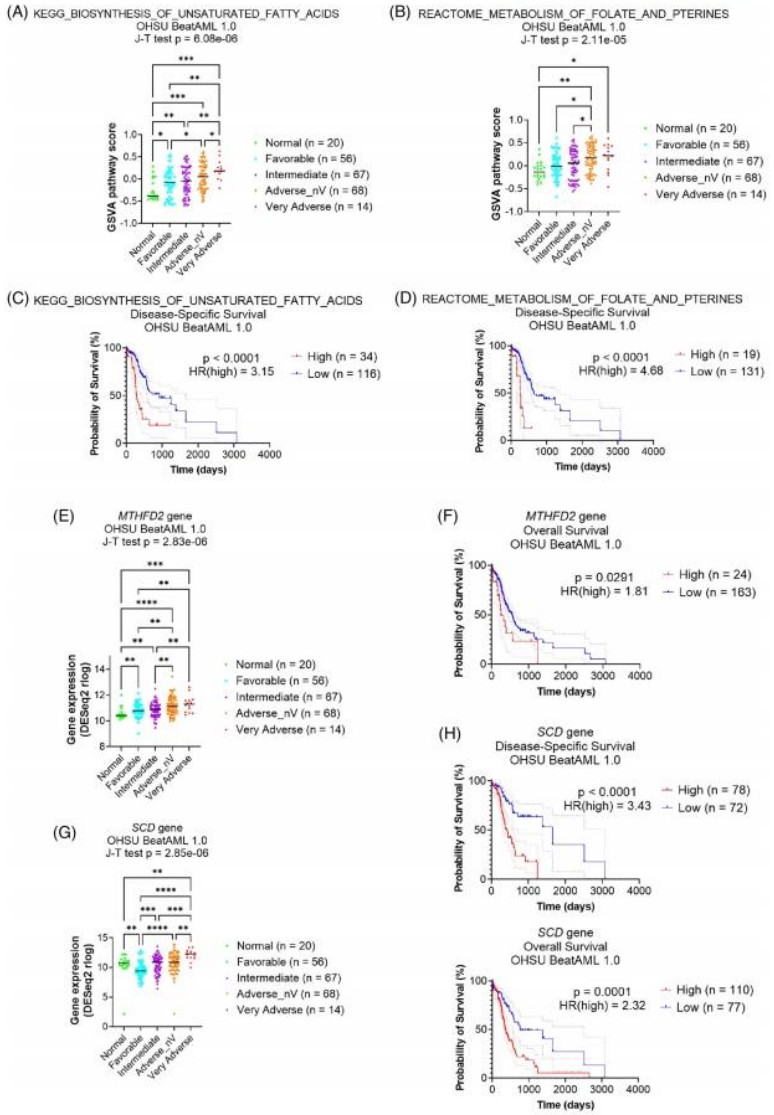


FIGURE 2 Identification of risk-correlated biological pathways. The distributions of GSVA pathway scores of (A) KEGG_BIOSYNTHESIS_OF_UNSATURATED FATTY_ACIDS pathway or (B) REACTOME_METABOLISM_OF_FOLATE_AND_PTERINES pathway, in each risk category of revised ELN2017 in OHSU BeatAML 1.0 database. Kaplan–Meier curves with 95% confidence intervals (dotted lines) for disease-specific survival of AML patients in OHSU BeatAML 1.0 database, for (C) KEGG_BIOSYNTHESIS_OF_UNSATURATED FATTY_ACIDS pathway or (D) REACTOME_METABOLISM_OF_FOLATE_AND_PTERINES pathway. The distributions of gene expression of (E) *MTHFD2* gene or (G) *SCD* gene in each risk category of revised ELN2017 in OHSU BeatAML 1.0 database. Kaplan–Meier curves with 95% confidence intervals (dotted lines) for (F) overall survival of AML patients in the OHSU BeatAML 1.0 database for *MTHFD2* gene or (H) disease-specific survival and overall survival of AML patients in OHSU BeatAML 1.0 database for *SCD* gene. For (A), (B), (E) and (G), 'Adverse_nV' refers to the patients in the 'Adverse' category but not in the 'Very Adverse' category. The black lines indicate medians for each group. p values are from the Jonckheere-Terpstra test. Post hoc analyses were performed with a two-stage linear step-up procedure. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$, $^{****}p < .0001$. For (C), (D), (F) and (H), p values are from the log-rank test. The stratification of two groups in each graph was based on the best risk separation approach. HR(high) refers to the hazard ratio of the group with high pathway scores or gene expression.

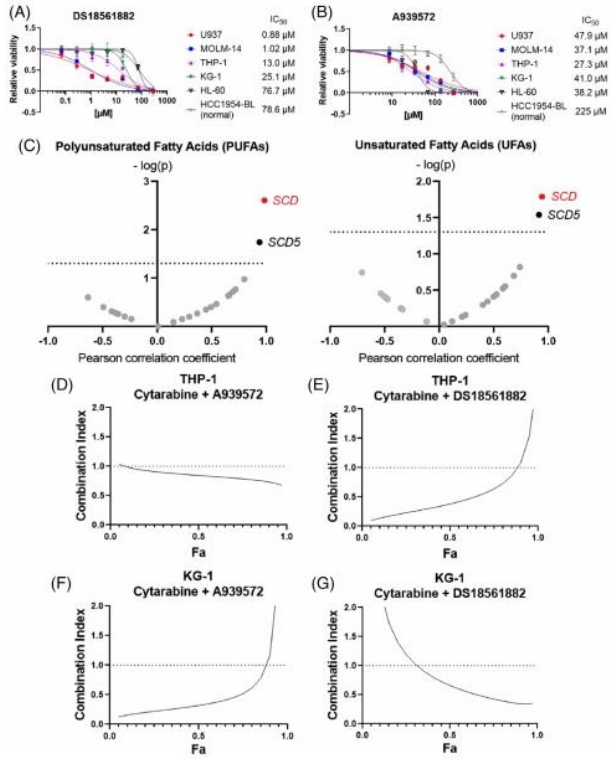


FIGURE 3 Experimental and functional validation of *SCD* and *MTHFD2* genes. Dose–response curves and IC_{50}s for (A) DS18561882 drug and (B) A939572 drug to five AML cell lines (U937, MOLM-14, THP-1, KG-1 and HL-60) and one normal cell line (HCC1954-BL) by CCK-8 assay. The drugs were treated for 48 h. (C) Volcano plots for PUFAs and UFAs in the 5 AML cell lines identical to (A) and (B). The *x*-axis refers to the Pearson correlation coefficient calculated by correlating gene expression of five AML cell lines in Depmap with measured PUFAs or UFAs by NMR for the genes only in the 'UFA_Synthesis' pathway. The *y*-axis refers to the *p* value of the Pearson correlation coefficient. Nonsignificant (*p* value .05 or higher) genes are indicated in grey. Combination index plots for THP-1 cell line for the combination of cytarabine with (D) A939572 or (E) DS18561882. Combination index plots for KG-1 cell line for the combination of cytarabine with (F) A939572 or (G) DS18561882. For (D)–(G), the combination index of less than 1 indicates synergy and Fa refers to fractions affected by particular dose of a drug, herein the fraction of dead cells compared to nontreated samples.

Additionally, SCD and MTHFD2 proteins' expression was higher in AML cell lines than normal PBMCs (Figure S8). However, no significant correlations were observed among AML cell lines (Figure S9). DS18561882 treatment reduced MTHFD2 protein levels (Figure S10A), and A939572 treatment reduced unsaturated fatty acid levels without affecting SCD protein levels (Figure S10B and C), confirming the targeted effects of both drugs on the target proteins. Due to limited references regarding SCD in AML, we further validated its functional relevance. Unsaturated fatty acids (PUFA and UFA) were measured in AML cell lines using nuclear magnetic resonance (NMR). Among the 22 genes, *SCD* highly correlated with unsaturated fatty acids (Figures 3C and S11; Table S3). Our findings suggest the proposed inhibitors' selectivity toward AML cells, indicating their potential use for high-risk AML groups.

The 'cell-cyclerelated' cluster was found significantly correlated with AML-risk groups. Cytarabine, a DNA replication inhibitor and standard-of-care drug for AML, is also used to high-risk AML patients, along with hypomethylating agents and venetoclax regimen. When tested on cell lines, it exhibited a large variation in sensitivity (Figure S12). Combining A939572 or DS18561882 with cytarabine for cells with high cytarabine IC_{50} (THP-1 and KG-1; Figure S13) resulted in synergistic inhibition of cell survival (Figure 3D–G). In addition, dose reduction of cytarabine was observed in all four combinations (Table S4), suggesting the potential of A939572, DS18561882 or their derivatives in alleviating cytarabine toxicity.

Overall, our study identified risk-associated pathways, target genes and potential synergistic drugs with cytarabine in AML through novel bioinformatic analysis on large multiomic datasets. Since not much is known about the roles of unsaturated fatty acids or folate metabolism in AML, our results could be further exploited to find a mechanistic relationship between those pathways and the malignancy of AML. More detailed discussion, including the significant advantage of our approach to finding new target genes/pathways (Table S5) or even in solid tumours, is in the Supplementary Information.

AUTHOR CONTRIBUTIONS

HSK, DK and JK established the methodology, obtained the experimental data, performed the formal analysis and reviewed the manuscript. SP and ACVG conceptualised the study, obtained funding, supervised the study and wrote the manuscript.

FUNDING

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea grants funded by the Korean government Ministry of Science and ICT (NRF-2018R1A3B1052328 to).

CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflicts of interest.

DATA AVAILABILITY STATEMENT

For the original data, please contact arviecamille@snu.ac.kr.

Han Sun Kim
Doyeon Kim
Jiwoo Kim
Sunghyouk Park
Arvie Camille V. de Guzman

Natural Products Research Institute, College of Pharmacy, Seoul National University, Seoul, South Korea

Correspondence
Sunghyouk Park and Arvie Camille V. de Guzman, Natural Products Research Institute, College of Pharmacy, Seoul National University, 1 Gwanak-Ro, Gwanak-gu, Seoul 08826, South Korea.
Email: psh@snu.ac.kr; arviecamille@snu.ac.kr

ORCID

Han Sun Kim https://orcid.org/0009-0007-8229-3933
Sunghyouk Park https://orcid.org/0000-0003-1981-3274
Arvie Camille V. de Guzman https://orcid.org/0000-0003-0962-4824

REFERENCES

1. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424-447.
2. Herold T, Rothenberg-Thurley M, Grunwald VV, et al. Validation and refinement of the revised 2017 European LeukemiaNet genetic risk stratification of acute myeloid leukemia. *Leukemia*. 2020;34(12):3161-3172.
3. Tyner JW, Tognon CE, Bottomly D, et al. Functional genomic landscape of acute myeloid leukaemia. *Nature*. 2018;562(7728):526-531.
4. Niu J, Peng D, Liu L. Drug resistance mechanisms of acute myeloid leukemia stem cells. *Front Oncol*. 2022;12:896426.
5. Vetrie D, Helgason GV, Copland M. The leukaemia stem cell: similarities, differences and clinical prospects in CML and AML. *Nat Rev Cancer*. 2020;20(3):158-173.
6. The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*. 2013;368(22):2059-2074.
7. Kramer MH, Zhang Q, Sprung R, et al. Proteomic and phosphoproteomic landscapes of acute myeloid leukemia. *Blood*. 2022;140(13):1533-1548.
8. Pikman Y, Puissant A, Alexe G, et al. Targeting MTHFD2 in acute myeloid leukemia. *J Exp Med*. 2016;213(7):1285-1306.
9. Bonagas N, Gustafsson NMS, Henriksson M, et al. Pharmacological targeting of MTHFD2 suppresses acute myeloid leukemia by inducing thymidine depletion and replication stress. *Nat Cancer*. 2022;3(2):156-172.
10. Lin JT, Tong WP, Trippett TM, et al. Basis for natural resistance to methotrexate in human acute non-lymphocytic leukemia. *Leuk Res*. 1991;15(12):1191-1196.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.